PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The spatio-temporal colonization and diversification across the Indo-Pacific by a 'great speciator' (Aves, *Erythropitta erythrogaster*)

Martin Irestedt, Pierre-Henri Fabre, Henrique Batalha-Filho, Knud A. Jønsson, Cees S. Roselaar, George Sangster and Per G. P. Ericson

Proc. R. Soc. B 2013 **280**, 20130309, published 3 April 2013

PROCEEDINGS OF THE ROYAL SOCIETY B

rspb.royalsocietypublishing.org

Research

Cite this article: Irestedt M, Fabre P-H, Batalha-Filho H, Jønsson KA, Roselaar CS, Sangster G, Ericson PGP. 2013 The spatio-temporal colonization and diversification across the Indo-Pacific by a 'great speciator' (Aves, *Erythropitta erythrogaster*). Proc R Soc B 280: 20130309.
http://dx.doi.org/10.1098/rspb.2013.0309

Received: 12 February 2013
Accepted: 8 March 2013

Subject Areas:
evolution, taxonomy and systematics, ecology

Keywords:
island biogeography, integrative taxonomy, speciation, phylogeny, Pleistocene climate changes

Author for correspondence:
Martin Irestedt
e-mail: martin.irestedt@nrm.se

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2013.0309 or via http://rspb.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

The spatio-temporal colonization and diversification across the Indo-Pacific by a 'great speciator' (Aves, *Erythropitta erythrogaster*)

Martin Irestedt[1], Pierre-Henri Fabre[3], Henrique Batalha-Filho[4], Knud A. Jønsson[3,5], Cees S. Roselaar[6], George Sangster[2,7] and Per G. P. Ericson[2]

[1]Department of Biodiversity Informatics and Genetics, and [2]Department of Vertebrate Zoology, Swedish Museum of Natural History, PO Box 50007, 10405 Stockholm, Sweden
[3]Center for Macroecology, Evolution and Climate at the Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[4]Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil
[5]Department of Life Sciences, Imperial College London, Silwood Park campus, Ascot SL5 7PY, UK
[6]Naturalis Biodiversity Center, Darwinweg 2, PO Box 9517, 2300 RA Leiden, The Netherlands
[7]Department of Zoology, Stockholm University, 10691 Stockholm, Sweden

The Indo-Pacific region has arguably been the most important area for the formulation of theories about biogeography and speciation, but modern studies of the tempo, mode and magnitude of diversification across this region are scarce. We study the biogeographic history and characterize levels of diversification in the wide-ranging passerine bird *Erythropitta erythrogaster* using molecular, phylogeographic and population genetics methods, as well as morphometric and plumage analyses. Our results suggest that *E. erythrogaster* colonized the Indo-Pacific during the Pleistocene in an eastward direction following a stepping stone pathway, and that sea-level fluctuations during the Pleistocene may have promoted gene flow only locally. A molecular species delimitation test suggests that several allopatric island populations of *E. erythrogaster* may be regarded as species. Most of these putative new species are further characterized by diagnostic differences in plumage. Our study reconfirms the *E. erythrogaster* complex as a 'great speciator': it represents a complex of up to 17 allopatrically distributed, reciprocally monophyletic and/or morphologically diagnosable species that originated during the Pleistocene. Our results support the view that observed latitudinal gradients of genetic divergence among avian sister species may have been affected by incomplete knowledge of taxonomic limits in tropical bird species.

1. Introduction

The Indo-Pacific archipelagos have played a fundamental role in the formulation of modern biogeographic and speciation theories. Observations of the complex distributions of Asian and Australian faunas in the centre of the region (present-day Wallacea) inspired Alfred Russell Wallace to develop his ideas on the connection between geography and animal distributions [1,2]. Further east in Melanesia, patterns of geographical variation in birds formed the basis of Ernst Mayr's highly influential allopatric model of speciation [3].

A dynamic geological history [4] and Pleistocene sea-level fluctuations [5] have generated complex patterns of spatio-temporal vicariance events and dispersal routes [6]. Today, the Indo-Pacific archipelagos comprise more than 20 000 islands ranging from small atolls to large, geologically complex, tropical islands, such as New Guinea, Sulawesi and Borneo. Several biodiversity hotspots have been recognized within the Indo-Pacific region [7], and the largest island (New Guinea) is estimated to host the third-largest 'high biodiversity' wilderness in the world, only surpassed by the Amazon and Congo forests [8].

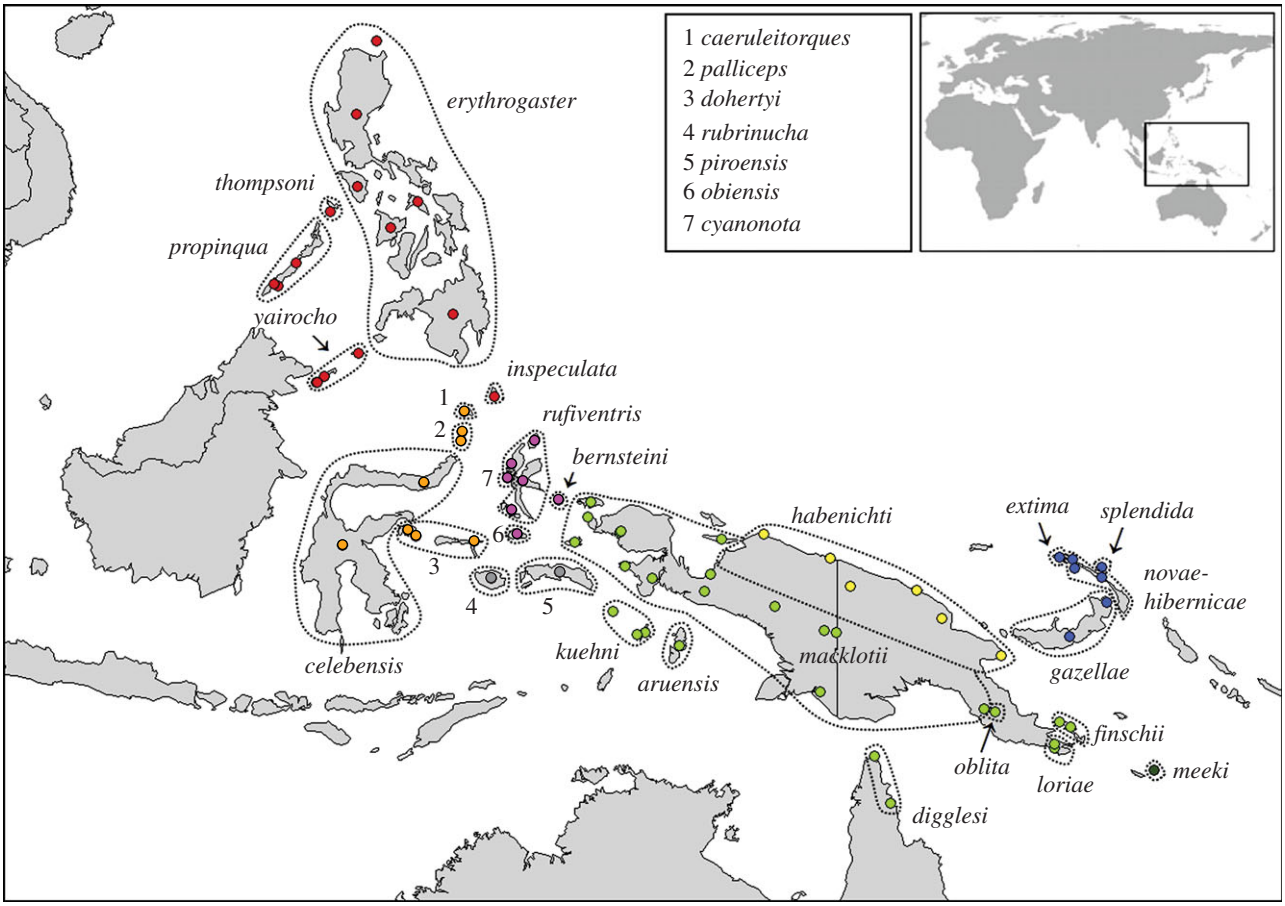

Figure 1. Range map of the *Erythropitta erythrogaster* complex, showing the ranges of all 28 taxa mentioned in the text. Sampling localities for the molecular analyses are indicated by coloured circles. Colours represent the eight major phylogroups identified in the phylogenetic analyses (figure 2).

Many islands in the region are truly oceanic. Consequently, the biota must have arrived by over-water dispersal and the importance of dispersal has been demonstrated for the distributions of many faunal groups in the region [6]. Pleistocene climate oscillations have also influenced the population structure of many organisms [9,10], and this influence is assumed to be particularly strong in the shallow parts of the Indo-Pacific, where lowering of sea levels connected islands that are currently separated by water barriers [5].

Despite the rich biodiversity of the Indo-Pacific and the fact that this region has played a major role in the development of modern biogeography, most phylogeographic analyses of single species have been restricted to subregions within the Indo-Pacific archipelagos, such as the Philippines, the Sunda Shelf islands, Australia and New Guinea [6]. Other studies have employed species-level phylogenies to investigate biogeographic patterns of birds across larger parts of the region [11,12]. However, studies based on complete geographical taxon sampling, allowing for more in-depth studies of colonization and diversification patterns, are still scarce [13]. Consequently, the understanding of Pleistocene dispersal and diversification in the Indo-Pacific is still largely based on taxonomic data based on morphological variation [14].

The red-bellied pitta *Erythropitta erythrogaster* has a wide distribution, spanning large parts of the Indo-Pacific (figure 1). It is a medium-sized passerine bird that inhabits lowland rainforests and monsoon forests, and most populations are believed to be strictly sedentary [15,16]. Many populations confined to specific archipelagos or islands are morphologically distinct, and up to 28 subspecies are currently recognized [15–20]. Several other widespread superspecies are known from the Indo-Pacific region [14], and *E. erythrogaster* is therefore ideal for studying avian dispersal and diversification across a large insular region. The species is regarded as one of the 'great speciators' [14], but species limits have never been addressed using modern approaches [21,22].

Here, we examine the population structure of *E. erythrogaster* based on molecular and morphological data within a spatio-temporal framework. We test alternative scenarios of the tempo and mode of colonization, and characterize the level of diversification among archipelagos and island populations. Ultimately, we aim to test (i) how external factors such as Pleistocene climate oscillation have influenced dispersal rates over time, and (ii) if dispersal followed a stepping stone pathway, as predicted by the theory of island biogeography [23]. Finally, we define the genetic and morphological variation of *E. erythrogaster* in a geographical context, and discuss the taxonomic status of distinct populations.

2. Material and methods

(a) Datasets

For the molecular analyses, we aimed for dense geographical coverage and sampled 139 individuals of *E. erythrogaster* (figure 1), including all 28 described subspecies [15–19]. Two closely related species, *E. kochi* and *E. venusta*, were also included to test the monophyly of the *E. erythrogaster* complex and to root the trees. For all samples, two mitochondrial regions were sequenced: the complete NADH dehydrogenase subunit 3 (ND3) and partial cytochrome *b* (*cyt-b*). For most samples, we also obtained partial sequences from two nuclear introns: Myoglobin intron 2 (MYO2) and CSDE1 (Cold shock domain containing E1, RNA-binding) intron 5.

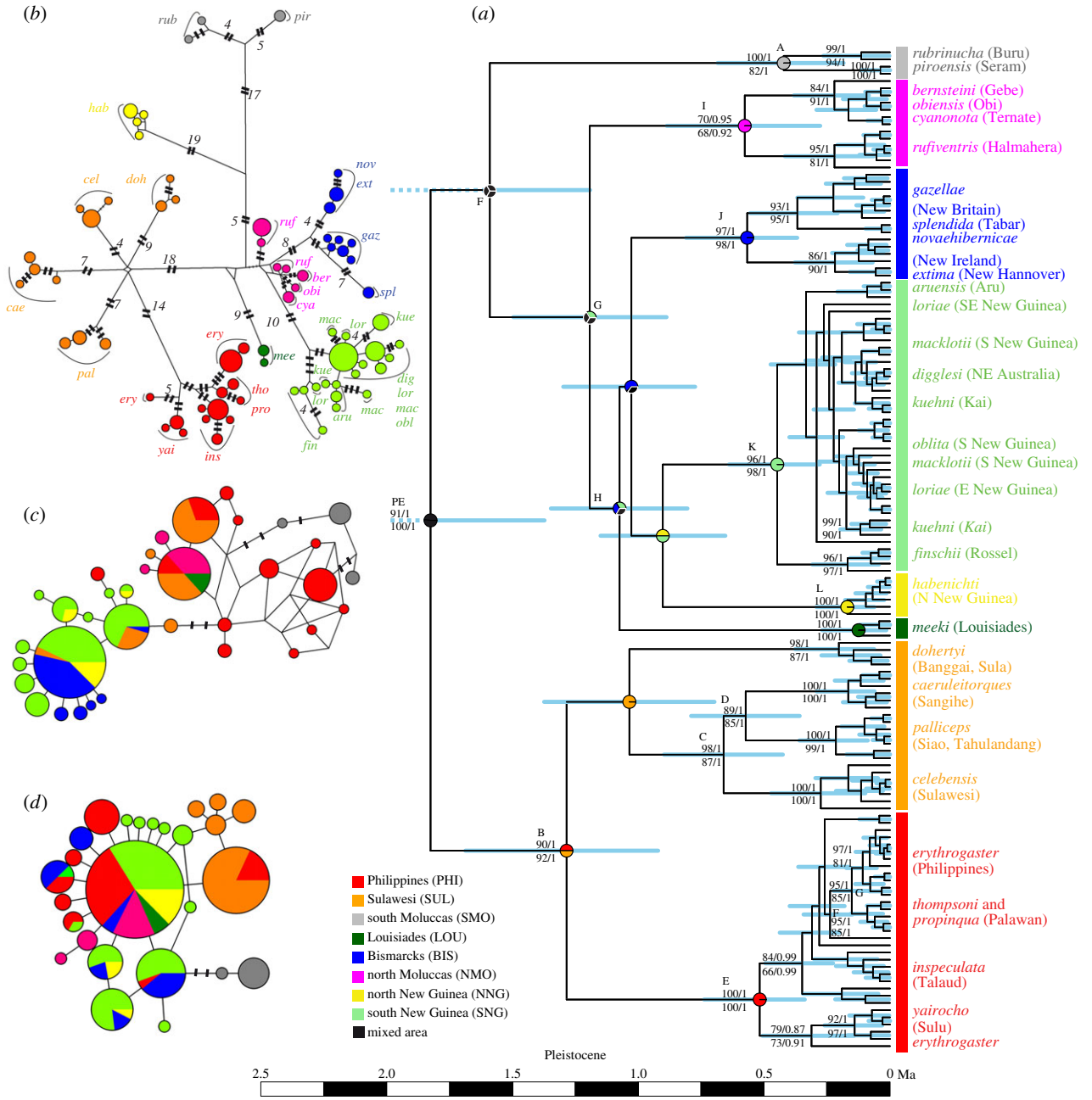

Figure 2. Estimates of phylogenetic relationships, divergence times and ancestral area reconstructions of *E. erythrogaster* complex. (*a*) Lagrange ancestral area reconstructions. The tree is a chronogram (strict molecular clock) based on a BEAST Markov chain Monte Carlo (MCMC) analysis of the combined dataset. Coloured pies indicate the origin of a given node based on eight pre-assigned geographical areas for the LAGRANGE analysis: Philippines (PHI, red), Sulawesi (SUL, orange), south Moluccas (SMO, grey), north Moluccas (NMO, magenta), Bismarcks (BIS, blue), north New Guinea (NNG, yellow), south New Guinea and north Australia (SNG, pale green), Louisiades (LOU, dark green). Black pies indicate mixed origin. Nodal support values indicated above the branches correspond to bootstrap support and posterior probability, respectively, from the analyses of the mitochondrial dataset, while nodal support values indicated below the branches correspond to bootstrap support and posterior probability obtained from the analyses of the concatenated dataset. (*b–d*) Median-joining haplotype networks of (*b*) mitochondrial, (*c*) CSDE and (*d*) MYO2 genes. Each circle corresponds to one haplotype and its size is proportional to its frequency; each line connecting the haplotypes refers to a mutational step; marks in the lines are indicated when more than one step was found. When more than three mutational steps were found, they are indicated by a number corresponding to the number of steps. Colours correspond to eight major geographical regions given in the legend. Names at the mitochondrial network correspond to subspecies of *E. erythrogaster* represented by their first three letters.

We also analysed an extensive morphometric dataset from 440 study skin specimens (143 females and 297 males) compiled by Johannes Erritzoe [15], and examined the plumage variation of 210 study skins.

(b) Phylogenetic analyses, molecular dating and ancestral area reconstruction

We used maximum-likelihood and Bayesian inference to estimate the phylogenetic relationships in the *E. erythrogaster* complex. The nucleotide substitution models were selected for each gene individually, and supported GTR + Γ for cytochrome *b*, HKY + I + Γ for ND3, HKY + I for myoglobin intron 2 and GTR + I + Γ for CSDE1 intron 5. We applied the models above for each partition for both a mitochondrial and a concatenated dataset in the Bayesian analysis (10 million generations) using MrBayes v. 3.1.2 [24] and the GTRMIX option in the maximum-likelihood analyses (1000 tree search replicates and non-parametric bootstrapping to estimate nodal support) using RAxML [25]. The burn-in and convergence diagnostics of

MrBayes analysis were estimated using the potential scale reduction factor [26] and Tracer [27].

We analysed the mitochondrial dataset using BEAST v. 1.7.3 [28] to estimate divergence times within *E. erythrogaster*, applying the models selected above and assuming a Yule speciation process for the tree prior. Because a clock-like evolution of the mitochondrial dataset could not be rejected, we applied both a strict and a relaxed uncorrelated lognormal molecular clock to our data partitioning. Because no fossils are available for this group, we used a rate of evolution of 1.05 per cent (± 0.35) sequence divergence per lineage per million years for cyt-*b* to obtain absolute dates [29]. The program Tracer [30] was used to assess convergence diagnostics.

We used Lagrange [31,32] to study the direction and sequence of colonization of the *E. erythrogaster* complex across the Indo-Pacific. We assigned eight geographical areas for the Lagrange analysis after considering geological events [4]: Philippines (PHI), Sulawesi (SUL), south Moluccas (SMO), north Moluccas (NMO), Bismarcks (BIS), north New Guinea (NNG), south New Guinea and north Australia (SNG), Louisiades (LOU). Analyses in Bayes–Lagrange [32] were run on 1000 trees (stochastically sampled and disregarding the burin), from the posterior distribution of the BEAST analysis, to estimate the frequency of probability of ancestral areas for each clade.

(c) Analyses of population structure in the molecular and morphometric datasets

We used multivariate statistics to analyse the population structure. To further complement the phylogenetic inferences, we obtained median-joining networks [33] for each gene to study the relationships between haplotypes and their geographical distribution. The level of population genetic structure [34] was calculated using Arlequin v. 3.5.1.2 [35], while a Mantel test was conducted to test for correlation between genetic and geographical distances [36].

A principal component analysis (PCA) was conducted from the covariance matrix of the log10-transformed variables from the morphometric dataset. We tested for significant mean differences among eight major phylogroups as well as for 17 species suggested by our integrative taxonomic approach (see below). ANOVA was employed for each morphometric variable and for the first four principal component axes. MANOVA analyses were performed on the principal component scores to assess the effects of different phylogenetic and biogeographic factors.

(d) Molecular species delimitation and diversification rate analyses

In an attempt to identify populations of *E. erythrogaster* that may merit species rank, we employed a sequence-based species delimitation test developed by Pons *et al.* [22], applying a general mixed Yule coalescent model. In addition, uncorrected intraspecific and interspecific distances in cytochrome *b* were calculated for the eight pre-assigned geographical areas and for the putative species detected in the molecular species delimitation test.

We used R [37], APE [38] and the LASER [27] packages to calculate the diversification rate. The analyses were computed on the maximum clade credibility tree as well as on 1000 phylogenies randomly sampled from the posterior distributions of trees (disregarding the burn-in). We tested for constant diversification rates over time using both the γ statistic [39] and the maximum-likelihood method of the ΔAIC_{RC} [27] by using a null distribution from 5000 simulated topologies of the same sample size ($n = 16$, based on the results from the molecular species delimitation test, see above and results below) under a Yule model. Testing for undetected or extinct species, we used three further levels of total species richness representing 75, 50 and 25 per cent of the true species number.

Using R [37] and the Geiger [40] and apTreeshape packages [41], we implemented two approaches to identify nodes that display significant shifts in diversification rates: the $\Delta 1$ statistic [42,43] and the relative cladogenesis test [44].

Detailed descriptions of datasets, laboratory procedures and analyses, as well as voucher and GenBank numbers, are given in the electronic supplementary material.

3. Results

(a) Phylogeneny, biogeography and population structure

The phylogenetic results, divergence time estimates, biogeographic results and the haplotype network are summarized in figure 2.

The spatio-temporal analyses of *E. erythrogaster* suggest that diversification took place in the Pleistocene, with the majority of regional diversification taking place within the last 400 000 years, and that colonization of the Indo-Pacific mainly had an easterly directionality. Our analyses also detect strong geographical structure in both the genetic and morphometric data. The multivariate analyses of morphometric data show that populations in eight pre-defined geographical areas are significantly separated. The PCA (figure 3) indicates that geographically closely located populations are generally morphometrically more similar than geographically distant populations. Most of the species identified by the integrative taxonomic approach (see below) were significantly differentiated in at least one of four principal components (see the electronic supplementary material). Our re-evaluation of the plumage variation within *E. erythrogaster* finds that 23 of the 28 described taxa show a unique combination of plumage characters (i.e. variation in plumage is not clinal across the range of the *E. erythrogaster* complex), because populations on adjacent archipelagos invariably differ by one or more diagnostic differences, and each archipelago also include multiple diagnosable taxa (see plumage diagnosis in the electronic supplementary material).

The genetic structuring of populations is largely congruent with the morphometric structuring (described above) in *E. erythrogaster*. Most of the genetic variation is explained by the variation among the pre-assigned geographical regions (table 1). Within geographical regions, we also found a strong genetic structure in the mitochondrial haplotype network (figure 2*b*), with genetically unique populations on many peripheral islands within larger island groups. The AMOVA analyses of the nuclear datasets (table 1) also recovered some geographical structure, and the frequencies of unique haplotypes showed differentiation among geographical regions (figure 2*c*,*d*). However, many populations clearly separated by the mitochondrial dataset were not recovered by the nuclear dataset. In addition, pairwise comparisons of Φ_{ST} between geographical regions also showed high genetic differences between these groups in both mitochondrial and nuclear datasets, and the neutrality and the population size change tests did not find any evidence of demographic expansion for most lineages, except for south New Guinea and south Bismarcks (see tables S4 and S5 in the electronic supplementary material for further details).

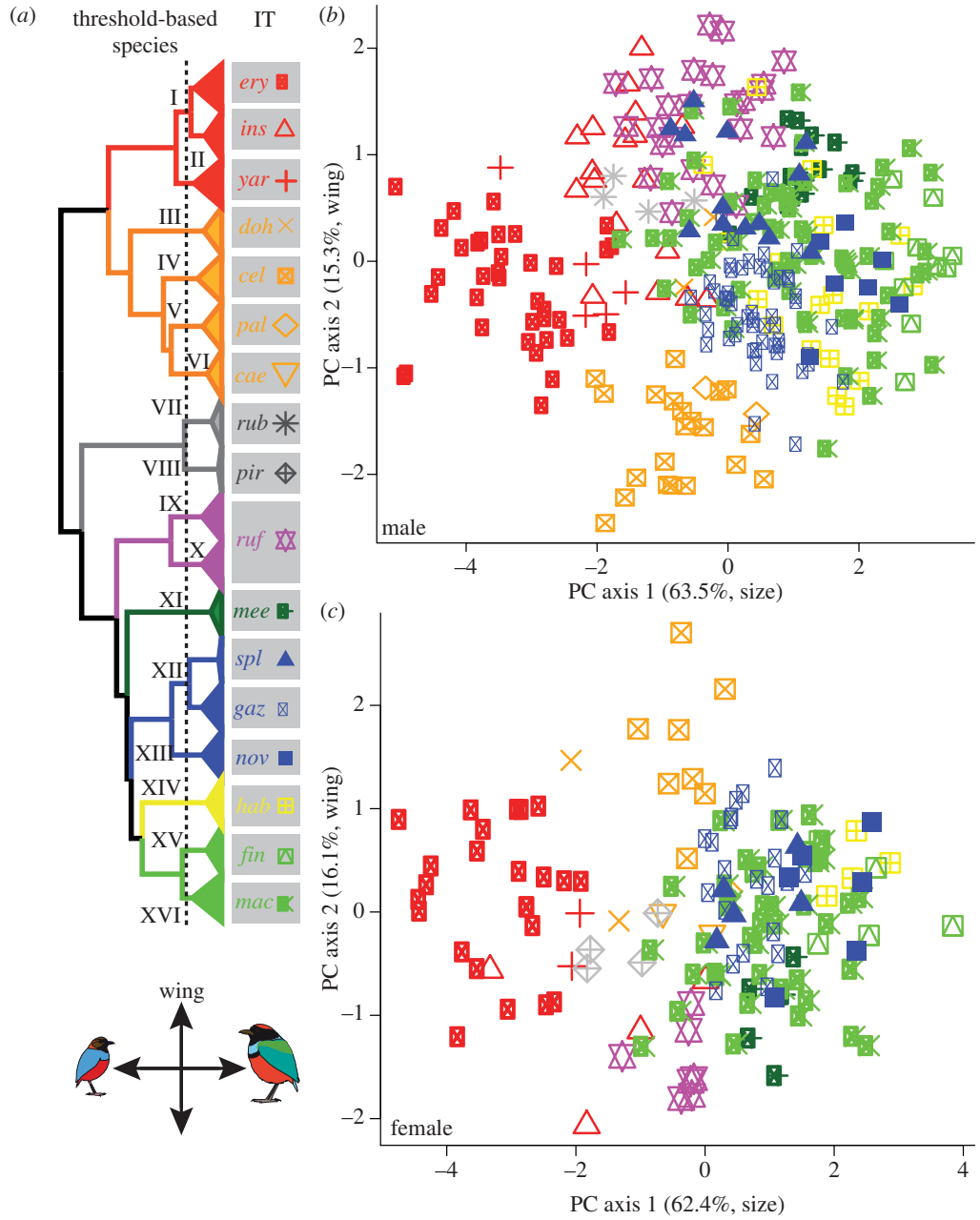

Figure 3. Results from principal components analysis and molecular species delimitation test. (*a*) The ultrametric tree obtained from the BEAST analysis of the *E. erythrogaster* complex, indicating cluster of specimens as putative species based on the results from the molecular species delimitation test [22]. Genetic clusters recognized as putative species are indicated on the right-hand side of the topology. The vertical bars group all sequences within each significant clusters, labelled I to XVIII. (*b–c*) Plot of principal components 1 (PC1) and 2 (PC2), based on covariance matrix of the log10-transformed variables from (*b*) males ($n = 297$) and (*c*) females ($n = 143$). Colours refer to geographical origin and symbols refer to lineages identified by the molecular species delimitation test.

(b) Molecular species delimitation and diversification rate analyses

The molecular species delimitation test suggested 16 putative species (figure 3). The uncorrected, intraspecific and interspecific genetic distances in cytochrome *b* for these potential species as well as the eight pre-assigned geographical areas showed few cases where intraspecific and interspecific distances overlap (one for the 8 pre-assigned geographical areas and three for the 16 putative species).

The macroevolutionary analyses did not recover any shifts in diversification rates within the *E. erythrogaster* complex (figure 4*a*), while the Mantel test showed a correlation between genetic and geographical distances, which provides evidence for isolation by distance (figure 4*b*). More detailed information of the results from the phylogenetic, biogeographic, population, diversification rate, Mantel test and species delimitation analyses can be found in the electronic supplementary material.

4. Discussion

(a) Stepping stone colonization and the influence of Pleistocene climate oscillation

The phylogenetic analyses and the corresponding biogeographic results (figure 2*a*) support that *E. erythrogaster* started to colonize the Indo-Pacific during the Pleistocene, from Asia (probably from the Sunda Islands or the Philippines, given the distribution of closely related species). Further dispersal via the Moluccas and New Guinea to the

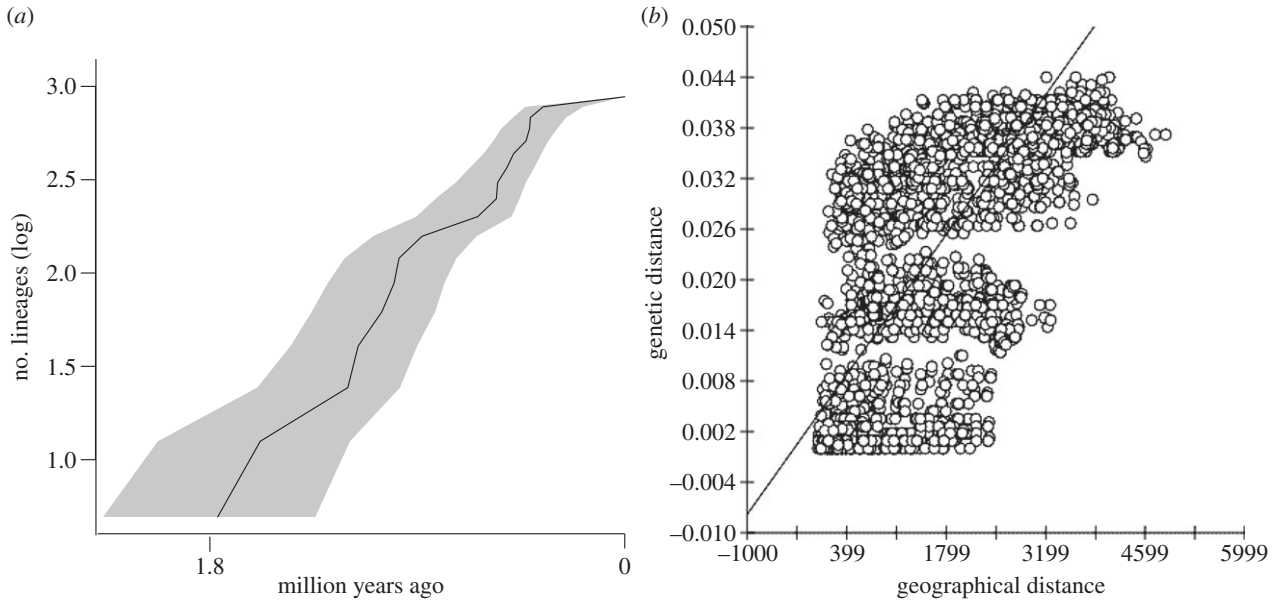

Figure 4. (*a*) Divergence rate plot. (*b*) Relationship between geographical distance (in kilometres) and genetic (uncorrected p) distance in the *E. erythrogaster* complex. The line represents the fitted linear regression. The Mantel test revealed a significant correlation ($r = 0.52$; $p < 0.0001$).

Table 1. Analyses of molecular variance (AMOVA) in *E. erythrogaster* for (*a*) mitochondrial, (*b*) CSDE and (*c*) MYO2 genes. DF, degrees of freedom; SQ, sum of squares; VC, variance components; PV, percentage of variance. All components of variance were significant at $p < 0.05$.

gene	source of variation	DF	SQ	VC	PV
(*a*)	among geographical regions	7	1523.639	9.42472	56.12
	among phylogroups within geographical regions	7	211.801	5.66606	33.74
	among individuals within phylogroups	124	211.013	1.70172	10.13
	total	138	1946.453	16.79251	
(*b*)	among geographical regions	7	512.542	1.94209	44.13
	among phylogroups within geographical regions	7	94.02	1.54952	35.21
	among individuals within phylogroups	191	173.666	0.90925	20.66
	total	205	780.228	4.40085	
(*c*)	among geographical regions	7	72.835	0.28043	30.7
	among phylogroups within geographical regions	7	12.775	0.19819	21.7
	among individuals within phylogroups	189	82.178	0.43481	47.6
	total	203	167.789	0.91343	

Bismarck archipelago and subsequent diversification took place largely within the second half of Pleistocene. Despite its contemporary sedentary nature, *E. erythrogaster* has dispersed across open expanses of water multiple times, leading to its present distribution on many isolated islands, which were never connected with other landmasses during Pleistocene sea-level fluctuations [6].

The theory of island biogeography assumes a simple model where the proximity of islands is the most important factor (size being the second most important factor) for successful dispersal and colonization [23]. According to this theory, one would expect to find a stepping stone pattern of colonization for taxa with wide distributions in oceanic archipelagoes. However, a number of empirical studies have recently challenged this simple model by revealing complex patterns of island colonizations [13,45,46]. Signatures of true stepping stone colonization may, as time passes, be obscured by extinctions due to erratic catastrophic events or because islands have been submerged during times of changing sea levels. It is also possible that external factors might promote a more rapid (or even the opposite: hinder) colonization of new islands, such as low sea levels in the Pleistocene, which have been found to be correlated with increased rates of avian dispersal in certain areas of the Indo-Pacific [6,47].

Because of the relatively young age of the *E. erythrogaster* complex, it is less likely that past extinctions have had a significant impact on the present phylogeographic structure, an assumption supported by the diversity-dependent diversification rate found for the complex (figure 4*a*). This result also rejects the hypothesis that Pleistocene climate changes or other external factors have had major effects on the dispersal rate of *E. erythrogaster*, in which case we would expect shifts in diversification rates associated with sea-level low stands. The haplotype networks, the phylogeny and the

ancestral area reconstruction (figure 2) overall support a successive colonization of the Indo-Pacific with a clear positive correlation between geographical distance between populations and relatedness. In addition, the Mantel test result (figure 4*b*) strongly supported the isolation by distance hypothesis, which also fits a model of a successive colonization of the Indo-Pacific by *E. erythrogaster*. Overall, we found good evidence for stepping stone colonization of the Indo-Pacific archipelagos by *E. erythrogaster*, in line with the general island biogeographic patterns predicted by MacArthur & Wilson [23].

For the populations in south New Guinea, north Australia and New Britain, we found significant support for population expansions. In this shallow part of the Indo-Pacific, Pleistocene land connections at times of low sea levels promoted dispersal between Australia and New Guinea for many taxonomic groups [47–50]. It is possible that the population structure of *E. erythrogaster* was locally affected [5,51] by Pleistocene sea-level fluctuations, but it cannot be ruled out that this pattern is the consequence of more recent dispersal, because *E. e. digglesi* in north Australia is the only taxon within *E. erythrogaster* that migrates across water [18].

Although most species within the family Pittidae are sedentary, some species within the genus *Pitta* are migratory [15,16]. It is thus possible that migration has enabled colonization by long-distance dispersal, for example, into the Indo-Pacific where several species of *Pitta* are resident. However, long-distance dispersal seems not to have been a major force in the colonization of the Indo-Pacific by *E. erythrogaster*, given the strong stepping stone signature found in this study.

Changes in habitat distributions as response to the climate change during the Pleistocene are also assumed to have affected population structuring of many organisms [9,10]. However, most of the islands inhabited by *E. erythrogaster* are larger than 1000 km^2 and include montane forest, probably leaving sufficient suitable habitat pockets for *E. erythrogaster* even during Pleistocene climate changes with associated shifts in annual temperatures. The deep seas surrounding most of the islands have also prevented significant changes in the distance between islands during times of sea-level fluctuations. We suggest that the combination of these factors and the relatively recent dispersal have been important factors in shaping and maintaining the stepping stone signature of island colonization in *E. erythrogaster*.

(b) Integrative taxonomy of a 'great speciator'

Several species concepts have been used to support taxonomic decisions of species delimitation, and the relative merits of these concepts have been debated vigorously [52,53]. Modern approaches emphasize that species are populations that extend through time (i.e. are lineages) and that most species 'concepts' merely provide alternative operational criteria for delimiting species in practice [54,55]. According to this view, species taxa are hypotheses of evolutionary lineages, which can be identified in multiple ways [21,54,56]. Consequently, species taxa are best documented using as many lines of evidence as possible. This evidence must be integrated to identify congruence and conflicts among datasets [21]. We used an integrative approach to identify species in the *E. erythrogaster* complex. Using a molecular species delimitation test [22], we first identified 16 lineages that may correspond to species taxa (figure 3).

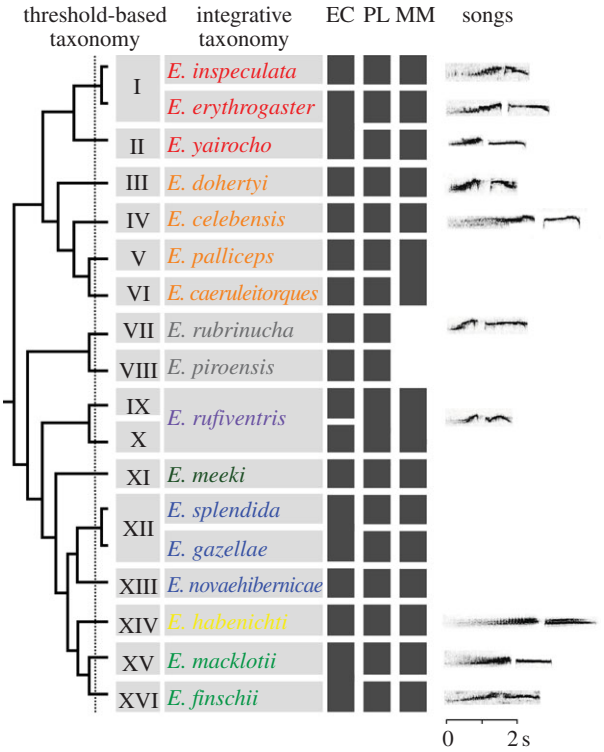

Figure 5. Comparison of threshold-based and integrative taxonomy of the *E. erythrogaster* complex. From left to right: the tree topology based on the BEAST analysis, species identified by the threshold-based approach, our proposed revision based on integrative taxonomy, groups characterized by exclusive coalescence in the ML analysis of the multi-locus dataset (EC), groups diagnosable on the basis of plumage differences (PL), groups differing statistically in morphometrics (MM) and sonagrams of songs. For two taxa, insufficient morphometric data were available for meaningful analyses.

Most of these taxa are distributed allopatrically, were reciprocally monophyletic in our ML phylogeny (see the electronic supplementary material, figures S1 and S2) and can be diagnosed by one or more plumage characters (figure 5; electronic supplementary material). Several of these taxa also differ in morphometrics (figure 5; electronic supplementary material). This suggests that these taxa have remained isolated for a considerable time, are on independent evolutionary trajectories and are best treated as separate species.

There are a few cases in which species delimitations based on geographical, phylogenetic or plumage data are in conflict (figure 5), and for which species status is debatable. However, based on all available evidence, we propose that the *E. erythrogaster* complex is composed of 17 species (figure 5; see also the electronic supplementary material for evaluation of species status in cases of conflicting data): (i) Philippine pitta *E. erythrogaster*, with subspecies *erythrogaster* and *propinqua* (the latter including '*thompsoni*'); (ii) Talaud pitta *E. inspeculata*; (iii) Sulu pitta *E. yairocho*; (iv) Sula pitta *E. dohertyi*; (v) Siao pitta *E. palliceps*; (vi) Sangihe pitta *E. caeruleitorques*; (vii) Sulawesi pitta *E. celebensis*; (viii) Buru pitta *E. rubrinucha*; (ix) Seram pitta *E. piroensis*; (x) Louisiade pitta *E. meeki*; (xi) Moluccan pitta *E. rufiventris*, with subspecies *rufiventris* (including '*obiensis*'), *cyanonota* and *bernsteini*; (xii) New Ireland pitta *E. novaehibernicae*, with subspecies *extima* and *novaehibernicae*); (xiii) New Britain pitta *E. gazellae*; (xiv) Tabar pitta *E. splendida*; (xv) Habenicht's pitta *E. habenichti*; (xvi) D'Entrecasteaux pitta *E. finschii*; and (xvii) Papuan pitta *E. macklotii*, with subspecies *macklotii* (including '*aruensis*' and '*kuehni*'), *loriae* (including '*oblita*') and *digglesi*.

The large number of distinct species, in combination with their relatively recent (Pleistocene) divergence, suggests that the *E. erythrogaster* complex is indeed a 'great speciator' [14]. To our knowledge, the *E. erythrogaster* complex includes one of the largest number of distinctive species previously grouped in a single 'polytypic' species of bird [19,20]. However, it is likely that similar cases may be revealed when other polytypic species in the region are thoroughly examined. Whether the species proposed herein are reproductively isolated is speculative, and requires insight into the behavioural, ecological and genetic mechanisms that prevent closely related parapatric and sympatric species of pittas from fusing into a single population. Such knowledge is currently lacking. If differences in plumage and vocalizations provide cues for species recognition, then most of the species that we recognize are likely to be reproductively isolated. The differences in plumage and songs among taxa in the *E. erythrogaster* complex are comparable with those observed among closely related species in the *E. granatina* and *Hydrornis guajana* complexes (see the electronic supplementary material), both of which have been revised recently [55,57].

Our study shows that molecular species delimitation methods [22] may provide a starting point for integrative taxonomic studies. However, our study also shows that such methods may be insufficient to accurately identify all species. The molecular species delimitation analysis did not identify *E. inspeculata* and *E. splendida* as species because their divergence from *E. erythrogaster* and *E. gazellae*, respectively, failed to exceed the inferred threshold level. In contrast, the integrative approach found both pairs to be distinct in plumage and morphometrics, and further documented reciprocal monophyly of *inspeculata* and *erythrogaster*. Furthermore, two sympatric lineages within *E. rufiventris* were identified as putative species using the method of Pons *et al.* [22], but in the light of other evidence, we consider it unlikely that these represent species. Thus, we recommend that the outcomes of molecular species delimitation methods are integrated with, and evaluated in combination with, other data to provide a comprehensive account of species-level diversity.

(c) Taxonomic limits and latitudinal gradient of genetic divergence

The recognition of 17 species in what was believed to be a single widespread species underscores that the task of the correct discrimination of species taxa in birds is not 'virtually completed', as proposed by Mayr [58] (see also [59]). Importantly, revisions of tropical species of birds may help to remove biases from macroecological studies resulting from incorrect taxonomies. For example, it has been suggested that genetic divergence among sister species is higher in tropical birds than in temperate regions. This in turn has been interpreted as evidence for higher rates of speciation and extinction in temperate regions than in the tropics [60]. However, such a pattern may simply represent an artefact caused by incomplete studies of species limits in tropical groups [61] and by a bias towards the tropics in the downgrading of distinctive taxa to subspecies rank during the first half of the twentieth century [62]. Our results support the view that there may be a taxonomic bias against recent speciation in tropical birds owing to incorrect taxonomy. We found that sister species in the *E. erythrogaster* complex typically differ by 1.1 per cent or less in cytochrome *b* (see the electronic supplementary material). These levels are in line with those found in temperate sister species, but are lower than those typically observed among tropical species [63–66]. Thus, latitudinal gradients of genetic divergence among sister species probably become flatter when taxonomic limits of tropical bird species become better known.

We are grateful to Johannes Erritzoe for kindly sharing his morphometric data, and to two anonymous reviewers for valuable comments on this paper. We thank the following institutions for providing tissue samples: American Museum of Natural History, New York, NY (Paul Sweet); Museum of Comparative Zoology, Cambridge, MA (Jeremiah Trimble); Naturalis Biodiversity Center, Leiden, the Netherlands (Steven van der Mije); Senckenberg Naturhistorische Sammlungen Dresden and Hessian State Museum Darmstadt, Germany (Martin Päckert); Zoological Museum, University of Copenhagen, Denmark (Jan Bolding Kristensen); Delaware Museum of Natural History, Wilmington, DE (Jean Woods); Museum für Naturkunde, Humboldt University Berlin, Germany (Sylke Frahnert); Australian Museum, Sydney, Australia (Walter Boles); Swedish Museum of Natural History, Stockholm, Sweden (Ulf Johansson); and Peabody Museum of Natural History, Yale University, New Haven, CT (Kristof Zyskowski). P.-H.F. and K.A.J. acknowledge the Danish National Research Foundation for funding to the Center for Macroecology, Evolution and Climate. K.A.J. acknowledges support from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. PIEF-GA-2011-300924. This work was supported by the Swedish Research Council (grant no. 621-2010-5321 to P.G.P.E.).

References

1. Wallace AR. 1869 *The Malay Archipelago*. London, UK: Macmillan.
2. Wallace AR. 1876 *The geographical distribution of animals*. New York, NY: Harper and Brothers.
3. Mayr E. 1942 *Systematics and the origin of species*. New York, NY: Columbia University Press.
4. Hall R. 2011 Australia–SE Asia collision: plate tectonics and crustal flow. In *The SE Asian gateway: history and tectonics of Australia–Asia Collision* (eds R Hall, MA Cottam, MEJ Wilson), pp. 75–109. London, UK: The Geological Society.
5. Voris HK. 2000 Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* **27**, 1153–1167. (doi:10.1046/j.1365-2699.2000.00489.x)
6. Lohman DJ, de Bruyn M, Page T, von Rintelen K, Hall R, Ng PKL, Shih HT, Carvalho GR, von Rintelen T. 2011 Biogeography of the Indo-Australian Archipelago. *Annu. Rev. Ecol. Evol. Syst.* **42**, 205–226. (doi:10.1146/annurev-ecolsys-102710-145001)
7. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858. (doi:10.1038/35002501)
8. Mittermeier RA, Mittermeier CG, Brooks TM, Pilgrim JD, Konstant WR, da Fonseca GAB, Kormos C. 2003 Wilderness and biodiversity conservation. *Proc. Natl Acad. Sci. USA* **100**, 10 309–10 313. (doi:10.1073/pnas.1732458100)
9. Avise JC. 2000 *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.
10. Cannon CH, Morley RJ, Bush AB. 2009 The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proc. Natl Acad. Sci. USA* **106**, 11 188–11 193. (doi:10.1073/pnas.0809865106)

11. Jønsson KA, Fabre PH, Ricklefs RE, Fjeldså J. 2011 Major global radiation of corvoid birds originated in the proto-Papuan archipelago. *Proc. Natl Acad. Sci. USA* **108**, 2328–2333. (doi:10.1073/pnas.1018956108)
12. Jønsson KA, Irestedt M, Fuchs J, Ericson PGP, Christidis L, Bowie RCK, Norman JA, Pasquet E, Fjeldså J. 2008 Explosive avian radiations and multi-directional dispersal across Wallacea: evidence from the Campephagidae and other Crown Corvida (Aves). *Mol. Phylogenet. Evol.* **47**, 221–236. (doi:10.1016/j.ympev.2008.01.017)
13. Cibois A, Beadell JS, Graves GR, Pasquet E, Slikas B, Sonsthagen SA, Thibault JC, Fleischer RC. 2011 Charting the course of reed-warblers across the Pacific islands. *J. Biogeogr.* **38**, 1963–1975. (doi:10.1111/j.1365-2699.2011.02542.x)
14. Mayr E, Diamond J. 2001 *The birds of Northern Melanesia: speciation, ecology, and biogeography*. New York, NY: Oxford University Press.
15. Erritzoe J, Erritzoe HB. 1998 *Pittas of the world: a monograph on the pitta family*. Cambridge, UK: Lutterworth Press.
16. Lambert F, Woodcock M. 1996 *Pittas, broadbills and asities*. Robertsbridge, UK: Pica Press.
17. Hachisuka M. 1935 *The birds of the Philippine Island: with notes on the mammal fauna. Accipitriformes to Passeriformes* (*Timaliidæ*). London, UK: Witherby.
18. Higgins PJ, Peter JM, Steele WK. 2001 *Handbook of Australian, New Zealand and Antarctic birds: volume 5: tyrant-flycatchers to chats*. Melbourne, Australia: Oxford University Press.
19. Dickinson EC. 2003 *The Howard and Moore complete checklist of the birds of the world*. London, UK: Christopher Helm.
20. Mayr E. 1979 Family Pittidae. In *Check-list of birds of the world* (ed. MA Traylor), pp. 310–329. Cambridge, MA: Museum of Comparative Zoology.
21. Padial JM, Miralles A, De la Riva I, Vences M. 2010 The integrative future of taxonomy. *Front. Zool.* **7**, 16. (doi:10.1186/1742-9994-7-16)
22. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006 Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* **55**, 595–609. (doi:10.1080/10635150600852011)
23. MacArthur RH, Wilson EO. 1967 *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
24. Huelsenbeck JP, Ronquist F. 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. (doi:10.1093/bioinformatics/17.8.754)
25. Stamatakis A. 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)
26. Gelman A, Rubin DB. 1992 Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–511. (doi:10.1214/ss/1177011136)
27. Rabosky DL. 2006 Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* **60**, 1152–1164.
28. Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)
29. Weir JT, Schluter D. 2008 Calibrating the avian molecular clock. *Mol. Ecol.* **17**, 2321–2328. (doi:10.1111/j.1365-294X.2008.03742.x)
30. Rambaut A, Drummond AJ. 2007 Tracer v. 1.4. See http://evolve.zoo.ox.ac.uk/software.html.
31. Ree RH, Smith SA. 2008 Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* **57**, 4–14. (doi:10.1080/10635150701883881)
32. Smith SA. 2009 Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifolieae. *J. Biogeogr.* **36**, 2324–2337. (doi:10.1111/j.1365-2699.2009.02160.x)
33. Bandelt HJ, Forster P, Rohl A. 1999 Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48. (doi:10.1093/oxfordjournals.molbev.a026036)
34. Excoffier L, Smouse PE, Quattro JM. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial-DNA restriction data. *Genetics* **131**, 479–491.
35. Excoffier L, Laval G, Schneider S. 2005 Arlequin (v. 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47–50.
36. Jensen JL, Bohonak AJ, Kelley ST. 2005 Isolation by distance, web service. *BMC Genet.* **6**, 13. (doi:10.1186/1471-2156-6-13)
37. R Development Core Team. 2011 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See www.R-project.org.
38. Paradis E, Claude J, Strimmer K. 2004 APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290. (doi:10.1093/bioinformatics/btg412)
39. Pybus OG, Harvey PH. 2000 Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. Lond. B* **267**, 2267–2272. (doi:10.1098/rspb.2000.1278)
40. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008 GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131. (doi:10.1093/bioinformatics/btm538)
41. Bortolussi N, Durand E, Blum M, Francois O. 2006 apTreeshape: statistical analysis of phylogenetic tree shape. *Bioinformatics* **22**, 363–364. (doi:10.1093/bioinformatics/bti798)
42. Chan KMA, Moore BR. 2005 SymmeTREE: whole-tree analysis of differential diversification rates. *Bioinformatics* **21**, 1709–1710. (doi:10.1093/bioinformatics/bti175)
43. Moore BR, Chan KMA, Donoghue MJ. 2004 Detecting diversification rate variation in supertrees. In *Phylogenetic supertrees: combining information to reveal the tree of life* (ed. ORP Bininda-Emonds), pp. 487–533. Dordrecht, The Netherlands: Kluwer Academic.
44. Nee S, Mooers AO, Harvey PH. 1992 Tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl Acad. Sci. USA* **89**, 8322–8326. (doi:10.1073/pnas.89.17.8322)
45. de Queiroz A. 2005 The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* **20**, 68–73. (doi:10.1016/j.tree.2004.11.006)
46. Heaney LR. 2007 Is a new paradigm emerging for oceanic island biogeography? *J. Biogeogr.* **34**, 753–757. (doi:10.1111/j.1365-2699.2007.01692.x)
47. Norman JA, Rheindt FE, Rowe DL, Christidis L. 2007 Speciation dynamics in the Australo-Papuan *Meliphaga* honeyeaters. *Mol. Phylogenet. Evol.* **42**, 80–91. (doi:10.1016/j.ympev.2006.05.032)
48. Rowe KC, Aplin KP, Baverstock PR, Moritz C. 2011 Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. *Syst. Biol.* **60**, 188–203. (doi:10.1093/sysbio/syq092)
49. Kearns AM, Joseph L, Omland KE, Cook LG. 2011 Testing the effect of transient Plio-Pleistocene barriers in monsoonal Australo-Papua: did mangrove habitats maintain genetic connectivity in the black butcherbird? *Mol. Ecol.* **20**, 5042–5059. (doi:10.1111/j.1365-294X.2011.05330.x)
50. Rawlings LH, Donnellan SC. 2003 Phylogeographic analysis of the green python, *Morelia viridis*, reveals cryptic diversity. *Mol. Phylogenet. Evol.* **27**, 36–44. (doi:10.1016/S1055-7903(02)00396-2)
51. Yokoyama Y, Purcell A, Lambeck K, Johnston P. 2001 Shore-line reconstruction around Australia during the Last Glacial Maximum and Late Glacial Stage. *Quat. Int.* **83–85**, 9–18. (doi:10.1016/s1040-6182(01)00028-3)
52. Ereshefsky M. 1991 *The units of evolution: essays on the nature of species*. Cambridge, MA: MIT Press.
53. Wheeler Q, Meier R. 2000 *Species concepts and phylogenetic theory: a debate*. New York, NY: Columbia University Press.
54. de Queiroz K. 2007 Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886. (doi:10.1080/10635150701701083)
55. Rozendaal FG. 1994 Species limits within Garnet Pitta-complex. *Dutch Birding* **16**, 239–245.
56. de Queiroz K. 1999 The general lineage concept of species and the defining properties of the species category. In *Species: new interdisciplinary essays* (ed. RA Wilson), pp. 49–89. Cambridge, MA: MIT Press.
57. Rheindt FE, Eaton JA. 2010 Biological species limits in the banded pitta *Pitta guajana*. *Forktail* **26**, 86–91.
58. Mayr E. 1980 Problems of the classification of birds: a progress report. *Proc. Int Ornithol. Congr.* **17**, 95–112.

59. Martin G. 1996 Birds in double trouble. *Nature* **380**, 666–667. (doi:10.1038/380666a0)
60. Weir JT, Schluter D. 2007 The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* **315**, 1574–1576. (doi:10.1126/science.1135590)
61. Tobias JA, Bates JM, Hackett SJ, Seddon N. 2008 Comment on the latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* **319**, 901. (doi:10.1126/science.1150568)
62. Sangster G. 2009 Increasing numbers of bird species result from taxonomic progress, not taxonomic inflation. *Proc. R. Soc. B* **276**, 3185–3191. (doi:10.1098/rspb.2009.0582)
63. Johnson NK, Cicero C. 2004 New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* **58**, 1122–1130.
64. Kerr KC, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PD. 2007 Comprehensive DNA barcode coverage of North American birds. *Mol. Ecol. Not.* **7**, 535–543. (doi:10.1111/j.1471-8286.2007.01670.x)
65. Kerr KCR, Lijtmaer DA, Barreira AS, Hebert PDN, Tubaro PL. 2009 Probing evolutionary patterns in Neotropical birds through DNA barcodes. *PLoS ONE* **4**, e4379. (doi:10.1371/Journal.Pone.0004379)
66. Lohman D *et al*. 2010 Cryptic genetic diversity in 'widespread' Southeast Asian bird species suggests that Philippine avian endemism is gravely underestimated. *Biol. Conserv.* **143**, 1885–1890. (doi:10.1016/j.biocon.2010.04.042)