Molecular Phylogeny, Biogeography, and Habitat Preference Evolution of Marsupials

Kieren J. Mitchell,*[,1] Renae C. Pratt,[2] Laura N. Watson,[3] Gillian C. Gibb,[4] Bastien Llamas,[1] Marta Kasper,[1] Janette Edson,[1] Blair Hopwood,[1] Dean Male,[1] Kyle N. Armstrong,[1] Matthias Meyer,[5] Michael Hofreiter,[6] Jeremy Austin,[1] Stephen C. Donnellan,[1,7] Michael S. Y. Lee,[1,7] Matthew J. Phillips,[8] and Alan Cooper*[,1]

[1]Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA, Australia

[2]Department of Evolution, Ecology and Genetics, The Australian National University, Canberra, ACT, Australia

[3]Department of Obstetrics and Gynaecology, Research Centre for Reproductive Health, University of Adelaide, Adelaide, SA, Australia

[4]Institute of Agriculture and Environment, Massey University, Palmerston North, New Zealand

[5]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[6]Department of Biology, University of York, York, United Kingdom

[7]South Australian Museum, Adelaide, SA, Australia

[8]School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, QLD, Australia

***Corresponding author:** E-mail: kieren.mitchell@adelaide.edu.au; alan.cooper@adelaide.edu.au.

Associate editor: Emma Teeling

Abstract

Marsupials exhibit great diversity in ecology and morphology. However, compared with their sister group, the placental mammals, our understanding of many aspects of marsupial evolution remains limited. We use 101 mitochondrial genomes and data from 26 nuclear loci to reconstruct a dated phylogeny including 97% of extant genera and 58% of modern marsupial species. This tree allows us to analyze the evolution of habitat preference and geographic distributions of marsupial species through time. We found a pattern of mesic-adapted lineages evolving to use more arid and open habitats, which is broadly consistent with regional climate and environmental change. However, contrary to the general trend, several lineages subsequently appear to have reverted from drier to more mesic habitats. Biogeographic reconstructions suggest that current views on the connectivity between Australia and New Guinea/Wallacea during the Miocene and Pliocene need to be revised. The antiquity of several endemic New Guinean clades strongly suggests a substantially older period of connection stretching back to the Middle Miocene and implies that New Guinea was colonized by multiple clades almost immediately after its principal formation.

Key words: **supermatrix, ancestral state reconstruction, mammal, mitochondrion.**

Article

Introduction

Modern marsupials originated in the Late Cretaceous and today include over 300 known species distributed across the Americas and Australasia (Wilson and Reeder 2005). These species vary from gliding possums to hopping kangaroos and include obligate carnivores, specialized insectivores, omnivores, and herbivores, and both diurnal and nocturnal forms. They also occupy diverse habitats ranging from the wet tropical rainforests of New Guinea and the Amazon to the arid Australian interior. Marsupials occupy virtually every terrestrial, nonvolant niche occupied by placental analogs on other continents. However, despite the great diversity of adaptations displayed by living marsupials, studies attempting to explain the drivers and constraints behind their evolution have been limited compared with their sister group. This is partly due to extensive gaps in the marsupial fossil record, which obscure several critical periods in their evolutionary history (Archer et al. 1999).

Where the fossil record is depauperate, molecular phylogenies can be particularly useful tools for reconstructing the tempo and mode of evolution. Time-calibrated phylogenies are increasingly used to reconstruct biogeographic history and test dispersal/vicariance hypotheses (Avise et al. 1987; Avise 2000). Such phylogenies can also be used to model the evolution of morphological and ecological traits in a phylogenetic context (Felsenstein 1985; Huey 1987; Pagel 1997). Researchers have even attempted to use phylogenies of living taxa to infer changes in the rate of speciation and extinction through time (Alfaro et al. 2009; Stadler 2011; Hugall and Stuart-Fox 2012). To fully investigate these issues and understand the origin of extant marsupial diversity, it is necessary to develop a comprehensive phylogeny as a template for exploring marsupial macroevolution.

Since the turn of the century, great progress has been made in resolving the relationships among marsupial taxa. Eomarsupialia (Archer 1984; Phillips et al. 2006; Beck 2008;

Meredith et al. 2011) has been identified as a clade comprising the four Australasian orders: Diprotodontia (the kangaroos, possums, wombats, and koala), Peramelemorphia (bandicoots and bilbies), Dasyuromorphia (including the Tasmanian devil and the thylacine), and Notoryctemorphia (marsupial moles). Eomarsupialia is nested within a paraphyletic collection of three endemic American lineages (Beck 2008; Meredith, Westerman, Case, et al. 2008; Meredith et al. 2011): Didelphimorphia (opossums), Paucituberculata (shrew-opossums), and Microbiotheria (the monito del monte). The South American monito del monte (*Dromiciops gliroides*) is particularly interesting as its close affinity with the Australasian orders (together forming Australidelphia) has implications for understanding the biogeography and dispersal of ancestral marsupials in the Late Cretaceous and Paleocene (Beck 2012).

Unfortunately, the majority of past molecular studies have only been able to test a limited range of hypotheses. Studies seeking to resolve relationships among marsupial lineages using nucleotide sequence data have either included disparate exemplar higher taxa but relatively few individual species (Phillips et al. 2006; Beck 2008; Meredith, Westerman, Case, et al. 2008; Meredith et al. 2011; Phillips and Pratt 2008; Nilsson et al. 2010) or comprehensive species sampling but only within restricted subclades (Meredith et al. 2008a, 2008b, 2010; Voss and Jansa 2009; Westerman et al. 2012). Often these individual data sets have involved different genetic loci, making their results difficult to combine when using supermatrix approaches. To circumvent this problem, some authors have turned to supertree methods (Cardillo et al. 2004; Bininda-Emonds et al. 2007). However, this approach has several disadvantages including difficulties in quantifying statistical support for inferred topologies (de Queiroz and Gatesy 2007; von Haeseler 2012). As a result, phylogenies focusing on marsupials have generally been rather poorly resolved and inadequately dated, at least compared with placental mammals.

We have constructed a comprehensive phylogeny of modern marsupials to better understand their evolution, biogeography, and ecological history. We sequenced 69 mitochondrial genomes using next-generation sequencing technology and synthesized this with existing nuclear and mitochondrial genomic nucleotide data to generate a supermatrix that includes 97% of extant genera and 58% of modern marsupial species.

Results and Discussion

Maximum likelihood (ML), Bayesian, and parsimony analyses of the supermatrix resulted in well-resolved and concordant phylogenetic trees (fig. 1 and supplementary fig. S1 and data set S1, Supplementary Material online) with 85% of nodes in the ML tree supported by ML bootstrap support (MLBS) values $\geq$ 70% (fig. 1 and Supplementary fig. S1, Supplementary Material online). Relationships among marsupial orders and families are consistent with most recent molecular studies and generally receive high statistical support (MLBS $\geq$ 97%). There is currently uncertainty surrounding the precise affinities of two major clades: Macropodiformes (the kangaroos and wallabies) and Notoryctemorphia (the marsupial moles). We robustly recover Macropodiformes and Petauroidea (petauroid possums) as sister taxa (MLBS = 85%), concordant with a recent study of major mammal lineages (Meredith et al. 2011); this placement is relatively novel, as several previous studies of mitochondrial (Phillips and Pratt 2008) and nuclear loci (Meredith, Westerman, Case, et al. 2008; Meredith et al. 2009), as well as some aspects of morphology (Szalay 1994), favor an alternative grouping of Macropodiformes with Phalangeroidea (phalangeroid possums). Additionally, our results suggest that Notoryctemorphia is sister to a clade consisting of Peramelemorphia and Dasyuromorphia, although support for this grouping is moderate (MLBS = 68%). Relationships below the family level are largely concordant with previous estimates (e.g., Raterman et al. 2006; Beck 2008; Meredith et al. 2008b, 2009, 2010; Phillips and Pratt 2008; Westerman et al. 2008; Voss and Jansa 2009; Westerman et al. 2012).

To estimate divergence times among marsupials, we employed a set of 14 established fossil-based node age constraints (supplementary table S1, Supplementary Material online) (see Meredith, Westerman, Case, et al. 2008; Meredith et al. 2008a, 2008b, 2009, 2010, 2011; Westerman et al. 2012) for which taxonomic assignment is well justified. These constraints range in age from a Late Cretaceous (71.2 Ma) maximum bound for the divergence between Eomarsupialia and Microbiotheria, to a Middle Pliocene (3.62 Ma) minimum bound for the divergence between the bandicoot genera *Isoodon* and *Perameles*. Divergence dates inferred in this study are broadly consistent with past estimates based on molecular data (Beck 2008; Meredith, Westerman, Case, et al. 2008, Meredith et al. 2008a, 2009, 2010, 2011; Westerman et al. 2012). The mean age of the marsupial crown group is 86.8 Ma (95% highest posterior density [HPD]: 79.7–94.6), whereas the four Australasian orders diverge within a relatively brief window closely associated with the KPg boundary: 67–64 Ma. The divergence between the Australasian stem lineage and Microbiotheria only slightly predates this period (~69 Ma) but is in turn preceded by an approximately 14 Ma internode. Such a tight clustering of divergences relatively deep in the tree is unlikely under a standard birth–death diversification model. This may indicate rapid diversification of ancestral australidelphians catalyzed by ecological opportunity following the KPg mass extinction or colonization of Antarctica/Australia from South America. The latter hypothesis implies that the current distribution of the monito del monte may reflect a back migration from Antarctica/Australia to South America in the early Cenozoic (Beck 2012).

Several previous studies have constructed comprehensive marsupial phylogenies with similar aims to this study (Cardillo et al. 2004; Bininda-Emonds et al. 2007). However, these used supertree approaches and do not recover several widely accepted clades that are well supported by our analyses (e.g., Potoroidae, Echymiperinae, and Eomarsupialia). Similarly, many marsupial genera in these previous studies are represented by polytomies, with resolution between congeneric species completely lacking. Lack of resolution makes the

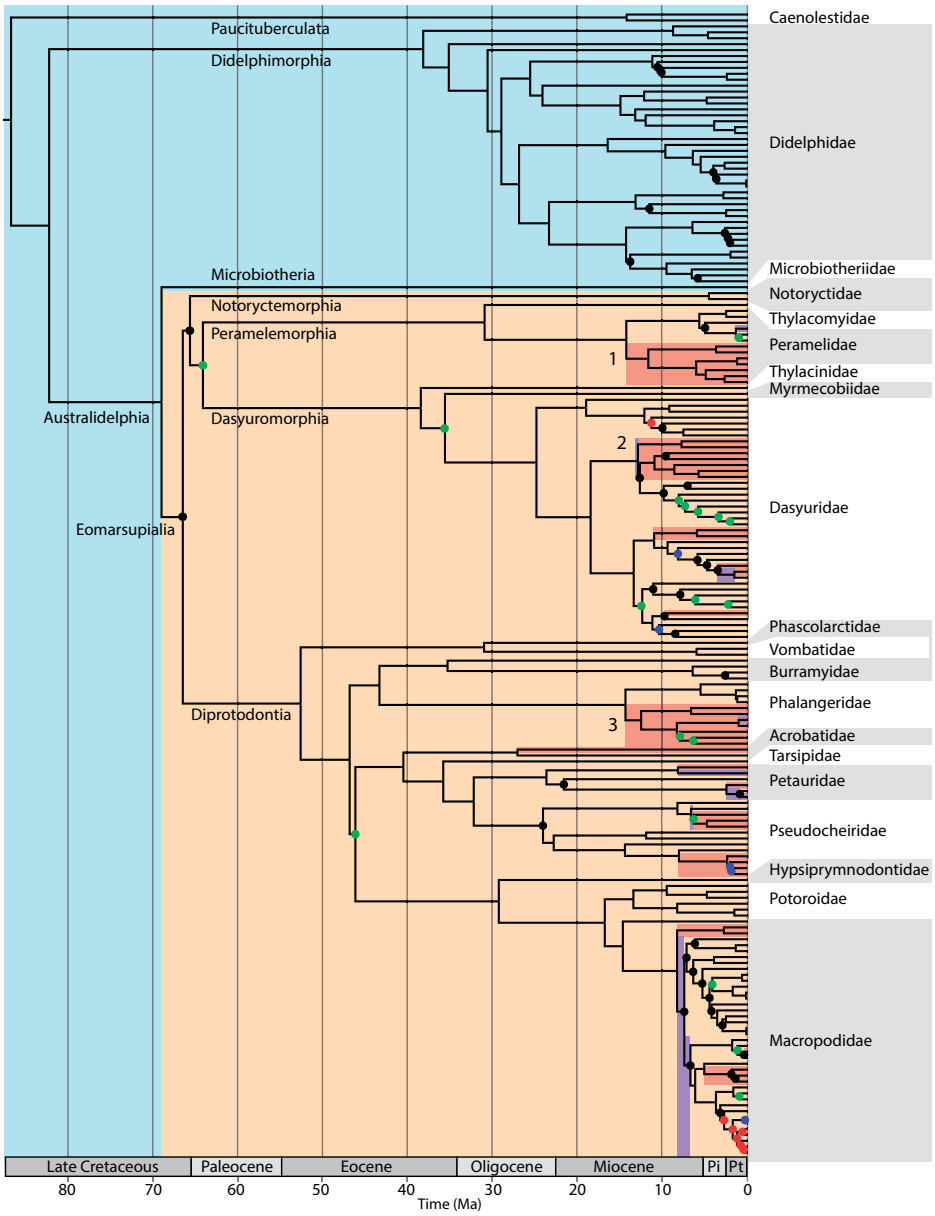

Fig. 1. Time-calibrated phylogeny of 193 marsupial species based on ML topology (supplementary data set S1, Supplementary Material online). All nodes are strongly supported by all three methods (ML bootstrap ≥90%, Bayesian posterior probability ≥0.95, and parsimony bootstrap ≥90%) unless indicated by a dot. Black dot indicates a node appears in trees inferred using all three methods but is not strongly supported in at least one tree. Blue dot indicates node absent in Bayesian tree, green dot indicates node absent in parsimony tree, and red dot indicates node absent in both Bayesian and parsimony trees. Shading behind branches reflects ancestral distribution as estimated using Lagrange: light blue for the Americas; orange for Australia; and red for New Guinea and Wallacea. Lineages that are inferred by Lagrange to be codistributed between Australia and New Guinea/Wallacea are shaded purple. Numbers indicate New Guinean/Wallacean "old endemic" clades (see main text): 1) New Guinean/Wallacean bandicoots (Echymiperinae and Peroryctinae), 2) "*Murexia*" dasyures, and 3) cuscuses (phalangerin possums). Geological epochs are marked on the *x*-axis: Late Cretaceous, Paleocene, Eocene, Oligocene, Miocene, Pliocene (Pi), and Pleistocene (Pt). The Holocene (~11 ka to the present day) is not visible at this scale. Parsimony optimization of distribution based on the ML topology gives comparable overall results (supplementary data set S1, Supplementary Material online). Individual species names and locus coverage are presented in supplementary figure S1, Supplementary Material online.

dating analyses of Bininda-Emonds et al. (2007) inaccurate for many clades, which is exacerbated by the fact that approximately 25% of node ages in their marsupial phylogeny were interpolated rather than empirically estimated. Perhaps as a result, mean node ages inferred by Bininda-Emonds et al. (2007) are substantially older for many clades (e.g., Didelphimorphia) than those inferred by more recent studies (e.g., Jansa et al. 2014), including our own. Consequently, the direct dating and high resolution of our phylogeny make it a more robust scaffold for testing macroevolutionary hypotheses than existing supertrees.

Extant marsupials inhabit a diversity of habitat types ranging from rainforest to desert. Figure 2 illustrates that habitat preference is highly phylogenetically conserved. Using Bayesian ancestral state reconstruction, we calculated the probability that each node on the phylogeny (representing

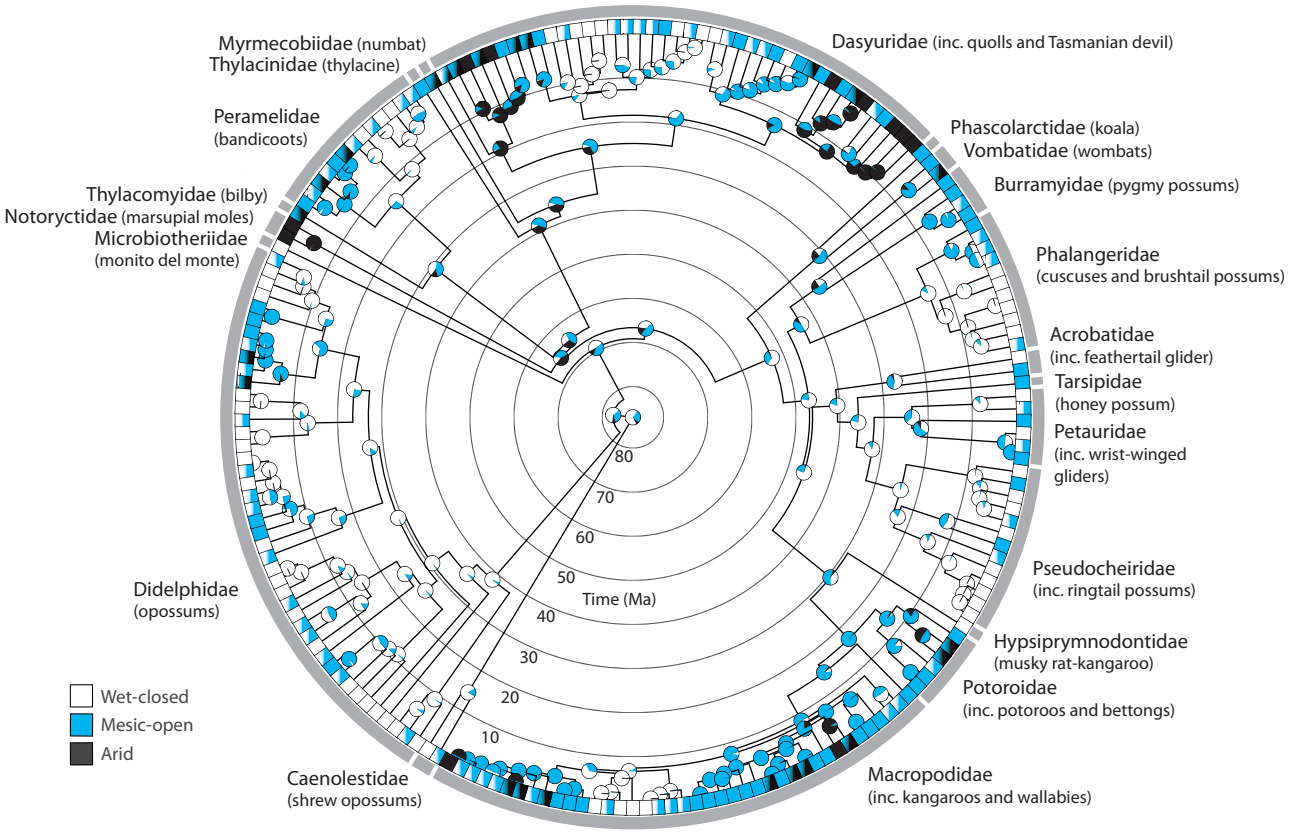

Fig. 2. Ancestral habitat preference reconstruction using Bayesian optimization methods on the ML topology (fig. 1 and supplementary fig. S1 and data set S1, Supplementary Material online). Boxes at tips reflect the modern state, whereas pie charts reflect the inferred probability of a particular state at each ancestral node. Modern taxa inhabiting multiple habitat types are represented by an appropriate color gradient. Parsimony optimization of habitat states gives comparable results (supplementary data set S1, Supplementary Material online).

an ancestral species) was associated with a given habitat type: wet-closed (e.g., rainforest), mesic-open (e.g., sclerophyll or grassland), or arid. A large proportion of the posterior probability for the most basal nodes in the phylogeny was associated with wet-closed environments (fig. 2). This suggests that most of the primary divergences among extant marsupial groups occurred in a rainforest context, consistent with many paleontological studies (Travouillon et al. 2009). However, the ancestral habitat on nodes at the base of the australidelphian radiation is unclear, with intermediate (mesic-open) habitat often having the highest probability.

Our results also suggest that adaptation to open and arid habitats has evolved independently numerous times among marsupials and that these events occurred relatively recently. Among South American marsupials, transitions from wet-closed to drier and more open environments appear unidirectional and begin no earlier than the Miocene (Jansa et al. 2014) (fig. 2). The timing of these transitions follows the Oligocene retraction of rainforest and expansion of wooded savanna and grassland in southern South America (Ortiz-Jaureguizar and Cladera 2006; Barreda and Palazzesi 2007; Iglesias et al. 2011). Concordantly, extant open-habitat-adapted opossum lineages (e.g., *Thylamys* and *Lestodelphys*) are distributed primarily in southern Brazil, Paraguay, Uruguay, Chile, and Argentina, whereas their respective sister taxa are generally distributed in rainforests to the north. In Australia, rainforest was the dominant habitat type across most of the continent during the early Cenozoic (Frakes et al. 1987): Open sclerophyll and grassland did not become widespread until after the Middle Miocene, and true arid zones did not develop until the Pliocene (Martin 2006; White 2006). Given this sequence of events, it is generally thought that the modern Australian arid biota evolved from ancestral mesic forms and that the ancestral state for these clades should consequently optimize as wet-closed (Byrne et al. 2008, 2011). However, departures from this expected direction of habitat evolution are observed for several Australasian taxa. For example, several clades and individual species appear to have undergone reversions from mesic-open to wet-closed habitats (fig. 2), particularly within Macropodidae and Dasyuridae. Additionally, several groups including Dasyuromorphia, Vombatiformes, and Trichosurini appear to have made the transition from wet-closed to mesic-open environments prior to the onset of widespread Miocene aridification. One explanation for this observation is that these older inferred transitions reflect lineages that survived from drier periods in the Oligocene when limited open forest habitats may have been available (Travouillon et al. 2009).

A limitation of molecular analyses is that they are based purely on modern species. This may lead to inference of artificially old wet-closed to mesic-open transitions through failure to sample extinct rainforest-adapted lineages. For

instance, support for an early origin of mesic-open habitats in dasyuromorphs decreases when their close relatives, the arid-adapted marsupial moles (Notoryctemorphia), are omitted from the analysis. The two extant marsupial mole species are associated exclusively with arid habitats, but extinct mesic notoryctids are known from Miocene Riversleigh layers (Archer et al. 2010). Additionally, extinct dasyurid and vombatiform taxa are known from Pleistocene rainforest assemblages (Cramb et al. 2009; Black et al. 2012). However, even if we were to assume that no single wet-closed to mesic-open transition is truly older than the Middle Miocene, these transitions are still temporally dispersed over a long period. This suggests that evolution toward mesic-open and arid habitats has been a gradual process rather than a concerted shift across many Australasian clades in association with some abrupt climatic event. This observation, along with evidence for reversions from mesic-open to wet-closed habitats, indicates a high level of ongoing plasticity among marsupials with regard to habitat adaptation.

A large proportion of Australasian wet-closed forest lineages inhabit New Guinea and Wallacea rather than Australia (Flannery 1995). The colonization timeline of these islands is uncertain; long-distance overwater dispersal is improbable for most marsupial taxa, and due to the geological complexity of the region, evidence for the presence and duration of land connections is equivocal (Metcalfe et al. 2001). Previous studies have attempted to use the age and distribution of marsupial lineages to infer periods of biotic connection between Australia and New Guinea/Wallacea (Aplin et al. 1993). We implemented a similar approach on our phylogeny, using biogeographic reconstructions to identify the temporal origin of New Guinean/Wallacean marsupial taxa (fig. 1). For taxa inferred to have colonized New Guinea/Wallacea from Australia, the oldest probable dispersal time is represented by the divergence between the ancestor of a New Guinean/Wallacean clade and its Australian sister taxon, whereas the most recent probable dispersal time is marked by the age of the New Guinean/Wallacean crown clade. Inferred dispersal dates for any individual clade may be artificially young or old due to extinction of Australian lineages or parallel colonization events, respectively. Thus, the power of our study for testing biogeographical hypotheses comes from the sheer number of independent dispersal events encompassed by our data set.

A vicariant origin of the New Guinean/Wallacean marsupial fauna has previously been advanced based on early geological reconstructions (Flannery 1988). Under this scenario, a substantial part of New Guinea was emergent and connected to Australia until the Oligocene when the intervening Papuan Basin was inundated (Dow 1976), thereby severing land connection and preventing dispersal. Our inferred dates are uniformly too young to be consistent with this hypothesis. Indeed, more recent geological reconstructions indicate that in fact no substantial part of New Guinea was emergent until the Miocene (van Ufford and Cloos 2005). Subsequently, it has been suggested that the ancestors of New Guinean/Wallacean marsupials dispersed no earlier than the latest Miocene (Westerman et al. 2012) when a substantial drop in global sea levels may have permitted overland dispersal (Hodell et al. 1986). Most of our inferred dispersal dates can be reconciled with this scenario and/or more recent periods of low sea level in the Pliocene and Pleistocene. However, the age of several clades is unexpectedly old given this hypothesis.

The crown ages of the cuscuses (phalangerin possums), New Guinean bandicoots (Peroryctinae and Echymiperinae), and "*Murexia*" dasyures suggest a period of accessibility between Australia and New Guinea/Wallacea at the beginning of the Late Miocene (fig. 1). The youngest inferred date for the presence of cucuses in New Guinea/Wallacea is 10.65 Ma (95% HPD minimum bound of the crown group); the analogous values for New Guinean bandicoots and dasyures are 9.84 and 9.28 Ma, respectively. Although inference of such early colonization dates could also be explained by multiple parallel dispersals within each clade at the end of the Miocene, the tight concordance among the crown ages of these three clades strongly suggests a dispersal window 11–9 Ma. Similarly, sweepstakes long-range overwater dispersal is unlikely to have generated three relatively simultaneous events. This suggests that dispersal has been at least intermittently possible between mainland Australia and New Guinea via diffuse connection since the Middle Miocene. A period of connection 11–9 Ma is broadly consistent with the origin of the bulk of modern New Guinea, which formed as a result of orogeny approximately 12 Ma (van Ufford and Cloos 2005), and is consistent with the suggestion of a similar period of dispersal based on microcomplement fixation (Aplin et al. 1993).

This study presents a comprehensive molecular phylogeny of marsupials and highlights substantial macroevolutionary heterogeneity. The widely hypothesized trend of ancestral wet-closed forest lineages evolving into open and arid forms since the Middle Miocene appears to be overly simplistic: Several lineages have evolved in the opposite direction. Additionally, the taxonomic breadth of our phylogeny allows us to compare the timing of numerous phylogenetically independent dispersals from Australia to New Guinea/Wallacea. The age and concordance of these events provides substantial evidence of biological connectivity between New Guinea and Australia coincident with New Guinea's principal formation approximately 12 Ma. Ultimately, our results suggest that marsupial evolution in the Cenozoic was dynamic and characterized by on-going ecological plasticity and opportunistic dispersal.

Materials and Methods

Data Set

We sequenced mitochondrial genomes (mitogenomes) for 69 extant marsupial species (supplementary table S2, Supplementary Material online). Sixty-three mitogenomes were sequenced according to a previously published protocol (Lerner et al. 2011) using a combination of multiplexed 454 pyrosequencing (Meyer et al. 2008) and traditional capillary electrophoresis. First, we used long-range polymerase chain reaction (see supplementary table S3, Supplementary Material online, for primer sequences) to amplify each

mitogenome in several fragments. Each fragment was then sheared, barcoded, and sequenced on a Roche/454 GS FLX. Resulting sequencing reads were de novo assembled using a previously published pipeline (Lerner et al. 2011): Sequences were demultiplexed using "untag" (http://bioinf.eva.mpg.de/pts/, last accessed June 5, 2014), and consensus sequences were created using "newbler" (v. 2.0.00.20), SeqMan Pro (Lasergene suite 8, DNASTAR), and MIA (http://sourceforge.net/projects/mia-assembler/, last accessed June 5, 2014). Remaining gaps and areas of low coverage were completed using big dye terminator chemistry on an ABI 3130 (see supplementary table S3, Supplementary Material online, for primer sequences). The remaining six mitogenomes were obtained via shotgun sequencing of raw extract on an Illumina Genome Analyzer II (one lane per sample). Sequencing reads were mapped to published mitogenomes of closely related species using TMAP v3.2.2 (https://github.com/nh13/TMAP, last accessed June 5, 2014) and samtools v1.4 (Li et al. 2009) according to a previously published pipeline (Mitchell et al. 2014); duplicate reads were removed with PicardTools v1.79 (http://picard.sourceforge.net, last accessed June 5, 2014), and a consensus sequence was generated in Geneious v6.1.2 (Biomatters; http://www.geneious.com, last accessed June 5, 2014).

New sequence data were combined with 32 previously published mitogenomes and data from 26 nuclear loci to form a single nucleotide supermatrix (supplementary fig. S1 and data set S2, Supplementary Material online). The final data set included 193 marsupial species and ten outgroup taxa: Two monotremes and eight placental mammals. The supermatrix (43,616 bp) was 39% complete (excluding outgroups), with the proportion of missing taxa per locus varying from 12.5% to 88% (0–15% at the family level and 0% at the order level). We divided this alignment into 72 discrete bins: Codon positions of each protein-coding nuclear locus, first and second codon positions of all H-strand mitochondrial protein-coding genes (concatenated into one locus), individual introns, and stem and loop positions of mitochondrial RNA-coding loci (all rRNAs and tRNAs were concatenated). Mitochondrial third codon positions were discarded to minimize bias in branch length estimation arising from saturation. These bins were analyzed using partitionfinder v0.9.2 (Lanfear et al. 2012) to determine the most appropriate partitioning scheme and substitution models (supplementary table S4, Supplementary Material online) for downstream analysis in RAxML v7.2.8, MrBayes v3.2.1, and MCMCtree (PAML v4.6).

Phylogenetics

Model-Based Methods

Tree topology was estimated under ML and Bayesian frameworks using RAxML v7.2.8 (Stamatakis 2006) and MrBayes v3.2.1 (Ronquist and Huelsenbeck 2003), respectively. Both RAxML and MrBayes gave very similar trees (fig. 1 and supplementary fig. S1 and data set S1, Supplementary Material online). All analyses were repeated with and without outgroup taxa (placentals and monotremes) to assess the sensitivity of the splits within marsupials to outgroup sampling; these analyses produced very similar trees (supplementary data set S1, Supplementary Material online). Our RAxML analysis comprised an ML search for the best-scoring tree from 1,000 bootstrap replicates. MrBayes analyses were undated (clock free), as dated (clock) analyses did not converge. There were three runs: Each individual run employed four Markov chains (one cold and three incrementally heated) with default priors. Each chain ran for 10^7 generations, sampling every 500. Convergence in topology was assessed using the average standard deviation of split frequencies (<0.02), whereas convergence in individual parameter values was assessed through broadly overlapping distributions in Tracer v1.5 (http://tree.bio.ed.ac.uk/software/tracer/, last accessed June 5, 2014) and effective sample sizes more than 200. Trees generated by the three runs were pooled before being summarized, and the first 25% of trees from each run were discarded as burn-in.

Parsimony Methods

Parsimony analyses used PAUP* (Swofford 2002), with most parsimonious trees found via heuristic searches involving 100 random addition searches followed by a strict consensus; bootstrapping was performed with 100 replicates of simple heuristic searches (due to time constraints), followed by a majority-rule consensus. All analyses were again repeated with and without outgroup taxa, with similar results (supplementary data set S1, Supplementary Material online).

Molecular Dating

We implemented 14 node calibrations across the marsupial tree (supplementary table S1, Supplementary Material online). Most constraints followed previous studies (e.g., Meredith, Westerman, Case, et al. 2008; Meredith et al. 2008a, 2008b, 2009, 2010, 2011; Westerman et al. 2012), with minimum bounds being chosen based on the oldest known occurrence of fossil crown-group taxa. Maximum bounds were determined using either stratigraphic bounding (Benton and Donoghue 2007) or phylogenetic bracketing (Reisz and Müller 2004; Müller and Reisz 2005). To allow for the patchiness of the marsupial fossil record, we followed the conservative philosophy of Meredith, Westerman, Springer (2008) when determining maxima: Stratigraphic bounding maxima were judged based on absence of the lineage of interest from two (rather than one) preceding fossil bearing layers, whereas phylogenetic bracketing maxima were judged based on the oldest known occurrence of at least the second closest (rather than first closest) outgroup to the target clade.

Molecular dating was performed using MCMCtree, within the PAML v4.6 software package (Yang 2007), on the RAxML tree topology. All calibrations (supplementary table S1, Supplementary Material online) were implemented as uniform priors with hard minima and soft maxima (97.5%). To maximize computational efficiency, we used the likelihood approximation approach implemented in MCMCtree and performed our analysis on ingroup taxa only. For the purposes of our MCMCtree analyses, we defined individual time units as 10 Ma, such that most node ages fell between 10 and 0.1

units. Among-partition rate variation was modeled using a gamma distribution with the parameter values $\alpha = 1$ and $\beta = 56$. After a burn-in of 10^4 iterations, each chain was run for 8×10^5 iterations sampling every 80 for a total of 10^4 samples. To ensure convergence and adequate sampling, two separate runs were performed, and each parameter was monitored in Tracer v1.5 (http://tree.bio.ed.ac.uk/software/tracer/, last accessed June 5, 2014). Parameter values for one partition (out of 24) consistently failed to converge due to a low number of informative sites; this partition was consequently omitted from the final analysis (see supplementary table S4, Supplementary Material online).

Ancestral State Reconstruction

Bayesian trait reconstruction was performed on the final dated phylogeny using the MCMC option for discrete state evolution in BayesTraits v1.0 (Pagel et al. 2004). The analysis was run for 10^6 iterations, sampling every 10^3. The first 10% of samples were discarded as burn-in. To ensure convergence and adequate sampling, each parameter was monitored in Tracer v1.5 (http://tree.bio.ed.ac.uk/software/tracer/, last accessed June 5, 2014). All modern taxa were coded as wet-closed, mesic-open, or arid adapted (or a combination of these) according to their distribution and habits (supplementary table S5, Supplementary Material online). An ordered (wet-closed to mesic-open to arid) single-rate model was used for evolution between the three character states. A two-rate, ordered model allowing a separate rate for transitions toward more arid environments and transitions toward more mesic environments provided no significant improvement in fit as determined by Bayes factor comparison ($\Delta BF < 0.5$); we thus used the simpler model in our analysis to avoid overfitting. Both ordered models provided a substantially better fit to the data than corresponding unordered models ($\Delta BF > 5$). To test the robustness of our results under different optimization methods, we performed strict parsimony state reconstructions (again using both ordered and unordered models) in Mesquite (Maddison WP and Maddison DR 2011) with concordant results (supplementary data set S1, Supplementary Material online).

Geographical Range Evolution

We used Lagrange v20120508 (Ree and Smith 2008) to reconstruct ancestral distribution and infer dispersal times. The model employed by this program attempts to account for dispersal, fusion, and fission of geographic regions; speciation; and local extinction. Distribution was coded as one (or a combination) of three states (supplementary table S5, Supplementary Material online): America, Australia, or New Guinea/Wallacea. Individual species (including ancestral species) were permitted to simultaneously inhabit the Americas and Australia (to reflect potential connection scenarios in the Paleocene), and Australia and New Guinea/Wallacea. The model we employed allowed range shifts to occur bidirectionally but only between adjacent areas (Australia and New Guinea/Wallacea, or America and Australia), because a direct connection between the Americas and New Guinea/Wallacea is not plausible. The most probable inherited range for each node on the phylogeny is reflected in figure 1. To test our results, we performed a strict parsimony state reconstruction in Mesquite (Maddison WP and Maddison DR 2011) with concordant results (supplementary data set S1, Supplementary Material online).

Supplementary Material

Supplementary tables S1–S5, figure S1, and data set S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This study was funded by the Australian Research Council. High-performance computing was provided by eResearch SA. For access to tissue and DNA samples, the authors thank: J. Patton and C. Cicero (Museum of Vertebrate Zoology); R. Voss, S. Jansa, and E. Westwig (American Museum of Natural History); J. Cook (Museum of Southwestern Biology); W. Foster (Adelaide Zoo); J. Graves (Australian National University); South Australian Museum; and L. Joseph and R. Palmer (Australian National Wildlife Collection).

References

Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci U S A.* 106:13410–13414.

Aplin K, Baverstock PR, Donnellan SC. 1993. Albumin immunological evidence for the time and mode of origin of the New Guinean terrestrial mammal fauna. *Sci New Guinea.* 19:131–145.

Archer M. 1984. The Australian marsupial radiation. In: Archer M, Clayton G, editors. Vertebrate zoogeography and evolution in Australia. Carlisle (Australia): Hesperian Press. p. 633–708.

Archer M, Arena R, Bassarova M, Black K, Brammall J, Cooke B, Creaser P, Crosby K, Gillespie A, Godthelp H, et al. 1999. The evolutionary history and diversity of Australian mammals. *Aust Mammal.* 21: 1–45.

Archer M, Beck R, Gott M, Hand S, Godthelp H, Black K. 2010. Australia's first fossil marsupial mole (Notoryctemorphia) resolves controversies about their evolution and palaeoenvironmental origins. *Proc R Soc B.* 278:1498–1506.

Avise JC. 2000. Phylogeography: the history and formation of species. Cambridge (MA): Harvard University Press.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography—the mitochondrial–DNA bridge between population-genetics and systematics. *Annu Rev Ecol Syst.* 18:489–522.

Barreda V, Palazzesi L. 2007. Patagonian vegetation turnovers during the Paleogene–Early Neogene: origin of arid-adapted floras. *Bot Rev.* 73: 31–50.

Beck RM. 2012. An 'ameridelphian' marsupial from the early Eocene of Australia supports a complex model of Southern Hemisphere marsupial biogeography. *Naturwissenschaften* 99:715–729.

Beck RMD. 2008. A dated phylogeny of marsupials using a molecular supermatrix and multiple fossil constraints. *J Mammal.* 89:175–189.

Benton MJ, Donoghue PCJ. 2007. Paleontological evidence to date the tree of life. *Mol Biol Evol.* 24:26–53.

Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. 2007. The delayed rise of present-day mammals. *Nature* 446:507–512.

Black K, Archer M, Hand S, Godthelp H. 2012. The rise of Australian marsupials: a synopsis of biostratigraphic, phylogenetic,

palaeoecologic and palaeobiogeographic understanding. In: Talent J, editor. Earth and life: Global biodiversity, extinction intervals and biogeographic perturbations through time. Dordrecht (The Netherlands): Springer. p. 983–1078.
Byrne M, Steane DA, Joseph L, Yeates DK, Jordan GJ, Crayn D, Aplin K, Cantrill DJ, Cook LG, Crisp MD, et al. 2011. Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *J Biogeogr.* 38:1635–1656.
Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ, Cooper S, Donnellan SC, Keogh JS, Leys R, et al. 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Mol Ecol.* 17:4398–4417.
Cardillo M, Bininda-Emonds ORP, Boakes E, Purvis A. 2004. A species-level phylogenetic supertree of marsupials. *J Zool (Lond).* 264:11–31.
Cramb J, Hocknull S, Webb GE. 2009. High diversity Pleistocene rainforest dasyurid assemblages with implications for the radiation of the Dasyuridae. *Austral Ecol.* 34:663–669.
de Queiroz A, Gatesy J. 2007. The supermatrix approach to systematics. *Trends Ecol Evol.* 22:34–41.
Dow DB. 1976. A geological synthesis of Papua New Guinea. Vol. 201. Canberra (Australia): Bureau of Mineral Resources, Geology and Geophysics. p. 41.
Felsenstein J. 1985. Phylogenies and the comparative method. *Am Nat.* 125:1–15.
Flannery T. 1995. Mammals of New Guinea. Chatswood (Australia): Reed Books.
Flannery TF. 1988. Origins of the Australo-Papuan land mammal fauna. *Aust Zool Rev.* 1:15–24.
Frakes LA, McGowran B, Bowler JM. 1987. Evolution of Australian environments. In: Dyne GR, Walton DW, editors. Fauna of Australia. Vol. 1A. Canberra (Australia): Australian Government Publishing Service. p. 1–16.
Hodell DA, Elmstrom KM, Kennett JP. 1986. Latest Miocene benthic $\delta^{18}O$ changes, global ice volume, sea-level and the Messinian salinity crisis. *Nature* 320:411–414.
Huey RB. 1987. Phylogeny, history and the comparative method. In: Feder ME, Bennett AF, Burgen W, Huey RB, editors. New directions in ecological physiology. Cambridge (United Kingdom): Cambridge University Press. p. 76–98.
Hugall AF, Stuart-Fox D. 2012. Accelerated speciation in colour-polymorphic birds. *Nature* 485:631–634.
Iglesias ARI, Artabe AE, Morel EM. 2011. The evolution of Patagonian climate and vegetation from the Mesozoic to the present. *Biol J Linn Soc.* 103:409–422.
Jansa SA, Barker FK, Voss RS. 2014. The early diversification of didelphid marsupials: a window into South America's "splendid isolation". *Evolution* 68:684–695.
Lanfear R, Calcott B, Ho SY, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution. *Mol Biol Evol.* 29: 1537–1719.
Lerner HR, Meyer M, James HF, Hofreiter M, Fleischer RC. 2011. Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Curr Biol.* 21: 1838–1844.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP. 2009. The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25: 2078–2079.
Maddison WP, Maddison DR. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75 [Internet]. [cited 2014 Jun 5]. Available from: http://mesquiteproject.org.
Martin HA. 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. *J Arid Environ.* 66:533–563.
Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T, et al. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524.
Meredith RW, Mendoza MA, Roberts KK, Westerman M, Springer MS. 2010. A phylogeny and timescale for the evolution of Pseudocheiridae (Marsupialia : Diprotodontia) in Australia and New Guinea. *J Mamm Evol.* 17:75–99.
Meredith RW, Westerman M, Case JA, Springer MS. 2008. A phylogeny and timescale for marsupial evolution based on sequences for five nuclear genes. *J Mamm Evol.* 15:1–36.
Meredith RW, Westerman M, Springer MS. 2008a. A phylogeny and timescale for the living genera of kangaroos and kin (Macropodiformes : Marsupialia) based on nuclear DNA sequences. *Aust J Zool.* 56:395–410.
Meredith RW, Westerman M, Springer MS. 2008b. A timescale and phylogeny for "Bandicoots" (Peramelemorphia : Marsupialia) based on sequences for five nuclear genes. *Mol Phylogenet Evol.* 47:1–20.
Meredith RW, Westerman M, Springer MS. 2009. A phylogeny of Diprotodontia (Marsupialia) based on sequences for five nuclear genes. *Mol Phylogenet Evol.* 51:554–571.
Metcalfe I, Smith JMB, Morwood M, Davidson I. 2001. Faunal and floral migrations and evolution in SE Asia-Australasia. Lisse (The Netherlands): Swets & Zeitlinger Publishers.
Meyer M, Stenzel U, Hofreiter M. 2008. Parallel tagged sequencing on the 454 platform. *Nat Protoc.* 3:267–278.
Mitchell KJ, Wood JR, Scofield RP, Llamas B, Cooper A. 2014. Ancient mitochondrial genome reveals unsuspected taxonomic affinity of the extinct Chatham duck (*Pachyanas chathamica*) and resolves divergence times for New Zealand and sub-Antarctic brown teals. *Mol Phylogenet Evol.* 70:420–428.
Müller J, Reisz RR. 2005. Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. *Bioessays* 27:1069–1075.
Nilsson MA, Churakov G, Sommer M, Van Tran N, Zemann A, Brosius J, Schmitz J. 2010. Tracking marsupial evolution using archaic genomic retroposon insertions. *PLoS Biol.* 8:e1000436.
Ortiz-Jaureguizar E, Cladera GA. 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. *J Arid Environ.* 66: 498–532.
Pagel M. 1997. Inferring evolutionary processes from phylogenies. *Zool Scr.* 26:331–348.
Pagel M, Meade A, Barker D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst Biol.* 53:673–684.
Phillips MJ, McLenachan PA, Down C, Gibb GC, Penny D. 2006. Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. *Syst Biol.* 55:122–137.
Phillips MJ, Pratt RC. 2008. Family-level relationships among the Australasian marsupial "herbivores" (Diprotodontia : Koala, wombats, kangaroos and possums). *Mol Phylogenet Evol.* 46: 594–605.
Raterman D, Meredith RW, Ruedas LA, Springer MS. 2006. Phylogenetic relationships of the cuscuses and brushtail possums (Marsupialia : Phalangeridae) using the nuclear gene BRCA1. *Aust J Zool.* 54: 353–361.
Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst Biol.* 57:4–14.
Reisz RR, Müller J. 2004. Molecular timescales and the fossil record: a paleontological perspective. *Trends Genet.* 20:237–241.
Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
Stadler T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl Acad Sci U S A.* 108:6187–6192.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sunderland (MA): Sinauer Associates.
Szalay FJ. 1994. Evolutionary history of the marsupials and an analysis of osteological characters. Cambridge (MA): Cambridge University Press.
Travouillon KJ, Legendre S, Archer M, Hand SJ. 2009. Palaeoecological analyses of Riversleigh's Oligo-Miocene sites: implications for Oligo-

Miocene climate change in Australia. *Palaeogeogr Palaeoclimatol Palaeoecol.* 276:24–37.
van Ufford AQ, Cloos M. 2005. Cenozoic tectonics of New Guinea. *AAPG Bull.* 89:119–140.
von Haeseler A. 2012. Do we still need supertrees? *BMC Biol.* 10:13.
Voss RS, Jansa SA. 2009. Phylogenetic relationships and classification of didelphid marsupials, an extant radiation of new world metatherian mammals. *Bull Am Mus Nat Hist.* 322:1–177.
Westerman M, Kear BP, Aplin K, Meredith RW, Emerling C, Springer MS. 2012. Phylogenetic relationships of living and recently extinct bandicoots based on nuclear and mitochondrial DNA sequences. *Mol Phylogenet Evol.* 62:97–108.
Westerman M, Young J, Krajewski C. 2008. Molecular relationships of species of *Pseudantechinus*, *Parantechinus* and *Dasykaluta* (Marsupialia: Dasyuridae). *Aust Mammal.* 29:201–212.
White ME. 2006. Environments of the geological past. In: Merrick JR, Archer M, Hickey GM, Lee MSY, editors. Evolution and biogeography of Australasian vertebrates. Sydney (Australia): Auscipub Pty Ltd.
Wilson DE, Reeder DM. 2005. Mammal species of the world: a taxonomic and geographic reference. Baltimore (MD): The Johns Hopkins University Press.
Yang Z. 2007. PAML4: a program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.