Molecular Phylogenetics and Evolution 118 (2018) 306–317

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Geographic isolation and elevational gradients promote diversification in an endemic shrew on Sulawesi

Ryan A. Eldridge[a,b,*], Anang S. Achmadi[c], Thomas C. Giarla[a,b,1], Kevin C. Rowe[d,e], Jacob A. Esselstyn[a,b]

[a] Museum of Natural Science, Louisiana State University, 119 Foster Hall, Baton Rouge, LA 70803, USA
[b] Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA
[c] Museum Zoologicum Bogoriense, Research Center for Biology, Cibinong, Bogor 16911, Indonesia
[d] Sciences Department, Museum Victoria, Melbourne, VIC 3001, Australia
[e] School of Biosciences, The University of Melbourne, Melbourne, VIC 3001, Australia

ARTICLE INFO

Keywords:
Areas of endemism
Biogeography
Crocidura
Population genetics
Soricidae
Wallacea

ABSTRACT

Phylogeographic research on endemic primates and amphibians inhabiting the Indonesian island of Sulawesi revealed the existence of seven areas of endemism (AoEs). Here, we use phylogenetic and population genetic analyses of one mitochondrial gene and 15 nuclear loci to assess geographic patterns of genetic partitioning in a shrew (*Crocidura elongata*) that is endemic to Sulawesi, but occurs across the island. We uncover substantial genetic diversity in this species both between and within AoEs, but we also identify close relationships between populations residing in different AoEs. One of the earliest divergences within *C. elongata* distinguishes a high-elevation clade from low-elevation clades. In addition, on one mountain, we observe three distinct genetic groups from low, middle, and high elevations, suggesting divergence along a single elevational gradient. In general, our results show that *C. elongata*, like several other Sulawesi endemic taxa, harbors extensive genetic diversity. This diversity is structured in part by known AoE boundaries, but also by elevational gradients and geographic isolation within AoEs.

1. Introduction

Insular Southeast Asia holds an important place in the history of biogeography and indeed, in the history of evolutionary biology (Wallace, 1869, 1876, 1880). Early naturalists documented sharp zones of faunal turnover and remarkable levels of endemism (e.g., Huxley, 1868; Wallace, 1869; Dickerson, 1928; Taylor, 1934). Despite these early discoveries, the fauna of this megabiodiverse region remains poorly known in many respects, and modern expeditions frequently uncover new species of vertebrates, including mammals (e.g., Heaney et al., 2011, 2014; Esselstyn et al., 2012, 2013; Rowe et al., 2016). It is perhaps unsurprising then that the biogeographical factors that shaped the region's current patterns of diversity and endemism remain incompletely understood (Lohman et al., 2011; Brown et al., 2013; Sheldon et al., 2015). Nevertheless, within-island diversification, facilitated by topographic complexity, is emerging as a significant contributor to the phylogenetic diversity of insular faunas (Heaney and Rickart, 1990; Heaney et al., 2011; Esselstyn et al., 2013; Hosner et al., 2013; Toussaint et al., 2014; Justiniano et al., 2015; Demos et al., 2016). Several studies have identified species endemic to single mountains or mountain ranges, with close relatives found in neighboring ranges (e.g., Heaney et al., 2011; Esselstyn et al., 2013; Justiniano et al., 2015). This suggests that unsuitable lowland habitats, competitive exclusion by closely related lowland species, or both represent long-term ecological barriers to dispersal for these taxa. Furthermore, elevational gradients on individual mountains may generate diversity if divergent natural selection promotes reproductive isolation and the formation of parapatric species. However, documented examples of this phenomenon are rare and often debatable (Caro et al., 2013; Demos et al., 2016).

The island of Sulawesi is large (11th largest globally), topographically complex (several mountains exceed 3000 m elevation), and shaped like a "K", with its four peninsulas together comprising a greater land area than the island's central core (Fig. 1). In addition, Sulawesi is a composite island, formed as no fewer than four (and perhaps seven or more) prehistoric landmasses collided between ~25 and 2 Ma (Hall,

* Corresponding author at: 119 Foster Hall, Baton Rouge, LA 70803, USA.
E-mail addresses: therealryaneldridge@gmail.com (R.A. Eldridge), gudelly@gmail.com (A.S. Achmadi), tgiarla@siena.edu (T.C. Giarla), krowe@museum.vic.gov.au (K.C. Rowe), esselstyn@lsu.edu (J.A. Esselstyn).
[1] Present address: Department of Biology, Siena College, Loudonville, NY 12211, USA.

http://dx.doi.org/10.1016/j.ympev.2017.09.018
Received 2 December 2016; Received in revised form 8 August 2017; Accepted 25 September 2017
Available online 27 September 2017
1055-7903/

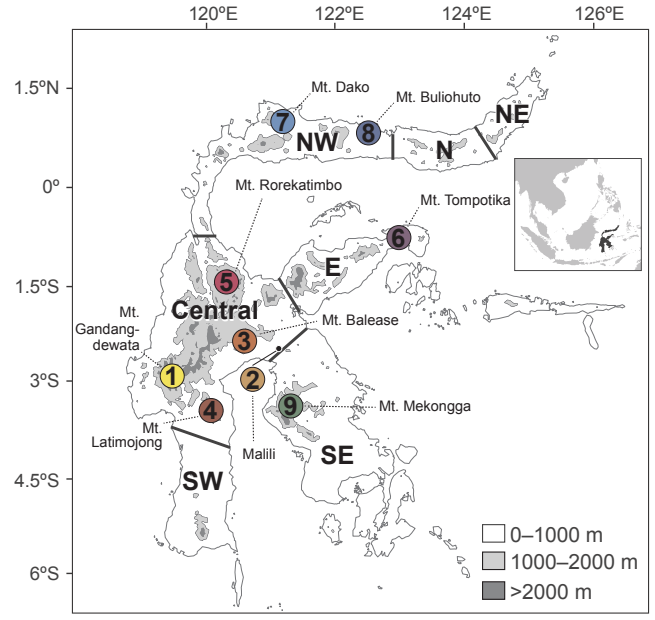

Fig. 1. Map of Sulawesi with black lines corresponding to AoE boundaries. Numbers indicate the locations of sampling localities as follow: 1 = Mt. Gandangdewata, 2 = Malili, 3 = Mt. Balease, 4 = Mt. Latimojong, 5 = Mt. Rorekatimbo, 6 = Mt. Tompotika, 7 = Mt. Dako, 8 = Mt. Buliohuto, and 9 = Mt. Mekongga. Bold text indicates the abbreviated name of each AoE. Inset shows the position of Sulawesi in SE Asia.

2002; Spakman and Hall, 2010; Hall, 2011). Moreover, because the individual landmasses constituting this proto-Sulawesi archipelago were divided further by periods of partial marine inundation (Whitten et al., 2002), the total number of prehistoric islands was even greater. As the centerpiece of the Wallacean biogeographic zone, Sulawesi has been colonized by faunal elements from up to three biologically diverse and distinctive areas: Sahul (Australia and New Guinea), Sunda (Borneo, Java, and Sumatra), and perhaps the Philippines (Stelbrink et al., 2012). Together, Sulawesi's geographical characteristics and geological history provide numerous plausible mechanisms for generating high levels of species diversity and concordant phylogeographic patterns.

Early taxonomic work on Sulawesi's vertebrate fauna showed that most species are endemic (e.g., Miller and Hollister, 1921a, 1921b), but it was not until much later that biologists began to notice geographic patterns of endemism within the island. Fooden (1969) first noted that macaque hybrid zones correspond to natural geographic boundaries separating the peninsulas from the central core, and individual components of the north peninsula from each other. More recent genetic comparisons demonstrated that these boundaries are consistent for various lineages of amphibians, reptiles, and mammals, suggesting that genetic diversity in these lineages has been influenced by shared mechanisms (Evans et al., 2003a, 2003b, 2008; McGuire et al., 2007; Shekelle et al., 2010; Setiadi et al., 2011). These studies defined seven areas of endemism (AoEs; Fig. 1) on Sulawesi that correspond to the island's central core, its east, southeast, and southwest peninsulas, and three areas on the north peninsula. However, some subsequent studies of Sulawesi taxa such as snails (von Rintelen et al., 2014), tarsiers (Driller et al., 2016), and bats (Campbell et al., 2007) have recorded geographic patterns only partially consistent with those observed in the monkeys and toads that defined Sulawesi's original AoEs (Evans et al., 2003a). Evans (2012) reviewed available evidence and noted that in addition to the original AoE boundaries, several lineages are partitioned within the central core area, and the geographic position of these boundaries is shared by some lineages. Because AoEs on Sulawesi have been defined solely by shared phylogeographic patterns, the roles played by various plausible mechanisms (such as paleo-island geography, sea incursions, ecological barriers, and peninsular effects) in the formation of Sulawesi's AoEs are not well understood. As such, the extent to which the AoE paradigm can be applied to the whole of Sulawesi's biota remains unknown, and testing for the AoE pattern in additional organisms will help assess the importance of shared mechanisms of diversification, while also offering improved guidance to conservation efforts.

Whereas AoE patterns are inherently based on geographic distances, which have dominated discussions of within-island diversification on Sulawesi, elevational gradients provide potential alternative mechanisms of generating diversity and local endemism, as has been documented on other islands (e.g. Justiniano et al., 2015; Demos et al., 2016). Because studies that have investigated the AoE pattern on Sulawesi have not generally considered the elevational distribution of their focal lineage or its potential effect on geographic isolation and diversification, it is unclear whether elevational gradients serve as a cause of some reported AoE patterns of diversity, whether they act as an additional cause of diversification, or whether they have little influence on diversity and endemism.

Shrews of the genus *Crocidura* are represented on Sulawesi by six endemic species derived from two independent colonizations (Ruedi, 1995; Ruedi et al., 1998; Esselstyn et al., 2009). Among these shrews, *Crocidura elongata* offers a particularly promising opportunity to assess phylogeographic structure relative to AoE boundaries and along elevational gradients because it is widespread on the island, occurs from near sea level to the highest mountain areas that have been sampled, and its members are easily distinguished from other species by external characters (Ruedi, 1995). Here, we perform multi-locus population genetic, phylogenetic, and species delimitation analyses on *C. elongata* sampled from across Sulawesi to test if the species exhibits genetic structure associated with AoE boundaries, with elevational gradients, with both, or with neither.

2. Materials and methods

2.1. Taxon sampling

Specimens of *Crocidura elongata* were collected from 15 sites contained in nine general localities. These localities are spread across Sulawesi and represent four of the island's seven AoEs. Specimens are vouchered at the Field Museum of Natural History, Chicago (FMNH); Museum Victoria, Melbourne (NMV); Louisiana State University Museum of Natural Science, Baton Rouge (LSUMZ); Museum of Wildlife and Fish Biology, Davis (WFB); and Museum Zoologicum Bogoriense, Bogor (MZB). Our sampling includes multiple localities within the Central Core and Northwest (NW) AoEs, with one locality sampled from each of the Southeast (SE) and East (E) AoEs (Fig. 1; Table 1). All shrews were collected in forested habitats, with the majority taken in relatively undisturbed primary forest. Only the lowest site on Mt. Gandangdewata (also known as Gandadiwata) and the Malili locality consisted of secondary forest and shrubby, regenerating vegetation. Several of the mountains were sampled at multiple elevations (see above), often encompassing the major forest types of lowland tropical evergreen rainforest and montane forest (*sensu* Musser, 2014). We generally found *C. elongata* co-occurring with four to five other species of *Crocidura*.

2.2. DNA sequencing

DNA was extracted from tissue samples using the Qiagen DNeasy protocol or Promega Wizard SV kits. We amplified one mitochondrial gene (cytochrome *b* [cyt *b*]) and 15 nuclear loci via PCR. Of the 15 nuclear loci, eight (apolipoprotein B [ApoB]; brain-derived neurotrophic factor [BDNF]; breast cancer susceptibility gene [BRCA]; growth hormone receptor [GHR]; mast-cell growth factor [MCGF]; prostaglandin E2 receptor [PTGER]; recombination-activating gene

Table 1
Geographic sampling of *C. elongata*; locality numbers indicate geographic locations shown in Fig. 1.

AoE	Locality number	Locality name	Elevation (m)	Number of individuals
Central Core	1a	Mt. Gandangdewata (low-elevation)	170	4
	1b	Mt. Gandangdewata (mid-elevation)	1535–1600	5
	1c	Mt. Gandangdewata (high-elevation)	2200–2600	15
	2	Malili	455	3
	3	Mt. Balease	830–1140	10
	4	Mt. Latimojong	2050	6
	5	Mt. Rorekatimbo	2020	9
East (E)	6	Mt. Tompotika	350	3
Northwest (NW)	7	Mt. Dako	512	3
			1600	6
	8	Mt. Buliohuto	480–580	5
			1200–1390	3
Southeast (SE)	9	Mt. Mekongga	150	1
			1515	1
			1899	1

[RAG-1]; and von Willebrand factor [vWF]) were amplified and sequenced following the protocols of Esselstyn et al. (2013). Primers for the remaining seven loci (ankyrin repeat domain [ANKRD]; cyclin-T [CCNT]; glutamate receptor, N-methyl D-Aspartate [GRIN]; methyl-CpG-binding domain [MBD]; nuclear receptor coactivator [NCOA]; SLIT and NTRK family [SLITRK]; and transmembrane protein [TMEM]) were designed using alignments of coding loci from *Sorex araneus*, *Bos taurus*, and *Mus musculus*, taken from the OrthoMam database (Douzery et al. 2014). General methods of amplification and cycle sequencing for these seven loci followed Esselstyn et al. (2013), but with locus-specific annealing temperatures (Supplementary Table 1) used during PCR cycles. All nuclear loci are entirely protein coding except MCGF, which spans exon-intron boundaries. PCR products were purified using ExoSAP-IT and prepared for sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit and a non-commercial ethanol clean up procedure. PCR products were sequenced in both directions at the Cornell University Biotechnology Resource Center. Sequences were edited by eye in Geneious 7.1.7 (Kearse et al., 2012) and aligned using the MUSCLE algorithm (Edgar, 2004) in Geneious. Alignments were examined by eye and verified to be free of premature stop codons. Degenerate bases were called as heterozygotes by Geneious. Chromatograms associated with each heterozygous site were subsequently checked by eye to avoid scoring sequencing errors as genetic variants.

All nuclear sequences were resolved into statistically probable haplotypes using PHASE 2.1.1 (Stephens et al., 2001; Stephens and Donnelly, 2003). The online application SeqPHASE (Flot, 2010) was used to convert FASTA files to PHASE input files, as well as convert PHASE output back to FASTA format.

2.3. Mitochondrial gene tree and divergence date estimation

We estimated a mitochondrial gene tree for cyt *b* using Bayesian and maximum likelihood (ML) frameworks. An HKY + Γ model of nucleotide substitution was chosen using the Bayesian Information Criteria (BIC) in jModelTest 2.1.7 (Guindon and Gascuel, 2003; Darriba et al., 2012), where we allowed three substitution schemes (JC/F81, K80/HKY, and SYM/GTR) and used a single ML-optimized tree for calculating model likelihoods.

We performed Bayesian phylogenetic analysis in BEAST 2.1.3 (Bouckaert et al., 2014) using a Yule model on the tree shape prior. Two independent runs of 2×10^9 Markov chain Monte Carlo (MCMC) generations were initiated with random starting trees. Samples were drawn every 2×10^5 generations. We estimated divergence dates using a mean substitution rate for cyt *b* of 0.01 substitutions per site per million years (Brown et al., 1979; see also Nabholz et al., 2007), with 95% of the prior probability between 0.001 and 0.02. We assumed a relaxed clock with a lognormal distribution. For the ucld mean prior we assigned a lognormal distribution. We assessed MCMC convergence and selected appropriate burn-in values by examining trace plots of the likelihood and other parameters, while verifying that adequate effective sample sizes (> 200) were obtained, in Tracer v1.6 (Rambaut et al., 2014). Trees produced by BEAST were summarized using TreeAnnotator v2.1.2 after discarding the burn-in (Bouckaert et al., 2014).

We conducted ML phylogenetic analyses using Garli v2.0 (Zwickl, 2006). We performed six independent analyses, each consisting of five replicates of 5×10^6 generations. Outside of specifying the model of nucleotide substitution, we kept all settings at their defaults. We used *Crocidura batakorum*, a Philippine endemic species inferred as sister to the main radiation of Sulawesi shrews (Esselstyn et al., 2009), as the outgroup. Bootstrap support values were obtained from 1000 bootstrap replicates and mapped onto the ML tree using Sumtrees v4.1.0 in DendroPy v4.1.0 (Sukumaran and Holder, 2010; Sukumaran and Holder, 2016).

2.4. Phylogenetic analysis of concatenated nuclear genes

We conducted a concatenated analysis using MrBayes v3.2.5 using all 15 unphased nuclear loci. This analysis was partitioned, with distinct models of nucleotide substitution chosen for each unphased gene using the BIC in jModelTest 2.1.7 (Supplementary Table 2; Guindon and Gascuel, 2003; Darriba et al., 2012) as above. We ran two independent MCMC analyses, each with four chains. Both analyses were run for 10^6 generations, drawing samples every 1000 generations, and we discarded the first 25% of generations as burn-in. We used Tracer v1.6 (Rambaut et al., 2014) to confirm MCMC stationarity and ensure that effective sample sizes > 200 were obtained. *Crocidura nigripes*, a distantly related species (Ruedi et al., 1998; Esselstyn et al., 2009), was used as the outgroup.

2.5. Assessment of geographic patterns of diversity

We used Mantel tests (Mantel, 1967) in the *ade4* package (Dray and Dufour, 2007) in R (R Core Team, 2016) to test for spatial autocorrelation between mitochondrial genetic distance (calculated with MEGAv6.0; Tamura et al., 2013) and both straight-line geographic distance and elevational difference. We based our results on 10^4 permutations.

We used the Bayesian approach implemented in Structure v2.3 (Pritchard et al., 2000) to cluster individuals according to genetic similarity and thereby ascertain geographic patterns within *C. elongata*. We used Structure's admixture model and assumed that allele frequencies were correlated between clusters (Falush et al., 2003). We chose not to use collection locality as a prior. Initial analyses, including the full set of 76 individuals and 15 phased nuclear loci, were run for 5×10^4 generations with 10^4 generations discarded as burn-into explore the potential number of clusters (K) at values of 1–10 (10 replicates each). Continuing to use the full dataset, we subsequently ran Structure longer (7.5×10^5 generations, 5×10^5 generations discarded as burn-in), 20 times with a narrower range of K values (1 −5) based on results from our initial analyses. As finer-scale population genetic partitions can be overshadowed by higher levels of genetic structuring (Evanno et al., 2005), additional runs of Structure were performed in a hierarchical manner with subsets of the original group of individuals included, similar to the approach used by Vähä et al. (2007). Following the identification of the highest-level split of clusters in the initial Structure analyses ("Level I"), Structure was run independently on each of the resultant clusters ("Level II"). This logic was extended to a third round ("Level III"), by which point most individuals

had sorted into clusters corresponding to collection localities.

In the majority of cases, we determined the best estimate of the number of clusters (K) using the method of Evanno et al. (2005) as implemented in the online application Structure Harvester (Earl and vonHoldt, 2012). Estimates of K are made independently of any direct consideration of biological reality and therefore blind acceptance of estimated K values has been discouraged (Meirmans, 2015). In this study, there was a single instance (Level IIIc) in which the K identified using the method of Evanno et al. (2005) seemed dubious on biological grounds (see Results).

All runs in Levels II and III were initially conducted for 5×10^4 generations (burn-in of 10^4 generations; 20 times for each K of 1–10), whereupon cluster assignment coefficients were compared between replicates to ensure consistency among runs at each K. The two analyses initially produced inconsistent results (3b and 3c) and were re-run for longer durations (3b was run for 7.5×10^5 generations with a burn-in of 5×10^5, 20 times for each K 1–10; after use of those settings in 3c continued to produce inconsistent results, we re-ran 3c for 2×10^6 generations with a burn-in of 7.5×10^5, 20 times for each K 1–10).

2.6. Delimiting diversity

Because our phylogenetic analyses showed substantial genetic divergence between populations, we tested the hypothesis that these populations represented independent evolutionary lineages (and possibly undiagnosed species) under the multispecies coalescent using BPP v3.1 (Yang and Rannala, 2010, 2014; Rannala and Yang, 2013; Yang, 2015). Although it is sometimes touted as useful for such purposes, BPP does not perform species delimitation *per se* (Sukumaran and Knowles, 2017). Instead, it delimits units of population structure that taxonomists can use as one type of evidence for the diagnosis of species. As such, we used BPP to provide a maximal estimate of the potential number of species. We included all phased nuclear loci (but not cyt *b*) in this analysis. We divided individuals among 12 potential species. Each sampling locality was represented by one potential species, with the exceptions of Mts. Gandangdewata and Rorekatimbo, which were divided into three and two potential species, respectively, on the basis of our observations from our mitochondrial gene tree analyses (see Results). We also excluded one Mt. Mekongga individual (WFB8203) from the analysis because it appeared alone at the end of a long branch in the cyt *b* tree and we wanted to avoid artifacts of small sample size (initial runs including that individual produced inconsistent results). We did not specify a guide tree, allowing the program to test for population structure under all possible topologies. Each BPP analysis was run for 10^5 steps after discarding a burn-in of 8000 steps. As prior selection influences the posterior probabilities for models (Yang and Rannala, 2010), BPP was run with three separate combinations of priors for ancestral population size (θ) and root age (τ_0), mirroring those applied by Leaché and Fujita (2010). We ran BPP twice with each combination of priors to ensure convergence and assess the effects of priors.

2.7. Gene flow estimation

We used the Bayesian coalescent model implemented in MIGRATE-N v3.6 (Beerli and Felsenstein, 2001; Beerli, 2006) to estimate rates of gene flow between various clusters and groups of clusters identified in our Structure analyses. Because we wanted to assess patterns of gene flow among different sets of populations, we conducted three rounds of analyses.

In each of the three analyses, we used (1) phased nuclear loci; (2) MIGRATE-N's built-in FST-like calculation (Beerli and Felsenstein, 1999) to generate initial parameter values; (3) a constant mutation rate across loci; (4) a finite sites mutation model; (5) an MCMC heating scheme with four heated chains (allowing swapping between chains); (6) uniform prior distributions; (7) a burn-in of 30,000 trees per chain; (8) MCMC with 8000 total steps and a sampling increment of 200 (total trees sampled per replicate thus equaled 1.6×10^5) for each of four replicates; and (9) posterior modes as our parameter estimates. In analysis A we allowed the mutation-scaled effective immigration rates, *M*, to be asymmetrical, while in analyses B and C, these rates were symmetrical. Additional, analysis-specific settings are given in Supplementary Table 3.

3. Results

3.1. Alignment contents and sequence variation

All new DNA sequences were published as GenBank Accessions KY771329–KY772426 (Supplementary Appendix). Our cyt *b* alignment included 46 unique haplotypes from 75 individuals and was 1131 bases long with 25.6% missing data (841.5 bases sampled per individual). The concatenated alignment of all unphased nuclear loci was 9480 bases long with 18.8% missing data. Sequence lengths for phased nuclear loci were 450–793 bases, with the number of unique alleles ranging from 19 to 64 (Supplementary Table 4). The percentage of missing characters per locus ranged from 0.3% to 26.3%. Uncorrected mitochondrial *p*-distances ranged from 0.017 to 0.121 between sampling sites, while intra-site distances were ≤ 0.012 (Table 2).

3.2. Mitochondrial tree and divergence time estimation

In our Bayesian reconstruction of the mitochondrial gene tree, the earliest divergence separated the NW AoE (Mts. Dako and Buliohuto) from all other localities (Fig. 2). Within the latter clade, the first divergence divided Mts. Latimojong and Rorekatimbo, plus the mid- and high-elevation sites on Mt. Gandangdewata, from the other localities. Bayesian posterior probabilities for many nodes exceeded 0.95, although several nodes had considerably less support. For example, the

Table 2
Uncorrected mitochondrial *p*-distances calculated in MEGA. Numbers in headings correspond to locality numbers (see Fig. 1). 1a, 1b, and 1c denote the low-, mid-, and high-elevation groups of individuals sampled from Mt. Gandangdewata; 8203 refers to WFB8203 from low elevation on Mt. Mekongga.

Locality	1a	1b	1c	2	3	4	5	6	7	8	9	8203
1a	0.003											
1b	0.095	0.001										
1c	0.102	0.041	0.002									
2	0.017	0.111	0.105	0								
3	0.012	0.100	0.100	0.011	0.004							
4	0.099	0.017	0.040	0.118	0.109	0.002						
5	0.093	0.031	0.031	0.110	0.103	0.030	0.012					
6	0.031	0.089	0.089	0.029	0.025	0.097	0.094	0.004				
7	0.100	0.106	0.116	0.117	0.105	0.110	0.109	0.100	0.007			
8	0.098	0.098	0.099	0.114	0.102	0.102	0.097	0.100	0.047	0.008		
9	0.045	0.110	0.098	0.050	0.048	0.122	0.109	0.048	0.118	0.103	0.00	
8203	0.037	0.114	0.106	0.044	0.041	0.119	0.116	0.031	0.115	0.107	0.045	n/a

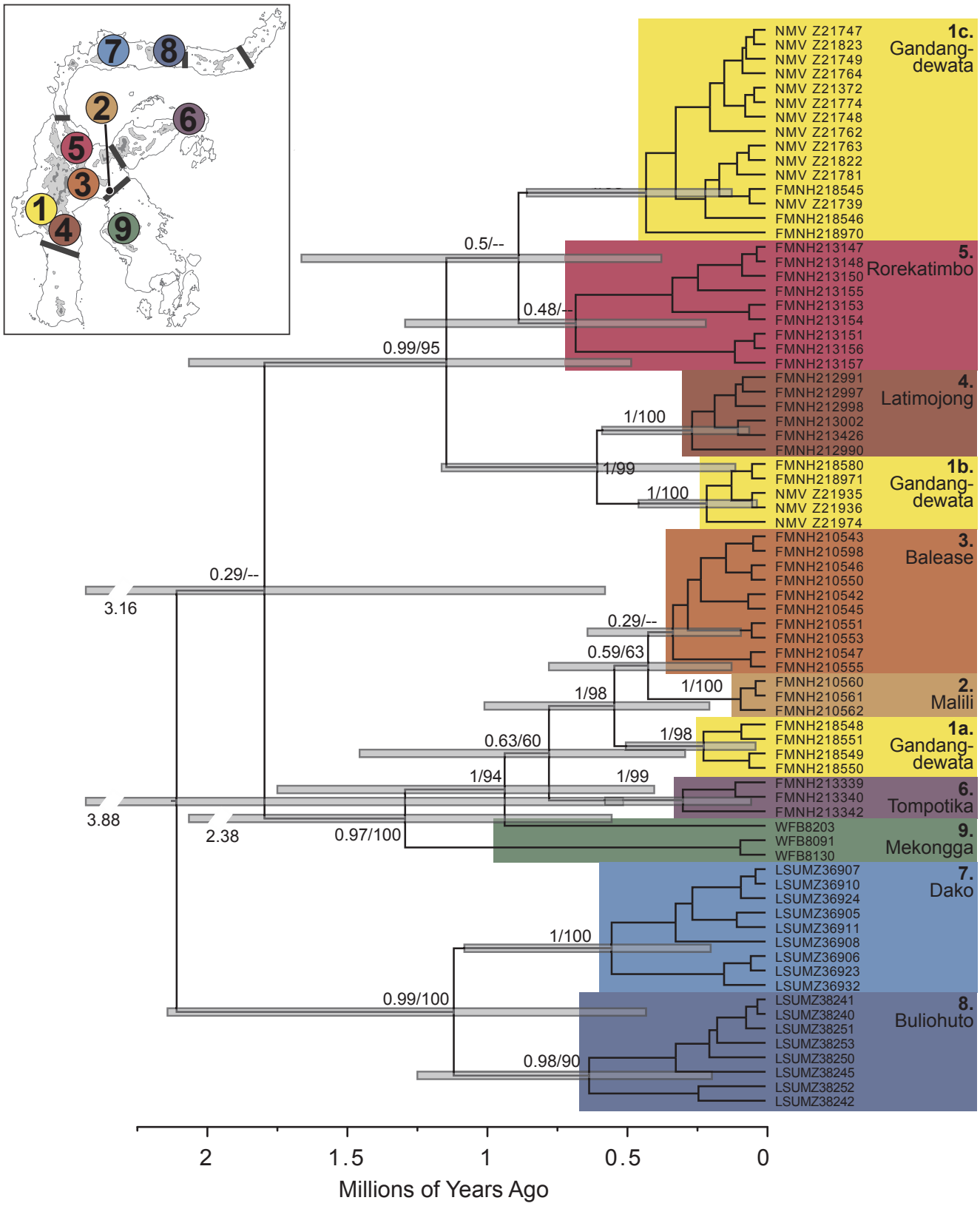

Fig. 2. Bayesian mitochondrial gene tree generated in BEAST illustrating the relationships between individuals of *Crocidura elongata*. Numbered boxes correspond to localities (see inset; 1a, 1b, and 1c denote the low-, mid-, and high-elevation groups of individuals, respectively, from Mt. Gandangdewata), as well as entities supported as species in our BPP analysis (see Table 3). Nodal support is given as [Bayesian posterior probability]/[maximum likelihood bootstrap value] and 95% highest posterior densities (HPD) on node age estimates (only for nodes at or above the population level) are represented by grey bars. Broken bars are truncated with the upper limit of the HPD written below the break.

node grouping all sites except Mts. Dako and Buliohuto received a posterior probability of 0.29. Most of the sampled general localities formed clades. However, the Mt. Gandangdewata and Mt. Mekongga samples were both paraphyletic. On Mt. Gandangdewata, samples grouped according to elevation (the low-, mid-, and high-elevation samples each formed a clade [mean intergroup *p*-distance was 0.067]); on Mt. Mekongga, one individual (WFB8203) from low elevation (150 m) was more closely related to samples from other mountains than to the other two individuals sampled from higher elevations (1515 and 1899 m; *p*-distance of 0.045). Individuals from Mt. Rorekatimbo were monophyletic on the gene tree, but two clades were separated by relatively long branches [*p*-distance of 0.021] and the node joining them was poorly supported (posterior probability of 0.48). However, individuals from the two Mt. Rorekatimbo mitochondrial clades were sampled sympatrically. Our estimate of the age of the oldest divergence within *C. elongata* had a mean of 2.10 Ma with a 95% HPD interval of 0.97–3.88 Ma (Fig. 2).

Our ML tree (Supplementary Fig. 1) largely supported the relationships inferred in the Bayesian analysis. However, individuals collected from Mt. Rorekatimbo were paraphyletic with respect to individuals from Mts. Gandangdewata and Latimojong, but with relatively weak support (bootstrap = 0.61). The two clades recovered from Mt. Rorekatimbo were identical to those recovered in the Bayesian analysis. Additionally, the clade of individuals collected from Malili was nested among the Mt. Balease specimens.

3.3. Phylogenetic analysis of concatenated nuclear genes

The relationships and groupings supported in our concatenated analysis (Fig. 3) were largely consistent with those supported by the mitochondrial gene tree analyses. For instance, the base of the tree is characterized by a polytomy between the NW AoE samples (Mts. Dako and Buliohuto), some of the Central Core AoE samples (Mts. Latimojong and Rorekatimbo, and the upper two sites on Mt. Gandangdewata), and the remaining sites in the Central Core (Mt. Balease, Malili, and the lowest site on Mt. Gandangdewata) combined with the E (Mt. Tompotika) and SE AoEs (Mt. Mekongga). In this analysis, we found that populations from Mt. Tompotika, low- and high-elevation Mt. Gandangdewata, Mt. Mekongga, Mt. Buliohuto, Mt. Dako, Mt. Rorekatimbo, and Mt. Latimojong were each monophyletic, but samples from the other localities (Balease, Malili, and mid-elevation Mt. Gandangdewata) were paraphyletic.

3.4. Assessment of geographic patterns of diversity

Our Mantel test indicated that both geographic distance (Mantel r statistic = 0.58; $p \ll 0.001$) and elevational difference (Mantel r statistic = 0.33; $p \ll 0.001$) exhibit significant correlation with mitochondrial genetic distance among individuals of *C. elongata*.

The hierarchical implementation of our Structure analyses revealed genetic partitions associated with both broad- and fine-scale geography. The initial set of all individuals (Fig. 4, Level I) was estimated by the Evanno et al. (2005) method to comprise two clusters with no evidence of admixture. These two clusters corresponded with the initial divergence in our ML estimate of the mitochondrial gene tree (Supplementary Fig. 1). One cluster consisted of all individuals from low-elevation Mt. Gandangdewata (Locality 1a), Mt. Balease (3), Malili (2), Mt. Tompotika (6), Mt. Dako (7), Mt. Buliohuto (8), and Mt. Mekongga (9); the other of all individuals from mid-high elevation Mt. Gandangdewata (1b and 1c), Mt. Latimojong (4), and Mt. Rorekatimbo (5). We observed a distinct elevational difference between the two clusters (Fig. 5a), with the mean elevation of the latter cluster (2169 m, range 1535–2600 m) nearly 1300 m above that of the former (864 m, range 150–1899 m).

Level II of the Structure analysis produced a similar pattern, with each of the two clusters identified in Level I splitting into two clusters with no signal of admixture (Fig. 4, Level II): individuals from low-elevation Mt. Gandangdewata (Locality 1a), Mt. Balease (3), Malili (2), Mt. Mekongga (9), and Mt. Tompotika (6) were split from those from the NW AoE (Mt. Dako (7) and Mt. Buliohuto (8)), and the individuals from Mts. Latimojong (4) and Rorekatimbo (5), plus mid-elevation Mt. Gandangdewata (1b), were split from high-elevation Mt. Gandangdewata (1c). In the latter split, we once again observed a distinct elevational difference between the two resultant clusters (Fig. 5b), with a mean elevation of 1917.5 (range 1535–2050) for the cluster comprising individuals from Mts. Latimojong, Rorekatimbo, and mid-elevation Gandangdewata, versus a mean elevation of 2366.7 (range 2200–2600) for the high-elevation Mt. Gandangdewata cluster.

Each cluster identified in Level II of the analysis was subjected to its own Structure analysis. The analysis conducted on the cluster comprising the higher-elevation group of Mt. Gandangdewata (Locality 1c) specimens (identified in Level IIa), assigned each individual by nearly equal amounts to each of seven clusters; we regarded these results as biologically unrealistic and discarded them, leaving three Level III analyses.

All individuals in Levels IIIa and IIIb sorted into clusters by locality with the exception that the mid-elevation cluster of Mt. Gandangdewata (Locality 1b) specimens grouped with the Mt. Rorekatimbo (5) specimens. In Level IIIc of the analysis, the most appropriate K was estimated as five by the Evanno et al. (2005) method. Individuals from low-elevation Mt. Gandangdewata (1a) and Mt. Tompotika (6) formed their own clusters, as did two individuals from Mt. Mekongga (9). The remaining individual from Mt. Mekongga (WFB8203) and all individuals from Malili (2) and Mt. Balease (3) were partially assigned, to varying degrees, to two clusters (WFB8203 and the Malili specimens always received the four highest membership coefficients for one of the two clusters). However, because of the extensive degree of admixture observed between the two clusters, the biological basis for this number of clusters may be weak. Lowering K to 4 produced somewhat inconsistent results, but in most replicates Mt. Balease individuals were consolidated into one cluster, with WFB8203 and the Malili individuals' assignments split between the Mt. Balease and Mt. Tompotika clusters. Because the individuals from low-elevation Mt. Gandangdewata and the two individuals from Mt. Mekongga form their own clusters among different estimates of K, we consider it sensible to view these as discrete groups. However, the extent to which WFB8203, as well as individuals from Mt. Tompotika, Mt. Balease, and Malili, represent defined entities is less clear.

3.5. Delimiting diversity

Our BPP analyses distinguished most of the potential species we specified based on our mtDNA topology (but see Sukumaran and Knowles, 2017). Across every replicate, with each combination of priors, all potential species other than the two from Mt. Rorekatimbo were delimited with posterior probabilities of 1 (Table 3). All replicates assigned a much higher posterior probability to a single species consisting of all individuals from Mt. Rorekatimbo than to a division of these individuals into two species. In other words, each locality was supported as containing at least one distinct species.

3.6. Gene flow estimation

For the most part, we estimated minimal gene flow among the Structure-defined populations included in our MIGRATE-N analyses (Fig. 6; Supplementary Table 3), but migration rates were higher among populations (Fig. 6) that were defined in Level III Structure analyses (Fig. 4). In general, however, convergence in the posterior distribution for *M* was poor when migration rates were higher. The migration rates estimated between the NW AoE (represented by Mts. Dako (Locality 7) and Buliohuto (8)) and populations elsewhere on the island never exceeded 0.02 immigrants per generation (Fig. 6A). Similarly, rates were

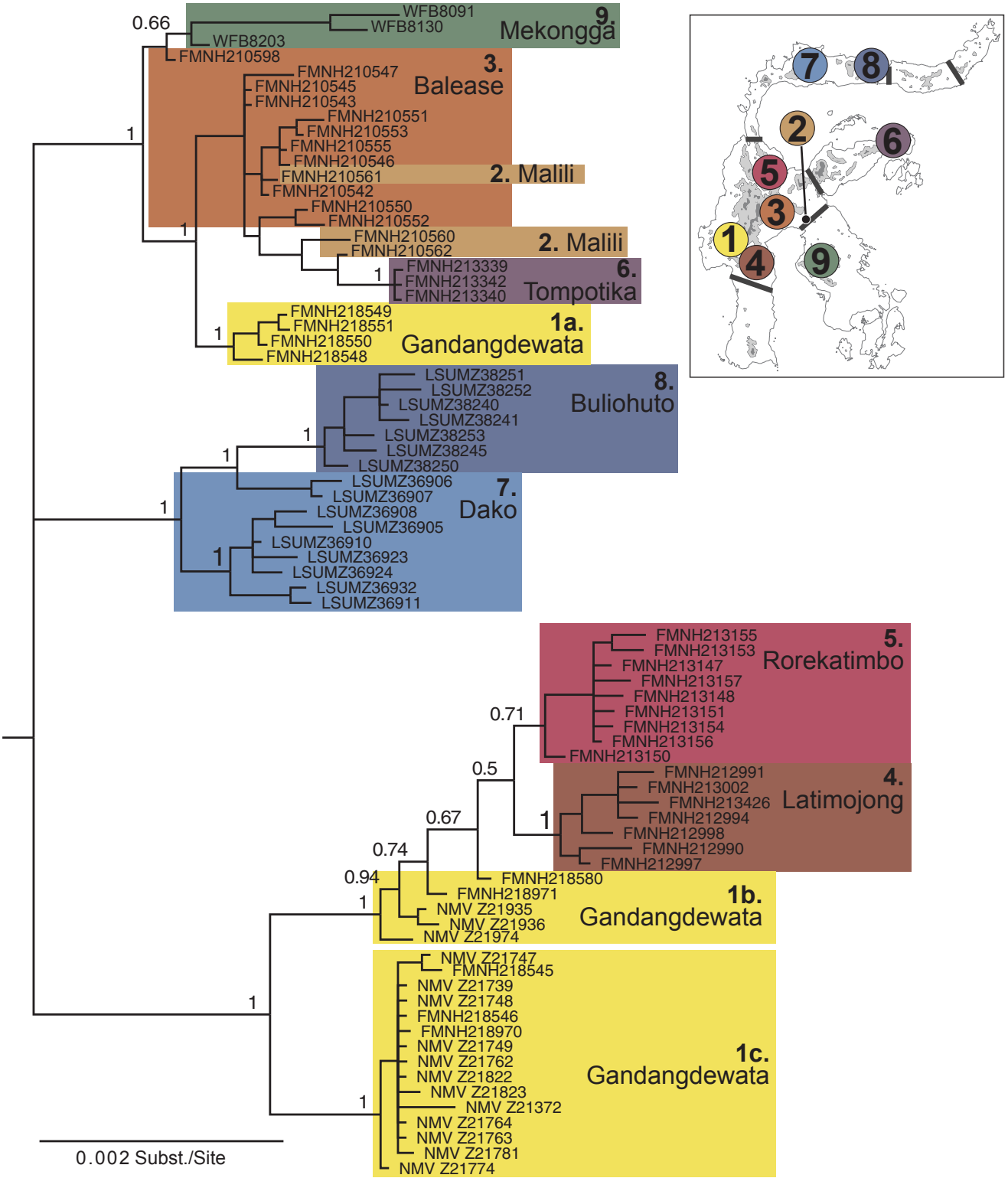

Fig. 3. Concatenated nuclear gene tree generated in MrBayes illustrating the relationships between individuals of *Crocidura elongata*. The outgroup in the analysis, *Crocidura nigripes*, is not shown in this figure. Numbered boxes correspond to localities (see inset; 1a, 1b, and 1c denote the low-, mid-, and high-elevation groups of individuals, respectively, from Mt. Gandangdewata).

low between localities included in Level IIa of the Structure analysis and those included in Level IIIc (≤0.01 immigrants per generation; Fig. 6B and C). Migration rates within each of those two groups were generally much higher. Among the populations defined in Level IIIc of the Structure analysis (Mt. Balease (3)/Malili (2)/WFB8203 (partial 9), low-elevation Mt. Gandangdewata (1a), and Mt. Tompotika (6)), migration rates ranged from 0.171 to 0.948 migrants per generation, though we note that the posterior distribution for *M* between low-elevation Gandangdewata and Mt. Tompotika failed to converge. Likewise, migration rates between the populations defined in Level IIIa in the Structure analysis (Mt. Latimojong (4), Mt. Rorekatimbo (5), and mid-elevation Mt. Gandangdewata (1b)) ranged from 0.208 to 0.691

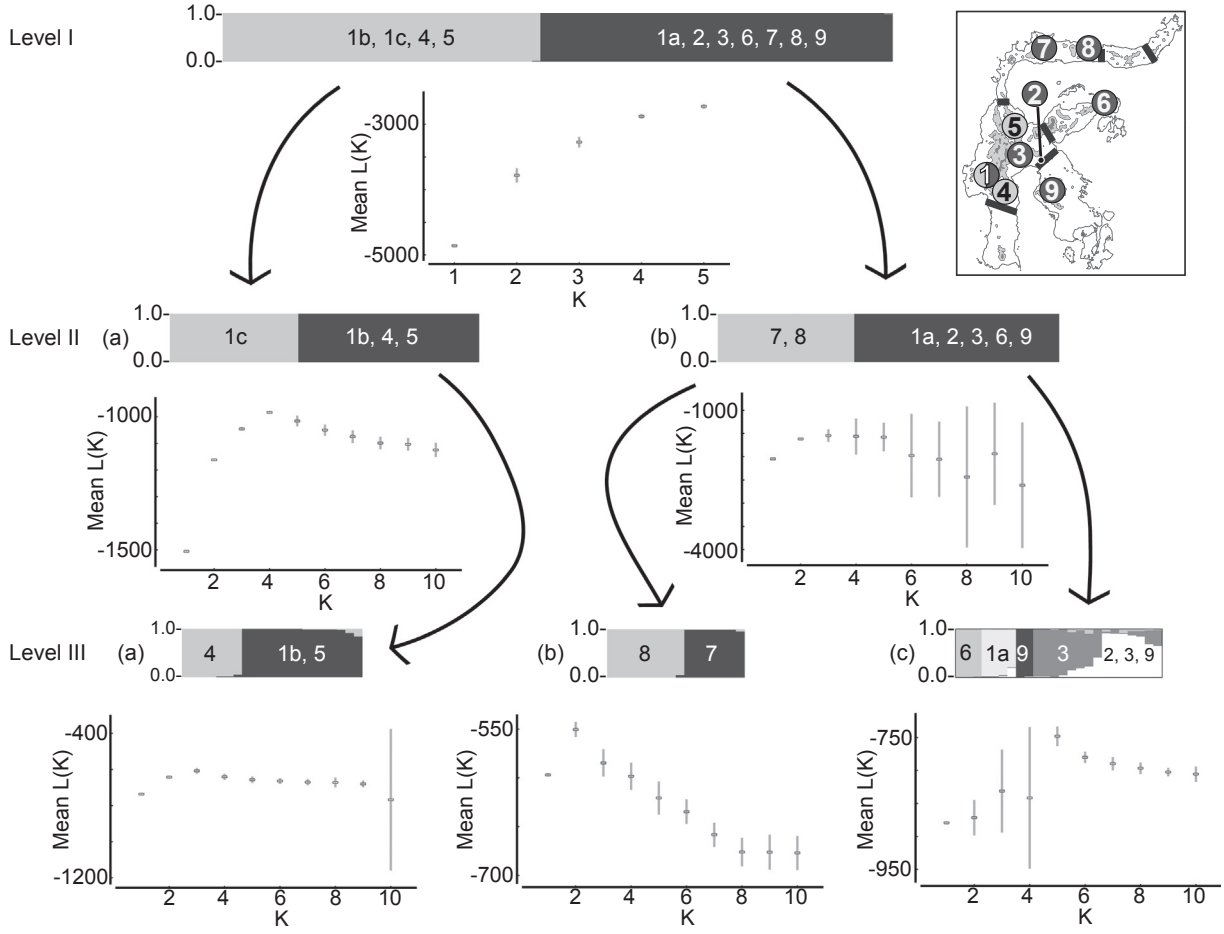

Fig. 4. Summary and likelihood plots from our hierarchical implementation of Structure. The x-axes on the summary plots are representative of individual specimens. A specific shade on a summary plot represents a cluster assignment; an individual's cluster assignment coefficient, the probability an individual belongs to a specific cluster, is expressed by the height of that shade. Numbers within summary plots indicate localities from which individuals constituting the numbered cluster were sampled (see inset). The analysis labeled "Level I" consists of all individuals from all localities. Arrows point from a cluster assignment in a previous analysis to an analysis including only the individuals assigned to that cluster. Below each summary plot is a plot showing the likelihoods (mean L(K)) for each number of cluster assignments (K) tested in that analysis. Error bars indicate standard deviation.

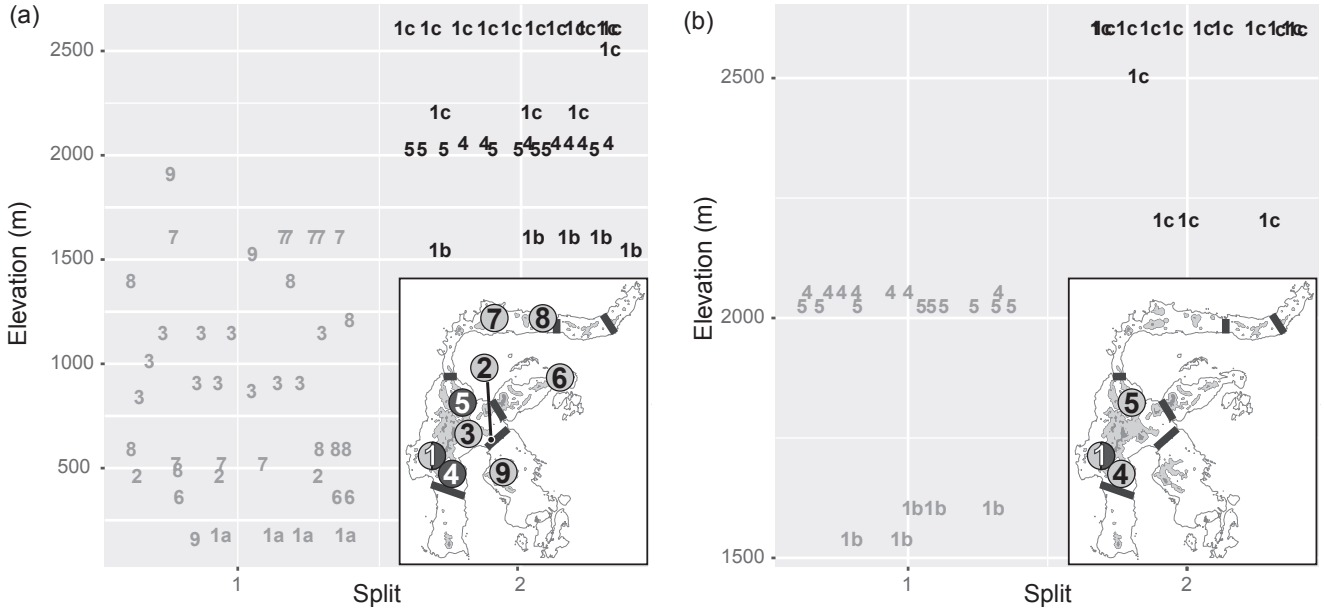

Fig. 5. Elevational comparison of clusters identified in Structure analyses. Numbered points represent individuals identified with their locality number (see insets). Individuals are grouped according to their split assignment in a specific Structure analysis; (a) illustrates the elevational difference between the two clusters identified in the Level I Structure analysis, and (b) illustrates the elevational difference between the two clusters identified in the Level IIa Structure analysis.

Table 3
Posterior probabilities of species delimitation under three combinations of ancestral population size and root age prior distributions. Each probability is averaged across two replicates. Locality names correspond to those listed in Fig. 1; 1a, 1b, and 1c denote the low-, mid-, and high-elevation groups of individuals from Mt. Gandangdewata. 5a and 5b are the two mitochondrial clades from Mt. Rorekatimbo.

Locality (locality number)	Mean posterior probability (± SD)		
	Θ, τ_0 = (1,10)	Θ, τ_0 = (2,2000)	θ = (1,10); τ_0 = (2,2000)
Mt. Gandangdewata (1a)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Gandangdewata (1b)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Gandangdewata (1c)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Malili (2)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Balease (3)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Latimojong (4)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Rorekatimbo (5a)	0.128 ± 0.118	0.092 ± 0.029	0.0005 ± 0.000
Mt. Rorekatimbo (5b)	0.128 ± 0.118	0.092 ± 0.029	0.0005 ± 0.000
Mt. Rorekatimbo (5b) + Mt. Rorekatimbo (5a)	0.872 ± 0.118	0.908 ± 0.029	0.999 ± 0.000
Mt. Tompotika (6)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Dako (7)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Buliohuto (8)	1.00 ± 0	1.00 ± 0	1.00 ± 0
Mt. Mekongga (9)	1.00 ± 0	1.00 ± 0	1.00 ± 0

immigrants per generation (Fig. 6C). The high-elevation Mt. Gandangdewata group (1c), however, appears largely isolated from all other populations, as the migration rate from this group to the mid-elevation group from the same mountain was 0.03 immigrants per generation (Fig. 6C) with no other migration rates to or from other groups exceeding 0.013 migrants per generation.

4. Discussion and conclusions

Our analyses identified deep divergences (relative to those expected within a single species) among mitochondrial and multi-locus clades of *Crocidura elongata*, demonstrating that the current single-species classification obscures significant genetic structure in this taxon. Although our results were not entirely concordant, the analyses agreed in many respects. For instance, clusters identified in the Structure analyses (Fig. 4) generally corresponded to mitochondrial clades and, with a few exceptions, to clades in our concatenated nuclear tree. Likewise, our estimates of migration rates between groups identified by Structure were generally lower than those estimated within groups identified by Structure (Fig. 6; Supplementary Table 3).

In our Bayesian mitochondrial gene tree analysis, we estimated the oldest divergence within *C. elongata* at < 4 Ma (95% HPD 0.97–3.88 Ma). Thus, although our estimates are rough, the oldest divergence within the species appears to post-date the fusion of all components of the proto-Sulawesi archipelago (ca. 5 Ma), with the exception of a fragment of the SE AoE that likely coalesced during the last 2 My (Hall, 2002). Therefore, we consider it unlikely that *C. elongata* colonized and diversified on multiple islands in proto Sulawesi. This finding contrasts with inferences for some other mammals, which may have arrived and diversified on Sulawesi much earlier (e.g., squirrels: Hawkins et al., 2016; tarsiers: Driller et al., 2016).

Although our geographic sampling is insufficient to thoroughly test the limits of all AoEs, the pattern of genetic partitioning among *C. elongata* groups is only partially consistent with the patterns observed in monkeys and toads (Evans et al., 2003a). For instance, consistent with the AoE paradigm, two localities representing the NW AoE (Mts. Dako and Buliohuto) grouped together as an early-diverging clade across our phylogenetic analyses (Figs. 2 and 3; Supplementary Fig. 1). However, some of our other observations are inconsistent with AoE geography dominating patterns of genetic diversity. For example, we observed both extensive divergence within AoEs (e.g., between samples from mid-high elevation Mt. Gandangdewata and Mt. Balease) and little divergence between populations in different AoEs (e.g., Mts. Mekongga and Balease; Table 2). Because *C. elongata* is common in lowland forests near sea level, we did not expect to find *p*-distances of ~10% in mitochondrial DNA between adjacent mountains located within a single AoE. However, other studies have identified similarly deep genetic differences between intraspecific populations in the Central Core AoE (e.g., Campbell et al., 2007; McGuire et al., 2007; Setiadi et al., 2011), which suggests that treating this AoE as a single, relatively uniform biological region may obscure genetic diversity in many taxa. Indeed, expanding the definition of AoEs on Sulawesi to incorporate, for instance, the Palu-Koro fault as an additional boundary (Evans, 2012) would better accord with our observed divergences between high-elevation Mt. Gandangdewata and Mt. Rorekatimbo, as well as between the Mt. Balease/Malili clade and low-elevation Mt. Gandangdewata. In contrast to these divergences, we estimated high rates of gene flow into and out of the E AoE (Fig. 6), indicating that, at least in the case of this area and this taxon, the boundary is permeable. In summary, the AoE pattern as defined by monkeys and toads does not appear to be dominant in *C. elongata*. Rather, our observations are better captured by an AoE definition that includes the ancient plate boundaries illustrated by Evans (2012) combined with elevation as a major factor that contributes to diversification.

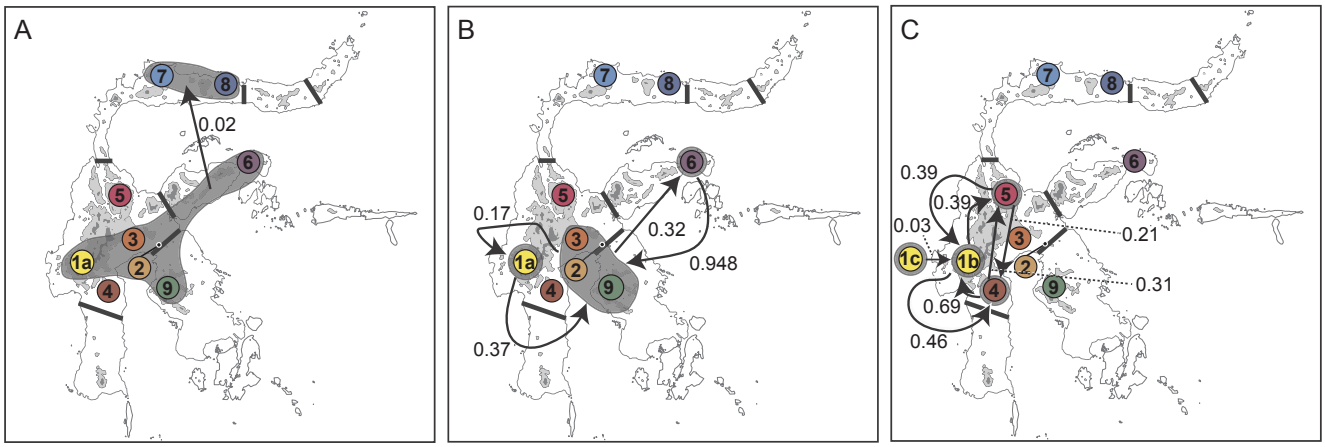

Fig. 6. Illustration of all gene flow rates ≥0.02 immigrants per generation estimated in Analyses 1 (panel A), 2 (panel B), and 3 (panel C) in MIGRATE-N. All sampling localities (identified by their locality number; see Fig. 1) within a shaded area were treated as one "population" in that analysis. Arrows signify gene flow direction. 1a, 1b, and 1c denote the low-, mid-, and high-elevation groups of individuals, respectively, from Mt. Gandangdewata; 9* refers to WFB8203 from low elevation on Mt. Mekongga.

Our results show that genetically distinct populations occur along a single elevational gradient, and that elevational differences across the island are correlated with genetic distance. On Mt. Gandangdewata, we obtained specimens from 170 m, 1535–1600 m, and 2200–2600 m, and we found that samples from each elevational band are as, or more, closely related to populations from other mountains than to those from the neighboring elevational bands on the same mountain. This pattern is evident in both our mitochondrial gene tree (Fig. 2) and our concatenated nuclear gene tree (Fig. 3). In addition, the initial split inferred by our Structure analyses (and weakly supported by our maximum-likelihood mtDNA gene tree) distinguishes individuals collected at high elevations from those taken at low elevations, with the boundary occurring around 1500–1600 m (Figs. 4, 5). Consistent with our Level I Structure result, our estimation of gene-flow rates between these two Structure populations was near 0 (Fig. 6; Supp. Table 3). On Sulawesi, the middle elevations around 1500 m are typically where natural habitats transition from tall-canopy evergreen rainforests at low elevation to short-canopy, much wetter montane forests at higher elevations where oaks and chestnuts are common (Whitmore, 1984; Musser et al., 2010). Elevation has long been posited as a factor in the diversification of terrestrial vertebrates of Southeast Asia (Heaney and Rickart, 1990; Ruedi, 1995; Musser et al., 2010; Justiniano et al., 2015) due to its potential to generate disparate ecological selection along a single gradient, or by isolating montane populations on distant mountain peaks. While examples of the latter are relatively common (e.g., murine rodents: Heaney et al., 2011; Justiniano et al., 2015; *Aethopyga* sunbirds: Hosner et al., 2013), our results from Mt. Gandangdewata provide a rare example of divergence along a single gradient, and the first example from Sulawesi as far as we are aware. That said, divergence along a single elevational gradient is probably not pervasive in *C. elongata*. We sampled a few other mountains at multiple elevational bands, which included both lowland and montane forest on Mts. Dako (sampled at 512 m and 1600 m), Buliohuto (480–580 m and 1200–1390 m), and Mekongga (150, 1515, 1899 m), but we found no clear evidence of genetic isolation within them. On Mt. Mekongga, we did find some mitochondrial difference between the lowest elevation sample and the others, but our sample size is severely limited (only three individuals from Mt. Mekongga) and we did not find many differences in nuclear DNA (Figs. 3 and 4). Unfortunately, since the advent of genetic sample collection, small mammal communities have not been thoroughly sampled across a complete elevational gradient on any Sulawesi mountain. Even on Mt. Gandangdewata where our sampling is most comprehensive, elevations between 200 and 1500 m and between 2600 and 3400 m remain entirely unexplored. As such, collection of additional voucher specimens is sorely needed to further explore the role of elevational gradients in generating diversity among Sulawesi's rich small mammal fauna. In particular, increased sampling would allow a test of whether the genetic divergences we observed along Mt. Gandangdewata's elevational gradient are the result of adaptation to local environments.

Crocidura elongata, as defined by current classification (Miller and Hollister, 1921a, 1921b; Ruedi, 1995), is a complex of well-defined phylogenetic lineages. Our current sampling indicates that these lineages are not partitioned exclusively by previously hypothesized AoE boundaries, but that diversity is instead associated with a combination of AoE boundaries, geographic distance, isolation among mountain ranges, and in one case, an elevational gradient within a single mountain. As such, biodiversity conservation strategies designed exclusively around AoEs will not preserve the full genetic diversity found in this, and perhaps other, species. While we posit that a combination of ecological gradients (elevation) and geographic isolation generated genetic divergence in *C. elongata*, denser geographic sampling is needed to test the relative importance of these and other mechanisms in *C. elongata* and to fully understand how pervasive these patterns are among Sulawesi's endemic species. Nevertheless, this study joins a growing body of work (e.g., Rickart et al., 2005; Esselstyn et al., 2013; Hosner et al., 2013; Toussaint et al., 2014; Justiniano et al., 2015; Demos et al., 2016) that recognizes the importance of geography and topography in generating within-island diversification and hence, maintaining biodiversity on the larger islands of the Indo-Australian Archipelago.

Acknowledgments

This work was funded by the National Science Foundation (DEB-1343517, DEB-1441634, and OISE-0965856) and National Geographic Society (9025-11). We thank Karen Rowe, James Patton, Carol Patton, Tri Haryoko, Heru Handika, Mardin Sarkam, and Alan Hitch for their contributions to specimen collection. We thank the Indonesian Ministry of Research and Technology (KEMENRISTEK), Lore Lindu National Park, and the BKSDA offices in Sulawesi Utara and Sulawesi Selatan for providing permits. The Field Museum of Natural History, Museum Victoria, Louisiana State University Museum of Natural Science, University of California at Davis Museum of Wildlife and Fish Biology, and Museum Zoologicum Bogoriense all contributed specimens and curatorial support to this study. We thank Frederick Sheldon, Jeremy Brown, and two anonymous reviewers for helpful comments during the production of this manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.09.018.

References

Beerli, P., 2006. Comparison of Bayesian and maximum likelihood inference of population genetic parameters. Bioinformatics 22, 341–345. http://dx.doi.org/10.1093/bioinformatics/bti803.

Beerli, P., Felsenstein, J., 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152, 763–773.

Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n populations by using a coalescent approach. Proc. Natl. Acad. Sci. USA 98, 4563–4568. http://dx.doi.org/10.1073/pnas.081068098.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Comput. Biol. 10. http://dx.doi.org/10.1371/journal.pcbi.1003537.

Brown, R.M., Siler, C.D., Oliveros, C.H., Esselstyn, J.A., Diesmos, A.C., Hosner, P.A., Linkem, C.W., Barley, A.J., Oaks, J.R., Sanguila, M.B., Welton, L.J., Moyle, R.G., Peterson, A.T., Alcala, A.C., 2013. Evolutionary processes of diversification in a model island archipelago. Annu. Rev. Ecol. Evol. Syst. 44, 411–435. http://dx.doi.org/10.1146/annurev-ecolsys-110411-160323.

Brown, W.M., George Jr., M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76, 1967–1971.

Campbell, P., Putnam, A.S., Bonney, C., Bilgin, R., Morales, J.C., Kunz, T.H., Ruedas, L.A., 2007. Contrasting patterns of genetic differentiation between endemic and widespread species of fruit bats (Chiroptera: Pteropodidae) in Sulawesi, Indonesia. Mol. Phylogenet. Evol. 44, 474–482.

Caro, L.M., Caycedo-Rosales, P.C., Bowie, R.C.K., Slabbekoorn, H., Cadena, C.D., 2013. Ecological speciation along an elevational gradient in a tropical passerine bird? J. Evol. Biol. 26, 357–374.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. http://dx.doi.org/10.1038/nmeth.2109.

Demos, T.C., Achmadi, A.S., Giarla, T.C., Handika, H., Rowe, K.C., Esselstyn, J.A., 2016. Local endemism and within-island diversification of shrews illustrate the importance of speciation in building Sundaland mammal diversity. Mol. Ecol. 25, 5158–5173.

Dickerson, R.E., 1928. Distribution of Life in the Philippines. Philippine Bureau of Science, Manila.

Douzery, E.P., Scornavacca, C., Romiguier, J., Belkhir, K., Galtier, N., Delsuc, F., Ranwez, V., 2014. OrthoMaM v8: a database of orthologous exons and coding sequences for comparative genomics in mammals. Mol. Biol. Evol. 31, 1923–1928. http://dx.doi.org/10.1093/molbev/msu132.

Dray, S., Dufour, A.B., 2007. The *ade4* package: implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20. http://dx.doi.org/10.18637/jss.v022.i04.

Driller, C., Merker, S., Perwitasari-Farajallah, D., Sinaga, W., Anggraeni, N., Zischler, H., 2016. Stop and go – waves of tarsier dispersal mirror the genesis of Sulawesi Island. PLoS ONE 10, e0141212. http://dx.doi.org/10.1371/journal.pone.0141212.

Earl, D.A., vonHoldt, B.M., 2012. Structure HARVESTER: a website and program for visualizing Structure output and implementing the Evanno method. Conserv. Genet.

Resour. 4, 359–361. http://dx.doi.org/10.1007/s12686-011-9548-7.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797. http://dx.doi.org/10.1093/nar/gkh340.
Esselstyn, J.A., Achmadi, A.S., Rowe, K.C., 2012. Evolutionary novelty in a rat with no molars. Biol. Lett. 8, 990–993. http://dx.doi.org/10.1098/rsbl.2012.0574.
Esselstyn, J.A., Maharadatunkamsi, Achmadi, A.S., Siler, C.D., Evans, B.J., 2013. Carving out turf in a biodiversity hotspot: multiple unrecognized shrew species co–occur on Java Island, Indonesia. Mol. Ecol. 22, 4972–4987. http://dx.doi.org/10.1111/mec.12450.
Esselstyn, J.A., Timm, R.M., Brown, R.M., 2009. Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. Evolution 63, 2595–2610. http://dx.doi.org/10.1111/j.1558-5646.2009.00743.x.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x.
Evans, B.J., 2012. Coalescent-based analysis of demography: applications to biogeography on Sulawesi. In: Gower, D.J., Johnson, K.G., Richardson, J.E., Rosen, B.R., Rüber, L., Williams, S.T. (Eds.), Biotic Evolution and Environmental Change in Southeast Asia. Cambridge University Press, Cambridge, pp. 270–289.
Evans, B.J., Brown, R.M., McGuire, J.A., Supriatna, J., Andayani, N., Diesmos, A., Iskandar, D., Melnick, D.J., Cannatella, D.C., 2003b. Phylogenetics of fanged frogs: testing biogeographcial hypotheses at the interface of the Asian and Australian faunal zones. Syst. Biol. 52, 794–819. http://dx.doi.org/10.1093/sysbio/52.6.794.
Evans, B.J., Supriatna, J., Andayani, N., Setiati, M.I., Cannatella, D.C., Melnick, D.J., 2003a. Monkeys and toads define areas of endemism on Sulawesi. Evolution 57, 1436–1443. http://dx.doi.org/10.1554/02-443.
Evans, B.J., McGuire, J.A., Brown, R.M., Andayani, N., Supriatna, J., 2008. A coalescent framework for comparing alternative models of population structure with genetic data: evolution of Celebes toads. Biol. Lett. 4, 430–433. http://dx.doi.org/10.1098/rsbl.2008.0166.
Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164, 1567–1587.
Flot, J.F., 2010. SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Mol. Ecol. Resour. 10, 162–166. http://dx.doi.org/10.1111/j.1755-0998.2009.02732.x.
Fooden, J., 1969. Taxonomy and Evolution of the Monkeys of Celebes (Primates: Cercopithecidae). S. Karger, Basel.
Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704. http://dx.doi.org/10.1080/10635150390235520.
Hall, R., 2002. Cenozoic geological and plate tectonic evolution of SE Asian and the SW Pacific: computer–based reconstructions and animations. J. Asian Earth Sci. 20, 353–434. http://dx.doi.org/10.1016/S1367-9120(01)00069-4.
Hall, R., 2011. Australia–SE Asia collision: plate tectonics and crustal flow. Geol. Soc. Lond. Spec. Publ. 355, 75–109. http://dx.doi.org/10.1144/SP355.5.
Hawkins, M.T.R., Leonard, J.A., Helgen, K.M., McDonough, M.M., Rockwood, L.L., Maldonado, J.E., 2016. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evol. Biol. 16, 80. http://dx.doi.org/10.1186/s12862-016-0650-z.
Heaney, L.R., Balete, D.S., Rickart, E.A., Alviola, P.A., Duya, M.R.M., Duya, M.V., Veluz, M.J., VandeVrede, L., Steppan, S.J., 2011. Seven new species and a new subgenus of forest mice (Rodentia: Muridae: *Apomys*) from Luzon Island, Fieldiana. Life Earth Sci. 5, 1–60.
Heaney, L.R., Balete, D.S., Veluz, M.J., Steppan, S.J., Esselstyn, J.A., Pfeiffer, A.W., Rickart, E.A., 2014. Two new species of Philippine forest mice (*Apomys*, Muridae, Rodentia) from Lubang and Luzon Islands, with a redescription of *Apomys sacobianus* Johnson, 1962. Proc. Biol. Soc. Wash. 126, 395–413. http://dx.doi.org/10.2988/0006-324X-126.4.395.
Heaney, L.R., Rickart, E.A., 1990. Correlations of clades and clines: geographic, elevational, and phylogenetic distribution patterns among Philippine mammals. In: Peters, G., Hutterer, R. (Eds.), Vertebrates in the Tropics. Museum Alexander Koenig, Bonn, pp. 321–332.
Hosner, P.A., Nyári, Á.S., Moyle, R.G., 2013. Water barriers and intra-island isolation contribute to diversification in the insular *Aethopyga* sunbirds (Aves: Nectariniidae). J. Biogeogr. 40, 1094–1106.
Huxley, T.H., 1868. On the distribution of the Alectoromorphae and Heteromorphae. Proc. Zool. Soc. Lond. 1868, 294–319.
Justiniano, R., Schenk, J.J., Balete, D.S., Rickart, E.A., Esselstyn, J.A., Heaney, L.R., Steppan, S.J., 2015. Testing diversification models of endemic Philippine forest mice (*Apomys*) with nuclear phylogenies across elevational gradients reveals repeated colonization of isolated mountain ranges. J. Biogeogr. 42, 51–64. http://dx.doi.org/10.1111/jbi.12401.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. http://dx.doi.org/10.1093/bioinformatics/bts199.
Leaché, A.D., Fujita, M.K., 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. B: Biol. Sci. 277, 3071–3077. http://dx.doi.org/10.1098/rspb.2010.0662.
Lohman, D.J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P.K.L., Shih, H.-T., Carvalho, G.R., von Rintelen, T., 2011. Biogeography of the Indo-Australian archipelago. Annu. Rev. Ecol. Evol. Syst. 42, 205–226.
Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209–220.
McGuire, J.A., Brown, R.M., Mumpuni, Riyanto, R., Andayani, N., 2007. The flying lizards of the *Draco lineatus* group (Squamata: Iguania: Agamidae): a taxonomic revision with descriptions of two new species. Herpetol. Monogr. 21, 179–212. http://dx.doi.org/10.1655/07-012.1.
Meirmans, P.G., 2015. Seven common mistakes in population genetics and how to avoid them. Mol. Ecol. 24, 3223–3231. http://dx.doi.org/10.1111/mec.13243.
Miller Jr., G.S., Hollister, N., 1921a. Twenty new mammals collected by H. C. Raven in Celebes. Proc. Biol. Soc. Wash. 34, 93–104.
Miller Jr., G.S., Hollister, N., 1921b. Descriptions of sixteen new murine rodents from Celebes. Proc. Biol. Soc. Wash. 34, 67–76.
Musser, G.G., 2014. A systematic review of Sulawesi *Bunomys* (Muridae, Murinae) with the description of two new species. Bull. Am. Mus. Nat. Hist. 392, 1–313. http://dx.doi.org/10.1206/863.1.
Musser, G.G., Durden, L.A., Holden, M.E., Light, J.E., 2010. Systematic review of endemic Sulawesi squirrels (Rodentia, Sciuridae), with descriptions of new species of associated sucking lice (Insecta, Anoplura), and phylogenetic and zoogeographic assessments of sciurid lice. Bull. Amer. Mus. Nat. Hist. 339, 1–255. http://dx.doi.org/10.1206/695.1.
Nabholz, B., Glémin, S., Galtier, N., 2007. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Mol. Biol. Evol. 25, 120–130. http://dx.doi.org/10.1093/molbev/msm248.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from < https://www.R-project.org/ > (accessed 21.05.16).
Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. Available from < http://beast.bio.ed.ac.uk/Tracer > (accessed 21.05.16).
Rannala, B., Yang, Z., 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics 194, 245–253. http://dx.doi.org/10.1534/genetics.112.149039.
Rickart, E.A., Heaney, L.R., Goodman, S.M., Jansa, S., 2005. Review of the Philippine genera *Chrotomys* and *Celaenomys* (Muridae) and description of a new species. J. Mammal. 86, 415–428. http://dx.doi.org/10.1644/BEL-124.1.
Rowe, K.C., Achmadi, A.S., Esselstyn, J.A., 2016. A new genus and species of omnivorous rodent (Muridae: Murinae) from Sulawesi, nested within a clade of endemic carnivores. J. Mammal. 97, 978–991. http://dx.doi.org/10.1093/jmammal/gyw029.
Ruedi, M., 1995. Taxonomic revision of shrews of the genus *Crocidura* from the Sunda Shelf and Sulawesi with description of two new species (Mammalia: Soricidae). Zool. J. Linn. Soc. 115, 211–265. http://dx.doi.org/10.1111/j.1096-3642.1995.tb02461.x.
Ruedi, M., Auberson, M., Savolainen, V., 1998. Biogeography of Sulawesian shrews: testing for their origin with a parametric bootstrap on molecular data. Mol. Phylogenet. Evol. 9, 567–571. http://dx.doi.org/10.1006/mpev.1998.0487.
Setiadi, M.I., McGuire, J.A., Brown, R.M., Zubairi, M., Iskandar, D.T., Andayani, N., Supriatna, J., Evans, B.J., 2011. Adaptive radiation and ecological opportunity in Sulawesi and Philippine fanged frog (*Limnonectes*) communities. Am. Nat. 178, 221–240. http://dx.doi.org/10.1086/660830.
Shekelle, M., Meier, R., Wahyu, I., Widateti, Ting, N., 2010. Molecular phylogenetics and chronometrics of Tarsiidae based on 12s mtDNA haplotypes: evidence for Miocene origins of crown tarsiers and numerous species within the Sulawesian clade. Int. J. Primatol. 31, 1083–1106. http://dx.doi.org/10.1007/s10764-010-9457-8.
Sheldon, F.H., Lim, H.C., Moyle, R.G., 2015. Return to the Malay Archipelago: the biogeography of Sundaic rainforest birds. J. Ornithol. 156, S91–S113. http://dx.doi.org/10.1007/s10336-015-1188-3.
Spakman, W., Hall, R., 2010. Surface deformation and slab-mantle interaction during Banda Arc subduction rollback. Nat. Geosci. 3, 562–566. http://dx.doi.org/10.1038/ngeo917.
Stelbrink, B., Albrecht, C., Hall, R., von Rintelen, T., 2012. The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? Evolution 66, 2252–2271. http://dx.doi.org/10.1111/j.1558-5646.2012.01588.x.
Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction. Am. J. Hum. Genet. 73, 1162–1169. http://dx.doi.org/10.1086/379378.
Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989. http://dx.doi.org/10.1086/319501.
Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26, 1569–1571. http://dx.doi.org/10.1093/bioinformatics/btq228.
Sukumaran, J., Holder, M.T. Sumtrees: Phylogenetic Tree Summarization. 4.1.0. Available at < https://github.com/jeetsukumaran/DendroPy > (accessed 21.05.16).
Sukumaran, J., Knowles, L.L., 2017. Multispecies coalescent delimits structure, not species. Proc. Nat. Acad. Sci. USA 114, 1607–1612. http://dx.doi.org/10.1073/pnas.1607921114.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. http://dx.doi.org/10.1093/molbev/mst197.
Taylor, E.H., 1934. Philippine land mammals. Monogr. Bur. Sci. Manila 30, 1–548.
Toussaint, E.F.A., Hall, R.A., Monaghan, M.T., Sagata, K., Ibalim, S., Shaverdo, H.V., Volger, A.P., Pons, J., Balke, M., 2014. The towering orogeny of New Guinea as a trigger for arthropod megadiversity. Nat. Commun. 5, 1–10. http://dx.doi.org/10.1038/ncomms5001.
Vähä, J., Erkinaro, J., Niemelä, E., Primmer, C., 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Mol. Ecol. 16,

2638–2654. http://dx.doi.org/10.1111/j.1365-294X.2007.03329.x.
von Rintelen, T., Stelbrink, B., Marwoto, R., Glaubrecht, M., 2014. A snail perspective on the biogeography of Sulawesi, Indonesia: origin and intra-island dispersal of the viviparous freshwater gastropod *Tylomelania*. PLoS ONE 9, e98917. http://dx.doi.org/10.1371/journal.pone.0098917.
Wallace, A.R., 1869. The Malay Archipelago, the Land of the Orang-Utan and the Bird of Paradise: A Narrative of Travel with Studies of Man and Nature. Macmillan, London.
Wallace, A.R., 1876. The Geographical Distribution of Animals; With a Study of the Relations of Living and Extinct Faunas as Elucidating the Past Changes of the Earth's Surface. Macmillan, London.
Wallace, A.R., 1880. Island Life, or the Phenomena and Causes of Insular Faunas and Floras, Including a Revision and Attempted Solution of the Problem of Geological Climates. Macmillan, London.
Whitmore, T.C., 1984. Tropical Rain Forests of the Far East, second ed. Clarendon Press, Oxford.
Whitten, A.J., Mustafa, M., Henderson, G.S., 2002. The Ecology of Sulawesi. Periplus, Singapore.
Yang, Z., 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61, 854–865. http://dx.doi.org/10.1093/czoolo/61.5.854.
Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. 107, 9264–9269. http://dx.doi.org/10.1073/pnas.0913022107.
Yang, Z., Rannala, B., 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 31, 3125–3135. http://dx.doi.org/10.1093/molbev/msu279.
Zwickl, D.J., 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation. The University of Texas at Austin.