Title

Phylogenomic Analysis Reveals Dispersal-Driven Speciation and Divergence with Gene Flow in Lesser Sunda Flying Lizards (Genus Draco)

Permalink

https://escholarship.org/uc/item/79c81353

Journal

Systematic Biology, 71(1)

ISSN

1063-5157

Authors

Reilly, Sean B
Stubbs, Alexander L
Arida, Evy
et al.

Publication Date

2021-12-16

DOI

10.1093/sysbio/syab043

Peer reviewed

Phylogenomic Analysis Reveals Dispersal-Driven Speciation and Divergence with Gene Flow in Lesser Sunda Flying Lizards (Genus *Draco*)

Sean B. Reilly[1],[*], Alexander L. Stubbs[1], Evy Arida[2], Benjamin R. Karin[1], Umilaela Arifin[3],[4], Hinrich Kaiser[5],[6], Ke Bi[1],[7], Djoko T. Iskandar[8], and Jimmy A. McGuire[1]

[1]*Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA;* [2]*Museum Zoologicum Bogoriense, Indonesian Institute of Sciences, Cibinong, Indonesia;* [3]*School of Life Sciences and Technology, Institut Teknologi Bandung, Jalan Ganesha 10, Tamansari, Bandung 40132, Indonesia;* [4]*Centrum für Naturkunde, Zoologisches Museum Hamburg, Universität Hamburg, Edmund-Siemers-Allee 1, 20146 Hamburg, Germany;* [5]*Department of Vertebrate Zoology, Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany;* [6]*Department of Biology, Victor Valley College, Victorville, CA 92395, USA;* [7]*Computational Genomics Resource Laboratory, California Institute for Quantitative Biosciences, University of California, Berkeley, CA 94720, USA; and* [8]*Indonesian Academy of Sciences, Jalan Medan Merdeka Selatan No. 11, Jakarta 10110, Indonesia*
[*]*Correspondence to be sent to: Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA; E-mail: sreilly@ucsc.edu.*

Abstract.—The Lesser Sunda Archipelago offers exceptional potential as a model system for studying the dynamics of dispersal-driven diversification. The geographic proximity of the islands suggests the possibility for successful dispersal, but this is countered by the permanence of the marine barriers and extreme intervening currents that are expected to hinder gene flow. Phylogenetic and species delimitation analyses of flying lizards (genus *Draco*) using single mitochondrial genes, complete mitochondrial genomes, and exome-capture data sets identified 9–11 deeply divergent lineages including single-island endemics, lineages that span multiple islands, and parapatrically distributed nonsister lineages on the larger islands. Population clustering and PCA confirmed these genetic boundaries with isolation-by-distance playing a role in some islands or island sets. While *gdi* estimates place most candidate species comparisons in the ambiguous zone, migration estimates suggest 9 or 10 species exist with nuclear introgression detected across some intra-island contact zones. Initial entry of *Draco* into the archipelago occurred at 5.5–7.5 Ma, with most inter-island colonization events having occurred between 1–3 Ma. Biogeographical model testing favors scenarios integrating geographic distance and historical island connectivity, including an initial stepping-stone dispersal process from the Greater Sunda Shelf through the Sunda Arc as far eastward as Lembata Island. However, rather than reaching the adjacent island of Pantar by dispersing over the 15-km wide Alor Strait, *Draco* ultimately reached Pantar (and much of the rest of the archipelago) by way of a circuitous route involving at least five overwater dispersal events. These findings suggest that historical geological and oceanographic conditions heavily influenced dispersal pathways and gene flow, which in turn drove species formation and shaped species boundaries. [Biogeography; genomics, Indonesia; lizards; phylogeography; reptiles]

The advent of high-throughput sequencing and the corresponding availability of genome-scale data sets offer systematists the potential to tackle ever more challenging phylogenetic and biogeographic questions. It is now possible to not only estimate phylogenies with greater precision using multispecies coalescent methods (Rannala and Yang 2003, 2017; Degnan and Rosenberg 2009; Ogilvie et al. 2017) but also to harness population genomic analyses that improve species delimitation and allow for more rigorous tests of cryptic diversification (Yang and Rannala 2010; Fujita et al. 2012; Smith and Carstens 2020). The importance of accounting for gene flow, both for phylogenomic and species delimitation analyses (Solís-Lemus et al. 2016; Wen et al. 2018; Jiao and Yang 2021), has never been more clear. This is particularly true when attempting to elucidate diversification processes in dynamic island archipelagos, as the complex tectonic processes that cause islands to arise, merge, fragment, or disappear create ample opportunities for lineage divergence, extinction, and fusion. Gene flow can occur between islands, within islands, and episodically depending on sea level fluctuations, volcanic activity, and changes in sea currents. Thus, disentangling the complexities of insular diversification and biogeography represents an outstanding challenge for systematists (Warren et al. 2015; Shaw and Gillespie 2016).

The biogeographical realm known as "Wallacea" encompasses a set of oceanic archipelagos that lie between the Asian (Sunda) and Australo-Papuan (Sahul) continental shelves, roughly delimited by Wallace's and Lydekker's biogeographic lines, respectively (Wallace 1860, 1863; Huxley 1868; Lydekker 1896; Mayr 1944). The emergence of Wallacean islands, driven by some of the most complex tectonic processes documented anywhere in the world, set in motion a dynamic process of bidirectional faunal exchange that has fascinated generations of biogeographers (Mayr 1944; Spakman and Hall 2010; Lohman et al. 2011; Reilly et al. 2019a,b; Ali and Heaney 2021). The Lesser Sundas are a distinct biogeographic subregion of Wallacea (Stresemann 1939; Turner et al. 2001; Carstensen and Olesen 2009; Van Welzen et al. 2011; Carstensen et al. 2012), though further division of the Lesser Sundas has been proposed (Muller 1846; Simpson 1977; Michaux 2010). The Lesser Sundas can be divided into four regional components formed via distinct geological processes over differing timescales (Fig. 1b). 1) The Sunda Arc (Lombok through Lembata)
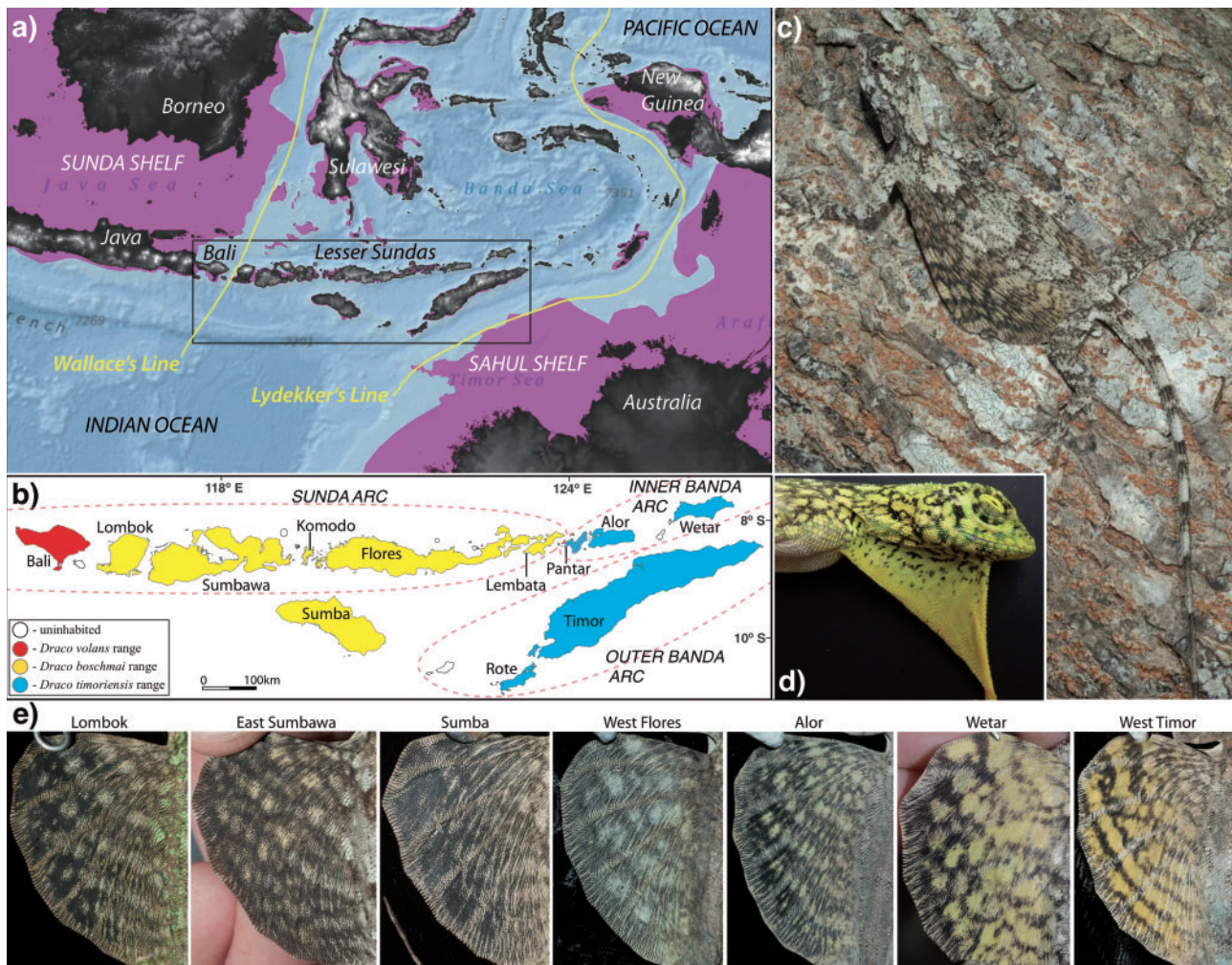
FIGURE 1.   a) Map of the southern Wallacea region. Areas colored in purple indicate depths < 120 m that become emergent during glacial maxima. b) The currently described ranges of *Draco boschmai* and *D. timoriensis*. c) Female *D. boschmai* from Sumba Island. d) Male *D. boschmai* from West Flores. e) Variation in male dorsal patagia pattern and coloration.

arose via subduction mediated volcanic activity and has been continuously emergent for up to ~11 Myr, representing the oldest islands in the Lesser Sundas (Hall 2009). 2) The volcanic Inner Banda Arc (Pantar, Alor, and Wetar), which has been emergent for the past ~5 Myr, extends eastward from the Sunda Arc and is a product of continent-island arc collision (Hall 2009, 2011). 3) The nonvolcanic Outer Banda Arc (Rote and Timor) formed as a result of accretionary build-up at this collision zone (Audley-Charles 2011; Rigg and Hall 2011). A westward progression of emergence is estimated with Eastern Timor having emerged 3.1–4.5 Ma (Haig and McCartain 2007; Nguyen et al. 2013; Tate et al. 2014) and Rote having emerged < 1 Ma (Roosmawati and Harris 2009; Harris 2011). 4) Sumba Island is an ancient continental plate fragment that, like the Outer Banda Arc islands, has been uplifted and emergent for < 3 Myr (Fortuin et al. 1997).

The individual islands are separated by straits of varying widths and depths (Goltenboth et al. 2006),

though the distance between islands may be less relevant for sweepstakes dispersal than the power of the currents passing between them (Kuhnt et al. 2004). These islands span part of the Indonesian Throughflow, where upper ocean waters from the Pacific Ocean are transferred to the Indian Ocean (Godfrey 1996) in one of the largest movements of water on the planet. The biogeographical interpretation of the Lesser Sundas as comprising a two-way filter zone reflects a clinical reduction in Asian-origin species together with a corresponding increase in Australo-Papuan species as one moves eastward through the archipelago, a pattern that is especially striking in terrestrial reptiles (Darlington 1957; Whittaker and Fernaìndez-Palacios 2007). Initial descriptions of reptile biogeography in the Lesser Sundas were based on patterns of species composition from individual islands (Darlington 1957; Auffenberg 1980; How and Kitchener 1997) and invoked stepping-stone colonization as the process most likely to have created such patterns (Hisheh et al. 1998). However, more recent analyses of this fauna

(Reilly 2016; Reilly et al. 2017, 2019a, 2019b, 2021; Blom et al. 2019; Jones et al. 2019;) and other fauna (Tänzler et al. 2016; Toussaint et al. 2020) have shown a wide range of patterns regarding the timing and sequence of island colonization.

Flying lizards of the agamid genus *Draco* originated in mainland SE Asia (McGuire and Heang 2001; Grismer et al. 2016) and have colonized nearly every major island within the Lesser Sundas (Fig. 1). These are small, arboreal, ant-specialists, which can be quite cryptically colored against their tree bark backgrounds (Fig. 1c), though males have brightly colored dewlaps (Fig. 1d) and variably colored patagia (Fig. 1e) that they employ during display (McGuire and Alcala 2000; McGuire et al. 2007; McGuire and Dudley 2011). *Draco boschmai* occurs on the Sunda Arc islands and Sumba, and *Draco timoriensis* from the Inner and Outer Banda Arc islands (Fig. 1b) and recently confirmed from Pantar (Reilly et al. 2020). Morphologically, *D. timoriensis* is distinguished from *D. boschmai* on the basis of its larger size, the presence of a row of enlarged, keeled scales along each side of the vertebral line, and a large spine-like tubercle in the nuchal region, though both species exhibit substantial intraspecific variation in color and scale characters between populations (Mertens 1930; Musters 1983; McGuire and Heang 2001). McGuire and Heang (2001) noted that both *D. timoriensis* and *D. boschmai* are composed of several diagnosable, allopatric lineages and further taxonomic changes would likely be necessary.

In this study, we utilized transcriptome-based exon-capture data from hundreds of orthologous loci to study the biogeographical history of Lesser Sunda flying lizards. Because inadequate sampling across a taxon's range can lead to over-splitting of species (Hillis 2019), we started with newly acquired mtDNA data from hundreds of *Draco* individuals densely sampled across every major island (along with sequences from McGuire and Heang 2001), which guided the selection of 100 samples for subsequent exon-capture data collection. Using both mitogenomic and genomic phylogenies we were able to delimit major lineage boundaries and assess the prevalence of any mitonuclear discordance, indicative of introgression. Using a time-calibrated genomic phylogeny, we then tested biogeographic hypotheses to determine whether the distance between regions, the arrangement of regions, prevalent oceanographic currents, or historical island connectivity best explained our results. Finally, we assessed the taxonomic diversity by corroborating species delimitation methods with estimates of population structure and gene flow.

Materials and Methods

Sample Collection and mtDNA Screening

Flying lizard specimens were field collected on Java and Bali (*Draco volans*), on Lombok, Sumbawa, Flores, Lembata, and Sumba (*D. boschmai*), and on Timor, Rote, Pantar, Alor, and Wetar (*D. timoriensis*;

Supplementary Table S1 available on Dryad at https://dx.doi.org/10.5061/dryad.gqnk98skg). Liver tissue was obtained from euthanized lizards and either stored in RNALater, 95% ethanol, or flash frozen in liquid nitrogen. Tissues and specimens were deposited in the Museum of Vertebrate Zoology at the University of California, Berkeley, the Museum Zoologicum Bogoriense in Cibinong, Indonesia, and in the United States National Museum of Natural History (Timor-Leste material).

DNA was extracted using standard salt extraction techniques or using the DNeasy Blood and Tissue kit (Qiagen). The *ND2* gene was PCR-amplified and sequenced for 362 *Draco* samples (GenBank numbers MT134466–134827) using standard conditions with the primers METf.1 and ALAr.2m (Macey et al. 1997). Raw sequence reads along with ten sequences previously deposited in GenBank (Supplementary Table S1 available on Dryad) were combined in Geneious Prime v2019.2 (www.geneious.com) and aligned with Muscle (Edgar 2004). The sequence alignment of 1092 bp was imported into jModelTest v2.1.4 (Darriba et al. 2012) to determine the best-fit model of sequence evolution (HKY + I + G). A maximum likelihood (ML) analysis was performed using RAxML (Stamatakis 2014) with 1000 bootstrap replicates and rooted using *Draco beccarii* (LSUMZ 81223) as the outgroup. Candidate species discovery was performed using the sequence-based program ABGD (Puillandre et al. 2012) and the phylogeny-based program MPTP (Zhang et al. 2013). Details on the settings and rationale for these analyses are found in the Supplementary methods available on Dryad. These candidate species were then validated using multilocus species delimitation and gene flow analyses in the "Population Structure and Species Delimitation" and "Inter-Island Demographics" sections below (see Hillis 2019; Chan et al. 2020).

Development of Exon-Capture Experiment

Genomic sequence data were obtained using transcriptome-based exon-capture methods (Bi et al. 2012, 2013). Orthologous exonic targets were derived from four *Draco* transcriptomes (Supplementary Table S2 available on Dryad), two of which were used to design 120-bp, 3X-tiled probes (44,964 unique probes) for an in-solution targeted capture experiment (MYbaits) targeting ~1 Mb of sequence data from 1634 exons representing 709 independent genes (sequence loci). The selection of 100 samples for genomic library preparation was guided by the *ND2* phylogeny with the goal of maximizing genetic diversity and unique localities (Supplementary Table S1 available on Dryad; Supplementary Fig. S1 available on Dryad). Sampling includes 90 Lesser Sunda samples, six *D. volans* samples, and four outgroup samples for rooting phylogenetic trees (*D. becarii, D. walkeri, D. sumatranus,* and *D. modiglianii*). Detailed methods describing the development of loci, sample library preparation, an

initial array-chip experiment, a follow-up targeted exon-capture experiment, and bioinformatics pipelines are available in Supplementary methods available on Dryad. Sequence reads are deposited in the NCBI sequence read archive (SRA) under BioProject accession PRJNA612457.

Mitogenomics

Mitochondrial genomes were assembled by mapping cleaned Illumina sequencing reads to the mitochondrial genome of *Draco maculatus* (GenBank number KY073263) and aligning individual assemblies in GENEIOUS PRIME v2019.2. Most reads were derived from the initial array-chip experiment, and thus this data set is missing 5/6 Timor-Leste samples that were only sequenced with the follow-up target capture experiment. Enough low coverage reads were recovered and supplemented with Sanger sequenced *ND2* data to include one Timor-Leste sample (USNM 579490). The control region was removed due to poor coverage and alignment gaps resulting in a 15,767 base pair alignment that was analyzed with 1) ML estimation using RAXML (Stamatakis 2014) under the GTRCAT model of sequence evolution with node support assessed by 1000 nonparametric bootstrap replicates, 2) Bayesian estimation using MRBAYES v3.2 (Ronquist et al. 2012) run for 20 million generations, sampled every 2000 generations, with convergence assessed (ESS > 1000) using TRACER v1.7 (Rambaut et al. 2018), and 3) a reduced sampling version of the alignment containing 17 samples (see Supplementary methods available on Dryad) chosen to date key nodes was analyzed with BEAST v2.4.8 (Bouckaert et al. 2014). Runs were set up under a Yule birth rate prior, the HKY mutation model, and a relaxed clock log-normal distribution with a clock rate of 0.02 substitutions/site/MY which is estimated as the mean mitochondrial mutation rate across vertebrates (Allio et al. 2017). After convergence was assessed (ESS > 200) using TRACER v1.7, a 25% burn-in was removed from each of two 100 million generation runs sampled every 10,000 generations and the remaining 15,000 trees were combined to create a 50% majority rule consensus tree. Candidate species were inferred using ABGD (Puillandre et al. 2012) and MPTP (Zhang et al. 2013), and levels of sequence divergence between lineages were estimated using DIVEIN (Deng et al. 2010).

Phylogenomics

The concatenated alignment consisting of both target and flanking regions was analyzed under ML estimation with RAXML (Stamatakis 2014) under the GTRCAT model of sequence evolution with 1000 bootstrap replicates. A summary multispecies coalescent analysis was undertaken using ASTRAL-III v5.7.5 (Zhang et al. 2018) where individual gene trees estimated for each locus served as input files. This analysis was run with 1) each individual treated as a "species" to allow a direct comparison with the topology of the mitogenomic and phylogenomic tree topologies and 2) a species tree with individual species designations for biogeographically relevant clades identified in the ML analysis. We estimated maximum likelihood gene trees for each locus (690 after filtering) using IQTREE v2.1.1 (Minh et al. 2020) with automatic model selection (Kalyaanamoorthy et al. 2017) and 1000 ultrafast bootstrap replicates (Hoang et al. 2018). Nodes with less than 10% bootstrap support were collapsed using NEWICK UTILITIES (Junier and Zdobnov 2010), as this has been shown to improve species tree inference (Zhang et al. 2018). Branch lengths and local posterior probabilities of nodes were computed on the resulting species tree (Sayyari and Mirarab 2016). Individual gene trees also served as the input for an alternative species tree method MP-EST v2.0 (Liu et al. 2010), which uses a pseudo-likelihood function to search for the maximum likelihood tree.

The concatenated alignment was analyzed using BEAST v2.4.8 (Bouckaert et al. 2014) to produce a time-calibrated tree for subsequent use in biogeographical model testing analyses. This analysis utilized the same 17 samples as the mitogenomic BEAST analysis and was run using the Yule birth rate prior, the JC69 mutation model, and a relaxed clock log-normal distribution with a clock rate of 0.001 substitutions/site/MY (Brandley et al. 2011; Blom et al. 2016; Allio et al. 2017). After confirmation of convergence (ESS > 200) using TRACER v1.7 a 10% burn-in was removed from each of two 100 million generation runs sampled every 10,000 generations, and the remaining 18,000 trees were combined to create a 50% majority rule consensus tree which also serves as the guide tree for all BIOGEOBEARS analyses (see below).

Population Structure and Species Delimitation

All SNPs from the targeted-capture data set (107,689 SNPs) served as the input for a principal component analysis (PCA) of genetic covariance using the ADEGENET package for R (Jombart 2008). The results of the top two components were plotted against one another to visualize spatial clustering of individual samples. A further PCA was conducted on the 34 *D. timoriensis* samples that clustered together in the initial PCA, as well as the Flores + Lembata group. A Mantel test was run on matrices of Edwards' genetic distances and Euclidean geographic distances to test for signatures of isolation-by-distance (IBD) using the mantel.randtest function in ADEGENET with significance assessed using 100,000 permutations. Geographic distance was plotted against genetic distance to visualize if IBD occurs as a continuous cline or as patches of geographically distant, divergent populations. Local point density was measured by 2D kernel density estimation with the kde2d function in the R package "MASS" (Ripley and Venables 2002) to overlay a colored cloud over the points. This IBD approach was applied to the following subsets of samples: 1) all Lesser Sundas, 2) Sunda Arc islands (Lombok, Sumbawa,

Flores, Lembata), 3) Flores + Lembata, 4) Timor + Rote, and 5) Timor.

Genetic structure was analyzed with STRUCTURE v2.3.2 (Pritchard et al. 2000), which uses a Bayesian algorithm to assign individuals to genetic clusters. The program was initially run on the two major groups identified from the PCA: 1) Sunda Arc populations and 2) Sumba + *D. timoriensis*. The program was also run on groups where admixed individuals were detected: 1) Sumbawa, 2) Flores + Lembata, 3) Timor + Rote, and 4) Pantar + Alor. One informative SNP per locus (excluding singleton SNPs) from within each group of interest was utilized and three alternative SNP sets were analyzed for each group to ensure that SNP selection did not overly influence the results (Supplementary Table S4 available on Dryad). All analyses were run without specifying prior location information under the admixture model with correlated allele frequencies. Ten replicates per K were performed with 200,000 burn-in and 200,000 retained generations over a range of K based on group size and previous hypotheses of population structure (Sunda Arc K = 1–10; Sumba + *D. timoriensis* K = 1–10; Sumbawa K = 1–5; Flores + Lembata K = 1–6; Timor + Rote K = 1–6; Pantar + Alor K = 1–5). The optimal K was determined by the delta-K analysis (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl and vonHoldt 2012). Ancestry bar plots for the optimal K and any other supported values of K within each group were generated from subsequent runs on all three alternative SNP data sets with 500,000 burn-in and 1,000,000 retained generations.

The software BPP v4.3.8 was used to test species boundaries on a fixed guide tree using model A10 (Yang and Rannala 2010; Flouri et al. 2018). Individual sequence alignments for all loci served as the input. The analysis was run separately for *D. boschmai* and *D. timoriensis* with tree topologies inferred from the RAXML and ASTRAL analyses of the exon data set. The lineages tested within *D. boschmai* include Lombok, West Sumbawa, East Sumbawa, Flores, Lembata, and Sumba, and the lineages tested within *D. timoriensis* include Timor, Rote, Wetar, Alor, and Pantar. Strict phylogenomic lineages within the Flores + Lembata clade were tested as a way to assess the prevalence of over splitting. The program was run with a 50,000 generation burn-in and a 500,000-generation run with inverse-gamma priors of theta ∼ (2, 0.001) and tau ∼ (2, 0.001) ensuring an acceptance rate between 20 and 80%. The program was also run with alternate priors (2, 2000; 1, 10).

Lineage status was further assessed by calculating *gdi* (genealogical divergence index) scores from the multispecies coalescent model (MSC) output in BPP following the recommendations of Jackson et al. (2017) and Leaché et al. (2019). Jackson et al. (2017) suggested that the *gdi* provides a more conservative (and thus putatively more reliable) assessment of species status than does the A10 model implemented in BPP. The index is based on the relationship between the effective population size (Θ) and time since divergence (τ), with $gdi = 1 - e^{-2\tau/\Theta A}$. The *gdi* score indicates the probability that two sequences from population A will coalesce

before species divergence (τ) when the genealogy is traced backward in time (Leaché et al. 2019). The *gdi* score thus approaches 1 when a candidate species is characterized by a longer branch length relative to its effective population size, with *gdi* > 0.7 suggested to indicate discrete species, *gdi* < 0.2 indicating conspecific status, and intermediate values representing an ambiguous zone. The *gdi* posterior distribution can be similarly obtained by summarizing the *gdi* across the posterior distributions obtained for these parameters.

The multispecies coalescent model with introgression in BPP 4.3.8 (MSci; Jiao and Yang 2021) provides a full likelihood implementation to estimate Θ, τ, and migration while also accommodating gene flow. Obtaining *gdi* under the MSci model requires a simulation approach using the extended Newick tree format to reflect both a fixed tree topology and one unidirectional or bidirectional migration band. The output from the analysis is then used to parameterize a simulation using BPP's MCcoal functionality to generate three sequences representing two lineages, *a* and *b* (if the candidate species is lineage *a*, two sequences (*a*1 and *a*2) are generated for that putative lineage and one sequence (*b*) is generated for the putative sister lineage). The frequency that the two candidate species samples (*a*1 and *a*2) are returned as monophyletic represents *P*1 in the equation $gdi = (3 \times P1 - 1)/2$. For our analyses we simulated 1 million trees for each comparison, which allowed for the calculation of a *gdi* point estimate under the MSci. We did not attempt to generate posterior distributions of *gdi* under the MSci given the computational burden. For all of these analyses, the program was run with a 16,000 generation burn-in and a 500,000-generation run with inverse-gamma priors of theta ∼ (3, 0.004), tau ∼ (3, 0.002), and phi ∼ (1,1).

Interisland Demographics

We performed demographic analyses using G-PHOCS (Gronau et al. 2011) to estimate effective population sizes (θ) of the extant and ancestral populations, population divergence times (τ), and bidirectional migration rates (*m*) between extant populations. Flanking and intronic sequences from 690 of the 709 loci were retained for analysis after filtering for missing data. We limited these analyses to flanking and intronic sequences because these regions are less likely to be under selection. Because of computational and software limitations, analyses were set up as pairwise comparisons for adjacently distributed lineages identified by the phylogenomic analysis with gene flow allowed in both directions. This simplified approach was also justified because demographic models with large numbers of migration bands can result in spurious inference of migration estimates (Freedman et al. 2014).

An initial run of 200,000 generations was used to assess convergence of the parameters, followed by one or

more runs of 500,000–1,000,000 generations depending on when ESS values reached 200 or greater. After removing 50,000–150,000 generations as burn-in based on parameter ESS values as well as visual confirmation of convergence by trace plots in TRACER, the remaining generations were summarized to assess the posterior distribution of the demographic parameters. A mutation rate (μ) of 1×10^{-9} mutations/site/year was used to convert unscaled parameter estimates into real world values (Brandley et al. 2011; Blom et al. 2019). We fixed $\alpha = 1$ and $\beta = 10,000$ for the gamma distribution set for priors on the θ and τ parameters, and $\alpha = 0.002$ and $\beta = 0.00001$ for the gamma distribution set for priors on the migration rate (m) parameters. Parameter estimates were converted to estimates of absolute effective population sizes (N_e), population divergence time in years (T), and migrants per generation from the source to target (M_{st}) (see supplementary materials in Gronau et al. 2011). These migration rate estimates are independent of our mutation rate assumption but do rely on an estimate of generation time (Gronau et al. 2011). Based on life history data from the closely related *Draco spilopterus* a generation time of 1 year was used to convert migration rates (Alcala 1967).

Cases where the 95% highest posterior density (HPD) of the total migration rate includes zero and the 95% HPD high estimate is less than 0.05 are considered to have experienced effectively no gene flow (see Freedman et al. 2014). Cases where $M < 0.1$ in either direction but do not meet the requirements for no gene flow will constitute "negligible" gene flow. When $0.1 \leq M < 0.5$ in either direction it will be considered "moderate" gene flow, and $M \geq 0.5$ in either direction represents "substantial" gene flow such that those lineages are expected to eventually merge (Shaffer and Thompson 2007). Cases where M is effectively zero or negligible in one direction and moderate to substantial in the other direction will be considered "unidirectional gene flow." While asymmetric migration has been used to describe opposing migration ratios around 1:1.5 (Wares et al. 2001), we take a more conservative approach and consider "bidirectional asymmetric gene flow" to constitute cases where the migration ratio is 1:5 or greater with no overlap in 95% HPDs.

Introgression Tests

We used the "ABBA-BABA" (Green et al. 2010; Durand et al. 2011) and D_{FOIL} tests (Pease and Hahn 2015) implemented in COMPD (Mussmann et al. 2020) to detect statistically significant imbalances of discordant biallelic site patterns across asymmetric four-taxon and symmetric five-taxon species tree hypotheses, respectively (Eaton and Ree 2013; Martin et al. 2013). These methods are effective at detecting introgression in targeted-sequence type data sets that summarize the genomic landscape (Pease and Hahn 2015; Lambert et al. 2019; Mussmann et al. 2020), though neither test

can detect introgression between sister taxa. The four-taxon method assumes three ingroup taxa (P) and one outgroup taxon (O) with the topology (((P1,P2),P3),O) where introgression can be inferred between P3 and either P1 or P2, though directionality is not given. The five-taxon D_{FOIL} method assumes four ingroup taxa with the topology (((P1,P2),(P3,P4)),O) where introgression and its directionality can be detected both between groups (P1+P2, P3+P4) and their ancestors (P12, P34). Pruned tree hypotheses used in the four- and five-taxon tests can be found in Supplementary Figures S16 and S17 available on Dryad, respectively. An alignment of the exon-capture data set (967,361 bp) was converted to biallelic site data and run as SNP frequencies where heterozygous sites are included (Eaton et al. 2015). Tests are performed on one individual per taxon and all possible combinations of individuals from each taxon were tested with 1000 bootstrap replicates per test used to estimate a population-wide Z-score and *P*-value. A Z-score above 3 or below -3 along with a *P*-value < 0.001 indicates that significant introgression has occurred (Durand et al. 2011; Pease and Hahn 2015; Mussmann et al. 2020). For the purposes of this study $1.5 < Z < 3$ or $-3 < Z < -1.5$ are considered nonsignificant signals of introgression and $-1.5 < Z < 1.5$ is considered a signal of no introgression.

Biogeographical Model Comparison

Dispersal models were compared using the software BIOGEOBEARS to estimate ancestral ranges (Matzke 2013, 2014). If a population was monophyletic in the phylogenomic analyses it was represented by a single sample (e.g., Lombok), while paraphyletic regions were represented as multiple samples (e.g., Alor). We designated 12 regions (the maximum number allowed by the software) to which samples were assigned corresponding to the following regions: Sunda Shelf, Lombok, West Sumbawa, East Sumbawa, Flores, Lembata, Sumba, Rote, Timor, Wetar, Alor, and Pantar. The program was initially run with an equal dispersal probability between all regions. Then we ran a "Distance" model ($+X$) where the minimum distances (in km) between regions is used to scale dispersal probability (Van Dam and Matzke 2016). We then tested four manual dispersal multipliers (see Supplementary Methods available on Dryad) to explore factors that may help explain the current biogeographical patterns: 1) a "Stepping-Stone" model where dispersal is highly favored between adjacent islands/regions, 2) a "Stepping-Stone +1 Leap" model where dispersals are also favored between two regions that are separated by one region, (3) an "Ocean Current" model where dispersal is favored between both adjacent regions as well as unidirectional dispersals following prevalent oceanic currents (Godfrey 1996; Gordon and Fine 1996; Kuhnt et al. 2004), and (4) a "Land-Bridge" model that gives the highest dispersal likelihood between connected regions (e.g., West and East Sumbawa) and regions that become periodically connected during glacial maxima

(e.g., Lombok + Sumbawa), an intermediate likelihood for adjacent regions that do not become connected, and the lowest likelihood between all other regions. These six analyses were run under the DEC, DEC + J, DIVALIKE, DIVALIKE + J, BAYAREALIKE, and BAYAREALIKE + J biogeographic models implemented in BIOGEOBEARS and the log likelihoods (LnL), Akaike information criterion (AIC), and corrected AIC (AICc) scores were compared to identify the most probable models.

RESULTS

Mitochondrial Screening

ML analysis of the *ND2* gene recovered multiple well-supported lineages that are either allopatric or parapatric (Supplementary Fig. S2 available on Dryad). Many relationships between major lineages are not well supported. The larger islands of Sumbawa, Flores, and Timor contain two or more parapatric nonsister lineages. Individual sample relationships and localities can be viewed in the supplementary materials (Supplementary Figs. S3–S6 available on Dryad). Preliminary *ND2* species delimitation analyses suggest as many as 19 species using the sequence-based program ABGD (Supplementary Fig. S7a–c available on Dryad) or 23 species with the phylogeny-based program MPTP (Supplementary Figs. S3–S6 available on Dryad).

Exon-Capture Data Characteristics

The concatenated alignment of both the targeted and flanking regions from all 709 loci is 967,361 bp, for an average locus length of 1364 bp. One library failed to sequence, resulting in ten outgroup and 89 ingroup samples composed of 34 *D. timoriensis* from five islands, and 55 *D. boschmai* from five islands. The average coverage for the targeted regions was >150×, while the flanking regions had approximately 100× coverage (Supplementary Fig. S8 available on Dryad). The average coverage for each individual library was variable (Supplementary Figs. S9 and S10 available on Dryad). The average number of samples per alignment was 98/99 (Supplementary Figs. S11a and S12a available on Dryad). Individual locus length varied from 60–5335 bp (F Supplementary Fig. S12b available on Dryad). The number of informative sites has a relatively linear relationship with the length of the loci, with very few outlier loci (Supplementary Fig. S11b available on Dryad), and on average the level of informative sites per alignment is ∼5% (Supplementary Fig. S12c available on Dryad). There is no clear relationship between the alignment length and the percentage of gaps in each alignment (Supplementary Fig. S11c available on Dryad), and the percentage of missing data for any one locus was never higher than 25% after filtering (Supplementary Fig. S12d available on Dryad).

Mitogenomics

ML estimation of the mitogenomic alignment produced a better-resolved phylogeny than the *ND2* gene alone (Fig. 2b). Preliminary species delimitation analyses suggest as many as 16 (ABGD) or 19 species (MPTP) in the Lesser Sundas (Fig. 2b, Supplementary Figs. S7d–f and S13 available on Dryad). The initial split within the Lesser Sundas is between a Lombok + West Sumbawa lineage, which are sister to each other, and all other populations. Within the remaining *D. boschmai* populations there are four major lineages: 1) central Flores + East Flores + Lembata, 2) East Sumbawa, 3) West Flores, and 4) Sumba. All major nodes at the base of *D. boschmai* lineages have bootstrap support of 100. *Draco timoriensis* is recovered as monophyletic and sister to the Sumba lineage with notable relationships that include 1) a Rote lineage nested within West Timor samples, 2) an Inner Banda Arc lineage sister to Timor-Leste, and 3) a weakly divergent Pantar lineage nested within Alor. Levels of divergence between lineages (Supplementary Table S3 available on Dryad) ranged from 2.9% (Rote/West Timor) to ∼11% (Flores lineages vs. Lombok or West Sumbawa). Mitogenomic divergence estimates (which should be interpreted with caution) suggest a Pliocene entry into the archipelago from the Sunda Shelf (∼4.3 Ma), with divergence of *D. boschmai* lineages occurring in the mid-Pleistocene to late Pliocene (∼1.5–3.2 Ma; Fig. 3b). *Draco timoriensis* is estimated to have diverged from the Sumba lineage ∼1.3 Ma with subsequent divergence between *timoriensis* lineages occurring ∼0.75–1 Ma.

Phylogenomics

The ML phylogeny of the exon-capture data set included similar lineages as recovered in the mitogenomic phylogeny but with important differences in the deeper relationships between lineages and with generally higher node support (Fig. 2a). The initial split is again between a Lombok + West Sumbawa lineage (BP =100) and all other populations (BP =100). However, the next major split is between Eastern Sumbawa and all other populations (BP =100), followed by a split between a Flores + Lembata lineage and all other populations (BP =100), in which Lembata is nested within Flores and most closely related to samples from adjacent East Flores. The remaining samples include Sumba as sister to *D. timoriensis* (BP =100). Within *D. timoriensis*, the major split (BP =100) is between an Inner Banda Arc assemblage (Wetar, Alor, and Pantar) and an Outer Banda Arc assemblage (Timor and Rote). Pantar is nested within the Alor samples, and these two island populations together represent a deeply divergent sister clade to Wetar. Rote is monophyletic but nested inside Timor.

The summary coalescent and pseudolikelihood species tree methods converged on the same topology as the ML tree with respect to the relationships of the major lineages (Fig. 2c; Supplementary Fig. S14a
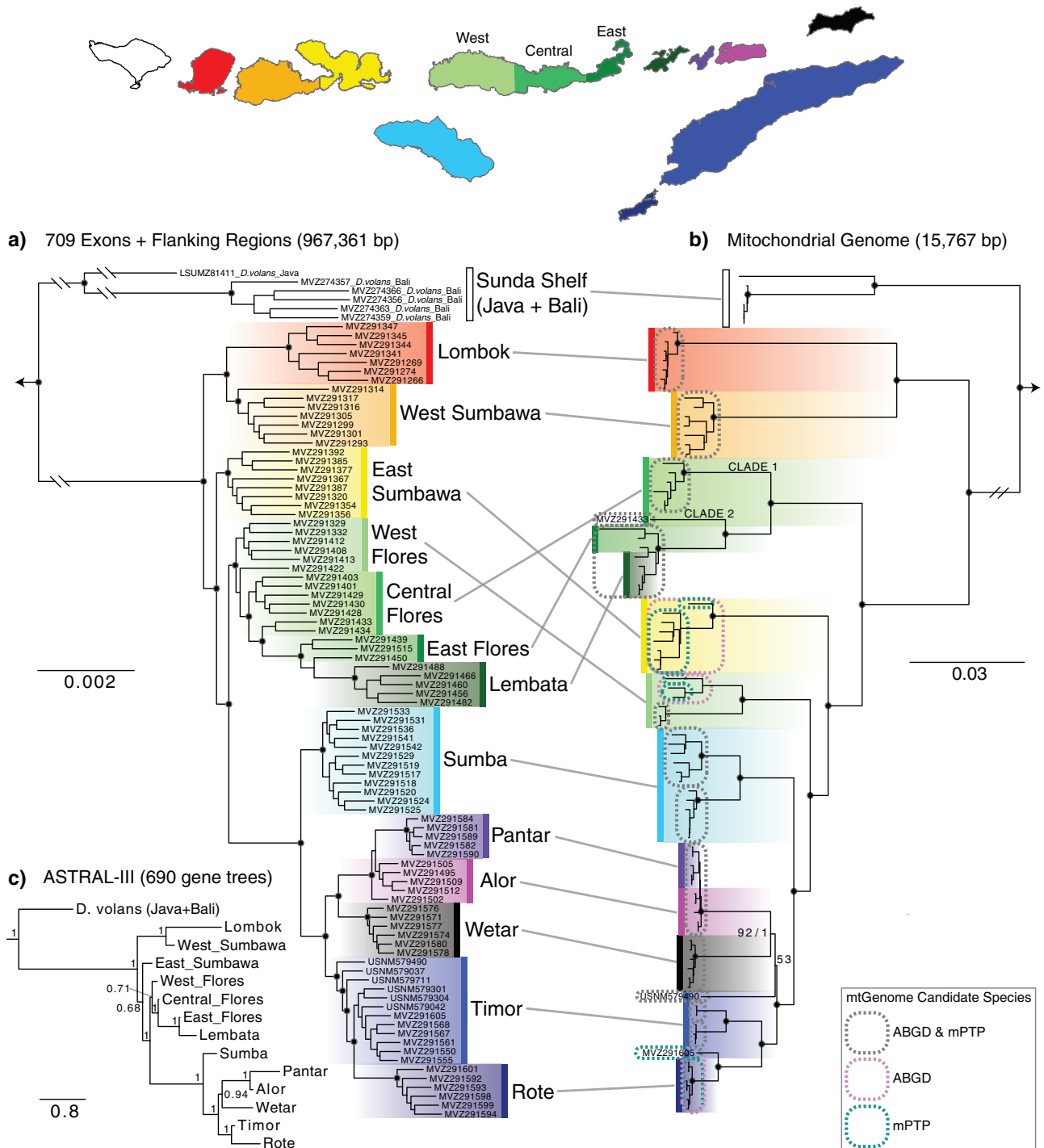
FIGURE 2.   Maximum Likelihood phylogenies of the a) exon-capture data set, b) the mitochondrial genome, and c) summary-coalescent species tree of the genomic data set. Node circles in (a) correspond to bootstrap proportions of 100. Node numbers in (b) represent bootstrap/Bayesian posterior probability values with black circles denoting 100/1 node support. Node numbers in (c) represent local posterior probability support. Solid gray lines connect corresponding clades to highlight topological differences. Scale bars for (a) and (b) are in units of mutations and the scale bar for (c) represents coalescent units.

and b available on Dryad)). The summary coalescent tree without species designations includes some topological differences from the concatenated analysis such as the placement of Lombok nested within West Sumbawa, the recovery of Alor and Pantar as sister lineages,

and the recovery of West Flores as monophyletic (Supplementary Fig. S14a available on Dryad)).

The time-calibrated Bayesian phylogenomic analysis converged on the same topology as the above methods and was used as the guide tree for biogeographical
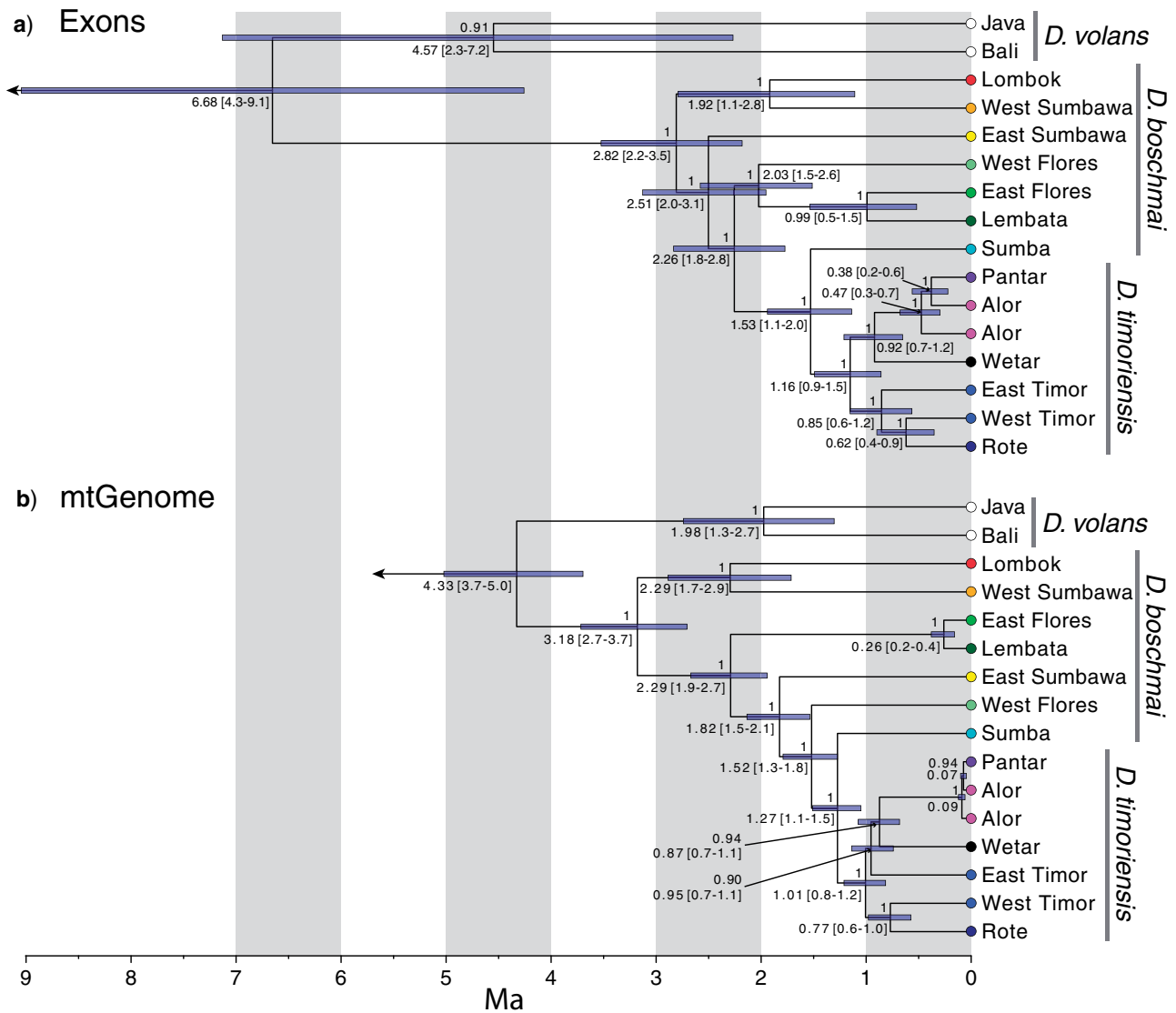
FIGURE 3. Time calibrated Bayesian phylogenies estimated with the a) exon capture data set and b) mitochondrial genome. Numbers above nodes represent posterior probability values and numbers below nodes represent node ages with 95% confidence intervals in brackets.

model testing analyses (Fig. 3a). These rough time estimates should be cautiously interpreted. This analysis estimated the initial entry into the archipelago around 6.7 Ma (95% CI = 4.3–9.1), with subsequent divergence between *D. boschmai* lineages occurring between 2.3 and 2.8 Ma (individual node 95% CIs range between 1.8 and 3.5 Ma). *Draco timoriensis* is estimated to have diverged from Sumba 1.5 Ma (95% CI = 1.1–2.0 Ma) with subsequent intraspecific lineage divergences occurring between 0.4 and 1.2 Ma (individual node 95% CIs range between 0.2 and 1.5 Ma).

Population Structure and Species Delimitation

PC1 (21.0% of variance) plotted against PC2 (8.0% of variance) recovered nine clusters with the Sunda Arc clusters of *D. boschmai* arranged along PC2 in the same arrangement as they are physically positioned longitudinally (Fig. 4a). Further PCA limited to *D. timoriensis* (PC1 = 22.4% of variance; PC2 = 14.1% of variance) delimits three major clusters corresponding to Wetar, Alor + Pantar, and Timor + Rote, though there are nonoverlapping clusters for Alor, Pantar, Timor, and Rote (Fig. 4b). The PCA focusing on Flores + Lembata (Fig. 4c) delimited each of the four suspected lineages with one West Flores sample (MVZ291422) that lies closest to the central lineage possibly representing a fifth genetic population.

The Mantel test indicated that significant signatures of IBD ($p < 0.001$ in all cases) are present in each of the subsets of samples we tested (Supplementary Fig. S15a–e available on Dryad)). The plot of geographic versus genetic distance for all Lesser Sunda samples reveals
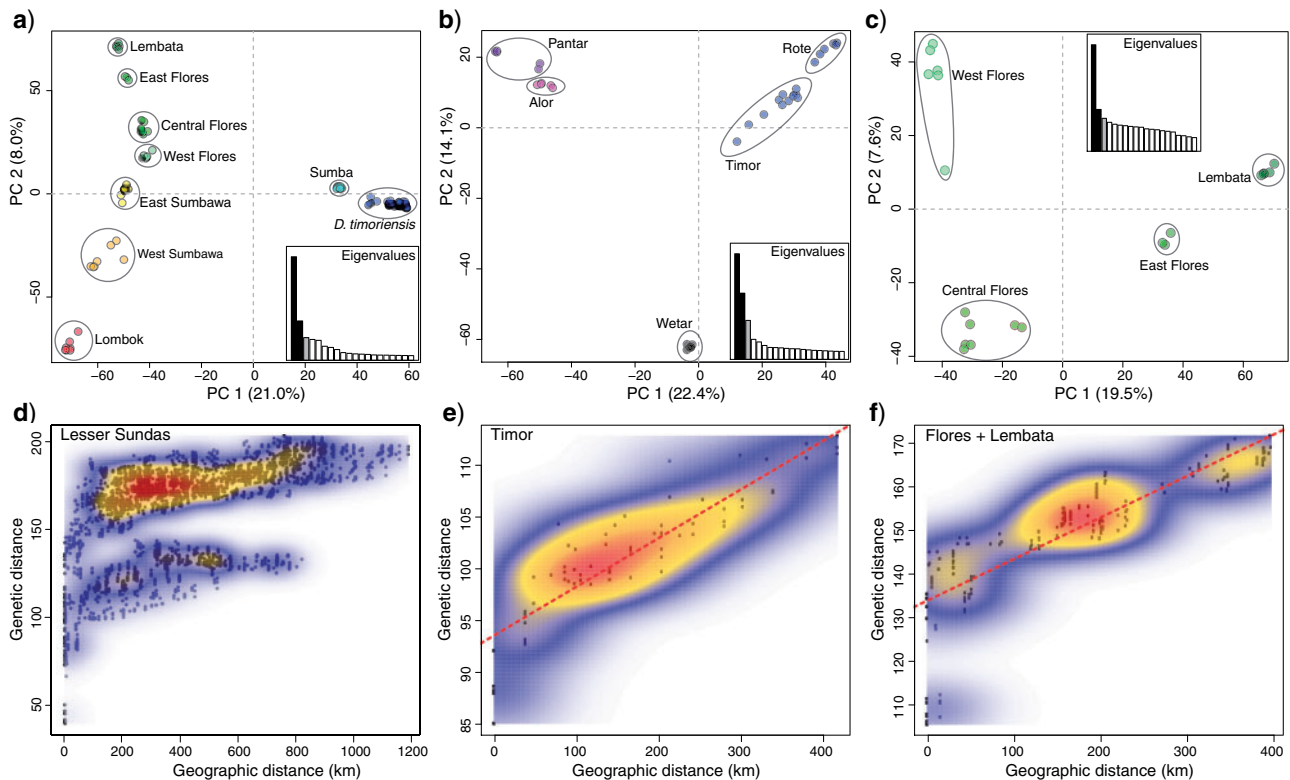
FIGURE 4.   Principal component analysis of genetic covariance for the exon data set consisting of 107,689 single nucleotide polymorphisms for a) all Lesser Sundas *Draco*, b) within *D. timoriensis*, and c) within Flores + Lembata. Plots of geographic versus genetic distance for d) all Lesser Sundas, e) Timor, and f) Flores + Lembata samples. Each point represents a two-sample comparison and colors correspond to point density estimation with warmer (red) colors indicating higher density. The red lines in panels e and f represent best-fit straight lines ($P < 0.001$ for both).

two distinct swaths of points corresponding to multiple geographically and genetically differentiated groups of samples (Fig. 4d). The Sunda Arc (Supplementary Fig. S15f available on Dryad), Timor + Rote (Supplementary Fig. S15g available on Dryad), and Flores + Lembata (Fig. 4f) plots recover point densities with discontinuities indicating distant and differentiated populations due to breaks in gene flow across the range. The Timor plot displays a single consistent cloud of points indicative of IBD due to range expansion (Fig. 4e).

The Evanno method results can be found in Supplementary Table S4 available on Dryad, bar plots for optimal K and other supported values of K can be viewed in Figure 5, and results from two alternative SNP sets are in Supplementary Figure S16 available on Dryad. The optimal K =2 within the Sunda Arc group corresponds to 1) Lombok + West Sumbawa and 2) East Sumbawa + Flores + Lembata, though K =3 and K =5 were also supported. Within Sumba + *D. timoriensis* the optimal K =5 corresponds to 1) Sumba, 2) Rote, 3) Timor, 4) Wetar, and 5) Alor + Pantar. Closer inspection of various groups recovered optimal K =2 for Sumbawa, K =2 for Flores + Lembata, K =2 for Timor + Rote, and K =2 for Pantar + Alor. Admixed individuals occur at contact zones within each of the larger islands of Sumbawa, Flores, and Timor (Fig. 5c–e) while Pantar and Alor lack admixed

individuals (Fig. 5f). Setting K =3 or 4 within Flores + Lembata revealed further structure corresponding to the previously described groups of West, Central, and East Flores, and Lembata. Setting K =3 within Rote + Timor defines the West and East Timor groups.

Species delimitation analyses using BPP software and the A10 model for both the *Draco boschmai* and *D. timoriensis* populations recovered each included lineage as a distinct species with posterior probability values of 1.0 for each node in the guide trees (Supplementary Fig. S17a and b available on Dryad). These species designations within *D. boschmai* delimit samples from Lombok, West Sumbawa, East Sumbawa, Flores (where West, Central, and East populations are all considered one candidate species), Lembata, and Sumba, and within *D. timoriensis* delimit samples from Timor, Rote, Wetar, Alor, and Pantar each as distinct species. The analysis focused on *D. boschmai* populations on Flores and Lembata, which was somewhat of a test of BPP for over splitting, recovered five species only lumping East Flores samples with Lembata (Supplementary Fig. S17c available on Dryad). We calculated *gdi* scores for 18 hypothetical species comparisons using output obtained from BPP 4.3.8 under both the MSC and MSci models. Jackson et al. (2017) proposed thresholds of 0.2 and 0.7, wherein a *gdi* score below 0.2 indicates a firm lack of species status, a score above 0.7 indicates solid species
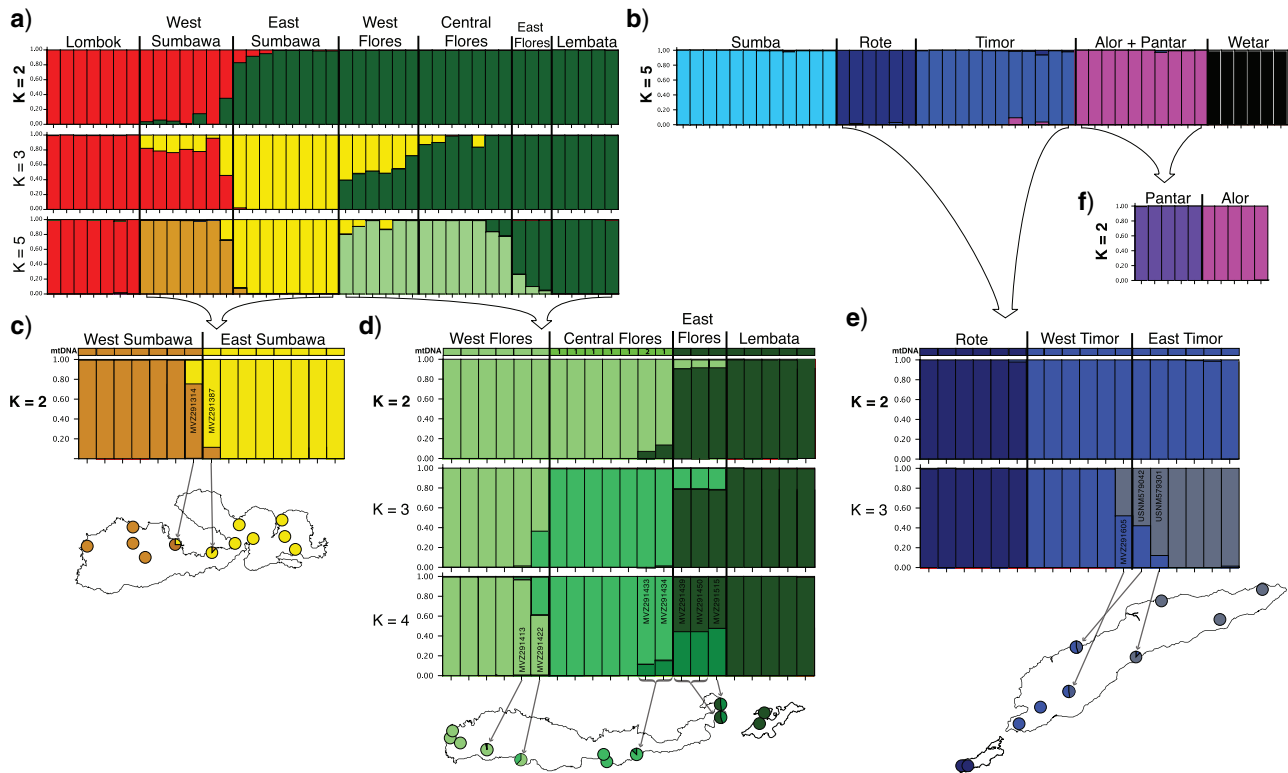
FIGURE 5.   Population structure bar plots of shared ancestry with the optimal K shown in bold for a) K = 2, 3, and 5 for the Sunda Arc populations, b) K = 5 for Sumba + *D. timoriensis*, c) K = 2 for Sumbawa, d) K = 2-4 for Flores + Lembata, e) K = 2-3 for Rote + Timor, and f) K = 2 for Pantar + Alor. Values on the y-axis of bar plots represent the proportion of ancestry from each population. Colored boxes above each bar plot in panels c-e indicate the mitochondrial lineage that sample belongs to, with Central Flores subclade numbers corresponding to Figure 2b.

status, and scores between 0.2 and 0.7 represent an ambiguous "gray zone." Many of our comparisons fell into this gray zone with at least one of the bidirectional comparisons (Fig. 6). The only comparisons that scored *gdi* values >0.7 under both the MSC and MSci were Bali–Lombok (*D. volans–D. boschmai*) and Lembata–Pantar (*D. boschmai–D. timoriensis*). The only comparison to score *gdi* values <0.2 under both the MSC and MSci was West Flores–Central Flores.

Demographic Analyses

The colonization of the Lesser Sundas from the Sunda Shelf is estimated to have occurred ~5.4 Ma, and a most recent common ancestor of all extant Lesser Sundas *Draco* is predicted to have existed 2.6–2.8 Ma (Fig. 7a). Population divergence times between lineages within *D. boschmai* range from 2.2 to 2.8 Ma, and between *D. timoriensis* lineages from 0.5 to 1.0 Ma. Effective population sizes are not discussed here but can be found in Supplementary Table S5 available on Dryad. Estimates of migrants per generation range from 0 to 2 (Fig. 7b). When considering mean migration rate estimates the adjacent regions that appear to have experienced moderate gene flow ($0.1 < M < 0.5$ in either direction) since divergence include West Sumbawa/East Sumbawa, East Sumbawa/Flores, Pantar/Alor, and,

to a much lesser extent, Timor/Rote. Adjacent regions with substantial gene flow ($M > 0.5$ in either direction) include West Flores/Central Flores, Central Flores/East Flores, East Flores/Lembata, and West Timor/East Timor. If the 95% HPD low migration estimates are used the Timor/Rote populations would be considered to have effectively no migration ($M < 0.1$), and if the 95% HPD high migration estimates are used the Pantar/Alor populations would be considered to have substantial gene flow ($M > 0.5$) (Supplementary Table S5 available on Dryad). Westward-dominant asymmetric gene flow is detected in the Flores and Lembata assemblage with unidirectional gene flow from Central to Western Flores and bidirectional asymmetric gene flow between East/Central Flores and Lembata/East Flores. Bidirectional asymmetric migration is detected between Pantar/Alor and Timor/Rote and unidirectional gene flow is detected from West to East Timor.

Introgression Tests

We detected significant signals of introgression with $Z < -3$ in tests that excluded the two contact zone individuals and only considered "pure" West and East Sumbawa lineages (Table 1; Supplementary Fig. S18 available on Dryad). Introgression was
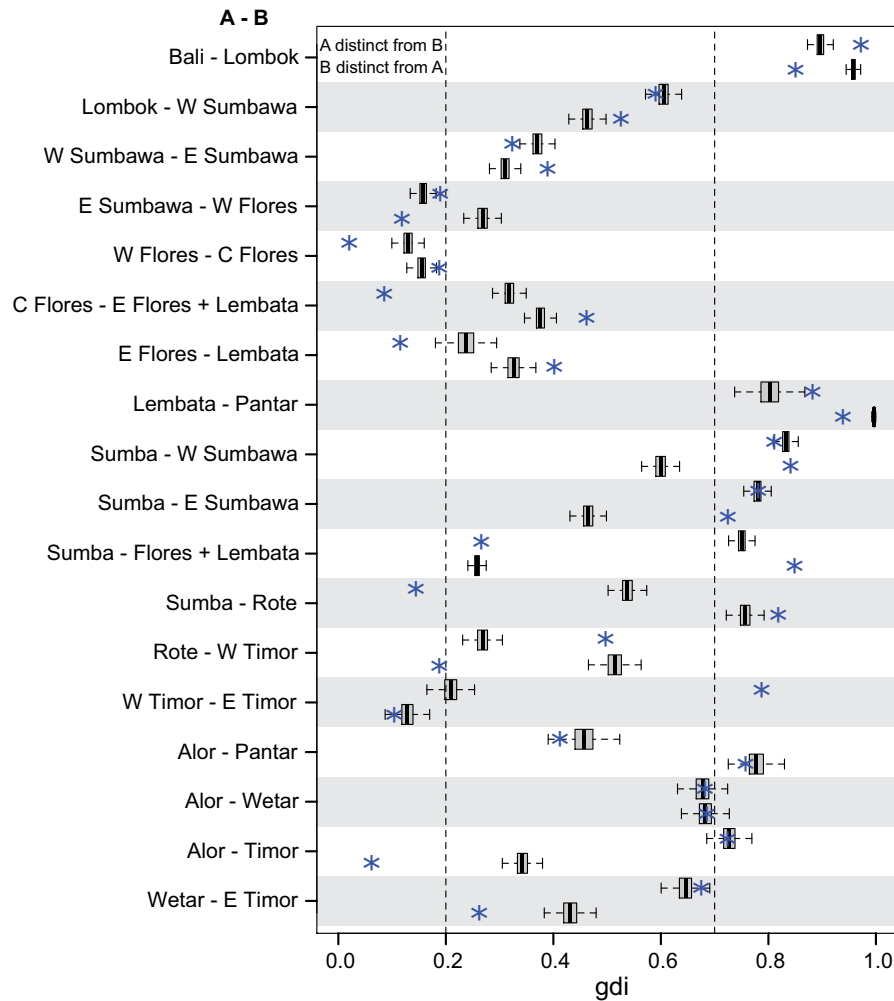
FIGURE 6.   Genealogical divergence index (*gdi*) estimates derived from the multispecies coalescent analyses without introgression (MSC) shown as box and whisker plots, and the multispecies coalescent analyses with introgression (MSci) shown as blue stars (point estimates). Values below 0.2 may indicate that the two groups are conspecific, values above 0.7 may indicate that they are distinct, and intermediate values are considered an ambiguous zone.

detected between East Sumbawa and the West Sumbawa contact zone (Z = 4.53) when treating each contact zone individual as a population. Tests for introgression within the Flores and Lembata lineages were nonsignificant but suggest possible low-levels of introgression between West and Central Flores (−3 < Z < −1.5). The five-taxon tests indicated nonsignificant levels of introgression between the East Sumbawa ancestor and West Sumbawa, and between the East Sumbawa/West Flores ancestor and West Sumbawa (Table 2; Supplementary Fig. S19 available on Dryad). Directionality of ancestral introgression cannot be inferred with the D_{FOIL} analysis.

Biogeographical Model Comparison

Models including the Jump parameter (J) were significantly preferred over models without it, with DIVALIKE + J as the top supported biogeographic

model within each of the six dispersal modifier scenarios (Supplementary Table S6 available on Dryad). The top three models according to the AIC scores are 1) DIVALIKE + J under "Stepping-Stone + 1 Leap" (LnL = −26.74; AIC = 59.48; AIC_c = 61.48), 2) DIVALIKE + J under "Land-Bridge"' (LnL = −27.13; AIC = 60.26; AIC_c = 62.26), and 3) DIVALIKE + J under 'Distance' (LnL = −26.18; AIC = 60.35; AIC_c = 63.99). The DIVALIKE + J "Stepping-Stone +1 Leap" model infers a best joint history whereby Lombok was first colonized from the Sunda Shelf, followed by colonization of East Sumbawa from Lombok and Flores from East Sumbawa, followed by southward dispersal from Flores to Sumba and eastward range expansion across Flores to Lembata (Fig. 8a). Sequentially, the next set of colonization events include Lombok to West Sumbawa, and Sumba to Timor. Both Rote and Wetar would have been colonized by the Timor population, with subsequent westward dispersal from Wetar to Alor and

then from Alor to Pantar. The DIVALIKE + J 'Land-Bridge' model infers that the initial region colonized from the Sunda Shelf is East Sumbawa, followed by westward range expansion into West Sumbawa, and dispersal from East Sumbawa to Sumba (Fig. 8b). The next set of movements include dispersal from Sumba north to Flores with eastward range expansion across Flores to Lembata, and a westward dispersal from West Sumbawa to Lombok. All colonization events of the Banda Arc islands are the same as for the top supported model. The DIVALIKE + J "Distance" model is nearly identical to DIVALIKE + J "Stepping-Stone +1 Leap" except that Sumba is colonized from East Sumbawa followed by dispersal from Sumba to Flores (Fig. 8c). While AIC_c scores place the DEC + J "Stepping-Stone +1 Leap," DEC + J "Land-Bridge," and BAYAREALIKE + J "Stepping-Stone +1 Leap" as the third, fourth, and fifth best supported models, respectively, the best joint histories of these are the same as the DIVALIKE + J scenarios for those same dispersal modifiers.

Discussion

Establishing a proper accounting of regional species diversity is a scale-dependent problem, with more geologically, topographically, and climatologically complex focal regions presenting ever increasing challenges for biodiversity researchers and conservation efforts. For this reason, relatively "simple" insular systems have played a particularly prominent role in establishing our understanding of speciation mechanisms (Grant and Grant 2008; Shaw and Gillespie 2016). The Lesser Sunda Archipelago is an example of an insular system that prior biogeographers have typically treated as a straightforward set of stepping-stone islands (Darlington 1957; How and Kitchener 1997; Hisheh et al. 1998). However, closer inspection reveals that this tropical archipelago is rich in the complexity that makes biodiversity accounting—and thus the downstream diversification studies dependent on that accounting—difficult.

Because the Lesser Sundas are a large archipelago composed of more than 500 oceanic islands, divergence in allopatry is presumed to have been rampant. Identifying species amongst allopatric lineages often inspires disagreement among taxonomists because, depending on one's species concept, they require subjective decisions regarding whether observed morphological divergences are sufficient to predict reproductive isolation (Coyne and Orr 2004). The Lesser Sundas present an additional layer of complexity because of the exceptionally dynamic tectonic history of the archipelago, resulting in islands that appeared at different times and through distinct geological processes over the last ~11 Myr (Hall 2009, 2011; Spakman and Hall 2010; Rigg and Hall 2011; Audley-Charles 2011), with some of the larger islands having likely first existed as archipelagos of smaller islands that later merged (Muraoka et al. 2002, 2005; Purwandono et al. 2019). Furthermore, much of the tectonic activity

has taken place in the last 2 Myr (Roosmawati and Harris 2009; Tate et al. 2017), suggesting the possibility of many young lineages including morphologically cryptic species, which poses particular challenges both for phylogenomic inference and for species delimitation (Shaffer and Thompson 2007; Borregaard et al. 2017). Thus, an analysis of species diversification and biogeography in the Lesser Sundas first requires a fine-scale delimitation of species across a heterogeneous archipelago wherein some lineages may have been isolated on islands for millions of years without experiencing subsequent gene flow, others may have been isolated on islands but with opportunities for some amount of migration, and still others co-occur with other lineages on large islands where gene flow is possible at points of secondary contact.

Phylogenetic Lineage Identification and Geographic Structure

We attempted to identify putative *Draco* species in the Lesser Sundas using phylogenetic analyses of a well-sampled mitochondrial *ND2* data set, as well as with mitochondrial genome and exome-capture data sets. We employed two species delimitation methods with the *ND2* data set, and these approaches inferred unexpectedly large numbers of *Draco* species in the Lesser Sundas, with ABGD and MPTP inferring 19 and 23 species, respectively. Based on phylogenetic analyses of the complete mitochondrial genome and exome-capture data sets, we identified a smaller number of well-supported, deeply divergent clades for consideration as possible species. There was substantial agreement between these data sets in terms of putative lineages, particularly for clades bounded by marine barriers. Notably, the points of disagreement between the mitogenomic and exome-capture phylogenetic analyses regarding putative lineages reflect the interpretation of structure present within the two largest islands in the archipelago, Flores (plus its land-bridge neighbor, Lembata) and Timor (plus its very close neighbor Rote), which likely result from within-island secondary contact and subsequent gene flow. Unlike the mitogenomic analysis, the exome-capture analysis does not support three discrete clades composed of Flores and Lembata samples, but rather presents as a west-to-east comb with as many as seven divergent subclades. A similar east to west comb is present for Timor as well. The nuclear phylogenetic estimate indicates a pattern of range expansion followed by isolation-by-distance driven divergence with gene flow, that the single locus mitogenomic data set is ill-suited to address.

Species Delimitation Using Genomic Clustering and Demographic Analyses

Traditional BPP species delimitation analyses returned nearly the maximum number of species in the guide
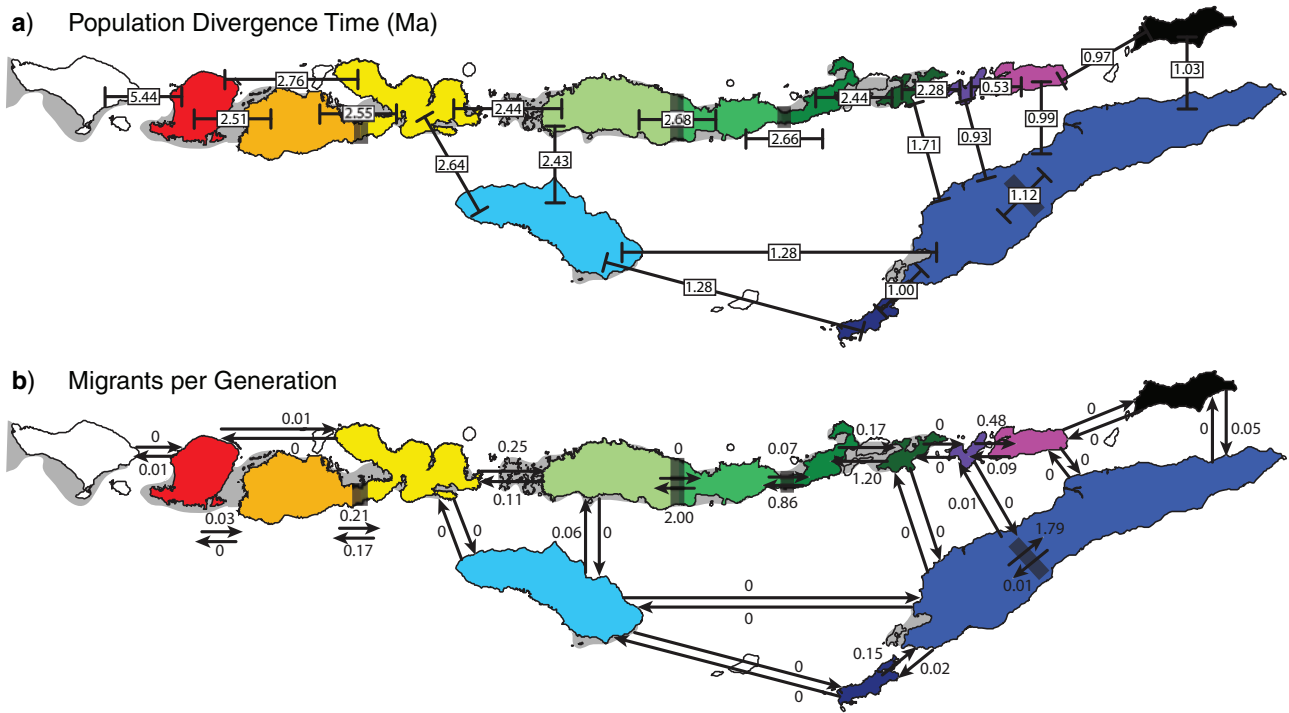
FIGURE 7.   Demographic estimates of a) population divergence time in millions of years, and b) migrants per generation between adjacent phylogenomic lineages. The light gray regions between islands represent the 120 m depth contour to depict islands that may become land bridged during glacial maxima.

TABLE 1.   Four-taxon tests for introgression with significant results ($Z < -3$ or $Z > 3$ and $P < 0.001$) in bold

N	(((P1, P2), P3), O)	ABBA	BABA	D	Z	P	Implication
1470	(((Lom, WSu), ESu), Bali)	140.1 ± 12.3	64.3 ± 7.1	0.37 ± 0.05	**−7.50**	**<10e−13**	**P2 <−> P3**
1260	(((WFl, ESu), WSu), Bali)	198.8 ± 20.4	92.4 ± 7.9	0.36 ± 0.06	**−6.32**	**<10e−9**	**P2 <−> P3**
210	(((WSu CZ, WSu), ESu), Bali)	108.9 ± 8.9	163.6 ± 11.5	-0.20 ± 0.04	**4.53**	**<10e−5**	**P1 <−> P3**
210	(((ESu CZ, ESu), WSu), Bali)	137.5 ± 13.6	175.1 ± 16.4	-0.12 ± 0.07	1.84	0.066	P1 <−> P3
882	(((EFl, CFl), WFl), ESu)	139.4 ± 22.7	111.4 ± 11.2	0.11 ± 0.05	−2.03	0.042	P2 <−> P3
630	(((Lem, EFl), CFl), WFl)	127.1 ± 19.3	130.8 ± 25.5	-0.01 ± 0.05	0.20	0.839	No Introgression

N = number of unique sample comparisons; **ABBA/BABA** = the number of allelic configurations and their standard deviation; **D** = Patterson's D-statistic; **Z** = statistical confidence score derived from the mean and standard deviation of all individual D-statistics; **P** = P-value; Lom = Lombok; WSu = West Sumbawa "pure"; WSu CZ = West Sumbawa contact zone; ESu = East Sumbawa "pure"; ESu CZ = East Sumbawa contact zone; WFl = West Flores; CFl = Central Flores; EFl = East Flores; Lem = Lembata.

TABLE 2.   Five-taxon tests for introgression with significant results ($Z < -3$ or $Z > 3$ and $P < 0.001$) in bold

N	((((P1, P2), (P3, P4)), O)	Z (DFO)	Z (DIL)	Z (DFI)	Z (DOL)	Implication
210	(((ESu, ESu CZ), (WSu CZ, WSu)), Bali)	1.58	1.82	0.64	0.54	P12<−> P4
3125	(((WFl, ESu), (WSu, Lom)), Bali)	−2.76	**−4.02**	1.54	0.77	P12 <−> P3

N = number of unique sample comparisons; **Z** = statistical confidence score derived from the mean and standard deviation of all individual D-statistics for the four D_{FOIL} statistics; Lom = Lombok; WSu = West Sumbawa "pure"; WSu CZ = West Sumbawa contact zone; ESu = East Sumbawa "pure"; ESu CZ = East Sumbawa contact zone; WFl = West Flores; P12 = ancestor of P1 and P2.

tree, a lack of discrimination that has been shown in a number of other studies (e.g., Jackson et al. 2017; Leaché et al. 2019). The analyses based on *gdi* scores returned a more nuanced assessment of species status with many populations falling into the gray area between what might be called definitely indistinct lineages (*gdi* < 0.2) and definitively distinct species (*gdi* > 0.7). Indeed, of the 18 pairwise comparisons undertaken, two were consistent with definitive species status in both of

the bidirectional comparisons (i.e., *D. volans* from Bali and *D. boschmai* from Lombok are distinct). The *gdi* scores obtained with and without accounting for gene flow (MSci vs. MSC, respectively) were often similar, but sometimes exhibited substantial disagreement (e.g., the Sumba versus Flores + Lembata comparison). Furthermore, our results showcase one of the concerns noted by Leaché et al. (2019)—the possibility for awkward scenarios in which species A is distinct at the
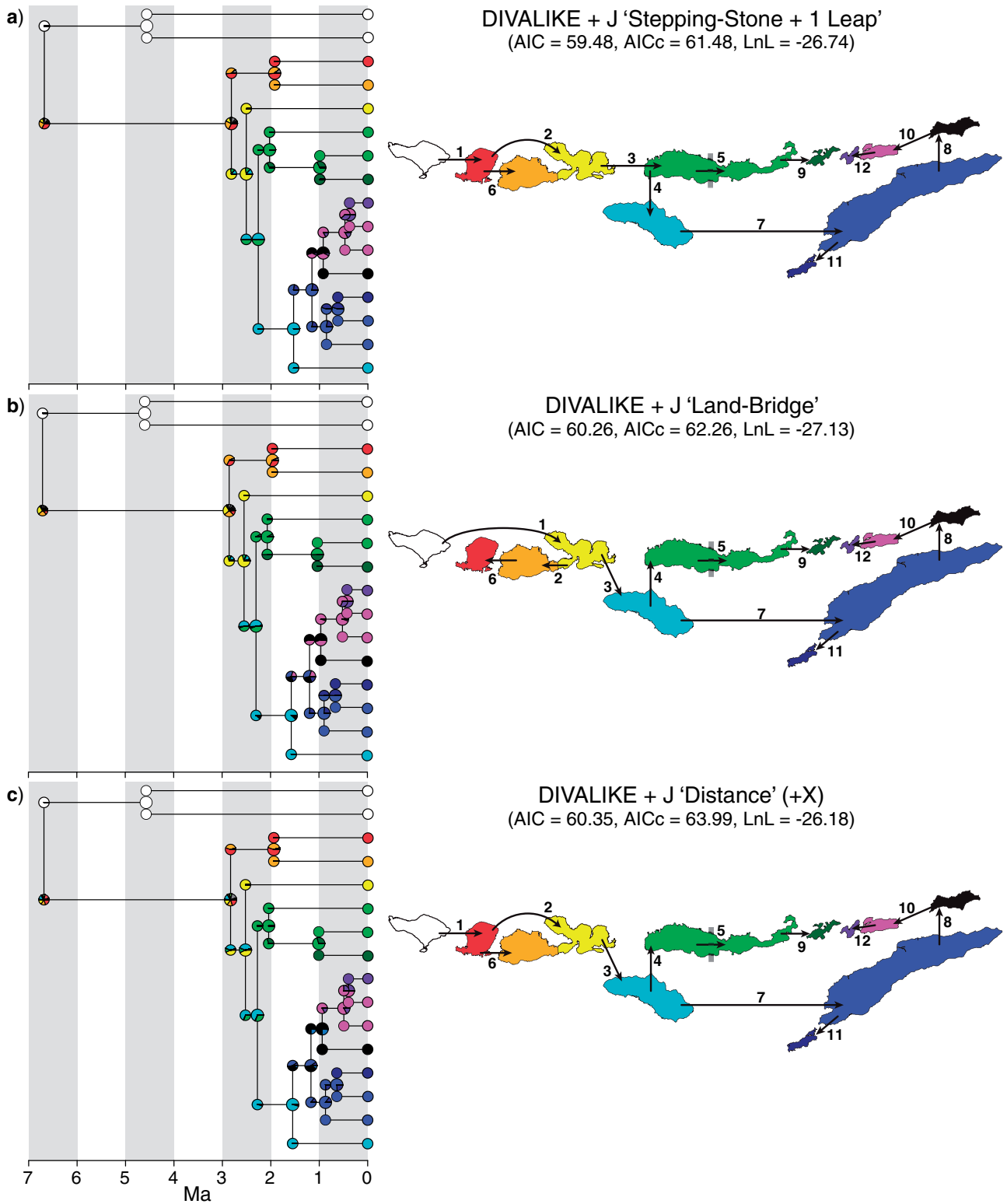
FIGURE 8. Biogeographical model testing results. The phylogenies show the relative probability of each ancestral range as pie charts at nodes and corners with colors corresponding to each range as colored on the maps. The maps display sequentially numbered dispersal sequences following the best joint history from the models.

species level from species B but species B is not distinct from species A. For example, our MSci analysis of the Alor and Timor Island populations returns a *gdi* of 0.72 for Alor versus Timor but 0.06 for Timor versus Alor. In summary, these analyses—particularly when considered in the context of our population demographic results—provide important additional information beyond that obtained with traditional BPP species delimitation analyses but still require downstream confirmation with more direct assessments of gene flow. Similarly, PCA and STRUCTURE analyses were more congruent with the mitogenomic and exome-capture phylogenetic analyses in terms of inferred lineage boundaries, suggesting six-eight (STRUCTURE) or eight (PCA) variably discrete *D. boschmai* clusters and four-five *D. timoriensis* clusters. While these 10–13 clusters may in fact represent cryptic as well as noncryptic species, documentation of genetic structure alone is insufficient evidence of a cryptic species boundary in our view, and confirmation requires obtaining additional evidence of a substantial barrier to gene flow between such lineages (Shaffer and Thompson 2007; Chan et al. 2020). In this case, the STRUCTURE analyses exhibited a clear signature of gene flow at each of the within-island lineage boundaries, though this appears to be restricted to zones of contact.

Population genetics theory suggests that migration rates greater than 0.5–1.0 migrants per generation are expected to result in lineages eventually merging when divergence is the result of genetic drift (Wright 1931; Slatkin 1987), and we consider this a useful threshold for assessing cryptic species boundaries. Here we apply the threshold of 0.5 migrants per generation following Shaffer and Thompson (2007), which suggests that there are ten discrete lineages within the Lesser Sundas with moderate or effectively no gene flow between most lineages (Fig. 7b). This threshold may be suboptimal for lineages that experienced cyclical bursts of gene flow followed by periods of complete isolation, in this case driven by fluctuating island connectivity and proximity due to sea level change and geological processes. However, use of the 0.5 migrant per generation threshold along with *gdi* values, introgression tests, and population clustering analyses offer the most complete picture of lineage distinctiveness in complex cases such as these.

The deep structure observed among Flores + Lembata samples and Timor samples is more nuanced exhibiting a strong signature of unidirectional and/or bidirectional asymmetric gene flow between lineages (Fig. 7b). Interestingly, the direction of migration in these cases is opposite of the inferred direction of range expansion, a phenomenon that warrants further explanation. Introgression tests of the Flores/Lembata lineages, which could only test for eastward introgression, were nonsignificant further supporting the migration estimates of near-zero eastward migration. Application of the 0.5 migrants per generation criterion results in Flores + Lembata and West + East Timor being treated as a single structured species or as incipient species in the process of merging or likely to merge following secondary contact (see Garrick et al. 2019). Given the challenges of accurately estimating rates of migration, as well as the fact that asymmetric gene flow between populations with different population sizes can obscure true species boundaries (Jiao and Yang 2021), we do not consider this inference conclusive. A fine-scaled assessment of *Draco* lineage boundaries on the islands spanning "Greater Flores" (Komodo, Rinca, Flores, Adonara, Solor, and Lembata) is both warranted and necessary to fully elucidate species limits in this sector of the Sunda Arc. Similarly, migration estimates from Pantar to Alor Islands are slightly below our threshold value of 0.5 migrants per generation (though the 95% HPD high value is 0.6), and we tentatively consider these insular populations to be valid cryptic species.

Secondary Contact Due to the Merger of Paleoislands Via Uplift

We initially considered the most surprising finding of this study to be the discovery of deep genetic boundaries within individual islands, as prior research on Lesser Sundas biogeography has focused almost entirely on marine barriers as drivers of differentiation (but see Reilly et al. 2019a). However, in retrospect, intra-island divergences should have been expected, particularly for the larger islands in the volcanic Sunda Arc that must have arisen as clusters of independent volcanoes that only amalgamated into large, elongated but contiguous islands after millions of years in isolation (Muraoka et al. 2005). The initial paleo-islands would have been colonized by overwater dispersal allowing lineages to diverge in allopatry, with the subsequent merger of these islands allowing for secondary contact and setting in motion dynamic interactions between once-isolated lineages.

The deep divergence between the *Draco* lineages of West and East Sumbawa is consistent with a long period of isolation on separate paleoislands followed by uplift in the intervening isthmus. Uplift rates are variable but similar across the Sunda Arc (Pirazzoli et al. 1993; Merritts et al. 1998; Suminto et al. 2008) and the uplift of Sumbawa is confirmed by series of coral terraces along its coastline (Borel-Best et al. 1985). This implies that the narrow, 60 m elevation isthmus comprising the Sumbawa *Draco* contact zone was an oceanic barrier until relatively recently (e.g., 200,000 years ago given a 0.3 mm/year rate of uplift), and long enough after their divergence (∼2.5 Ma) to allow for allopatric speciation. These lineages, which are not sister taxa, now experience gene flow corresponding to approximately one migrant every five generations in each direction. While clustering analyses suggest that this gene flow is limited to the contact zone, introgression tests suggest that lineage-wide introgression has also occurred.

The tectonics of the Sunda Arc between Flores and Lembata are extremely complex, with volcanism and

uplift being the two dominant factors influencing the region's topography (Suminto et al. 2008). A subduction zone in the west transitions to a collision zone in the east with a major geological break occurring at the "Bajawa Volcanic Rift Zone," where the alignment of volcanoes changes from a west–east orientation in West Flores to a series of NE–SW oriented en echelon (parallel structures formed by shearing/collision) shaped volcanic segments in the east (Muraoka et al. 2005; Pacey et al. 2013; Purwandono et al. 2019). Central Flores has uplifted nearly 800 m over the last 2.5 Myr (~0.32 mm/year), and the elevation of a series of paleoshorelines indicates that a marine barrier existed until recently in this area (Muraoka et al. 2002, 2005). This happens to be where the West and Central Flores *Draco* lineages meet. Similarly, the break between the Central and East Flores populations occurs near Maumere, a narrow region that lies below 300 m in elevation and resides between two en echelon volcanic segments that also may have been submerged until recently (Muraoka et al. 2002, 2005). The Maumere region was also shown to be a lineage boundary in the fanged frog *Limnonectes kadarsani* (Reilly et al. 2019a) and the scincid lizard *Sphenomorphus melanopogon* (Reilly 2016). Lembata represents yet another distinct en echelon volcanic complex separate from the one in East Flores (Muraoka et al. 2002, 2005), but here uplift in the region has not yet created a permanent subaerial connection and two shallow marine barriers remain when they are not connected during glacial lowstands (Heaney 1991).

Historical Biogeography and Patterns of Colonization

Perhaps the key finding of our biogeographical analyses is that several distinct alternative biogeographical scenarios should be considered similarly plausible, which illustrates the complexity of biogeographical inference for this region. Unsurprisingly, the top models include initial colonization of the archipelago by way of the westernmost islands in the chain, Lombok and Sumbawa. However, whereas most prior studies have assumed a standard stepping-stone model for the Lesser Sundas, this is the one possibility that is not supported in our analyses. Rather, we find support for models that include initial colonization of Lombok followed by independent colonizations of West and East Sumbawa from Lombok, or an initial colonization of East Sumbawa followed first by colonization of West Sumbawa and then Lombok. Similarly, there is disagreement among models for the next step in the colonization sequence, which involves either colonization of Flores to the east (followed by a jump from Flores to Sumba) or Sumba to the south (followed by a jump from Sumba to Flores). All other well-supported models support a stepping-stone route from West Flores to Lembata and a circuitous series of overwater dispersal events from Sumba into the Banda Arc islands.

One instance where hypothetical island connectivity does not match our results involves Lombok and West Sumbawa populations, which, though separated only by the shallow Alas Strait (depth < 120 m), are among the most divergent *Draco* lineages in the Lesser Sundas and show effectively zero gene flow since divergence, a pattern also seen in frogs of the genus *Limnonectes* (Reilly et al. 2019a). We suspect that the depth of the Alas Strait was deeper than 120 m until recently when uplift and/or sediment fill may have reduced its depth to less than 120 m. Thus, this hypothetical connection between Lombok and Sumbawa may be somewhat inconsequential for early arriving species but more important for the cohesion of recently arrived species (see Reilly et al. 2019b).

Biogeographical analyses all suggest that Pantar was colonized from Alor, a scenario supported by the fact that Pantar is nested within Alor in most of our phylogenetic estimates. Given the rates of uplift in Alor (~1 mm/year; Hantoro et al. 1994) and sea level fluctuations, Alor and Pantar have either never been connected or may not have become connected until the late Pleistocene (~40,000 years ago) depending on different geological models (see Fig. 4 in Kealy et al. 2020). One hypothesis is that our genomic estimate of population divergence of ~0.5 Ma ago may represent the initial divergence resulting from overwater dispersal, while the much more recent mitochondrial divergence (~70,000 years ago) may represent mitochondrial introgression of the Alor haplotype into Pantar following a second wave of migration once the channel separating these islands narrowed or perhaps the islands even became physically connected.

Insular Species Formation Within the Lesser Sundas

Hypotheses regarding candidate species within what is now *D. boschmai* comprise five lineages, including 1) Lombok, 2) West Sumbawa, 3) East Sumbawa, 4) Sumba, and 5) Flores + Lembata. The divergence times estimated for these candidate species range between 2.2–2.8 Myr based on demographic estimates and 1.9–2.8 Myr based on calibrated phylogenetic estimates, which is similar to divergence times estimated for *Limnonectes kadarsani* lineages (1.9–3.4 Myr) on the Sunda Arc islands (Reilly et al. 2019a). Within *D. timoriensis*, the data support five candidate species, corresponding to (1) Rote, (2) Timor, (3) Wetar, (4) Alor, and (5) Pantar. Divergence times within this complex are much younger than within *D. boschmai* and range from 0.5–1.0 Myr based on genomic demographic estimates and 0.4–1.2 Myr based on calibrated phylogenomic estimates. Taxonomic treatment and morphological divergence of these Lesser Sunda *Draco* lineages is in progress and will be presented elsewhere.

CONCLUSIONS

Analyses of the mitogenomic and exome-capture data sets show that Lesser Sunda flying lizards are composed

of multiple, deeply divergent, allopatric or parapatric lineages. Thus, the current two-species taxonomy substantially understates the diversity of this Lesser Sunda endemic clade. Furthermore, our data indicate that the lineage boundaries do not always correspond to marine barriers, with some deeply divergent breaks located within individual islands and likely reflecting a once more fragmented archipelago. *Draco* entered the western Sunda Arc islands during the late Miocene or early Pliocene and subsequently dispersed east, mostly according to stepping-stone dispersal with distance and island connectivity during glacial maxima having greatly influenced movement through the archipelago. Surprisingly, rather than making the short jump from Lembata to Pantar *Draco* appears to have taken the long way around by embarking on at least 4–5 over-water dispersals through Sumba and the Outer Banda Arc all the way to Wetar before dispersing west again to Pantar. Species formation in this system is driven by overwater dispersal and subsequent isolation, and even within-island divergences may be the consequence of allopatric divergence followed by secondary contact due to the coalescence of paleo-islands. Our findings provide a model for Lesser Sundas biogeography that can serve as an important comparative framework for biogeographical analyses of other widespread terrestrial organisms that have colonized this archipelago.

Supplementary Material

Data available from the Dryad Digital Repository: https://dx.doi.org/10.5061/dryad.gqnk98skg.

References

Alcala A.C. 1967. Population biology of the "flying" lizard, *Draco volans*, on Negros Island, Philippines. Nat. Appl. Sci. Bull. 20:335–372.

Ali J.R., Heaney L.R. 2021. Wallace's line, Wallacea, and associated divides and areas: history of a tortuous tangle of ideas and labels. Biol. Rev. 96:922–942.

Allio R., Donega S., Galtier N., Nabholz B. 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. 34:2762–2772.

Audley-Charles M.G. 2011. Tectonic post-collision processes in Timor. Geol. Soc. Spec. Publ. 355:241–266.

Auffenberg W. 1980. The herpetofauna of Komodo, with notes on adjacent areas. Bull. Fla. Mus. Nat. Hist. 25:39–156.

Bi K., Vanderpool D., Singhal S., Linderoth T., Moritz C., Good J.M. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13:403.

Bi K., Linderoth T., Vanderpool D., Good J.M., Nielsen R., Moritz C. 2013. Unlocking the vault: next-generation museum population genomics. Mol. Ecol. 22:6018–6032.

Blom M.P., Horner P., Moritz C. 2016. Convergence across a continent: adaptive diversification in a recent radiation of Australian lizards. Proc. R. Soc. B. 283:20160181.

Blom M.P., Matzke N.J., Bragg J.G., Arida E., Austin C.C., Backlin A.R., Carretero M.A., Fisher R.N., Glaw F., Hathaway S.A., Iskandar D.T., McGuire J.A., Karin B.R., Reilly S.B., Rittmeyer E.N., Rocha S., Sanchez M, Stubbs A.L., Vences M., Moritz C. 2019. Habitat preference modulates trans-oceanic dispersal in a terrestrial vertebrate. Proc. R. Soc. B. 286: 20182575.

Borel-Best M., Moll H., Boekschoten G.J. 1985. Investigations of recent and fossil coral reefs in eastern Indonesia (Snellius-II Expedition): a preliminary report. Proc. Fifth Int. Coral Reef Cong. 6:311–316.

Borregaard M.K., Amorim I.R., Borges P.A., Cabral J.S., Fernández-Palacios J.M., Field R., Heaney L.R., Kreft H., Matthews T.J., Olesen J.M., Price J. 2017. Oceanic island biogeography through the lens of the general dynamic model: assessment and prospect. Biol. Rev. 92:830–853.

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLOS Comput. Biol. 10(4):p.e1003537.

Brandley M.C., Wang Y., Guo X., Montes de Oca A.N., Feria-Ortiz M., Hikida T., Ota H. 2011. Accommodating heterogenous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon* (*Eumeces*) lizards. Syst. Biol. 60:3–15.

Carstensen D.W., Olesen J.M. 2009. Wallacea and its nectarivorous birds: nestedness and modules. J. Biogeogr. 36:1540–1550.

Carstensen D.W., Dalsgaard B., Svenning J.C., Rahbek C., Fjeldså J., Sutherland W.J., Olesen, J.M. 2012. Biogeographical modules and island roles: a comparison of Wallacea and the West Indies. J. Biogeogr. 39:739–749.

Chan K.O., Hutter C.R., Wood P.L. Jr, Grismer L.L., Das I., Brown R.M. 2020. Gene flow creates a mirage of cryptic species in

a Southeast Asian spotted stream frog complex. Mol. Ecol. 29:3970–3987.

Coyne J.A., Orr H.A. 2004. Speciation (Vol. 37). Sunderland, MA: Sinauer Associates.

Darlington P.J. 1957. Zoogeography: the geographical distribution of animals. Cambridge: Harvard University Press.

Darriba D., Taboada G.L., Doallo R., Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9:772–772.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Deng W., Maust B.S., Nickle D.C., Learn G.H., Liu Y., Heath L., Pond S.L.K., Mullins J.I. 2010. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. Biotechniques 48:405–408.

Durand E.Y., Patterson N., Reich D., Slatkin M. 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28:2239–2252.

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.

Eaton D.A., Ree R.H. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). Syst. Biol. 62:689–706.

Eaton D.A., Hipp A.L., González-Rodríguez A., Cavender-Bares J. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution 69:2587–2601.

Edgar R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

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.

Flouri T., Jiao X., Rannala B., Yang Z. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol. Biol. Evol. 35:2585–2593.

Fortuin A.R., Van der Werff W., Wensink H. 1997. Neogene basin history and paleomagnetism of a rifted and inverted forearc region, on-and offshore Sumba, Eastern Indonesia. J. Asian Earth Sci. 15:61–88.

Freedman A.H., Gronau I., Schweizer R.M., Ortega-Del Vecchyo D., Han, E., et al. 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 10:e1004016.

Fujita M.K., Leaché A.D., Burbrink F.T., McGuire J.A., Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27:480-488.

Garrick R.C., Banusiewicz J.D., Burgess S., Hyseni C., Symula R.E. 2019. Extending phylogeography to account for lineage fusion. J. Biogeogr. 46:268–278.

Godfrey J.S. 1996. The effect of the Indonesian throughflow on ocean circulation and heat exchange with the atmosphere: a review. J. Geophys. Res. Oceans 101:12217–12237.

Goltenboth F., Timotius K.H., Milan P.P., Margraf J. 2006. Ecology of insular Southeast Asia: the Indonesian archipelago. Amsterdam: Elsevier.

Gordon A.L., Fine R.A. 1996. Pathways of water between the Pacific and Indian oceans in the Indonesian seas. Nature 379:146–149.

Grant B.R., Grant P.R. 2008. Fission and fusion of Darwin's finches populations. Philos. Trans. R. Soc. B 363:2821–2829.

Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Pääbo S. 2010. A draft sequence of the Neandertal genome. Science 328:710–722.

Grismer J.L., Schulte J.A., Alexander A., Wagner P., Travers S.L., Buehler M.D., Welton L.J. Brown R.M., 2016. The Eurasian invasion: phylogenomic data reveal multiple Southeast Asian origins for Indian Dragon Lizards. BMC Evol. Bio. 16:43.

Gronau I., Hubisz M.J., Gulko B., Danko C.G., Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. 43:1031–1034.

Haig D.W., McCartain E. 2007. Carbonate pelagites in the post-Gondwana succession (Cretaceous – Neogene) of East Timor. Aust. J. Earth Sci. 54:875–897.

Hall R. 2009. Southeast Asia's changing paleogeography. Blumea 54:148–161.

Hall R. 2011. Australia-SE Asia collision: plate tectonics and crustal flow. Geol. Soc. Spec. Publ. 355:75–109.

Hantoro W.S., Pirazzoli P.A., Jouannic C., Faure H., Hoang C.T., Radtke U., Causse C., Best M.B., Lafont R., Bieda S., Lambeck K. 1994. Quaternary uplifted coral reef terraces on Alor Island, East Indonesia. Coral Reefs 13:215e223.

Harris R. 2011. The nature of the Banda arc-continent collision in the Timor region. In: Brown D., Ryan P.D., editors. Arc continent collision. Berlin, Heidelberg: Springer. p. 163–211.

Heaney L.R. 1991. A synopsis of climatic and vegetational change in Southeast Asia. Clim. Change 19:53–61.

Hillis D.M. 2019. Species delimitation in herpetology. J. Herpetol. 53:3–12.

Hisheh S., Westerman M., Schmitt L.H. 1998. Biogeography of the Indonesian archipelago: mitochondrial DNA variation in the fruit bat, *Eonycteris spelaea*. Biol. J. Linn. Soc. 65:329–345.

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

How R.A., Kitchener D.J. 1997. Biogeography of Indonesian snakes. J. Biogeogr. 24:725–735.

Huxley T.H. 1868. On the classification and distribution of the Alectoromorphae and Heteromorphae. Proc. Zool. Soc. Lond. 1868:294–319.

Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species delimitation with gene flow. Syst. Biol. 66:799–812.

Jiao X., Yang Z. 2021. Defining species when there is gene flow. Syst. Biol. 70:108-119.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Jones B.K., Saviola A.J., Reilly S.B., Stubbs A.L., Arida E., Iskandar D.T., McGuire J.A., Yates III J.R., Mackessy S.P. 2019. Venom composition in a phenotypically variable pit viper (*Trimeresurus insularis*) across the Lesser Sunda Archipelago. J. Proteome Res. 18:2206–2220.

Junier T., Zdobnov E.M. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. Bioinformatics 26:1669–1670.

Kalyaanamoorthy S., Minh B.Q., Wong T.K.F., von Haeseler A., Jermiin L.S. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14:587–589.

Kealy S., O'Connor S., Mahirta, Sari D.M., Shipton C., Langley M.C., Boulanger C., Kaharudin H.A.F., Patridina E., Algifary M.A., Irfan A., Beaumont P., Jankowski N., Hawkins S., Louys J. 2020. Forty-thousand years of specialized maritime subsistence near a changing shoreline on Alor Island (Indonesia). Quat. Sci. Rev. 249:106599.

Kuhnt W., Holbourn A., Hall R., Zuvela M., Käse R. 2004. Neogene history of the Indonesian throughflow. In: Cliff P.D., editor. Continent-ocean interactions within East Asian marginal seas. Washington (DC): American Geophysical Union. p. 1–22.

Lambert S.M., Streicher J.W., Fisher-Reid M.C., Méndez de la Cruz F.R., Martínez-Méndez N., García-Vázquez U.O., Nieto Montes de Oca A., Wiens, J.J. 2019. Inferring introgression using RADseq and DFOIL: power and pitfalls revealed in a case study of spiny lizards (Sceloporus). Mol. Ecol. Resour. 19:818–837.

Leaché A.D., Zhu T., Rannala B., Yang Z. 2019. The spectre of too many species. Syst. Biol. 68:168–181.

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

Lohman D.J., de Bruyn M., Page T., von Rintelen K., Hall R., Ng P.K., Shih H., Carvalho G.R., von Rintelen T. 2011. Biogeography of the Indo-Australian archipelago. Annu. Rev. Ecol. Evol. Syst. 42:205–226.

Lydekker R. 1896. A geographical history of mammals. Cambridge: Cambridge University Press.

Macey J.R., Larson A., Ananjeva N.B., Fang Z., Papenfuss T.J. 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol. Biol. Evol. 14:91–104.

Martin S.H., Dasmahapatra K.K., Nadeau N.J., Salazar C., Walters J.R., Simpson F., Blaxter M., Manica A., Mallet J., Jiggins C.D. 2013.

Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res. 23:1817–1828.

Matzke N.J. 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. R package, version 0.2.1.

Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63:951–970.

Mayr E. 1944. Wallace's line in the light of recent zoogeographic studies. Q. Rev. Biol. 19:1–14.

McGuire J.A., Alcala A.C. 2000. A taxonomic revision of the flying lizards (Iguania: Agamidae: *Draco*) of the Philippine Islands, with a description of a new species. Herp. Monogr. 14:81–138.

McGuire J.A., Heang K.B. 2001. Phylogenetic systematics of Southeast Asian flying lizards (Iguania: Agamidae: *Draco*) as inferred from mitochondrial DNA sequence data. Biol. J. Linn. Soc. 72:203–229.

McGuire J.A., Brown R.M., Riyanto A., Andayani N. 2007. The flying lizards of the *Draco lineatus* group (Squamata: Iguania: Agamidae): a taxonomic revision with descriptions of two new species. Herp. Monogr. 21:179–212.

McGuire J.A., Dudley R. 2011. The biology of gliding in flying lizards (Genus *Draco*) and their fossil and extant analogs. Integr. Comp. Biol. 51:983–990.

Merritts D., Eby R., Harris R.A., Edwards R.L., Chang H. 1998. Variable rates of Late Quaternary surface uplift along the Banda Arc–Australian plate collision zone, eastern Indonesia. In: Steward I.S., Vita-Finzi C., editors. Coastal tectonics. Special Publication, vol. 146. London: Geological Society. p. 213–224.

Mertens R. 1930. Die amphibien und reptilien der inseln Bali, Lombok, Sumbawa und Flores. Abh. Senckenberg. Naturf. Ges. 42:115–344.

Michaux B. 2010. Biogeology of Wallacea: geotectonic models, areas of endemism, and natural biogeographical units. Biol. J. Linn. Soc. 101:193–212.

Minh B.Q., Schmidt H.A, Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37:1530–1534.

Muller S. 1846. Uber den Charakter der Thierwelt auf den Inseln des indischen Archipels. Arch. Naturg. 12:109–128.

Muraoka H., Nasution A., Urai M., Takahashi M., Takashima I., Simanjuntak J., Sundhoro H., Aswin D., Nanlohy F., Sitorus K., Takahashi H., Koseki T. 2002. Tectonic, volcanic and stratiagraphic geology of the Bajawa geothermal field, central Flores, Indonesia. Bull. Geol. Surv. Japan 53:109–138.

Muraoka H., Nasution A., Simanjuntak J., Dwipa S., Takahashi M., Takahashi H., Matsuda K., Sueyoshi Y. 2005. Geology and geothermal systems in the Bajawa volcanic rift zone, Flores, Eastern Indonesia. Proc. World Geoth. Cong. 2005:1–13.

Mussmann S.M., Douglas M.R., Bangs M.R., Douglas M.E. 2020. Comp-D: a program for comprehensive computation of D-statistics and population summaries of reticulated evolution. Conserv. Genet. Resour. 12:263–267.

Musters, C.J.M. 1983. Taxonomy of the genus *Draco* L. (Agamidae, Lacertilia, Reptilia). Zool. Verh. 199:1–120.

Nguyen N., Duffy B., Shulmeister J., Quigley M. 2013. Rapid Pliocene uplift of Timor. Geology 41:179–182.

Ogilvie H.A., Bouckaert R.R., Drummond A.J. 2017. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. Mol. Biol. Evol. 34:2101–2114.

Pacey A., Macpherson C.G., McCaffrey K.J. 2013. Linear volcanic segments in the central Sunda Arc, Indonesia, identified using Hough Transform analysis: Implications for arc lithosphere control upon volcano distribution. Earth Planet. Sci. Lett. 369:24–33.

Pease J.B., Hahn, M.W. 2015. Detection and polarization of introgression in a five-taxon phylogeny. Syst. Biol. 64:651–662.

Pirazzoli P.A., Radtke U., Hantoro W.S., Jouannic C., Hoang C.T., Causse C., Best M.B. 1993. A one million-year-long sequence of marine terraces on Sumba Island, Indonesia. Mar. Geol. 109:221–236.

Pritchard J.K., Stephens M., Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Puillandre N., Lambert A., Brouillet S., Achaz G. 2012. ABGD, automatic barcode gap discovery for primary species delimitation. Mol. Ecol. 21:1864–1877.

Purwandono A.F., Bonte D., Utami P., Pramumijoyo S., Harijoko A., Beekman F. van Wees J. 2019. Tectonic and compositional variation in Flores Island, Indonesia: implication for volcanic structure and geothermal occurrences. Eur. Geoth. Congress 2019:1–4.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67:901–904.

Rannala B., Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164:1645–1656.

Rannala B., Yang Z. 2017. Efficient Bayesian species tree inference under the multispecies coalescent. Syst. Biol. 66:823–842.

Reilly S.B. 2016. Historical biogeography of reptiles and amphibians from the Lesser Sunda Islands of Indonesia. Doctoral Dissertation, University of California, Berkeley, CA.

Reilly S.B., Wogan G.O.U., Stubbs A.L., Arida E., Iskandar D.T., McGuire J.A. 2017. Toxic toad invasion of Wallacea: A biodiversity hotspot characterized by extraordinary endemism. Glob. Chang. Biol. 23:5029–5031.

Reilly S.B., Stubbs A.L., Karin B.R., Bi K., Arida E., Iskandar D.T., McGuire J.A. 2019a. Leap-frog dispersal and mitochondrial introgression: Phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. J. Biogeogr. 46:757–769.

Reilly S.B., Stubbs A.L., Karin B.R., Arida E., Iskandar D.T., McGuire J.A. 2019b. Recent colonization and expansion through the Lesser Sundas by seven amphibian and reptile species. Zool. Scr. 48:614–626.

Reilly S.B., Stubbs A.L., Arida E., Arifin U., Bloch L., Hamidy A., Harmon K., Hykin S., Karin B.R., Ramadhan G., Iskandar D.T., McGuire J.A. 2020. New island records for anurans and squamates from the Lesser Sunda Archipelago. Herpetol. Rev. 51:787–791.

Reilly S.B., Stubbs A.L., Karin B.R., Arifin U., Arida E., Iskandar D.T., McGuire J.A. 2021. Genetic divergence of the Sunda ratsnake (*Coelognathus subradiatus*) across the Lesser Sunda Islands (Squamata: Colubridae). Amphibia-Reptilia 42:269–273.

Rigg J.W., Hall R. 2011. Structural and stratigraphic evolution of the Savu Basin, Indonesia. In: Hall R., Cottam M.A., Wilson M.E.J., editors. The SE Asian gateway: history and tectonics of the Australia–Asia collision. Geol. Soc. Spec. Publ. 355:225–240.

Ripley B.D., Venables W.N. 2002. Modern applied statistics with S. New York, NY: Springer.

Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.

Roosmawati N., Harris R. 2009. Surface uplift history of the incipient Banda arc-continent collision: geology and synorogenic foraminifera of Rote and Savu Islands, Indonesia. Tectonophysics 479:95–110.

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

Shaffer H.B., Thompson R.C. 2007. Delimiting species in recent radiations. Syst. Biol. 56:896–906.

Shaw K.L., Gillespie R.G. 2016. Comparative phylogeography of ocean archipelagos: Hotspots for inferences of evolutionary process. Proc. Natl. Acad. Sci. USA 113:7986–7993.

Simpson G.G. 1977. Too many lines: the limits of the Oriental and Australian zoogeographic regions. Proc. Am. Philos. Soc. 121:107–120.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

Smith M.L., Carstens B.C. 2020. Process-based species delimitation leads to identification of more biologically relevant species. Evolution 74:216-229.

Solís-Lemus C., Yang M., Ané C. 2016. Inconsistency of species tree methods under gene flow. Syst. Biol. 65:843-851.

Spakman W., Hall R. 2010. Surface deformation and slab–mantle interaction during Banda arc subduction rollback. Nat. Geosci. 3:562–566.

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

Stresemann E. 1939. Die vogel von Celebes. Zoogeographie. J. Ornithol. 87:312– 425.

Suminto, Morwood M.J., Kurniawan I., Aziz F., van den Bergh G.D., Hobbs D.R. 2008. Geology and fossil sites of the Soa Basin, Flores, Indonesia. In: Geology, paleontology and archeology of the Pleistocene Soa Basin, Central Flores, Indonesia. Geological Survey Centre, Bandung, Indonesia. p. 19–40.

Tänzler R., Van Dam M.H., Toussaint E.F., Suhardjono Y.R., Balke M., Riedel A. 2016. Macroevolution of hyperdiverse flightless beetles reflects the complex geological history of the Sunda Arc. Sci. Rep. 6:18793.

Tate G.W., McQuarrie N., van Hinsbergen D.J.J., Bakker R.R., Harris R., Willett S., Reiners P.W., Fellin M.G., Ganerød M., Zachariasse W.J. 2014. Resolving spatial heterogeneities in exhumation and surface uplift in Timor-Leste: constraints on deformation processes in young orogens. Tectonics 33:1089–1112.

Tate G.W., McQuarrie N., Tiranda H., van Hinsbergen D.J., Harris R., Zachariasse W.J., Fellin M.G., Reiners P.W., Willett S.D. 2017. Reconciling regional continuity with local variability in structure, uplift and exhumation of the Timor orogen. Gondwana Res. 49:364-386.

Toussaint E.F., Müller C.J., Morinière J., Tänzler R., Balke M. 2020. A glide over the Indo-Australian geological maze: repeated transgressions of Lydekker's and Wallace's Lines in archdukes, barons and dukes (Nymphalidae: Limenitidinae: Adoliadini). Biol. J. Linn. Soc. 129:810–821.

Turner H., Hovenkamp P., Van Welzen P.C. 2001. Biogeography of Southeast Asia and the West Pacific. J. Biogeogr. 28:217–230.

Van Dam M.H., Matzke N.J. 2016. Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. J. Biogeogr. 43:1514–1532.

Van Welzen P.C., Parnell J.A., Slik J.F. 2011. Wallace's Line and plant distributions: two or three phytogeographical areas and where to group Java? Biol. J. Linn. Soc. 103:531–545.

Wallace A.R. 1860. On the zoological geography of the Malay Archipelago. J. Proc. Linn. Soc. Lond. Zool. 14:172–184.

Wallace A.R. 1863. On the physical geography of the Malay Archipelago. J. R. Geogr. Soc. 33:217–234.

Wares J.P., Gaines S., Cunningham C.W. 2001. A comparative study of asymmetric migration events across a marine biogeographic boundary. Evolution 55:295–306.

Warren B.H., Simberloff D., Ricklefs R.E., Aguilée R., Condamine F.L., Gravel D., Morlon H., Mouquet N., Rosindell J., Casquet J., Conti E., Cornuault, J., Fernández-Palacios J.M., Hengl T., Norder S.J., Rijsdijk K.F., Sanmartín I., Strasberg D., Triantis K.A., Valente L.M., Whittaker R.J., Gillespie R.G., Emerson B.C., Thébaud C. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecol. Lett. 18:200-217.

Wen D., Yu Y., Zhu J., Nakhleh, L. 2018. Inferring phylogenetic networks using PhyloNet. Syst. Biol. 67:735-740.

Whittaker R.J., Fernaìndez-Palacios J.M. 2007. Island biogeography: ecology, evolution, and conservation. Oxford: Oxford University Press.

Wright S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. USA 107:9264–9269.

Zhang J., Kapli P., Pavlidis P., Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placement. Bioinformatics 29:2869–2876.

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