ABSTRACT

Aim Unique amongst birds, megapodes (family Megapodiidae) have exchanged the strategy of incubating eggs with the warmth of their bodies for incubation behaviours that rely entirely on environmental heat sources. Typically, mound-builders capture heat released from the decomposition of organic materials, while burrow-nesters lay their eggs in geothermal or solar-heated soils. The evolutionary path towards novel incubation behaviours has led to ecological and physiological adaptations unique to megapodes. Here, we present a species tree for all extant megapodes that settles long-standing debates about megapode evolution: namely, their biogeographical origins and ancestral nesting behaviour.

Location Australasia.

Methods A time-calibrated multilocus species tree for all extant megapodes was constructed using *BEAST. We estimated and compared divergence dates for megapodes obtained from molecular rates, fossils, and a combination of fossils and rates. Using this tree, Bayesian estimation of ancestral nesting behaviour was conducted in BayesTraits and ancestral ranges were estimated in BioGeoBEARS.

Results Recent dispersal has led to the recolonization of mainland Australia and New Guinea by Megapodius. Bayesian estimation of ancestral states indicates that mound building is the most probable ancestral nesting behaviour in megapodes (posterior probability = 0.75). Burrow nesting was acquired early in the diversification of the family (at least 14 Ma), followed by a single switch back to mound building.

Main conclusions Divergence dates and biogeographical reconstructions strongly suggest that dispersal, and not vicariance, led to the isolation of megapodes in Australasia. We propose that flight-mediated dispersal to environmentally variable islands is responsible for the behavioural lability in nesting behaviours observed in some Megapodius species today.

Keywords *BEAST, Australian biogeography, BayesTraits, BioGeoBEARS, dispersal, fossil calibration, Megapodius, nesting behaviour, phylogeography.
Hatchlings of all species must dig themselves out of the incubation site: a daunting undertaking that can take anywhere from hours to days. Once at the surface, there is no parent waiting and hatchlings are immediately able to run, forage and fly for short bursts (Jones et al., 1995).

In addition to their peculiar nesting behaviour, the early divergence of megapodes, their distribution in Australasia, and the distribution of extant Galliformes in the Southern Hemisphere have led to conflicting theories about their geographical origin (Olson, 1978; Crowe et al., 2006). Multiple molecular studies of galliform divergence times have found support for a Gondwanan origin of megapodes and divergence from other Galliformes in the Cretaceous (Crowe et al., 2006; Pereira & Baker, 2006a). However, these studies used contentious fossil calibrations that conflict with a palaeontological record devoid of Neornithes until the Late Cretaceous (Mayr, 2008; Ksepka, 2009). Thus, the geographical origins of megapodes remain equivocal. Recent studies have shown that many clades originally thought to be the result of Gondwanan vicariance are too young to be caused by continental breakup (de Queiroz, 2005). In such cases, oceanic dispersal is the preferred explanatory process for current species distributions.

The earliest known fossil megapode, *Ngawupodius minya*, demonstrates that megapodes were present on Australia during the Oligocene (Boles & Ivison, 1999). During the Pliocene, when the Australian plate moved closer to Asia and sea levels changed, megapodes are thought to have dispersed outwards through the Indonesian archipelago and eastwards towards Polynesia, using islands as stepping stones (Dekker, 2007). Isolated on new islands, ancestral megapodes presumably diversified and evolved the reproductive strategies observed today.

The evolutionary path of megapodes towards alternative incubation strategies has led to ecological and physiological adaptations unique to this group (Jones et al., 1995; Booth & Jones, 2001). Nearly all stages of megapode development have adaptations specific to their peculiar nesting behaviours. Eggs and chicks, which are buried up to half a metre or more underground during incubation, must survive the extraordinary conditions this environment imposes, including extreme humidity, variable incubation temperatures, high carbon dioxide and low oxygen levels. Megapode eggs have multiple adaptations that facilitate gas exchange, including thin eggshells and egg pores that broaden throughout the course of development. Megapode eggs have unusually large eggs, high yolk content and long incubation periods. As a result, the emerging megapode chicks are the most developmentally advanced of any bird and are completely independent from their parents. However, the conspicuous nature of mounds

<table>
<thead>
<tr>
<th>Species</th>
<th>Status</th>
<th>Nest type</th>
<th>Heat Source</th>
<th>'Ancestral' biogeographical range</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aepypodius arfakianus</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Aepypodius bruijni</em></td>
<td>EN</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Alectura lathami</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Leipoa ocellata</em></td>
<td>VU</td>
<td>M</td>
<td>C, S</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Talegalla cuvieri</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Talegalla fuscirostris</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Talegalla johiensis</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td>'Scrubfowl' clade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Eulipoa wallacei</em></td>
<td>VU</td>
<td>B</td>
<td>S</td>
<td>Wallacea</td>
</tr>
<tr>
<td><em>Megapodius bernsteinii</em></td>
<td>VU</td>
<td>M</td>
<td>C</td>
<td>Wallacea</td>
</tr>
<tr>
<td><em>Megapodius cunningii</em></td>
<td>LC</td>
<td>B, M</td>
<td>C, R</td>
<td>PHL + Palawan, Wallacea</td>
</tr>
<tr>
<td><em>Megapodius decollatus</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Megapodius eremita</em></td>
<td>LC</td>
<td>B, M</td>
<td>C, G, R, S</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Megapodius forstenii</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>Wallacea</td>
</tr>
<tr>
<td><em>Megapodius freycinet</em></td>
<td>LC</td>
<td>M</td>
<td>C</td>
<td>Wallacea</td>
</tr>
<tr>
<td><em>Megapodius geelvinkiana</em></td>
<td>VU</td>
<td>?</td>
<td>?</td>
<td>AUS + PNG</td>
</tr>
<tr>
<td><em>Megapodius laperause</em></td>
<td>EN</td>
<td>B, M</td>
<td>C, G, S</td>
<td>Islands</td>
</tr>
<tr>
<td><em>Megapodius layardi</em></td>
<td>VU</td>
<td>B, M</td>
<td>C, G, R, S</td>
<td>Islands</td>
</tr>
<tr>
<td><em>Megapodius nicobariensis</em></td>
<td>VU</td>
<td>M</td>
<td>C</td>
<td>Nicobar</td>
</tr>
<tr>
<td><em>Megapodius pritchardii</em></td>
<td>EN</td>
<td>B</td>
<td>G, R? S?</td>
<td>Islands</td>
</tr>
<tr>
<td><em>Megapodius reinwardt</em></td>
<td>LC</td>
<td>M</td>
<td>C, S</td>
<td>AUS + PNG, Wallacea</td>
</tr>
<tr>
<td><em>Megapodius tenimberensis</em></td>
<td>NT</td>
<td>M</td>
<td>C</td>
<td>Wallacea</td>
</tr>
<tr>
<td><em>Macrocephalon maleo</em></td>
<td>EN</td>
<td>B</td>
<td>G, S</td>
<td>Wallacea</td>
</tr>
</tbody>
</table>

*Species represented by museum toe-pad samples only and excluded from the ‘full coverage’ data matrix.
†Species represented by fresh tissue and museum toe-pad samples.

Status: LC, least concern; EN, endangered; VU, vulnerable; NT, near threatened. Nest type: M, mound; B, burrow; N, typical nest; ?, unknown. Heat source: C, compost; S, sun; H, body heat; G, geothermal; R, decaying roots; ?, unknown. Biogeographical range: AUS, Australia; PHL, Philippines; PNG, Papua New Guinea.
and communal burrowing grounds make megapodes highly susceptible to predation on eggs and chicks, especially by humans (Jones et al., 1995). It is likely that recent extinction is high due to the impact of human colonization (Steadman, 1993).

The evolutionary sequence of nesting behaviours in megapodes remains unresolved, and previous studies have proposed two alternative hypotheses. If nesting behaviours evolved with increasing complexity, then burrow nesting would be the original megapode incubation strategy (Frith, 1956; Immelmann & Sossinka, 1986). In this scenario, ancestral megapodes built nests on top of naturally occurring heat sources, and over evolutionary time became dependent on those heat sources (e.g. solar radiation or geothermal heat). Alternatively, mound building could have evolved from a habit of covering eggs with debris to hide them from predators (Booth & Jones, 2001). Assuming a hot moist habitat, ambient temperatures protected eggs from chill and promoted decomposition of the organic material left on top. With increasing amounts of debris, eggs could be left unattended for longer periods of time, and eventually these megapodes stopped using body heat to incubate, having become entirely reliant upon mounds (Booth & Jones, 2001). Yet, mound building and burrow nesting are not mutually exclusive, and given that some species within the genus Megapodiidae currently utilize both incubation strategies, it is possible that the ancestral megapode was also behaviourally labile.

To test these hypotheses regarding the sequence of megapode incubation strategies, biogeographical origins and colonization history, we constructed a multilocus time-calibrated species tree (see Appendix S1 in Supporting Information), including six individuals representing all species from all seven genera (see Table 1). We included out-groups representing all major families of Galliformes: two from the family Cracidae (Ortalis wagleri and Penelope purpurascens), one from Phasianidae (Phasianus colchicus), one from Odontophoridae (Callipepla californica), and one from Numididae (Numida meleagris). We included one species from the order Anseriformes (Branta bernica), the sister order to the Galliformes, to aid calculation of divergence times. Taxon sampling included 28 species (Table 1) and 81 individuals (Appendix S1).

**Materials and Methods**

**Taxon sampling**

Our study represents the first complete phylogenetic analyses of the 22 currently recognized extant species of megapodes (Jones et al., 1995). We sampled a total of 75 megapode individuals representing all species from all seven genera (see Appendix S1 in Supporting Information), including six megapode species never before included in phylogenetic analysis: the recently re-discovered Aepyornis bruijnii (Heiji & Post, 2001; Mauro, 2005), two species from the genus Talegalla (T. jobiensis and T. cumingii), and four species from the genus Megapodus (M. nicobariensis, M. bernsteinii, M. geevinkianus and both subspecies of M. laperouse). Whenever possible, multiple individuals of each species were included to increase the information available for the coalescent-based species tree analysis (Appendix S1). To account for variation within species, we sampled across species’ ranges and included all available subspecies.

We used fresh tissue samples or whole genomic DNA when available. However, owing to the small populations, remote distributions, endangered status and cryptic nature of many megapodes, it was impossible to get fresh samples from some species (Table 1). In such cases, we used toe-pad samples from voucher specimens in museum collections.

Most recent classifications place Megapodiidae as sister to all other families in the Galliformes (Crowe et al., 2006; Cox et al., 2007; Wang et al., 2013). Therefore, we included out-groups representing all major families of Galliformes: two from the family Cracidae (Ortalis wagleri and Penelope purpurascens), one from Phasianidae (Phasianus colchicus), one from Odontophoridae (Callipepla californica), and one from Numididae (Numida meleagris). We included one species from the order Anseriformes (Branta bernica), the sister order to the Galliformes, to aid calculation of divergence times. Taxon sampling included 28 species (Table 1) and 81 individuals (Appendix S1).

**Data collection**

For fresh tissue samples, genomic DNA was extracted using the DNeasy Kit (QIAGEN, Valencia, CA, USA) following the manufacturer’s protocol. A total of 14 nuclear and two mitochondrial (mtDNA) genes were included (Table 2). Loci were PCR amplified and sequenced using standard procedures. Cycle-sequence products were cleaned using ExoSAP-IT (USB Corporation, Cleveland, OH, USA) before being analysed on an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) at the University of Washington Comparative Genomics Center. We confirmed the identity of each locus using a BLAST search (Altschul et al., 1990) against the chicken (Gallus gallus) genome.

DNA from toe-pad samples was extracted in a degraded-DNA laboratory with physical separation and air-handling isolation at Cornell University. Toe-pad extractions and PCR reactions were conducted at a laminar flow clean bench with sterilized surfaces. We interspersed negative controls at both the extraction and PCR steps to increase the chance of identifying contamination. For these degraded samples, we amplified short (150–300 bp), overlapping regions of the ND2 and cytB gene, as well as highly variable 250 bp regions of four nuclear introns (see Appendix S2).

We aligned sequences with MAFFT (Katoh et al., 2002) and visually edited sequences using Geneious Pro 5.3.6 (Drummond et al., 2010). Indels were present in 13 of 14 nuclear loci and gaps were coded as missing data. Ambiguity codes represent polymorphisms in heterozygous individuals.

**Gene tree analyses**

Gene trees were estimated using MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001) and RAXML 7.2.8 (Stamatakis et al., 2005, 2008). We estimated gene trees for each individual
nuclear locus separately, all 14 nuclear loci concatenated, the combined mtDNA data, and all 16 genes concatenated. The best fitting substitution model was selected for each gene using jModelTest 0.1.1 (Posada, 2008) under the Akaike information criterion (AIC) and AIC corrected for small sample size (AICc) (Table 2). Each analysis used six Markov chains that were run for 10 million generations for the analysis of separate genes and 15 million generations for the combined datasets (samplefreq = 5000). To ensure convergence, analyses were run until the potential scale reduction factor for estimated parameters was approximately 1.0 and the average standard deviation of split frequencies was < 0.01. Chain convergence was further assessed using Tracer 1.5 (Rambaut & Drummond, 2007) and convergence of the cumulative posterior probabilities of clades was assessed using the online program awty (Nylander et al., 2008). Posterior probability (PP) values were obtained by summarizing the posterior distribution of trees with a 50% majority rule consensus tree and 50% burn-in.

A maximum likelihood RAxML analysis estimated the optimal tree by conducting a search for the best scoring maximum-likelihood tree using the GTR+G model. We assessed nodal support using a nonparametric bootstrap analysis with 1000 replicates (Felsenstein, 1985). To help RAxML distinguish amongst trees with similar likelihood scores, the likelihood epsilon was set to \(-1.0 \times 10^{-5}\).

### Species tree analyses

We estimated a time-calibrated species tree using *BEAST* (Heled & Drummond, 2010), with time-calibration information incorporated as outlined in McCormack et al. (2010). We estimated and compared divergence dates for megapodes obtained from three separate analyses that utilized: (1) fossil information only; (2) molecular rate calibrations only; and (3) a combination of fossils and rates. To explore the impact of missing data, we analysed two datasets: (1) a near ‘full coverage’ matrix containing the 24 out of 28 species with both nuclear and mtDNA data; and (2) an ‘all species’ data matrix containing all 28 species, which contained missing data at some loci for four species (Table 1). See Appendix S1 for detailed information on sampling coverage at nuclear loci. In all analyses, six independent runs of the ‘all species’ dataset were performed for one billion generations, sampling every 100,000 generations. Each locus was treated as a separate data partition by unlinking gene trees, rates and substitution models (Table 2). However, the gene trees were linked for the mtDNA genes (*ND2* and *cytb*). We used random starting gene trees under the coalescent model, a birth–death process for the species-tree prior, and a constant population size over the entire time interval, both with default priors. Because the monophyly of the megapodes is well established (Birks & Edwards, 2002), we enforced the monophyly of the group. The mean clock (ucld.mean) was set to an exponential distribution with a mean of 10.

We assessed convergence of model parameters and node heights by plotting the marginal posterior distributions in TRACER, checking for overlapping posterior distributions across independent analyses, and ensuring high effective sample size (ESS) values \(\geq 200\). The post-burn-in samples (25%) were combined using LogCombiner 1.7.5 (Drummond & Rambaut, 2007). A maximum clade credibility (MCC) tree was constructed using TreeAnnotator 1.7.5 (Drummond & Rambaut, 2007) with median node heights.

### Estimating divergence time

Choosing an appropriate way to time-calibrate the megapode species tree is a difficult problem: in addition to a sparse galliform fossil record, previous Galliformes molecular dating analyses have used controversial fossils to estimate the rate of mtDNA sequence evolution (Mayr, 2008; Ksepka, 2009). The

---

**Table 2 Loci used in the study. Descriptive values are calculated from megapodes (outgroups excluded).**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primers (Fwd, Rev)</th>
<th>Source</th>
<th>Length (bp)</th>
<th>Model</th>
<th>Identical Sites (%)</th>
<th>% GC</th>
<th>GenBank Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGF</td>
<td>AIINGF5', AIINGF3'</td>
<td>Kimball et al. (2009)</td>
<td>737</td>
<td>HKY</td>
<td>80.2</td>
<td>51.0</td>
<td>KF814728-4780</td>
</tr>
<tr>
<td>ARNTL</td>
<td>ARNTL1.12F, ARNTL1.13R</td>
<td>Kimball et al. (2009)</td>
<td>784</td>
<td>HKY</td>
<td>79.8</td>
<td>40.6</td>
<td>KF833401-3461</td>
</tr>
<tr>
<td>CLTCL1</td>
<td>CLTCL1.17F, CLTCL1.e8Ralt</td>
<td>Kimball et al. (2009)</td>
<td>526</td>
<td>HKY</td>
<td>83.0</td>
<td>39.6</td>
<td>KF833462-3523</td>
</tr>
<tr>
<td>CSDE1</td>
<td>CSDE.5F, CSDE.6R</td>
<td>Kimball et al. (2009)</td>
<td>495</td>
<td>HKY</td>
<td>80.5</td>
<td>33.6</td>
<td>KF833524-3583</td>
</tr>
<tr>
<td>cytb</td>
<td>H15646, L14996</td>
<td>Sorenson et al. (1999)</td>
<td>866</td>
<td>HKY+G</td>
<td>41.3</td>
<td>49.4</td>
<td>KF833584-3639</td>
</tr>
<tr>
<td>EEF2</td>
<td>EEF2.6F, EEF2.7R</td>
<td>Kimball et al. (2009)</td>
<td>410</td>
<td>K80+G</td>
<td>81.1</td>
<td>49.4</td>
<td>KF833640-3698</td>
</tr>
<tr>
<td>FGB</td>
<td>Fib5, Fib6</td>
<td>Kimball et al. (2009)</td>
<td>596</td>
<td>HKY+G</td>
<td>76.6</td>
<td>37.4</td>
<td>KF833699-3756</td>
</tr>
<tr>
<td>IRF2</td>
<td>IRF2.2F, IRF2.3R</td>
<td>Kimball et al. (2009)</td>
<td>660</td>
<td>HKY+G</td>
<td>83.5</td>
<td>38.8</td>
<td>KF833757-3817</td>
</tr>
<tr>
<td>NAT15</td>
<td>NAT.4F, NAT.5R</td>
<td>Kimball et al. (2009)</td>
<td>1024</td>
<td>GTR+G</td>
<td>61.9</td>
<td>58.5</td>
<td>KF833818-3868</td>
</tr>
<tr>
<td>NADH-2</td>
<td>L5145, H6394</td>
<td>Birks and Edwards (2002)</td>
<td>1077</td>
<td>GTR+G</td>
<td>15.1</td>
<td>47.0</td>
<td>KF854310-4340</td>
</tr>
<tr>
<td>PARK7</td>
<td>PARK.2F, PARK3R</td>
<td>Kimball et al. (2009)</td>
<td>728</td>
<td>HKY</td>
<td>84.1</td>
<td>36.1</td>
<td>KF833869-3921</td>
</tr>
<tr>
<td>PAXIP1</td>
<td>PAX.20F, PAX.21R</td>
<td>Kimball et al. (2009)</td>
<td>463</td>
<td>HKY</td>
<td>75.5</td>
<td>31.9</td>
<td>KF833922-3978</td>
</tr>
<tr>
<td>PER2</td>
<td>PER.9F, PER.10R</td>
<td>Kimball et al. (2009)</td>
<td>578</td>
<td>HKY</td>
<td>68.5</td>
<td>23.9</td>
<td>KF833979-4035</td>
</tr>
<tr>
<td>RDP1</td>
<td>RDP1.1U, RDP1.1L</td>
<td>Birks and Edwards (2002)</td>
<td>901</td>
<td>HKY+G</td>
<td>80.7</td>
<td>46.3</td>
<td>KF834036-4073</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>TGFBR2.5F, TGFBR6.6R</td>
<td>Kimball et al. (2009)</td>
<td>586</td>
<td>HKY</td>
<td>81.9</td>
<td>42.0</td>
<td>KF834074-4130</td>
</tr>
<tr>
<td>TXNDC12</td>
<td>TXN.6F, TXN.7R</td>
<td>Kimball et al. (2009)</td>
<td>488</td>
<td>HKY+G</td>
<td>80.7</td>
<td>41.6</td>
<td>KF834131-4185</td>
</tr>
</tbody>
</table>
Galiformes fossil record lacks undisputed, phylogenetically placed fossils (Mayr & Weidig, 2004; Mayr, 2005, 2008; Ksepka, 2009). Therefore, we used the Late Cretaceous fossil Anseriformes, Vegavis iaai, as the upper bound for the Galiformes–Anseriformes split (Clarke et al., 2005), applying a gamma prior G(2, 9) with an offset of 66 Ma to the root Galiformes–Anseriformes split.

Recent studies have shown estimated rates of molecular evolution to be dependent on the time-scale of measurement (Ho et al., 2011). The widely used avian mtDNA clock of 2.1% is consistent over the past 12 Myr (Weir & Schluter, 2008); however, Galiformes are an early diverging group with stem fossils placed in the Eocene (Mayr, 2005) and estimates of molecular rates have consistently been well below the standard 2.1% (Kimball et al., 1997; Crowe et al., 2006; Pereira & Baker, 2006a). Here, we conducted analyses using three rates of cytb sequence evolution in Galiformes: 0.7% (Kimball et al., 1997), 1.2% (Pereira & Baker, 2006b), and 2.1% (Weir & Schluter, 2008). All genes had their own uncorrelated lognormal relaxed clock model and a mean rate of 0.0001 was chosen for a smoother starting point for the nuclear clock rate estimation. In addition, we conducted three analyses using a combination of each rate and the Vegavis iaai fossil (Table 3).

The posterior distribution of species trees is a product of both priors defined by the user and the data. To explore the influence of our priors on our posterior, we ran all *BEAST analyses with empty alignments (Drummond et al., 2006).

**Biogeography**

To investigate the biogeographical origins of megapodes, ancestral ranges were inferred using BioGeoBEARS implemented in R (Matzke, 2013a). This program allows for both probabilistic inference of ancestral geographical ranges and statistical comparisons of different models of range expansion [e.g. dispersal–extinction–cladogenesis (DEC), dispersal–vicariance analysis (DIVA) and BayArea] using a time-calibrated species tree. Additionally, BioGeoBEARS incorporates founder-event speciation (+), a process important in island systems, into all models (Matzke, 2013b). BioGeoBEARS is fully parameterized and includes a number of novel features that increase the flexibility of hypothesis testing. Here, we defined an ‘areas allowed’ matrix, a supplement to the dispersal matrix used in DEC models, such as the widely used program LAGRANGE. Prior to the development of BioGeoBEARS, users of DEC models specified the relative probability of dispersal between areas at different time strata; however, this is not the same as specifying the disappearance of areas in the past and can potentially be translated as long-term persistence. We compared models using the likelihood ratio test and AIC.

We defined eight geographical areas based on our knowledge of the biology and geographical history of the birds: (1) Africa, (2) New World, (3) Eurasia, (4) Australia + New Guinea, (5) Wallacea, (6) Nicobar Islands, (7) Philippines + Palawan, and (8) islands (Vanuatu + Tonga + Palau + Northern Marianas) (Fig. 1). Until approximately 5 Ma, Australia and New Guinea were intermittently connected (Hall, 1998), so we grouped these two areas together in our analysis. Despite their markedly different biogeographical histories, we defined Palawan and the Philippines as a single area in our analyses (Siler et al., 2012; Megapodius cumingii is the only resident species, and delineating between the two islands would not be informative. We included the Solomon Islands with New Guinea, as the only species in the Solomon Islands, M. eremita, is known to hybridize with M. reinwardtii, a widespread species resident to Australia, New Guinea and Wallacea. Within megapodes, species were assigned to their corresponding geographical area(s), whereas outgroup species were used as exemplars for all members of that family or order (Table 1). We set the maximum number of areas equal to eight because Anseriformes are present in all regions defined in the analysis. Two areas were constrained using maximum age based on information from the geological record: (1) Wallacea was submerged before 45 Ma (Hall, 1998); and (2) the oldest ‘island’ was not formed until c. 10 Ma (Hall, 2009).

Based on an earlier phylogeny of Megapodiidae (Birks & Edwards, 2002), Dekker (2007) presents the most recent biogeographical hypothesis for species in the genus **Table 3** Divergence time estimates (in million years ago, Ma) of megapodes from time-calibrated *BEAST analyses, using: fossil (Vegavis iaai) information only; three rates of cytb sequence evolution in Galiformes (0.7%, 1.2% and 2.1%); and a combination of fossils and molecular rates.

<table>
<thead>
<tr>
<th>Time-calibration</th>
<th>Galliformes/Anseriformes</th>
<th>Brush-turkey/scrubfowl</th>
<th>Megapodius</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean from prior (95% HPD)</td>
<td>Mean from posterior sample (95% HPD)</td>
<td>Mean from posterior sample (95% HPD)</td>
</tr>
<tr>
<td>Fossil</td>
<td>80.9 (66.3–101.9)</td>
<td>70.9 (66.1–80.7)</td>
<td>14.4 (11.9–17.2)</td>
</tr>
<tr>
<td>0.7%</td>
<td>–</td>
<td>132.0 (98.8–170.4)</td>
<td>29.6 (22.1–37.7)</td>
</tr>
<tr>
<td>1.2%</td>
<td>–</td>
<td>75.1 (54.5–96.9)</td>
<td>17.6 (13.0–23.0)</td>
</tr>
<tr>
<td>2.1%</td>
<td>–</td>
<td>42.2 (32.4–54.2)</td>
<td>10.0 (7.7–13.0)</td>
</tr>
<tr>
<td>Fossil + 0.7%</td>
<td>82.4 (66.4–105.0)</td>
<td>110.1 (85.1–135.7)</td>
<td>27.8 (22.0–34.6)</td>
</tr>
<tr>
<td>Fossil + 1.2%</td>
<td>82.5 (66.4–105.6)</td>
<td>74.2 (66.7–85.8)</td>
<td>18.2 (15.2–21.9)</td>
</tr>
<tr>
<td>Fossil + 2.1%</td>
<td>82.3 (67.1–104.4)</td>
<td>70.4 (66.1–78.7)</td>
<td>13.4 (10.8–16.8)</td>
</tr>
</tbody>
</table>

HPD, highest posterior density.
Figure 1 Biogeographical analysis of megapodes using BioGeoBEARS. The six biogeographical areas include: (1) New World, (2) Africa, (3) Eurasia, (4) Australia + New Guinea (AUS + PNG), (5) Wallacea, (6) Philippines (PHL) + Palawan, (7) Nicobar islands, and (6) islands north and east of Papua New Guinea. Outgroups are not shown. Pie charts at nodes indicate support for respective areas. Tips are labelled with present-day species distributions. PP, posterior probability.

Figure 2 Distribution of Megapodius species and four putative waves of colonization (Dekker, 2007). Colours indicate the current species distributions corresponding to each wave of colonization: orange, wave A; green, wave B; blue, wave C; and pink, wave D. Grey indicates present-day distributions of species not included in Dekker’s (2007) hypothesis.
suggesting that their current distribution is the result of four major waves of colonization – all originating from Australia. Species relationships can provide an explicit test of Dekker’s (2007) hypothesis (Fig. 2). Because the proposed waves are based on relatedness and the direction of dispersal from Australia (rather than biogeography), we were unable to test this hypothesis using ancestral area reconstruction. Instead, we summarized the posterior distribution of species trees to quantify the posterior probabilities for each of the proposed waves.

Ancestral state reconstruction

The nest building behaviour and heat source used by each species of megapode is well documented in the literature (Jones et al., 1995; Sinclair et al., 2002; Bowen, 2010) (Table 1). We coded nesting behaviour as an unordered multistate character with three states: standard nest, mound building or burrow nesting. We fixed the root state to ground nesting. We conducted ancestral state estimation using the module MultiState in BayesTraits 2.0 (Pagel et al., 2004). This program fits a continuous-time Markov model to discrete character data, allowing for changes between states at any given time over small intervals and estimating the posterior probability for character states at ancestral nodes by integrating over a posterior distribution of phylogenetic trees (Pagel, 1994). For each rendition of our analyses, we selected 1000 trees from the post-burn-in posterior distribution of species trees. Priors were seeded in similar topologies (Fig. 3, Appendix S3), suggesting that in this case missing data did not influence coalescent-based species tree inference. Both the Bayesian and ML concatenated gene tree topology show similar relationships between species (Appendix S3). Our species trees confirm Megapodius...
tailed as sister to all other galliform families, followed by the Cracidae. Within Megapodiidae, we found a deep split between the clade containing the brush-turkeys (Alectura and Aeypodius), tallgalas (Talegalla) and malleefowl (Leipoa), which we refer to here as the 'brush-turkey clade' for simplicity, and the clade containing Endlipoa and Megapodius (historically known as 'scrubfowl'), which is consistent with previous phylogenetic estimates (Birks & Edwards, 2002). A key feature of the tree topology is the placement of maleo (Macrocephalon maleo) at the base of the scrubfowl clade with high support, rather than with the brush-turkeys, as the mtDNA gene tree (Appendix S3) and previous molecular phylogeny would suggest (Birks & Edwards, 2002). An endangered, charismatic species confined to Sulawesi, the maleo is the only burrow-nesting megapode not closely related to the scrubfowl; it has a unique morphology including a dramatic knob-like head casque and distinct webbing between its toes. Thus, its phylogenetic placement has been of particular interest, but hampered by the fact that the divergence of its lineage came early within the megapodes, and that there are currently no closely related species. Within the brush-turkeys, we resolve Leipoa as sister to (Alectura, Aeypodius) with high support. With the addition of Talegalla cuvieri, T. jobiensis and Aeypodius bruijini, the genera Talegalla and Aeypodius remain monophyletic. The relationships within Megapodius Generally resemble those found by Birks et al. (2005). Both studies found an initial split off of M. cumingii and M. tenimberensis, strong support for M. layardi as sister to M. pritchardii (but low support for the clade’s position in the tree), and high support for the clade formed by M. decollatus, M. forstenii and M. freycinet (Fig. 3).

The addition of four megapode species sampled from only toe-pads (T. cuvieri, M. nicobariensis, M. geelvinkianus and M. bernsteini) with short fragmentary sequences and considerable missing data did not change the species tree topology. Overall, these species are placed with relatively low support, and each contributes to a ladder-shaped (pectinate) species tree. The placement of M. nicobariensis has been of particular interest, as it occupies the Nicobar Islands at the far western edge of the megapode distribution, disjunct from all other species. Based on geographical proximity, Dekker (2007) predicted a close relationship with M. cumingii, while Jones et al. (1995) predicted a close relationship with M. pritchardii based on morphology. Interestingly, M. nicobariensis diverged early in the genus and is not sister to either of these species, suggesting that geographical proximity may be a poor predictor of phylogenetic relationships in this genus (see below).

**Divergence times**

The species tree provided different divergence times depending on the priors used for time calibration (Table 3). Using the cytb molecular rate of 0.7%, we date the split between Galliformes and Anseriformes at c. 130 Ma, which is around the age of Early Cretaceous fossil sites, such as the Chinese Jehol Biota, which have yet to produce a single Neornithes fossil (Mayr, 2008), and only 20 Myr after Archaeopteryx – a highly unlikely scenario. The standard avian rate of 2.1% is no more realistic, as it dates this split at c. 42 Ma, which post-dates the well-established Vegavis iaai fossil by over 25 Myr. A molecular rate of 1.2%, however, places the Galliformes/Anseriformes split at c. 75 Ma, within the 95% confidence interval of our fossil prior calibration that we assumed in other analyses.

The disparity in our divergence time estimates using different molecular rates highlights the need for well-placed fossils. Without such information, divergence time estimates can be profoundly overestimated. Divergence dates estimated from the root fossil and the 1.2% rate calibrations alone, and in combination, are consistent with each other (Table 3). Therefore, in all figures, we present the results for the tree estimated using both the fossil and rate of 1.2%.

**Biogeography**

Our biased sampling towards megapodes does not permit us to make inferences about the geographical origin of Galliformes, and inferring the route ancestral megapodes took to Australasia is beyond the scope of this study. Regardless, ancestral megapodes reached the Australasian region – colonizing islands in Wallacea, as well as mainland Australia and New Guinea. Lineages represented in the brush-turkey clade continued on mainland Australia and New Guinea, while lin-

**Table 4** Megapode ancestral range reconstruction inferred in BroGroBEARS and the relative probabilities of each model calculated from the AICc weights. Models include dispersal–extinction–cladogenesis (DEC), dispersal–vicariance analysis (DIVA) and BayArea, as well as these three models allowing for founder-event speciation (+J).

<table>
<thead>
<tr>
<th>Model</th>
<th>Lnl</th>
<th>No. of parameters</th>
<th>d</th>
<th>e</th>
<th>j</th>
<th>AIC</th>
<th>AICc</th>
<th>Relative Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC</td>
<td>−59.1</td>
<td>2</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
<td>122.2</td>
<td>124.6</td>
<td>11%</td>
</tr>
<tr>
<td>DEC+J</td>
<td>−54.6</td>
<td>3</td>
<td>0.004</td>
<td>0</td>
<td>0.025</td>
<td>115.2</td>
<td>121.2</td>
<td>61%</td>
</tr>
<tr>
<td>DIVA</td>
<td>−64.8</td>
<td>2</td>
<td>0.006</td>
<td>0</td>
<td>0</td>
<td>133.6</td>
<td>136</td>
<td>0%</td>
</tr>
<tr>
<td>DIVA+J</td>
<td>−55.4</td>
<td>3</td>
<td>0.004</td>
<td>0</td>
<td>0.021</td>
<td>116.8</td>
<td>122.8</td>
<td>27%</td>
</tr>
<tr>
<td>BayArea</td>
<td>−86.9</td>
<td>2</td>
<td>0.013</td>
<td>0.055</td>
<td>0</td>
<td>177.8</td>
<td>180.2</td>
<td>0%</td>
</tr>
<tr>
<td>BayArea+J</td>
<td>−60.1</td>
<td>3</td>
<td>0.003</td>
<td>0</td>
<td>0.041</td>
<td>126.2</td>
<td>132.2</td>
<td>0%</td>
</tr>
</tbody>
</table>

Lnl, log-likelihood; d, rate of dispersal; e, rate of extinction; j, relative probability of founder-event speciation at cladogenesis; AIC, Akaike’s information criterion; AICc, Akaike’s information criterion corrected for small sample size.
eages ancestral to *Macrocephalon* and the scrubfowl propagated in the islands of Wallacea (Fig. 1). *Megapodius* appears to have had a predominantly island distribution, extending from the Indian Ocean to the South Pacific. Recent dispersal has led to the recolonization of *Megapodius* to mainland Australia and New Guinea. Model comparisons show that the three models implemented (DIVA, DEC and BayArea) with founder-speciation (+f) are favoured over the standard implementation (Table 4). We show the model with the highest relative probability (Table 4), DEC+J, in Fig. 1. Ancestral area reconstruction did not depend on time calibration and we present the results from the species tree estimated from the fossil and the 1.2% molecular rate.

Using the posterior distribution of species trees, we found low support for Dekker’s wave C (PP = 0.22), and no support for wave A (PP = 0.0). Given the results of our *BEAST* analyses (Fig. 3), the high support for waves B (PP = 0.98) and D (PP = 1.0) was unsurprising.

**Ancestral state reconstruction**

We reconstruct mound building as the most probable ancestral nesting behaviour in megapodes and find that burrow nesting evolved independently three times (Fig. 3). The ancestral megapode to the brush-turkey clade was a mound builder (PP = 0.87), as all species in this clade share this nesting strategy. The node separating *Macrocephalon* from the scrubfowl is reconstructed as burrow nesting, albeit with equivocal support (PP = 0.56).

Within *Megapodius*, mound building is the ancestral strategy, but unlike in other megapode genera, nesting strategy is variable. Four species are polymorphic for mound and burrow nesting but are not closely related and do not have overlapping distributions. There is high support for burrow nesting as the ancestral behaviour for *M. pritchardii* and *M. lanyardii*, the two species that have distributions in the eastern South Pacific on small, remote islands (Figs 2 & 3).

**DISCUSSION**

**Biogeographical origins**

The new time-calibrated species tree for megapodes (Fig. 3) indicates that they have been restricted to Australasia throughout their evolutionary history. This finding supports what was previously inferred from the palaeontological record. The absence of megapode fossils from New Zealand suggests that ancestral megapodes were not yet present on the Australian landmass at the time of New Zealand’s separation (c. 85 Ma) and the abundance of fossil phasianids and absence of fossil megapodes in the Miocene deposits of mainland Asia suggests that megapodes were confined to the Indo-Australian region at that time (Cheneval et al., 1991; Steadman, 1999). However, we note that the absence of fossils could result from poor conditions for fossilization or insufficient fieldwork, and should not be considered as proof that megapodes were absent for that locality.

The current megapode distribution was previously believed to be a result of Gondwanan vicariance, with ancestral megapodes originating in Australia. However, we date the split between megapodes and all other Galliformes at 70–75 Ma, well after Australia had begun its movement northwards (Hall, 2002), which strongly suggests that dispersal, and not vicariance, led to the isolation of megapodes in Australasia.

We reconstruct Australia, New Guinea and Wallacea as the ancestral range of all megapodes (c. 18 Ma). At this time, the Philippines were connected to mainland Asia via a shallow sea, the Australian continent was continuous with the southern portion of New Guinea, and part of Sulawesi was above water (Hall, 1998).

Approximately 5 Ma, New Guinea experienced a large climatic shift, which resulted in new habitats and occurred in conjunction with a rapid increase in land area and range elevations (Hall, 2002). Within the past 2 Ma, the islands of the Banda Arc formed and the rest of New Guinea emerged (Lohman et al., 2011). The formation of more islands during this time would have led to increased opportunity for dispersal and, consequently, allopatric speciation. These events coincide with the observed *Megapodius* radiation (Fig. 1).

**Evolution of nesting strategy**

We found new evidence for multiple origins of burrow nesting in megapodes (Fig. 3). Historically, megapode incubation strategy has been viewed as a dichotomous trait (but see Booth & Jones, 2001) and with the exception of *Megapodius*, all megapode genera are strictly mound or burrow nesters. Species within *Megapodius* show variation in nesting strategy and heat source usage (Table 1). *Megapodius* has a widespread, predominantly island distribution, whereas most other genera are restricted to mainland Australia and New Guinea and/or have limited ranges. This distribution may be due to flight ability: whereas *Megapodius* and *Eulipoa* species are strong fly- ers, other megapodes are generally heavy-bodied and poor fly- ers, incapable of long-distance water crossings (Dekker, 2007).

In adaptive radiations, ecological opportunity can directly promote changes in behaviour, morphology and physiology (Price et al., 2003). In megapodes, temporal and spatial variation in heat sources available for nesting should favour the evolution of behavioural lability, whereby species can change incubation strategy when faced with an immediate change in environment. We propose that strong flight capability allowed *Megapodius* to follow a distinct evolutionary trajectory: by dispersing throughout the newly formed islands of Australasia c. 2 Ma, these species encountered variable nesting substrates and heat sources, which led to the evolution of phenotypic lability in incubation strategy. Mound building may have been favoured in environments with ample vegetative material and rainfall and no alternative heat sources, whereas burrow nesting probably would have been favoured in areas with nesting substrates heated by sun or...
geothermal activity. Mound building is a much more energy-consuming process that involves not only building but tending to the mound throughout incubation, so there may be strong selection against megapodes to abandon mounds when there is an alternative environmental heat source available. Because at least some species of *Megapodius* appear to be monogamous, with females contributing to mound building (Crome & Brown, 1979), a shift to the energetically less costly strategy of burrow nesting may have freed females to invest more in eggs and produce more young. Thus, incubation behaviour (mound versus burrow) may largely be dependent on environment in *Megapodius*, rather than being phylogenetically constrained.

Evidence from extant *Megapodius* species, as well as palaeontological remains, supports this idea. For example, *M. laperouse* nests on atolls, volcanically active islands, and islands lacking in geothermal heat sources. Populations of *M. laperouse* residing on different islands use different heat sources: sun, geothermal heat, and a combination of sun and decomposition. Furthermore, individuals of *M. cumingii*, which harness the heat generated by decomposition, nest at multiple sites and vary incubation strategy depending on density of understorey, canopy cover and size of trees. These incubation strategies fall on a continuum from burrows in decaying roots to mounds away from trees (Sinclair et al., 2002). Currently, *M. pritchardii* is found on volcanically active islands and utilizes geothermal heat, but fossils have been found on other islands where this strategy would be impossible (Steadman, 1999). These observations and the results of ancestral reconstruction of incubation strategies suggest that within *Megapodius* nesting behaviour is phenotypically labile, and, along with strong flight, may have aided rapid colonization of the islands of Australasia.

### Extinction

Megapode incubation strategies may make them particularly vulnerable to extinction (Bennett & Owens, 1997): most species have high nest site fidelity and both mounds and burrows are extremely conspicuous. High extinction rate may have had impacts that have not been accounted for in our species tree estimation, biogeographical reconstruction and ancestral state estimation (Barracough & Nee, 2001). The birth–death model we employed assumes that the rate of speciation and extinction is constant through time, but it has been suggested that more than half of megapode species may have recently been extirpated (Steadman, 1999). Species with island distributions are well known to be more susceptible to extinction and since the arrival of humans in the South Pacific, megapodes have experienced range reductions and widespread extinction due to exploitation and habitat destruction (Steadman, 1993, 1999). Currently, nine megapode species are listed as threatened (Table 1) and at least four extinct species are known from bones found at cave sites. For example, no megapode species currently inhabits New Caledonia, but bones of the extinct *Megapodius molistructor* have been found in the caves of early Polynesian invaders, and even as recently as Captain Cook’s voyages, the extinct *Megapodius andersoni* was described as occurring there. New sequencing technologies allowing for the acquisition of DNA from bones may allow future researchers to determine the degree to which extinction has occurred within *Megapodiidae*.

### ACKNOWLEDGEMENTS

We thank the Leaché Lab, Klicka Lab, S. Billerman, N. Sly, S. Rohwer, B. Stelbrink and anonymous referees for comments on this manuscript. We thank A. Talaba for assistance with lab work. We are grateful to the following individuals and institutions for providing tissues: N. Rice (ANSP); P. Sweet, T. Trombone, P. Capainolo (AMNH); R. Palmer, L. Joseph (ANWC); W. Boles, J. Sladek (AM); J. Dumbacher, M. Flannery (CAS); R. Dekker (Leiden Museum); C.W. Moeliker, C.J. Heij and C. Rompas (NMR); L. Christidis and W. Longmire (NMV); P. Uniti (SDNHM); K. Burns (SDSU Museum of Biodiversity); San Diego Zoo; R. Moyle and M. Robbins (KUNHM); R. Johnstone and C. Stevenson (WAM), R. Prum and K. Zyskowski (YPM); C. Milensky (USNM); C. Kessler (USFWS); R. Sinclair, G. Waters, D. Jones, A. Göth and E. Curio. Part of this work was carried out at the Fuller Lab of Evolutionary Biology at the Cornell Lab of Ornithology. Support for this research was provided by the University of Washington, the Burke Museum and the UW Sargent Award.

### REFERENCES


Dekker, R.W.R.J. (2007) Distribution and speciation of
© 2014 John Wiley & Sons Ltd
Journal of Biogeography
41
Hall, R. (2002) Cenozoic geological and plate tectonic evolu-
Hall, R. (1998) The plate tectonics of Cenozoic SE Asia and
Frith, H.J. (1956) Breeding habits of the family Megapodi-
Heled, J. & Drummond, A.J. (2010) Bayesian inference of
Huelsenbeck, J. & Ronquist, F. (2001) MrBayes: Bayesian
Immelmann, K. & Sossinka, R. (1986) Parental behaviour in
A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolu-
loides wyomingensis—a stem group representative of Galli-


SUPPORTING INFORMATION

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

Appendix S1 Specimen IDs, vouchers and abbreviated sampling localities.

Appendix S2 Nuclear and mitochondrial ancient DNA primer information.

Appendix S3 Gene trees and ‘full coverage’ data matrix species tree not included in the manuscript.

BIOSKETCHES

Rebecca B. Harris is interested in the evolution of traits in avian systems and the biological factors influencing species tree inference.

Sharon M. Birks is interested in the evolution of birds, with an emphasis on reproductive behaviour and mating systems.

Adam D. Leaché studies phylogenetics, phylogeography and species delimitation, mostly in lizards.

Editor: Michael Patten