Journal of Heredity, 2020, 392–404
doi:10.1093/jhered/esaa014
Original Article
Advance Access publication June 3, 2020

Original Article

Mitogenomes Reveal Multiple Colonization of Mountains by *Rattus* in Sundaland

Miguel Camacho-Sanchez and Jennifer A. Leonard

Conservation and Evolutionary Genetics Group, Estación Biológica de Doñana (EBD-CSIC), Av. Américo Vespucio 26, 41092, Sevilla, Spain (Camacho-Sanchez and Leonard)

Address correspondence to Jennifer A. Leonard at the address above, or e-mail: jleonard@ebd.csic.es

Received August 21, 2019; First decision December 19, 2019; Accepted May 25, 2020.

Corresponding Editor: Mark Springer

Abstract

Tropical mountains are cradles of biodiversity and endemism. Sundaland, tropical Southeast Asia, hosts 3 species of *Rattus* endemic to elevations above 2000 m with an apparent convergence in external morphology: *Rattus korinchi* and *R. hoogerwerfi* from Sumatra, and *R. baluensis* from Borneo. A fourth one, *R. tiomanicus*, is restricted to lowland elevations across the whole region. The origins of these endemics are little known due to the absence of a robust phylogenetic framework. We use complete mitochondrial genomes from the 3 high altitude *Rattus*, and several related species to determine their relationships, date divergences, reconstruct their history of colonization, and test for selection on the mitochondrial DNA. We show that mountain colonization happened independently in Borneo (<390 Kya) and Sumatra (~1.38 Mya), likely from lowland lineages. The origin of the Bornean endemic *R. baluensis* is very recent and its genetic diversity is nested within the diversity of *R. tiomanicus*. We found weak evidence of positive selection in the high-elevation lineages and attributed the greater nonsynonymous mutations on these branches (specially *R. baluensis*) to lesser purifying selection having acted on the terminal branches in the phylogeny.

Subject areas: Molecular systematics and phylogenetics, Population structure and phylogeography
Keywords: selection, tropical mountain, adaptation, Rattini, endemism

Mountains are world biodiversity hotspots (Perrigo et al. 2019). Their greater endemism at higher elevations may be explained by topographic isolation (Steinbauer et al. 2016). The biogeographical region of Sundaland, in tropical Southeast Asia, is one of the most biodiverse of the world hotspots (Myers et al. 2000), where a considerable proportion of its biodiversity is associated with mountains (Merckx et al. 2015; Sheldon et al. 2015). This area is part of the Indo-Pacific, a center of diversification for murines. They are an interesting model to study evolution because in their radiation across the Indo-Pacific they have occupied diverse habitats across the different archipelagos (Fabre et al. 2013; Rowe et al. 2019), developing many cases of remarkable adaptive changes, such as those associated to carnivory (Fabre et al. 2017; Martinez et al. 2018).

Sundaland is home to 3 species of *Rattus* endemic to high elevations, above 2000 m: *Rattus baluensis* Thomas, 1894, only known from Sabah, northern Borneo (Musser 1986; Camacho-Sanchez et al. 2018), *R. korinchi* Robinson and Kloss, 1916; and *R. hoogerwerfi* Chasen, 1939, from Sumatra (Robinson and Kloss 1918, 1919; Miller 1942; Musser and Newcomb 1983; Musser 1986; Musser and Carleton 2005) (Figure 1). A fourth native *Rattus* species in Sundaland, *R. tiomanicus* Miller, 1900, is restricted to lowlands across this whole region. The 3 montane species inhabit similar habitats and share similarities in their external morphology, mainly long dark fur with woolly underfur, which has been suggested to be adaptive to cold montane environments (Musser 1986) (Supplementary Figure S1). These external similarities have misled

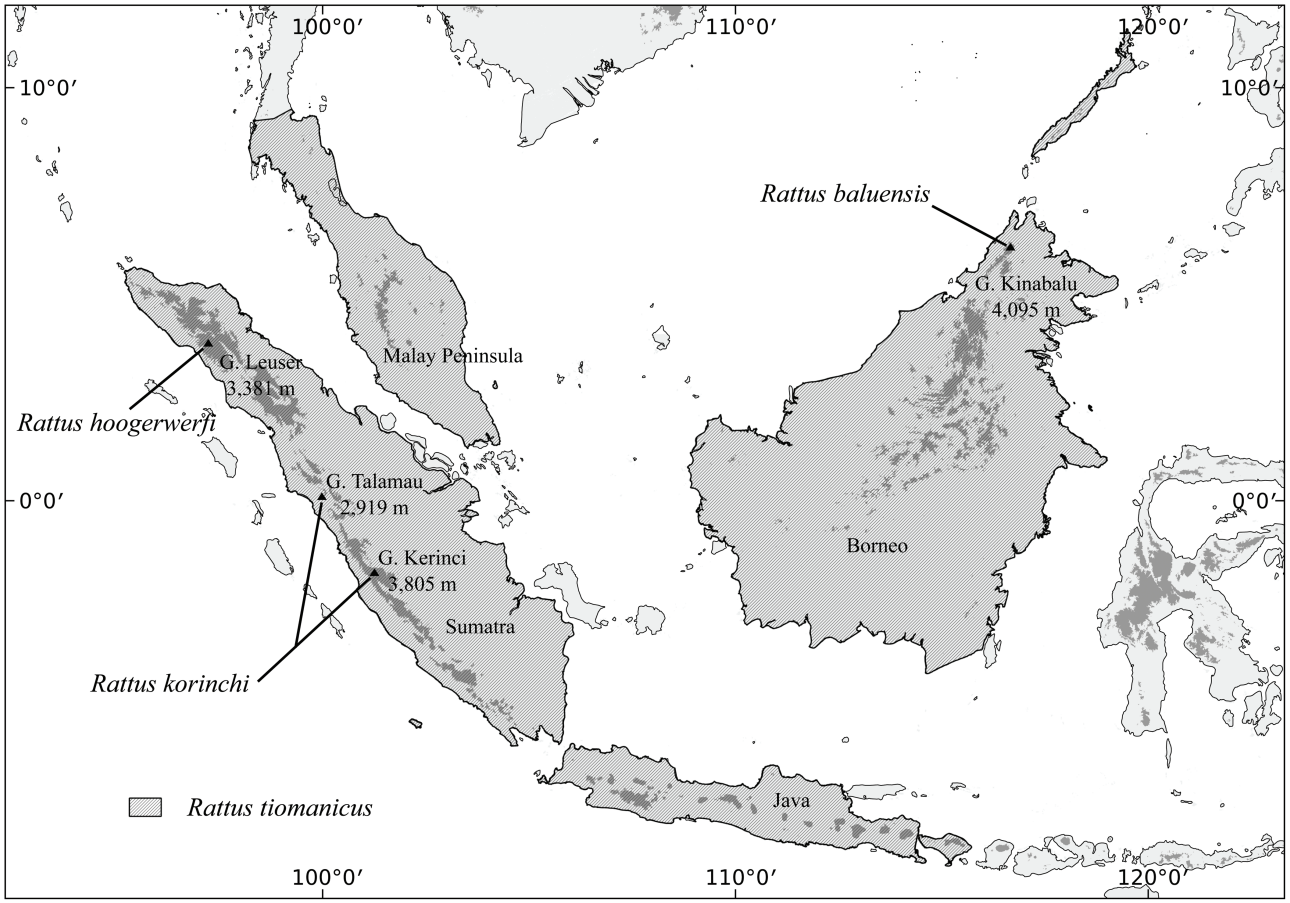

Figure 1. Distribution of the 4 *Rattus* species native to Sundaland. Dark gray corresponds to elevations above 1000 m. Data for *R. tiomanicus* were downloaded from the IUCN (2016).

taxonomists; until 1986, *R. korinchi* was considered a subspecies of *Rattus baluensis* due to their external resemblance (Musser 1986). The systematics of *Rattus* is complex. Even after a profound morphological review of Sundaland *Rattus*, the origin and position of these lineages within *Rattus* remain unresolved (Musser 1986). The evolutionary affinities at the molecular level between some of these linages and to other *Rattus* have been assessed with cytochrome b (*cyt b*) (Aplin et al. 2011; Thomson et al. 2018), providing a comprehensive framework for molecular systematics of *Rattus*. However, they missed important taxa native to Sundaland (e.g., *R. korinchi*), and they relied on *cyt b* alone, providing an incomplete evolutionary framework for *Rattus* native to Sundaland and their closest relatives in the Asian *Rattus* clade.

The origin of these mountain lineages has not been explored explicitly. A likely scenario is that montane endemics derive from local lowland ancestors. This has already been suggested for many taxa on Mount Kinabalu (Merckx et al. 2015), the tallest peak in Sundaland. In most occasions, it is difficult to guess the origin of the mountain lineages, which seem to fall into 2 categories: colonization from distant pre-adapted lineages or in situ speciation from lowlands (Merckx et al. 2015), although the sympatry of sister taxa is likely preceded by divergence in allopatry followed by secondary contact instead of "ecological speciation" along the elevational gradient (Moyle et al. 2017). In any case, demonstrating the polarity of the speciation, from lowland-to-highland habitats, or vice versa, is not trivial and can give clues about the origin of biodiversity in this region.

Robust phylogenies are essential to answer some of the questions about the polarity of these colorizations (Heath et al. 2008; Fabre et al. 2013; Stevens et al. 2019). Additionally, the physical conditions in mountain habitats (lower oxygen partial pressure, colder conditions) can exert important selection and drive adaptive genomic and physiological changes (Cheviron and Brumfield 2012), which can be investigated through the phylogeny. The mitochondrial genome contains 13 protein-coding genes that are part of the respiratory chain, producing most of the energy in mammals: 95% of adenosine triphosphate. The coding sequences have evolutionary constraints related to the metabolic requirements of different processes, such that the sequences of mitochondrial protein-coding genes have been identified as targets for adaptation to high elevations as less oxygen and colder conditions impose more stringent energy requirements (Fontanillas et al. 2005; Scott et al. 2011; Yu et al. 2011; Zhou et al. 2014).

We aimed to resolve the phylogenetic affinities of the montane *Rattus* native to Sundaland and use the resulting phylogeny to investigate the diversification to the mountain habitats in these lineages by assessing selection on the mitochondrial genome and reconstructing the distribution of the ancestral lineages.

Methods

Study System

We include all endemic Sundaland *Rattus* species recognized in Musser and Newcomb (1983), except for *R. annandalei*, which has

been moved to *Sundamys annandalei* (Camacho-Sanchez et al. 2017). At present, other lowland generalists in Sundaland, *R. argentiventer*, *R. rattus*, *R. norvegicus*, *R. tanezumi*, and *R. exulans*, are considered invasive. The *R. tiomanicus* complex includes several insular forms (*R. burrus*, *R. simalurensis*, *R. adustus*, *R. palmarum*, *R. mindorensis*, and *R. lugens*), which are recognized as different species (Musser and Carleton 2005). The molecular position has been evaluated for *R. burrus*, *R. mindorensis*, and *R. lugens*, which situates them within the Asian *Rattus* and close but outside *R. tiomanicus* (Thomson et al. 2018). Another species within the *R. tiomanicus* complex is *Rattus blangorum*. It is known from only 2 specimens (ANSP 20348 and 20349) from the Aceh region, northern Sumatra. Initially described as a singular species (Miller 1942), it was later placed in the *R. tiomanicus* complex (Musser and Califia 1982; Musser and Carleton 2005).

The three montane species (*R. baluensis*, *R. hoogerwerfi*, and *R. korinchi*) inhabit similar mountain habitat (Musser 1986). The summit rat, *R. baluensis*, is found above 2000 m in northern Borneo, from mossy forest, mountain scrubland, and up to subalpine vegetation in Kinabalu (Musser 1986). Musser and Carleton (2005) report 1524 m as its lowest elevation, probably from a misidentified museum specimen reported in Nor (2001). In support of this statement, the lower distribution of *R. baluensis* matches firmly the lower delineation of the cloud forest in Kinabalu Park at around 2000 m, as has been reported in wide surveys (Camacho-Sanchez et al. 2019). Musser (1986) also reports the lower limit of the range of *R. baluensis* to be at 7000 ft (2134 m), after extensive examination of museum specimens and literature review, and no other study has reported its presence below 2000 m. Korinch's rat, *Rattus korinchi*, is only known from the holotype collected on Mount (Mt.) Kerinci, at 2164 m (Robinson and Kloss 1918; BM 19.11.5.81, collector no. 442/14), and from a second specimen from montane or moss forest on Mt. Talamau, at 2773 m (Robinson and Kloss 1919; RMNH 23151, collector no. 351). Both of these mountains are in Sumatra and there are no other records for this species in museums (Musser 1986). The other Sumatran mountain endemic, the Hoogerwerf's rat, *R. hoogerwerfi*, is known from 29 specimens collected on Mt. Leuser, northern Sumatra, from 2133 to 2835 m, also inhabiting a similar montane habitat as its Kinabalu counterpart, described as "moss forest, with trees averaging only 5–40 feet in height, very hard, knotted, and twisted. Everywhere the ground and the branches of the trees were covered with a deep carpet of moss and ferns" (Miller 1942, p. 118) and also "mostly bare or covered with grass interspersed with patches of bushes and low trees" (Miller 1942, p. 108), and 7 specimens reported in Sody (1941, p. 300), also from the same area. The 800 m in elevation where the holotype was collected (Chasen 1939; also cited in Musser and Carleton 2005) is probably an error as this species is confined to higher elevations (Miller 1942). The high trapping success of *R. hoogerwerfi* compared with other small mammals suggests this species is present at high densities in its habitat, probably in the range of its Bornean counterpart *R. baluensis* on Mt. Kinabalu (Nor 2001; Camacho-Sanchez et al. 2019).

Taxonomic and Gene Sampling in the Phylogenetic Reconstruction With Mitogenomes

We sequenced complete mitochondrial genomes from modern and historical samples from the Sundaland montane species *R. korinchi* and *R. hoogerwerfi*, the lowland species *R. tiomanicus*, and other representative species of *Rattus* obtained from museum collections and the field (Table 1; Supplementary Table S1). One historical sample from the Sabah Museum labeled as *Lenothrix canus* (NH 2015) was reassigned to *R. tiomanicus* based on *cyt b* barcoding. The historical specimen NH 2147 labeled *Sundamys muelleri* was reassigned to *Rattus* sp. R3 sensu Pagès et al. (2010) based on *cyt b* barcoding (Table 1).

We also included 32 mitogenomes from *R. baluensis* (KY611359–KY611390) (Camacho-Sanchez et al. 2018) and other *Rattus* for which mitogenomes were available in GenBank (Australo-Papuan *Rattus*: *R. lutreolus* GU570661, *R. sordidus* GU570665, *R. pretor* NC 012461, *R. villosissimus* NC 014864, *R. tunneyi* NC 014861, *R. leucopus* GU570659, *R. niobe* KC152486, and *R. pretor* NC_012461; Asian *Rattus*: *R. tanezumi* EU273712, *R. rattus* NC_012374, *R. nitidus* KU200226, *R. exulans* EU273711, *R. norvegicus* AJ428514, and *R. fuscipes* NC_014867). As outgroups, we added some of the closest species from the *Rattus* division: *Sundamys muelleri* KY464175, *Bandicota indica* KT029807, and *Bunomys penitus* KY464167. Additional outgroups for dating were incorporated from 6 murines belonging to 2 molecular tribes in the *Mus* branch of the phylogeny (*Apodemus chejuensis* HM034867, *A. latronum* NC_019585, *A. peninsulae* NC_016060; *Mus cervicolor* KJ530560, *M. cookii* KJ530561, *M. spretus* NC_025952) (Fabre et al. 2013; Pagès et al. 2016).

DNA Extraction and Sequencing

We extracted DNA with DNeasy Blood and Tissue Kit (Qiagen). Museum tissue samples from historical specimens were processed in an isolated ancient DNA laboratory. Illumina libraries were constructed following a double indexing protocol with enrichment of complete mitochondrial genomes as in Camacho-Sanchez et al. (2017). They were sequenced on an Illumina HiSeq 2500 with 150 PE chemistry at the Genetics Resources Core Facility at John Hopkins University.

Mitogenome Assembly and Alignment

Adaptors were trimmed with cutadapt 1.8.3 (Martin 2011) using paired-end mode (details in github.com/csmiguel/rattus-highlands). Forward and reverse reads were paired with PEAR v0.9.6 (Zhang et al. 2014) using default parameters. The resulting assembled and unassembled forward and reverse reads were concatenated into a unique FastQ file for each sample. We mapped the reads from each sample to the circularized mitogenomes of *R. baluensis* KY611361 and *R. exulans* KJ530564 (for *R. exulans* BOR577 only), in Geneious 8 (Kearse et al. 2012) using default parameters and 3 iterations. SAMtools 1.3 (Li et al. 2009) was used to remove PCR duplicates. We called consensus sequences in Geneious with a minimum of 2× coverage and a 75% base calling threshold. Positions not passing this threshold were filled with ambiguities. As there were not long stretches of ambiguous positions and given the phylogenetic proximity of the reference, we assumed the lengths of these stretches with ambiguous positions to be the same as in the reference. Then, we used the MAFFT v7.017 (Katoh et al. 2002) plugin in Geneious for multiple sequence alignment using the *--auto* parameter. The alignments were visually inspected and the genes were translated into amino acids and inspected for stop codons in Geneious.

Phylogenetic Reconstructions and Molecular Dating

We used the protein-coding genes from the mitogenomes to evaluate the evolutionary relationships between the four Sundaland endemics and other *Rattus* species in a Maximum Likelihood framework with RAxML v8.2.10 (Stamatakis 2014).

Table 1. Field samples and museum specimens sequenced

Sample	*Rattus* species	Date collected	Tissue	Elev. (m)[a]	Locality	Lat/Lon	Collector
ANSP 20348	*R. blangorum*	4 April 1939	Old skin	1097	Sumatra: Mt. Leuser: Blangnanga camp	4.04, 97.13	F. A. Ulmer, Jr
BOR577[b]	*R. exulans*	16 March 2013	Fresh liver	357	Borneo: Sabah: Monggis substation	6.2, 116.75	Miguel C.
ANSP 20309	*R. hoogerwerfi*	27 April 1939	Old skin+dry tissue skull	2408	Sumatra: Mt. Leuser: Bivouac 5	3.87, 97.13	F. A. Ulmer, Jr
ANSP 20315	*R. hoogerwerfi*	5 May 1939	Old skin	2423	Sumatra: Mt. Leuser: Bivouac 6	3.87, 97.15	F. A. Ulmer, Jr
ANSP 20319	*R. hoogerwerfi*	8 May 1939	Old dry tissue skull	2621	Sumatra: Mt. Leuser: Bivouac 8	3.86, 97.14	F. A. Ulmer, Jr
BM 19.11.5.81	*R. korinchi*	26 April 1914	Old	2225	Sumatra: Mt. Kerinci: Sungai Kering	−1.73, 101.25	Robinson and Kloss
RMNH 23151	*R. korinchi*	14 June 1917	Old	2800	Sumatra: Mt. Talamau (=Talakmau)	0.08, 99.98	E. Jacobson
NH 2147	*R.* sp. R3[c]	1 February 1980	Old	16	Borneo: Sabah: Lahad Datu: Madai	4.72, 118.18	—
BOR260[b]	*R. tanezumi*	25 February 2013	Fresh liver	1538	Borneo: Sabah: Mt. Kinabalu: Kin. Park HQ	6.01, 116.55	M.T.R. Hawkins
NH 2015	*R. tiomanicus*[d]	22 August 1971	Old	126	Borneo: Sabah: Ulu Tuaran: Kg. Lebodon	6.15, 116.37	H. Tsen
USNM 590332	*R. tiomanicus*	19 January 2005	Fresh	22	Borneo: Sarawak: Ulu Kakas: Bukit Sarang	2.65, 113.05	Helgen, K. M.
USNM 590720	*R. tiomanicus*	24 January 2007	Fresh	22	Borneo: Sarawak: Ulu Kakas: Bukit Sarang	2.65, 113.05	Helgen, K. M.

ANSP, Academy of Natural Sciences of Drexel University, Philadelphia; BM, Natural History Museum, London; NH, Sabah Museum, Kota Kinabalu; RMNH, Naturalis Biodiversity Center; USNM, National Museum of Natural History, Smithsonian Institution.

[a]Extracted from field reports, museum labels, and inferred from coordinates.

[b]Field code (Camacho-Sanchez et al. 2019).

[c]Originally labeled *Sundamys muelleri*, but reassigned based on *cyt b* barcoding.

[d]Originally labeled *Lenothrix canus*, but reassigned based on *cyt b* barcoding.

The nonprotein coding genes were removed in Geneious and the gene *nd6*, which is in the light strand, was reverse-complemented. This mitogenome matrix had 57 individuals, with 21 species, 11 339 nucleotides (~69% of the mitogenome) and 0.04% missing data, as calculated with AMAS (Borowiec 2016). The best partition scheme was determined with PartitionFinder 2.1.1 (Lanfear et al. 2016) using the *rcluster* algorithm (Lanfear et al. 2014). The output partition scheme was specified as input in RAxML. It arranged the data into one partition for first and second codon positions and a second partition with the third codon position for all protein-coding genes, except codon position 2 of ATPase8 (54 sites), which fell in its own partition and was removed from RAxML analysis. The rapid bootstrapping algorithm was run on RAxML using the model of evolution GTR+ Γ. It converged after 350 replicates following the extended majority-rule stopping criterion.

We also reconstructed the phylogenetic tree in a Bayesian framework with BEAST 2.4.4 (Bouckaert et al. 2014) to date the nodes. To meet the assumptions of the tree (Yule speciation process) only one sample per species was kept (n = 21 mitogenomes; Supplementary Table S1). The *Rattus-Mus* split was used as a calibration point to date the tree. The final DNA matrix had 27 species, a length of 11 339 nucleotides with 0.01% missing data. We ran PartitionFinder 2.1.1 with the greedy algorithm and branch lengths unlinked. The best scheme was used to split the alignment into 3 sets which corresponded to codon positions 1, 2, and 3, for all genes, except *nd6* codon position 3 (165 positions) which had its own partition. We removed it from the dataset to avoid estimating extra parameters in BEAST as a good trade-off since that region was not very informative. The alignment was then split by codon positions 1 (3784 sites), 2 (3779 sites), and 3 (3611 sites), with AMAS. In BEAUTi, we set a GTR + G + I model to codon positions 1 and 3, and the HKY + G + I model to codon position 2, with estimated base frequencies, as suggested by PartitionFinder. We linked a relaxed clock model with frequencies sampled from a lognormal distribution to all partitions, but a relative substitution rate was estimated for each codon position. We used the 11.81 Mya (95% confidence interval: 11.11–12.68 Mya) suggested in Kimura et al. (2015) as a prior for the *Rattus-Mus* split. In BEAST, the prior was specified with a lognormal distribution (Morrison 2008). We ran 2 chains of 50 million generations, sampled every 10 000 generations. The 2 chains converged for each of the parameters in the combined log file after 10% burn-in, being the estimated sample sizes > 200, for all parameters. We generated a maximum clade credibility with TreeAnnotator after discarding the first 10% of the trees from each chain.

Mitochondrial DNA Structure in the *R. tiomanicus* Complex

We evaluated the phylogenetic relationships of the closely related high-elevation *R. baluensis* and the widespread, lowland *R. tiomanicus* lineages with *cyt b*, a widely used mitochondrial marker for which there was better geographical representation of *R. tiomanicus* samples in GenBank. For *Rattus baluensis*, we extracted *cyt b* from 32 mitogenomes, KY611359–KY611390, and added JN675495 (Aplin et al. 2011). For *R. tiomanicus*, we included 5 samples from Thailand (KC010165–KC010168 and HM217391) (Pagès et al. 2010; Latinne et al. 2013), plus *cyt b* extracted from mitogenome KP876560 from Peninsular Malaysia, one sample from Java (JN675515; Aplin et al. 2011), and six Bornean individuals, which included USNM 590332 and 590720, from Bintulu Division, Sarawak, two samples from Sungai Asap, Belaga, Sarawak

(JF436975 and JF436986; Tamrin and Abdullah 2011), one from Sabah, NH 2015 (Table 1), and one from Kalimantan (JN675516; Aplin et al. 2011). A total of 33 sequences from *R. baluensis* and 16 from *R. tiomanicus* (Supplementary Table S1) were aligned with MAFFT plugin in Geneious 8.1.5 with the *--auto* option. The alignment contained 1140 positions and 7% of missing data, as calculated with AMAS. A TCS haplotype network was built in PopART (Leigh and Bryant 2015) using 766 valid positions (with no missing data for any of the samples).

Reconstruction of Ancestral Distributions

We reconstructed the ancestral distribution of Asian *Rattus* on a mitochondrial tree to investigate the origin of the mountain distribution for *R. baluensis*, *R. hoogerwerfi*, and *R. korinchi*. We used the mitogenome tree from RAxML as a reference. Outgroups and nearly identical mitogenomes were dropped. The reconstruction of ancestral states improves with denser taxonomic sampling of the terminal taxa (Salisbury and Kim 2001; Heath et al. 2008) (Supplementary Figure S2). Therefore, we placed in the mitogenome tree other Asian *Rattus* for which only *cyt b* is available (Supplementary Table S1). *cyt b* sequences were aligned to the existing mitogenome alignment using MAFFT *--add* (Katoh and Frith 2012). Then, an Evolutionary Placement Algorithm in RAxML was used to place these lineages into the reference mitogenome tree (Berger et al. 2011). The reconstruction of the ancestral states was done with *phytools::rerootingMethod*, which re-roots the tree at all internal nodes and computes the marginal likelihoods for the ancestral states (Yang et al. 1995; Revell 2012). The native distribution of the lineages in the tree were classified into 4 categories according to their native distribution in Musser and Carleton (2005), and considering bioregions for Rattini described in Fabre et al. (2013): continental Asia north of Kra, Sundaland (except mountain endemics), mountains above 2000 m, and southeast of Wallace's Line (Wallacea and Sahul).

Selection on Mitochondrial DNA

The ratio of nonsynonymous to synonymous substitutions ($dN/dS = \omega$) can be used to assess selection on coding genes in phylogenetic trees (Nei and Gojobori 1986; Yang 1998). Values of $\omega = 1$, $\omega > 1$, and $\omega < 1$, indicate neutral, positive, and purifying selection, respectively. We estimated ω with Maximum Likelihood using CodeML in PAML 4.9 (Yang 2007) to evaluate signs of selection (different ω) associated to the mountain lineages. Based on the annotated multiple sequence alignment used for phylogenetic reconstructions with RAxML, we used trimAl 1.4 (Capella-Gutiérrez et al. 2009) to remove stop codons and keep the correct translation frame for all protein-coding genes. We carried selection analysis using the concatenated alignment and on a per-gene analysis. For the "per-gene" analysis, the alignment was partitioned per gene using AMAS. The sample *R. korinchi* BM19.11.5.81 was dropped because its large amounts of missing data restricted the number of useful positions for the analysis. Most nonsynonymous mutations are deleterious causing ω in branches to be below one in most cases. For that reason, the methodological approach for assessing selection is not to get the absolute value of ω, but to compare the likelihood of alternative models which accommodate different values of ω across the phylogeny (Jeffares et al. 2015). Accordingly, we ran a null model to estimate an average ω for the tree, and a 2-ratio branch model to estimate a ω for the background branches, ω0 (all the tree except foreground branches), and other ω for the foreground branches (high-elevation taxa): ω1 for *R. baluensis* and ω2 for *R. hoogerwerfi*–*R. korinchi* ($\omega 0 \neq \omega 1 \neq \omega 2$; Supplementary Figure S3). The likelihoods of the different models were contrasted with a likelihood ratio test (LRT) and the *P*-values were obtained from χ^2 distributions with a custom script in R (github.com/csmiguel/rattus-highlands). The differences in ω derived from this contrast can be mainly driven by purifying selection related to evolutionary constraints in the mitochondrial genes, and not necessarily positive selection (Yang and Nielsen 2002; Jeffares et al. 2015). Since positive selection often happens on specific amino acids and not across the whole gene, extended models called branch-site models were developed to accommodate a proportion of the amino acids to have positive selection in the foreground branches: null model is model A1 (*NSsites = 2*, *model = 2*, *fixomega = 1*) and the alternative model is model A (*NSsites = 2*, *model = 2*, *fixomega = 1*) (Yang and Nielsen 2002; Yang and Dos Reis 2011; Jeffares et al. 2015). We applied these models to detect positively-selected codons shared in the foreground branches (highland taxa*; R. hoogerwerfi*, *R. korinchi* and *R. baluensis*). The *P*-values from the LRT were calculated considering the null distribution is the 50:50 mixture of point mass 0 and χ^2_1 (Yang and Dos Reis 2011).

Results

Mitogenomes were successfully reconstructed from 9 of the 12 individuals attempted, which covered 6 of the 7 species. Successful genomes were 96.8–100% complete, with a per-sample coverage ranging from 9.2 to 163×. Unsuccessful genomes were 0.6–0.7% complete and had coverage of around 0.1×. The single species that was not successfully sequenced was *Rattus blangorum* (Table 2). Newly sequenced mitogenomes have been deposited in GenBank under accession numbers MN126561–MN126569.

Phylogenetic Relationships and Molecular Dating

All nodes within *Rattus* in the maximum likelihood tree and in the Bayesian maximum clade credibility tree were highly supported (bootstrap support/posterior probability: most support values near or equal to 100/1.00, respectively; Figures 2 and 3). The four *Rattus* species native to Sunda (*R. hoogerwerfi*, *R. korinchi*, *R. tiomanicus*, and *R. baluensis*) were in a clade inside Asian *Rattus*, well differentiated from the Australo-Papuan *Rattus*. However, they did not form a monophyletic clade. Both *Rattus tiomanicus* and *R. baluensis* are inside the *Rattus rattus* complex (sensu Aplin et al. 2011), which also includes the widespread Asian *R. tanezumi*, *R. rattus*, and one individual of the *Rattus* sp. lineage R3 sensu Pagès et al. (2010). The Sumatran montane *R. korinchi* and *R. hoogerwerfi* form a well-supported clade (87/1.00) and are the closest sequenced lineages to the *Rattus rattus* group. These Sumatran montane endemics are not sister lineages to the Bornean montane *R. baluensis*. The diversity of *R. baluensis* is nested within the diversity of *R. tiomanicus*.

According to the molecular dating, *Rattus* started to radiate about 3.3 Mya, (95% high posterior density, HPD: 2.82–3.85), and Asian *Rattus* at 2.95 Mya (2.5–3.45). The split of the 2 Sumatran montane rats (*R. korinchi* and *R. hoogerwerfi*) occurred at approximately 1.3 Mya (1.04–1.51). Their divergence from their presumably closest lowland ancestor is relatively deep, 1.38 Mya (1.15–1.63), compared with the shallow 0.31 Mya (0.23–0.39) of coalescent time estimated between a representative mitogenome from the highland *R. baluensis* and the widespread lowland *R. tiomanicus* from Sarawak.

Table 2. Information on mitochondrial genomes generated for this study

Code	*Rattus* species	Coverage (*x*)	Length (bp)	Mitogenome assembled (%)	GenBank
ANSP 20348	*R. blangorum*	0.1	16 311	0.6	—
BOR577[a]	*R. exulans*	107.4	16 303	100	MN126569
ANSP 20309	*R. hoogerwerfi*	0.13	16 311	0.7	—
ANSP 20315	*R. hoogerwerfi*	0.1	16 311	0.5	—
ANSP 20319	*R. hoogerwerfi*	62.7	16 314	100	MN126561
BM 19.11.5.81	*R. korinchi*	9.2	16 313	96.8	MN126567
RMNH 23151	*R. korinchi*	33.8	16 312	99.8	MN126568
NH 2147	*R.* sp. R3	79.2	16 308	99.6	MN126565
BOR260[a]	*R. tanezumi*	29	16 306	100	MN126566
NH 2015	*R. tiomanicus*	43.1	16 309	100	MN126562
USNM 590332	*R. tiomanicus*	121.4	16 312	100	MN126563
USNM 590720	*R. tiomanicus*	163	16 313	100	MN126564

[a]Field code (Camacho-Sanchez et al. 2019). Specimens at the Doñana Biological Station, Spain, not yet catalogued.

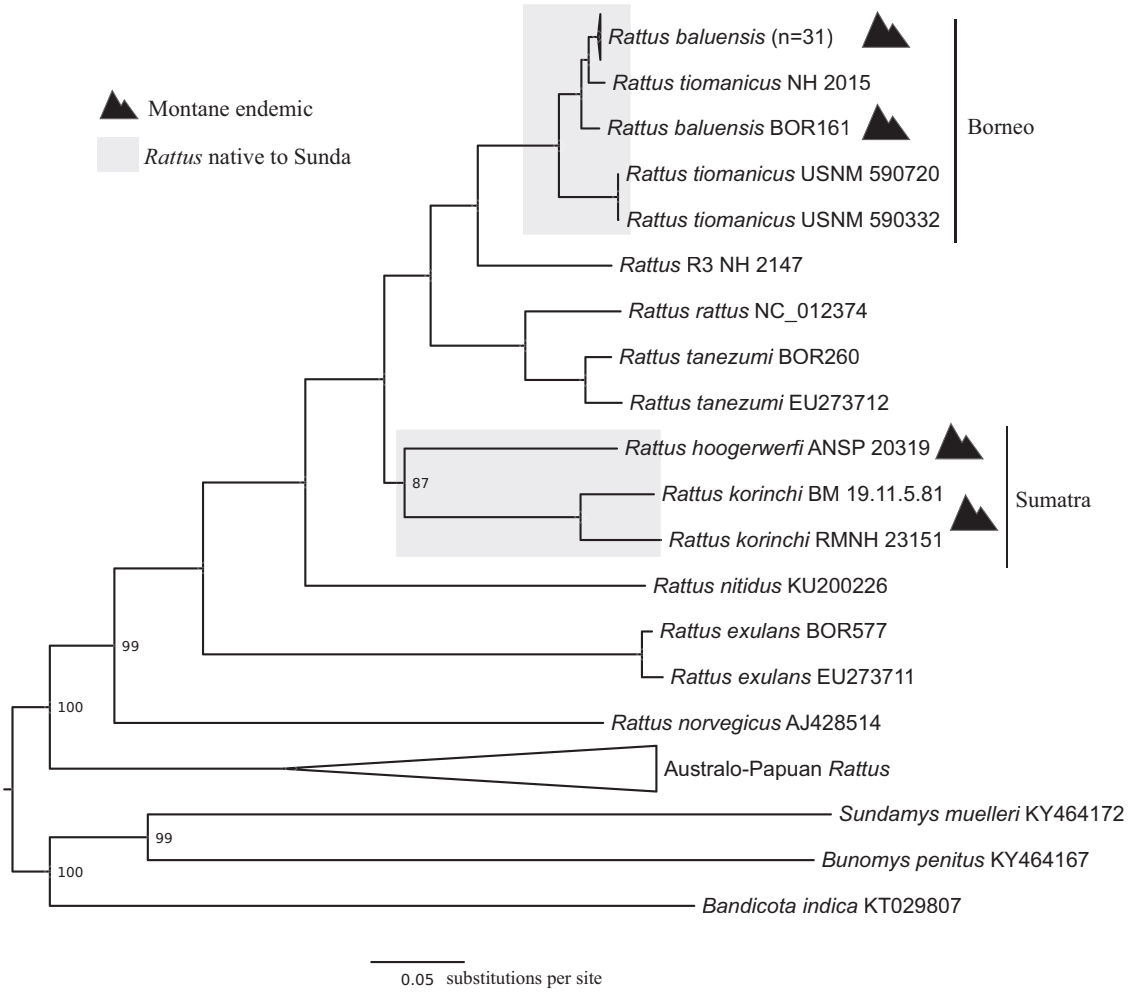

Figure 2. RAxML consensus tree from ML phylogenetic inference with protein-coding genes of mitogenomes. Diamonds represent 100% of bootstrap support.

Cytochrome *b* Diversity in *Rattus tiomanicus* and *R. baluensis*

A *cyt b* haplotype network showed *cyt b* diversity from the montane *R. baluensis* is much lower than and nested within the greater diversity of the widespread, lowland *R. tiomanicus* (Figure 4; Supplementary File S1). Haplotypes of *R. tiomanicus* from northern Borneo were closer to *R. baluensis* than to any other Bornean or western Sunda individuals (Figure 4). *Rattus baluensis* was monophyletic except for one divergent haplotype, Rba_3, identified in 1 of the 33 individuals sequenced. This haplotype had higher similarity to some haplotypes of *R. tiomanicus* (*R. tiomanicus* from Sarawak, Rti_6, 5 mutations) than to the core diversity of *R. baluensis* (Rba_1 and Rba_2, 9–10 mutations).

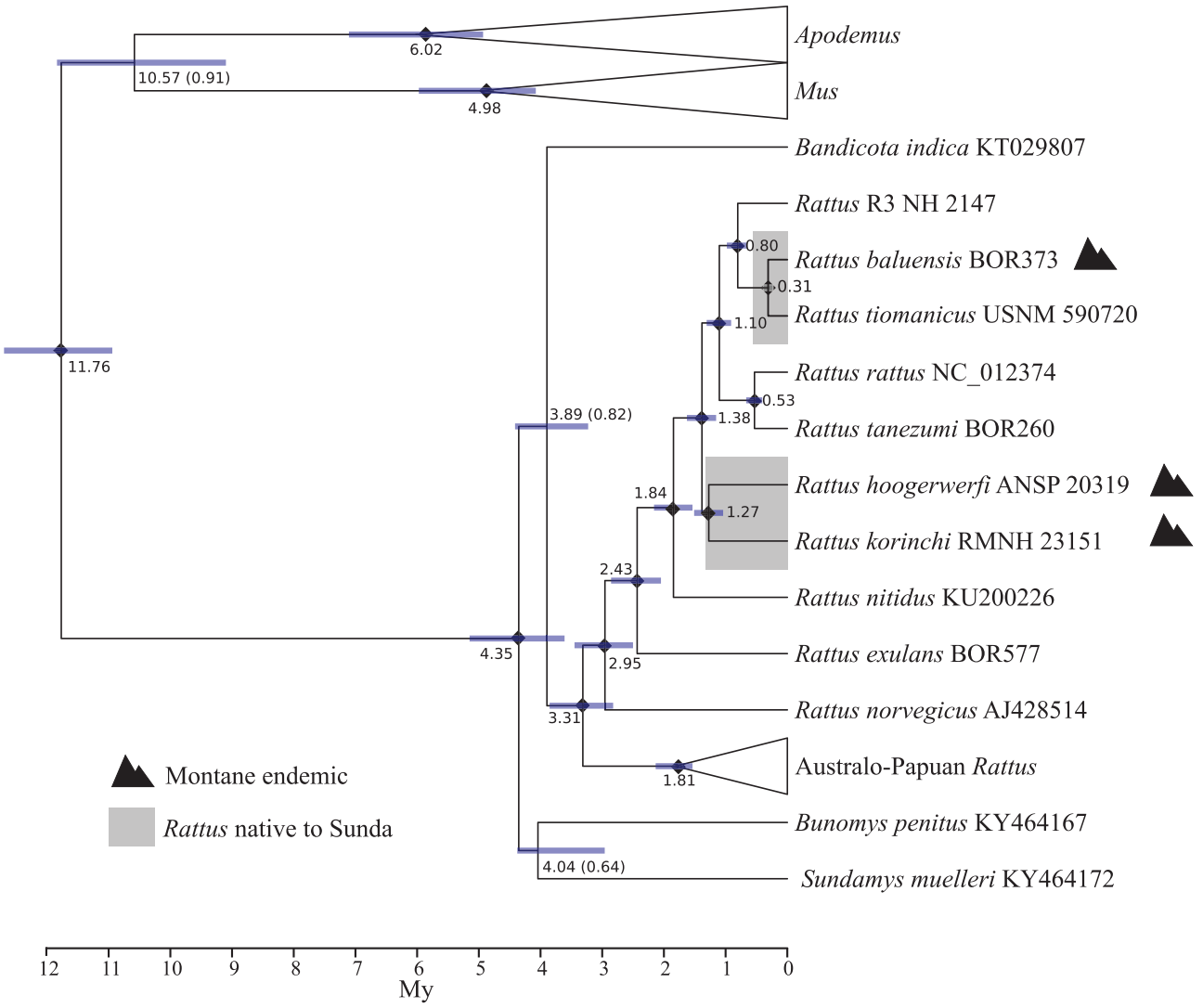

Figure 3. Maximum clade credibility tree from BEAST analysis using protein-coding genes of mitogenomes. Node ages in millions of years ago (My) with their 95% HPD are represented in each node. Diamonds represent PP = 1.00. PP below 1.00 are indicated in parenthesis.

Origin of Mountain Lineages

We detected at least 2 invasions of Asian *Rattus* to Sundaland from continental Asia. According to our reconstructions, the older invasion is represented in Sumatra, by the common ancestor of highland *R. korinchi* and *R. hoogerwerfi* (clade A, Figure 5). The second invasion is represented across all Sundaland by the *R. tiomanicus* complex, from which the montane *R. baluensis* diverged very recently in northern Borneo (clade B, Figure 5). The polarity of the origin of the mountain species on Borneo and Sumatra adhered to a lowland-to-highland colonization. Many of the *cyt b* sequences added to the tree came from historical specimens (mainly from Thomson et al. 2018) and are very short, which leads to uncertainty in their phylogenetic placement (e.g., *R. lugens* ACAD10905) (Matsen et al. 2012). We also detected a second crossing of Wallace's line by *R. hoffmanni* to Wallacea.

Selection on Mitochondrial DNA

The model describing a single ω for the whole tree was rejected (P = 0.008) in favor of a branch model that described a different ω for the background branches compared to the highland lineages (Supplementary Table S2). Accordingly, ω was greatest for *R. baluensis* (ω1 = 0.049), followed by the Sumatran montane lineages (*R. hoogerwerfi* and *R. korinchi*; ω2 = 0.040), and the background ω (ω0 = 0.37). The same contrast applied "per-gene" revealed significant differences in 2 of the 13 mitochondrial protein-coding genes, *atp6* (P = 0.002) and *cox3* (P = 0.006), where again the greatest values of ω were observed in *R. baluensis* (Figure 6; Supplementary Table S2). The genes *atp6* and *cox* had the lowest ω values, while *atp8* and *nd1/nd2* had the greatest. The low background value of ω in *cox3* (0.018) contrasted with the markedly greater value in the branch of *R. baluensis* (0.116). This gene (*cox3*) was the only one with signatures of positive selection according the branch-site models (P = 0. 028; Supplementary Table S2), although the BEB approximation (Yang 2005) did not identify any codon as significantly under positive selection (P > 0.95).

Discussion

Evolutionary Implications of the Molecular Phylogeny

Our estimation for the origin of *Rattus* based on mitogenomes, 2.82–3.85 Mya, is consistent with other representative studies, which

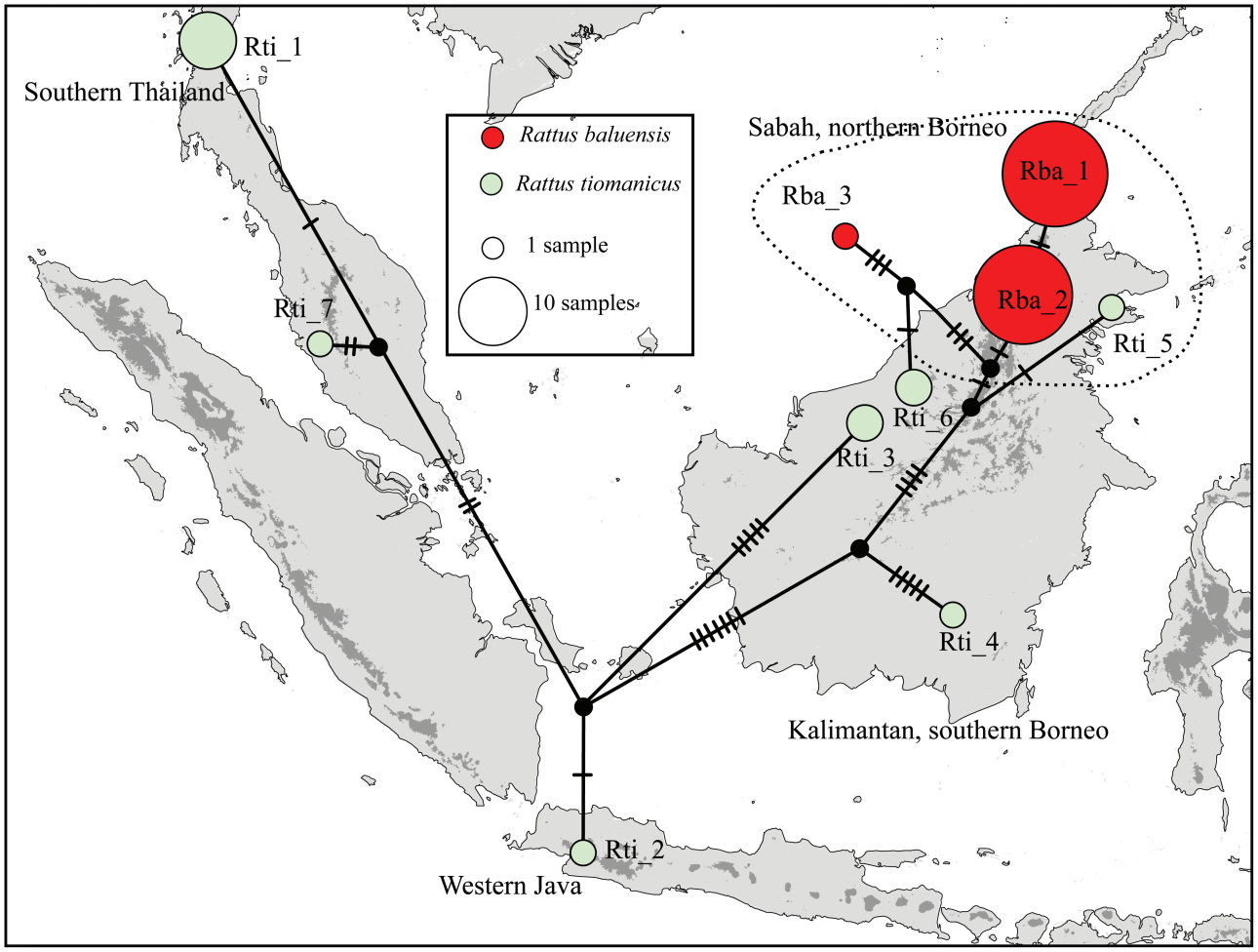

Figure 4. TCS haplotype network of *cyt b* sequences from *Rattus baluensis* and *R. tiomanicus*. The haplotypes are placed in their approximate geographic origin. A dashed line encircles all haplotypes from Sabah, northern Borneo. In the network, the circle size proportional to the number of sequences for the haplotype, black dots represent missing haplotypes, and perpendicular lines mutations between haplotypes.

included additional nuclear markers: 2.9–3.6 Mya in Steppan and Schenk (2017), ~2.6 Mya in Rowe et al. (2019) and 2.5–3.3 Mya in Fabre et al. (2013). The highland endemics *Rattus hoogerwerfi* and *R. korinchi* are sister taxa coalescing on relatively long branches of the tree, with a common ancestor at around 1.3 Mya, and they are peripheral to the *Rattus rattus* complex (Figures 2 and 3). These results resolve the uncertain taxonomic position (see Musser and Carleton 2005) of these high-elevation Sumatran species and are consistent with the taxonomic placement after morphological descriptions in Musser (1986). The pattern of high-elevation lineages on long branches has also been recorded in other mammals in this region, such as for *Sundamys infraluteus* (Camacho-Sanchez et al. 2017), *Sundasciurus everetti* (Hawkins et al. 2016), *Sundasciurus altitudinis* (den Tex et al. 2010), and *Rattus niobe* in the nearby New Guinea (Rowe et al. 2011).

Their deep divergence within Asian *Rattus* contrasts with the young origin of *Rattus baluensis* from *Rattus tiomanicus* (<0.39 Mya). The low mitochondrial diversity in *R. baluensis* derives from the local diversity of *R. tiomanicus* (Figures 2 and 4). These two species have not reached reciprocal monophyly at the mitochondrial level. The genetic structure of *R. baluensis* with respect to *R. tiomanicus* illustrates the predicted genetic consequences of vicariance among populations with different sizes, in which the smaller population (*R. baluensis* in this case) will become monophyletic first, while the larger one (*R. tiomanicus*) will remain paraphyletic for some longer time before reaching reciprocal monophyly (Zink and Barrowclough 2008). The nested position of *R. baluensis* with respect to *R. tiomanicus* has already been pointed out by Aplin et al. (2011) and Thomson et al. (2018), from fewer samples and a more limited geographical distribution. The retention of ancestral polymorphism or introgression from *R. tiomanicus* could explain this pattern. Both processes seem common at shallow evolutionary scales although, and they are difficult to disentangle (Peters et al. 2007; Hailer et al. 2012; Pagès et al. 2013).

Although mitogenomes provide robust support at different evolutionary depths in phylogenetic inference of Rattini (Robins et al. 2008; Camacho-Sanchez et al. 2017; Wei et al. 2017), they might not represent reliably the evolutionary history of these taxa, and unlinked nuclear markers should be evaluated to confirm the interpretations from mitogenomes (Brito and Edwards 2009; Pagès et al. 2013).

Origin of Montane Lineages

Rattus seems to have colonized the Sunda Shelf twice from its center of origin in Continental Asia (Figure 5). An older colonization event

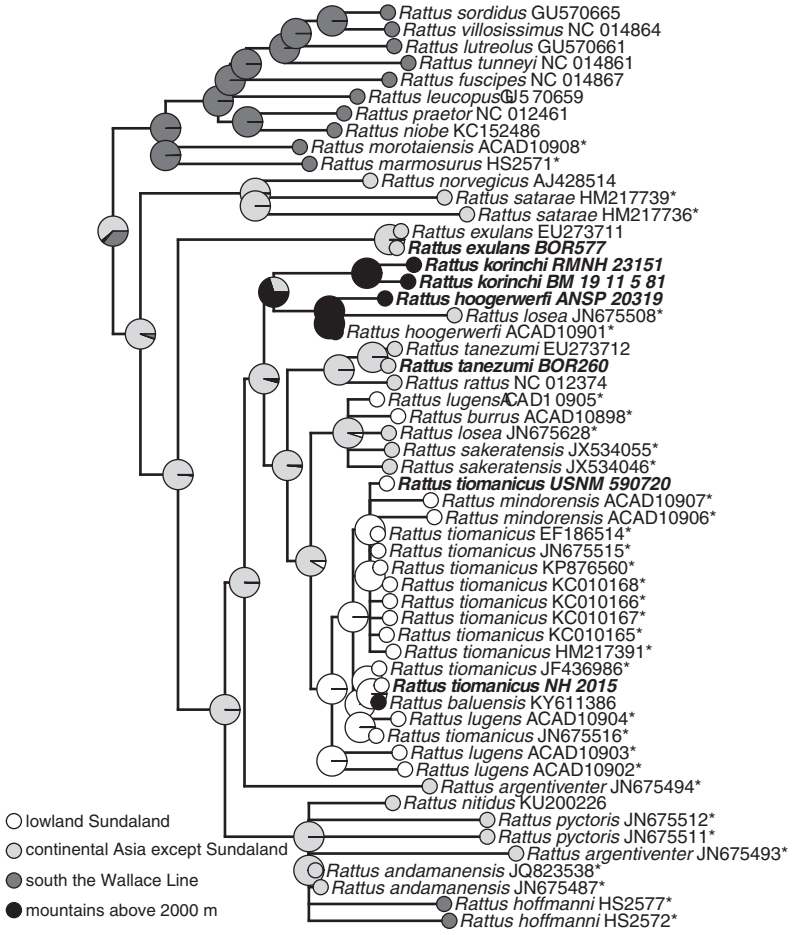

Figure 5. Reconstruction of the ancestral distributions on tree from RAxML based on mitogenome sequences in which *cyt b* sequences were placed using an Evolutionary Placement Algorithm (taxa marked with "*"). The samples sequenced in this study are in bold. The pies in the nodes and tips represent the marginal likelihoods of being native to different biogeographical regions.

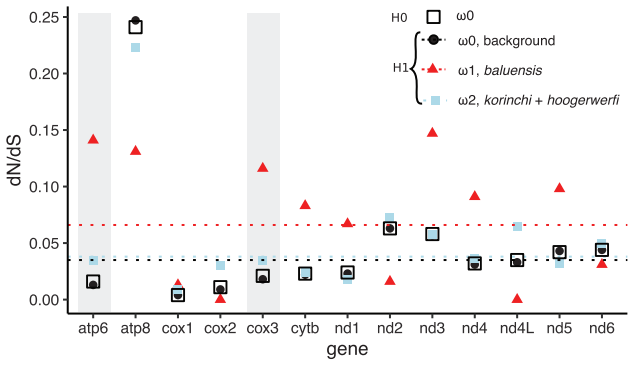

Figure 6. Ratio of the nonsynonymous to synonymous substitutions (ω) for the 13 mitochondrial protein-coding genes estimated from branch models for the concatenated sequences (dotted lines) and per-gene (points). ω values from the null model (a unique ω0 for the tree; black squares) are shown against the ω values for the alternative model (ω0 ≠ ω1 ≠ ω2; points). Grey boxes mark genes for which H1 ≠ H0 ($P < 0.05$) (extended results on Supplementary Table S2).

at around 1.3 Mya gave rise to the Sumatran montane endemics *R. hoogerwerfi* and *R. korinchi*, seemingly from a lowland ancestor from continental Asia. Inferring extinction in phylogenies is challenging (Sanmartín and Meseguer 2016), but this reconstruction requires extinction at least of lowland *Rattus* on Sumatra. A more recent colonization of Sundaland by *Rattus* less than 1 Mya diversified across all Sundaland, including the smaller islands, giving rise to the *R. tiomanicus* complex (Figure 5). This is supported by the molecular affinities between *R. tiomanicus*, *R. burrus*, *R. mindorensis* and *R. lugens*, and confirms the morphological affinities between these lineages (Musser 1986). A lineage from this complex, *R. mindorensis*, even seems to cross Huxley's line into Mindoro Island, Philippines (Figure 5). Another highlight, already identified in Fabre et al. (2013) and Rowe et al. (2019), is the origin of *R. hoffmanni* in Wallacea from continental Asia without any extant Sundaic stepping lineage, likely also indicating extinctions of other *Rattus* in Sundaland and/or the presence of unidentified species. The *cyt b* sequences available for some of these taxa contained much missing data leading in some cases to their ambiguous placement in the phylogeny. The inclusion of nuclear data and other island lineages not assessed in this study (i.e., *R. palmarum*, *R. simalurensis*, *R. adustus*; see Musser 1986) would help to gain understanding in the colonization dynamics of *Rattus* in Sundaland. Despite these limitations, we depict a complex phylogeographic history of *Rattus* as a result of their capacity to have crossed multiple times sharp biogeographical barriers. Rowe et al. (2019) described *Rattus* as the Indo-Pacific murine with the highest diversification rates and the group with most transitions

across biogeographical barriers. The origin of Rattini seems to be in continental Asia, and most Sundaic Rattini have ancestors from this area, similar to *Rattus* (i.e. *Niviventer*, *Leopoldamys*, *Lenothrix*, some species of *Maxomys*). Others lineages, however, seem to have arrived secondarily to Sundaland after having diverged in Wallacea (i.e., *Sundamys*, several *Maxomys* species) or continental Asia (Rowe et al. 2019).

In Borneo, our ample molecular sampling supports a likely peripatric speciation in which *R. baluensis* originated from a lineage of *R. tiomanicus* that colonized the mountain habitat. This makes *R. baluensis* another example of a Mt. Kinabalu endemic which originated from a lowland taxon (Merckx et al. 2015), and it is also consistent with the low mitochondrial genetic diversity reported in *R. baluensis* from hypothetical founder events (Figure 4; Camacho-Sanchez et al. 2018). The species status of these two lineages, *R. baluensis*–*R. tiomanicus*, is surprising given their evolutionary and spatial proximity: they have been reported on Mt. Kinabalu only 500 m apart in elevation (Musser and Califia 1982; Musser 1986; Camacho-Sanchez et al. 2019). This seems to be a unique case of Bornean mammals in which recently diverged sister taxa are sympatric (but they are not syntopic). Similar scenarios are found in several Bornean birds (Moyle et al. 2017). Considering mountains in Sumatra and Borneo, our results suggest a polarity of lowland-to-highland divergence of the montane lineages.

A possible genetic mechanisms for the origin of *R. baluensis* could be linked to founder events in the colonization of mountain habitats ("sky islands," McCormack et al. 2009) analogous to those described in mammals arriving to islands (Berry 1996; Frankham 1997; Abdelkrim et al. 2005): divergence in allopatry followed by secondary contact (Moyle et al. 2017), or adaptive divergence to different lowland/highland habitats in the presence of reduced gene flow (McCormack and Smith 2008; Linck et al. 2020). Genomic scale evaluation of the *R. baluensis*–*R. tiomanicus* complex from northern Borneo is needed to elucidate the possible mechanisms of speciation in *R. baluensis*.

Morphology in Montane *Rattus*

We did not explicitly gather quantitative data on the morphology of Sundaic native *Rattus*. However, the 3 montane species seem to share external morphological traits: larger bodied, darker, and thicker pelage (Musser 1986 and our own observations). Few characters differentiate the skulls of *R. baluensis* and *R. tiomanicus*; only the frequency of occurrence of cusps t3 in first and second molars and size of skull, which is larger in *R. baluensis* (Musser 1986). This is a low level of interspecific differentiation among congeneric Rattini (Musser and Newcomb 1983; Musser 1986). Little divergence in skulls contrasts with more differences in external morphology between these 2 species. The differences in external appearance are similar to those observed between the montane *R. korinchi* or *R. hoogerwerfi* to other lowland *Rattus*. The thick dark fur and larger skull and body measurements are common morphological traits in Sundaland montane *Rattus*, likely reflecting convergence to mountain habitats. This kind of morphological convergence is not unique to *Rattus*. For instance, it also hindered the taxonomic position of the codistributed ground squirrel, *Dremomys everetti*, which was recently moved to *Sundasciurus everetti* (Hawkins et al. 2016), and the so described montane external morphology seems also to be shared by other codistributed montane small mammals such as *Maxomys hylomyoides* and *Sundamys infraluteus*.

The larger size of *R. baluensis* as opposed to its lowland sister species *R. tiomanicus* (mean ± SD in mm for *R. baluensis*/*R. tiomanicus*; head body, HB: 170 ± 8.4 for $n = 23$/157.8 ± 14.6 for $n = 5$; greatest length of skull, GLS: 40.8 ± 1.3 for $n = 24$/37.4 ± 1.3 for $n = 12$; Musser and Califia 1982; Musser 1986) may suggest an "island" effect on its reduced mountain habitat. Particularly, in rodents, there is a general negative correlation between island size and body size, probably as a convergence to a better physiological efficiency, which is allowed by reduced predation and interspecific competition, while food availability does not become a limiting factor for small mammals at these small areas (Heaney 1978; Lomolino 1985). This larger size pattern in small islands is particularly marked in the larger insular populations of *R. tiomanicus* in many islands of eastern Borneo (most GLS around 40 mm vs. 37.4 mm in mainland Borneo; Musser and Califia 1982; Musser 1986). The even larger size of the Sumatran *R. hoogerwerfi* (HB: 182.7 ± 6.7, $n = 20$; GLS: 42.9 ± 0.8, $n = 16$) and *R. korinchi* (HB: 166 and 169; GLS: 41 and 41.8) could also be related to this island effect, but we lack a close lowland ancestor for proper comparison. This could also be an expression of Bergmann's (1848) rule, which describes a pattern of larger body size associated with colder climate. However, comparative analysis with a solid morphological dataset is needed to support the observations based on the external morphology.

Selection

We detected greater ω in *R. baluensis* and the Sumatran highlands than on the background tree. The signal seemed to be driven mainly by *atp6* and *cox3*. These inferences from branch models average the effect of purifying and positive selection acting on the branches. Thus, the large values of ω for *R. baluensis* could be the consequence of slightly deleterious mutations, which have not been yet eliminated by purifying selection on the terminal branches (Yang and Nielsen 2002; Elson et al. 2004; Ho et al. 2005; Kivisild et al. 2006). The greater deviation of ω1 across genes in *R. baluensis* from ω0 contrasts with the little deviation in the Sumatran highlands (ω2) (Figure 6), and could also reflect the stochasticity of few mutations in the branch of *R. baluensis* compared with the larger number of mutations in the older Sumatran lineages. The constraints in molecular evolution we describe for the mitochondrial protein-coding genes within *Rattus* (Figure 6; Supplementary Table S2) mirror that described for other mammals, with *cox* genes having the most stringent constraints together with *atp6*, compared with a relaxation in the evolution of *atp8* and some genes from the NADH dehydrogenase (da Fonseca et al. 2008). The branch-site models did find positive selection in *cox3* considering the highland lineages together (Supplementary Table S2). A study on geese also found *cox3* as a target gene for adaptation to high altitude (Scott et al. 2011). In our results, it is difficult to correlate selection with adaption to high elevations because the signal we detected was weak and no residues had statistically supported positive values of ω. The gene(s) and residue(s) on which selection can act in response to high-elevation adaptation varies between studies, so that a clear signal of positive selection and molecular predictions are often needed to support adaptation in mitochondrial genes (Scott et al. 2011; Yu et al. 2011; Zhou et al. 2014).

Supplementary Material

Supplementary material is available at *Journal of Heredity* online.
Table S1. Samples and sequences used in this study.
Table S2. Results from selection analysis on mitogenomes.
Figure S1. External morphology of Sundaic *Rattus*.

Figure S2. Ancestral reconstructions based on the RAxML with only mitogenome sequences.
Figure S3. Hypothesis for branch models in CodeML.
Figure S1. Dorsal view of the skins of the lowland R. tiomanicus, R. blangorum, and the 3 montane endemics, R. baluensis, R. hoogerwerfi, and R. korinchi, with a (2.2x) detail of the woolly underfur of R. korinchi.
Figure S2. Reconstruction of the ancestral distributions on the Maximum Likelihood tree from RAxML based on mitogenome sequences. The samples sequenced in this study are in bold. The pies in the nodes and tips represent the marginal likelihoods of being native to: blue, continental Asia north of Kra; green, southeast the Wallace Line (Wallacea and Sahul); yellow, mountains above 2000 m; red, lowland Sundaland.
Figure S3. Hypothesis tested for assessing ω on the protein-coding genes from the mitochondrial genome. The average ω on the tree, or on the background branches (ω0) was contrasted against a 2-site branch model in which the mountain lineages from Borneo (*R. baluensis* plus *R. tiomanicus* NH2015) and Sumatra had different ω (ω1 and ω2, respectively).
File S1: *cyt b* haplotypes in the *Rattus tiomanicus–R. baluensis* complex.

Data availability

Code and data for all analysis are deposited at github.com/csmiguel/rattus-highlands, with a stable release at Zenodo, DOI: 10.5281/zenodo.3883085. Mitogenomes have been deposited in GenBank under accession numbers MN126561–MN126569.

Funding

This work was supported by the Spanish Ministry of Science and Innovation (CGL2010-21524, CGL2014-58793-P, CGL2017-86068-P to J.A.L. and BES-2011–049186 to M.C.-S.); and the SYNTHESYS Project (http://www.synthesys.info/) which is financed by the European Community Research Infrastructure Action under the FP7 Integrating Activities Program (SYNTHESYS ACCESS NL-TAF-5588 and GB-TAF-5303 to M.C.-S.).

Acknowledgments

We thank the curators and managers that facilitated access to different mammal collections: N. Gilmore and T. Daeshler, ANSP, Philadelphia; Kristofer Helgen, N. Edminson, and the Mammal Division at the NMNH, Washington DC; S. van der Mije and P. Kamminga, NBC, Leiden; R. Portela, NHM, London; A. Lo, Sabah Museum, Kota Kinabalu. Kristofer Helgen provided valuable insight in the conception of this work and S.Y.W. Ho advised on molecular dating. Logistical support was provided by Laboratorio de Ecología Molecular, Estación Biológica de Doñana, CSIC (LEM-EBD) and by the infrastructures offered by Doñana's Singular Scientific-Technical Infrastructure (ICTS-EBD). We also thank Sabah Parks for research permits (TS/PTD/5/4 Jld. 45 (33) and TS/PTD/5/4 Jld. 47 (25)) and various kind of support, the Economic Planning Unit (reference: 100–24/1/299), and export permits from the Sabah Wildlife Department (JHL.600–3/7 Jld.7/19 and JHL.600–3/7 Jld.8/) and Sabah Biodiversity Council (Ref: TK/PP:8/8Jld.2). We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

References

Abdelkrim J, Pascal M, Samadi S. 2005. Island colonization and founder effects: the invasion of the Guadeloupe islands by ship rats (*Rattus rattus*). *Mol Ecol*. 14:2923–2931.

Aplin KP, Suzuki H, Chinen AA, Chesser RT, Ten Have J, Donnellan SC, Austin J, Frost A, Gonzalez JP, Herbreteau V, *et al.* 2011. Multiple geographic origins of commensalism and complex dispersal history of black rats. *PLoS One*. 6:e26357.

Berger SA, Krompass D, Stamatakis A. 2011. Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst Biol*. 60:291–302.

Bergmann C. 1848. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien*. 3:595–708.

Berry RJ. 1996. Small mammal differentiation on islands. *Philos Trans R Soc Lond B Biol Sci*. 351:753–764.

Borowiec ML. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ*. 4:e1660.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 10:e1003537.

Brito PH, Edwards SV. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*. 135:439–455.

Camacho-Sanchez M, Hawkins MTR, Tuh Yit Yu F, Maldonado JE, Leonard JA. 2019. Endemism and diversity of small mammals along two neighboring Bornean mountains. *PeerJ*. 7:e7858.

Camacho-Sanchez M, Leonard JA, Fitriana Y, Tilak M-K, Fabre P-H. 2017. The generic status of *Rattus annandalei* (Bonhote, 1903) (Rodentia, Murinae) and its evolutionary implications. *J Mamm*. 98:1340–1355.

Camacho-Sanchez M, Quintanilla I, Hawkins MTR, Tuh FYY, Wells K, Maldonado JE, Leonard JA. 2018. Interglacial refugia on tropical mountains: novel insights from the summit rat (*Rattus baluensis*), a Borneo mountain endemic. *Divers Distrib*. 24:1252–1266. doi:10.1111/ddi.12761

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25:1972–1973.

Chasen FN. 1939. Two new mammals from North Sumatra. *Treubia* 17:207–208.

Cheviron ZA, Brumfield RT. 2012. Genomic insights into adaptation to high-altitude environments. *Heredity (Edinb)*. 108:354–361.

da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A. 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*. 9:119.

den Tex RJ, Thorington R, Maldonado JE, Leonard JA. 2010. Speciation dynamics in the SE Asian tropics: putting a time perspective on the phylogeny and biogeography of Sundaland tree squirrels, *Sundasciurus*. *Mol Phylogenet Evol*. 55:711–720.

Elson JL, Turnbull DM, Howell N. 2004. Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet*. 74:229–238.

Fabre PH, Herrel A, Fitriana Y, Meslin L, Hautier L. 2017. Masticatory muscle architecture in a water-rat from Australasia (Murinae, Hydromys) and its implication for the evolution of carnivory in rodents. *J Anat*. 231:380–397.

Fabre P-H, Pagès M, Musser GG, Fitriana YS, Fjeldså J, Jennings A, … Helgen, K. M. (2013). A new genus of rodent from Wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for biogeography and Indo-Pacific Rattini systematics. *Zool J Linn Soc*. 169:408–447.

Fontanillas P, Dépraz A, Giorgi MS, Perrin N. 2005. Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, *Crocidura russula*. *Mol Ecol*. 14:661–670.

Frankham R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity (Edinb)*. 78 (Pt 3):311–327.

Hailer F, Kutschera VE, Hallström BM, Klassert D, Fain SR, Leonard JA, Arnason U, Janke A. 2012. Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science*. 336:344–347.

Hawkins MTR, Helgen KM, Maldonado JE, Rockwood LL, Tsuchiya MTN, Leonard JA. 2016. Phylogeny, biogeography and systematic revision of plain long-nosed squirrels (genus *Dremomys*, Nannosciurinae). *Mol Phylogenet Evol*. 94:752–764.

Heaney LR. (1978). Island area and body size of insular mammals: evidence from the tri-colored squirrel (*Callosciurus prevostii*) of Southeast Asia. *Evolution*. 32:29–44.

Heath TA, Hedtke SM, Hillis DM. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *J Syst Evol*. 46:239–257.

Ho SY, Phillips MJ, Cooper A, Drummond AJ. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol Biol Evol*. 22:1561–1568.

IUCN. 2016. *The IUCN Red List of Threatened Species*. Version 2016. https://www.iucnredlist.org

Jeffares DC, Tomiczek B, Sojo V, dos Reis M. 2015. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. *Methods Mol Biol*. 1201:65–90.

Katoh K, Frith MC. 2012. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics*. 28:3144–3146.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. doi:10.1093/nar/gkf436

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, *et al*. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647–1649.

Kimura Y, Hawkins MTR, Mcdonough MM, Jacobs LL, Flynn LJ. 2015. Corrected placement of Mus-Rattus fossil calibration forces precision in the molecular tree of rodents. *Sci Rep* 5:14444. doi:10.1038/srep14444

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A, *et al*. 2006. The role of selection in the evolution of human mitochondrial genomes. *Genetics*. 172:373–387.

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 PB, Wright AM, 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.

Latinne A, Waengsothorn S, Rojanadilok P, Eiamampai K, Sribuarod K, Michaux JR. 2013. Diversity and endemism of Murinae rodents in Thai limestone karsts. *Syst Biodivers*. 11:323–344.

Leigh JW, Bryant D. 2015. POPART: full-feature software for haplotype network construction. *Methods Ecol Evol*. 6:1110–1116.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.

Linck E, Freeman BG, Dumbacher JP. 2020. Speciation and gene flow across an elevational gradient in New Guinea kingfishers. *BioRxiv*. doi:10.1101/589044v3

Lomolino MV. 1985. Body size of mammals on islands: the island rule reexamined. *Am Nat*. 125:310.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J*. 17:10–12.

Martinez Q, Lebrun R, Achmadi AS, Esselstyn JA, Evans AR, Heaney LR, Miguez RP, Rowe KC, Fabre PH. 2018. Convergent evolution of an extreme dietary specialisation, the olfactory system of worm-eating rodents. *Sci Rep*. 8:17806.

Matsen FA, Hoffman NG, Gallagher A, Stamatakis A. 2012. A format for phylogenetic placements. *PLoS One*. 7:e31009.

McCormack JE, Huang H, Knowles LL. 2009. Sky Islands. In:Gillespie R, Clague D, editors. *Encyclopedia of islands*. Berkeley (CA): University of California Press. p. 839–843.

McCormack JE, Smith TB. 2008. Niche expansion leads to small-scale adaptive divergence along an elevation gradient in a medium-sized passerine bird. *Proc Biol Sci*. 275:2155–2164.

Merckx VS, Hendriks KP, Beentjes KK, Mennes CB, Becking LE, Peijnenburg KT, Afendy A, Arumugam N, de Boer H, Biun A, *et al*. 2015. Evolution of endemism on a young tropical mountain. *Nature*. 524:347–350.

Miller GS. 1942. Zoological results of the George Vanderbilt Sumatran Expedition, 1936–1939. Part V: mammals collected by Frederick A. Ulmer, Jr. on Sumatra and Nias. *Proc Natl Acad Sci Philadelphia*. 94:107–165.

Morrison DA. 2008. How to summarize estimates of ancestral divergence times. *Evol Bioinform Online*. 4:75–95.

Moyle RG, Manthey JD, Hosner PA, Rahman M, Lakim M, Sheldon FH. 2017. A genome-wide assessment of stages of elevational parapatry in Bornean passerine birds reveals no introgression: implications for processes and patterns of speciation. *PeerJ*. 5:e3335.

Musser GG. 1986. Sundaic Rattus: definitions of *Rattus baluensis* and *Rattus korinchi*. *Am Mus Novit*. 2862:1–24.

Musser GG, Califia D. 1982. Identities of rats from Pulau Maratua and other islands off East Borneo. *Am Mus Novit*. 2726:1–30.

Musser GG, Carleton MD. 2005. Superfamily Muroidea. In: Wilson DE, Reeder DM, editors. *Mammal species of the world: a taxonomic and geographic reference*, 3rd ed. Baltimore (MD): The Johns Hopkins University Press. p. 894–1531.

Musser GG, Newcomb C. 1983. Malaysian murids and the giant rat from Sumatra. *Bull Am Mus Nat Hist*. 174:327–598.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature*. 403:853–858.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 3:418–426.

Nor SM. 2001. Elevational diversity patterns of small mammals on Mount Kinabalu, Sabah, Malaysia. *Global Ecol Biogeogr*. 10:41–62.

Sody HJV. 1941. On a collection of rats from the Indo-Malayan and Indo-Australian regions. *Treubia* 18:255–325.

Steppan SJ, Schenk JJ. 2017. Muroid rodent phylogenetics: 900-Species tree reveals increasing diversification rates. *PLoS ONE* 12. doi:10.1371/journal.pone.0183070

Pagès M, Bazin E, Galan M, Chaval Y, Claude J, Herbreteau V, Michaux J, Piry S, Morand S, Cosson JF. 2013. Cytonuclear discordance among Southeast Asian black rats (*Rattus rattus* complex). *Mol Ecol*. 22:1019–1034.

Pagès M, Chaval Y, Herbreteau V, Waengsothorn S, Cosson JF, Hugot JP, Morand S, Michaux J. 2010. Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. *BMC Evol Biol*. 10:184.

Pagès M, Fabre P-H, Chaval Y, Mortelliti A, Nicolas V, Wells K, … Lazzari, V. (2016). Molecular phylogeny of South-East Asian arboreal murine rodents. *Zool Scripta*. 45:349–364.

Perrigo A, Hoorn C, Antonelli A. 2019. Why mountains matter for biodiversity. *J Biogeogr*. 47:315–325.

Peters JL, Zhuravlev Y, Fefelov I, Logie A, Omland KE. 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution*. 61:1992–2006.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 3:217–223.

Robins JH, McLenachan PA, Phillips MJ, Craig L, Ross HA, Matisoo-Smith E. 2008. Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *Mol Phylogenet Evol*. 49:460–466.

Robinson H, Kloss C. 1918. Results of an expedition to Korinchi Peak, 12,400 ft., Sumatra. 1. Mammals. *J Fed Malay State Mus*. 8:1–80.

Robinson H, Kloss C. 1919. On Mammals chiefly from the Ophir district, West Sumatra. *J Fed Malay State Mus*. 7:299–323.

Rowe KC, Achmadi AS, Fabre PH, Schenk JJ, Steppan SJ, Esselstyn JA. 2019. Oceanic islands of Wallacea as a source for dispersal and diversification of murine rodents. *J Biogeogr*. 46:2752–2768.

Rowe KC, Aplin KP, Baverstock PR, Moritz C. 2011. Recent and rapid speciation with limited morphological disparity in the genus Rattus. *Syst Biol*. 60:188–203.

Salisbury BA, Kim J. 2001. Ancestral state estimation and taxon sampling density. *Syst Biol*. 50:557–564.

Sanmartín I, Meseguer AS. 2016. Extinction in phylogenetics and biogeography: from timetrees to patterns of biotic assemblage. *Front Genet*. 7:35.

Scott GR, Schulte PM, Egginton S, Scott AL, Richards JG, Milsom WK. 2011. Molecular evolution of cytochrome C oxidase underlies high-altitude adaptation in the bar-headed goose. *Mol Biol Evol*. 28:351–363.

Sheldon FH, Lim HC, Moyle RG. 2015. Return to the Malay Archipelago: the biogeography of Sundaic rainforest birds. *J Ornithol*. 156:91–113.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30:1312–1313.

Steinbauer MJ, Field R, Grytnes JA, Trigas P, Ah-Peng C, Attorre F, Birks HJB, Borges PAV, Cardoso P, Chou CH, *et al*. 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob Ecol Biogeogr*. 25:1097–1107.

Stevens RD, Rowe RJ, Badgley C. 2019. Gradients of mammalian biodiversity through space and time. *J Mamm*. 100:1069–1086.

Tamrin NAM, Abdullah MT. 2011. Molecular phylogenetics and systematics of five genera of Malaysian murine rodents (*Maxomys*, *Sundamys*, *Leopoldomys*, *Niviventer* and *Rattus*) inferred from partial mitochondrial cytochrome *c* oxidase subunit I (COI) gene. *J Sci Technol Tropics*. 7:75–86.

Thomson V, Wiewel A, Chinen A, Maryanto I, Sinaga MH, How R, Aplin K, Suzuki H. 2018. A perspective for resolving the systematics of Rattus, the vertebrates with the most influence on human welfare. *Zootaxa*. 4459:431–452.

Wei H, Li F, Wang X, Wang Q, Chen G, Zong H, Chen S. 2017. The characterization of complete mitochondrial genome and phylogenetic relationship within *Rattus* genus (Rodentia: Muridae). *Biochem Syst Ecol*. 71:179–186.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol*. 15:568–573.

Yang Z. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol*. 22:1107–1118.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24:1586–1591.

Yang Z, Dos Reis, M. 2011. Statistical properties of the branch-site test of positive selection. *Mol Biol Evol*. 28:1217–1228.

Yang Z, Kumar S, Nei M. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics*. 141:1641–1650.

Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol*. 19:908–917.

Yu L, Wang X, Ting N, Zhang Y. 2011. Mitogenomic analysis of Chinese snub-nosed monkeys: evidence of positive selection in NADH dehydrogenase genes in high-altitude adaptation. *Mitochondrion*. 11:497–503.

Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*. 30:614–620.

Zhou T, Shen X, Irwin DM, Shen Y, Zhang Y. 2014. Mitogenomic analyses propose positive selection in mitochondrial genes for high-altitude adaptation in galliform birds. *Mitochondrion*. 18:70–75.

Zink RM, Barrowclough GF. 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol Ecol*. 17:2107–2121.