Oligo-Miocene colonisation and radiation within South-West Pacific arc terranes underpinned repeated upstream continental colonisations in pigeons (Columbiformes).

Paul Oliver[1], Andrew Hugall[1], Audrey Prasetya[1], Alex Slavenko[1], and Sabin Zahirovic[1]

[1]Affiliation not available

October 11, 2022

Abstract

Upstream colonisations from islands to continents have played an important role in the evolution of at least two major global bird radiations – the oscine passerines and the pigeons. Here we investigate the dynamics of insular diversification and upstream dispersal of pigeons(Columbiformes) within the Indo-Australian Archipelago. Based on a supermatrix fossil-calibrated phylogeny and model-based biogeographic analyses islands of Melanesia, now centred on New Guinea and considered separately from Australia, have been centre of pigeon diversification since the Eocene-Oligocene transition (~ 34 Ma). Geological reconstructions are concordant in suggesting arc terranes and continental ribbon fragments that underpin the contemporary Melanesian region may have formed extensive archipelagos separate from the main Australian landmass for much of the Oligocene and Miocene. For most of the last 35 million years these islands are inferred to have functioned as a net source of pigeon lineages for Asia, and especially Australia. However, the probability and evolutionary outcomes of upstream colonisation vary greatly across lineages and regions. Arboreal fruit-eating pigeons have colonised nearby continents on multiple occasions yet show little evidence of subsequent radiation. In contrast, insular terrestrial pigeons have been largely unable to colonise Asia, while a single early colonisation of Australia preceded an endemic radiation.

Abstract content goes here

Currently under peer-review.

Paul M. Oliver[1*], Andrew F. Hugall[2], Audrey Prasteya[3], Alex Slavenko[4], Sabin Zahirovic[5]

[1]Centre for Planetary Health and Food Security, Griffith University, 170 Kessels Rd, Nathan, Queensland 4111, Australia and Biodiversity and Geosciences Program, Queensland Museum, South Brisbane, Queensland 4101, Australia

[2]Department of Sciences, Museums Victoria, Melbourne Victoria 3001, Australia.

[3] Division of Ecology and Evolution, Research School of Biology, and Centre for Biodiversity Analysis, The Australian National University, Building 116, Daley Road, Acton, ACT 2601, Australia

[4]Macroevolution and Macroecology Group, Fenner School of Environment & Society, The Australian National University, Acton, Canberra, ACT 0200, Australia

[5]School of Geosciences, Faculty of Science, The University of Sydney, NSW, 2006.

[*]Corresponding author: p.oliver@griffith.edu.au)

Introduction

Theoretical predictions (MacArthur & Wilson, 1967) and empirical data (Crottini *et al.* , 2012; Economo *et al.* , 2015) suggest that islands are rarely a source of diversity for continents. This broad pattern is thought to be underpinned by varying combinations of the typically lower diversity of species on islands (Matthews *et al.* , 2019), tendency for insular taxa to lose adaptations for dispersal or competition (Bellemain & Ricklefs, 2008), and putatively higher rates of extinction in less stable island systems (Whittaker *et al.* , 2017). Because biotic communities on larger continents are often more species-rich they are also predicted to generate higher numbers of outwards colonists (MacArthur & Wilson, 1967) and act as a strong ecological filter against upstream colonisation (White *et al.* , 2021). However, running counter to these broad predictions, phylogenetic analyses and ancestral state reconstructions have demonstrated that some continental lineages have insular origins (upstream colonisation or reverse colonisation) (Carine *et al.* , 2004; Filardi & Moyle, 2005; Bocek & Bocak, 2019), especially in the Caribbean region (Nicholson *et al.* , 2005; Bellemain & Ricklefs, 2008). Understanding where, when and how islands contribute diversity to continents has the potential to provide broader insights into the factors that shape biotic dispersal and colonisation (Bellemain & Ricklefs, 2008).

The Indo-Australian Archipelago (IAA) spans from the Sundaland continental promontory to New Guinea, and is sometime also referred to as Malesia (as opposed to Melanesia). This region provides a prominent example of putative upstream colonisation with subsequent radiation in the global radiation of Oscine songbirds, which comprises of over 5000 extant species, all hypothesised to be descended from ancestors which occurred in islands of the IAA (Jønsson *et al.* , 2010; Moyle *et al.* , 2016). Upstream colonisation of south-east Asia from the IAA has also been recently inferred in two insect lineages (Letsch *et al.* , 2020; Bank, Cumming, & Bradler, 2021). The original geographic source of these upstream colonisations has been variably linked to terranes accreted to the northern edge of the Australian continent, often referred to as 'the proto-Papuan Archipelago' (Jønsson *et al.* , 2010), or the islands of "Wallacea" (Moyle *et al.* , 2016) (see more discussion of the geology of these regions below). These results emphasise the potential significance of the islands of eastern IAA as a source for diversity on nearby continents. However, the dynamics of insular diversification and island-continent dispersal across this region remains poorly understood.

Pigeons and doves (Columbiformes) are a moderately diverse global radiation of fruit and seed-eating birds (~350 nominal species), that also show a strong signature of insular diversification (Cibois *et al.* , 2014, 2017) and upstream colonisation (Lapiedra *et al.* , 2021). The islands of the IAA, and especially New Guinea and surrounds, are a world centre of pigeon ecological and lineage diversity, and were even more diverse pre-human settlement across the Pacific (Steadman, 2006; Steadman & Takano, 2020). Conversely, the oldest known Columbiform fossils are from early Miocene deposits in Australia (Worthy 2012) and New Zealand (Worthy et al. 2009), and on this basis it has been suggested that pigeons may have originated in Australia, or nearby islands (Low, 2014). To better understand the temporal and spatial dynamics of pigeon dispersal and diversification between islands and continents in the IAA, we assembled a supermatrix dated phylogeny for the pigeons, and from this we estimate, 1) the timing and geographic foci of insular pigeon diversification, with particular focus on the potential role of island arc terranes in the IAA in the early radiation of the group, and 2) estimate the relative frequency and direction of dispersal events between major island groups, and nearby continents.

Methods

Geological context

The geological history of the IAA is complex. Here we focus on summarising geological inference on the history of three broad clusters of putative island systems that may have formed to the north of the Australian continent, and which have been hypothesized to have a role in regional biotic diversification and dispersal through the Oligo-Miocene. Respectively these are: island arc and continental ribbon terranes that may underpin contemporary archipelagos spanning from the Philippines and to Melanesia (Oliver *et al.* , 2018a): the proto-Papuan Archipelago (Jønsson *et al.* , 2010); and the islands now clustered into the contemporary region of Wallacea (Moyle *et al.* , 2016). Due to the complex history of the IAA, some components of these clusters are not mutually exclusive. For example the southward migrating Caroline Arc has accreted onto the northern edge of New Guinea (e.g. Zahirovic et al. 2016). Similarly, the Halmahera Arc has migrated west from the New Guinea region into the contemporary Wallacean region (e.g. Hill & Hall 2003).

In this paper, when we refer to the region of Melanesia we include the island of New Guinea - reflecting the broader contemporary conceptualisation of the region across biotic and cultural spheres. We note that this broad definition differs from some key literature on birds which uses a more restrictive definition for Melanesia that excludes New Guinea (e.g. Mayr & Diamond, 2001),

For the broad history of island terranes in the IAA we here focus on Zahirovic et al. (2016) plate reconstructions but recast into a pure paleo-magnetic reference frame from Merdith et al. (2021) using GPlates (www.gplates.org). This allows us to preserve relative plate motions but remain true to "true" geographic paleo-latitudes that are relevant to climatically-sensitive analyses, such as those required for paleobiogeography. The distribution of shallow marine, emergent land, and elevated topographic regions are modified from paleogeographic reconstructions presented in Cao et al. (2017). Geological reconstructions will be uploaded to Dryad, and can also be accessed here - https://www.dropbox.com/t/nqL4mrvMdPlCxjQX

Island arc terranes and continental "ribbon terranes": Continental ribbon terranes probably had their origin in the convergence zone between the Eurasian, Indo-Australian, and Pacific tectonic plates. The oldest parts of the Philippine Archipelago likely formed 160 million years ago along the northern margin of Australia (Deng *et al.* , 2015; Zahirovic *et al.* , 2016). In addition to these Australia-derived terranes, younger volcanic arcs formed from ongoing subduction further to the north, including the Halmahera-Caroline, the Izu-Bonin-Mariana, and the Melanesian arc systems. For a broad period spanning from around 45 to 25 Ma tectonic reconstructions suggest that the Philippine, Caroline, and Melanesian arc systems may have broadly aligned to form a chain extending thousands of kilometres across the tropical south-west Pacific (Fig. 1). Within this chain, the Philippine arc terrane was proximal to Sundaland, and potentially provided a pathway for colonisation of the more distant islands of the Vitiaz Arc (itself comprised of the eastward continuation of the Melanesian and Caroline arc systems).

The proto-Papuan Archipelago : Within the IAA a persistent area of geological uncertainty in biogeographic analyses has been around the extent of islands in the present-day region of New Guinea, and whether these now-accreted fragments were contiguous or separate from the main Australian continent. This is further complicated by geologically long- and medium-term sea level fluctuations of up to ˜250 m since the Cretaceous (Haq, Hardenbol, & Vail, 1987) that may periodically isolate or connect these landmasses. Geological constructions suggest that the Sepik terrane likely collided with the Australian continental margin ˜50–30 million years ago, and generated uplift and topographic relief over the collisional timeframe (Schellart & Spakman, 2015; Zahirovic *et al.* , 2016; Mahoney *et al.* , 2019). The Caroline Arc likely collided between 15 and 5 million years ago (Zahirovic *et al.* , 2016; Mahoney *et al.* , 2019). Progenitor islands uplifted by these collisions were likely either partially or entirely isolated from the Australian mainland by a shallow sea over most of the last ˜50 million years (Norvick, 2003; Golonka *et al.* , 2006; Harrington *et al.* , 2017). This flooding of the northern Australian continental margin has been ascribed to the role of sinking tectonic plates (Harrington *et al.* , 2017), causing "dynamic subsidence" of the margin and north-eastward tilting of the Australian continent. The proto-Papuan archipelago itself was emergent only due to the collisions (e.g. the Sepik and Caroline arc accretions) occurring on this continental margin (Mahoney *et al.* , 2019), leading to crustal shortening and uplift on the leading edge of the advancing Australian continent (Fig. 1).

Wallacea : This region denotes the collection of islands bound to the west by Wallace's Line and to the east

by Lydekker's Line, including the Lesser Sundas, Sulawesi and Maluku (Fig. 1). It forms the contemporary geological and biotic bridge between the Sundaland (Eurasian) and Sahul (Gondwanan) bioregions. The islands of Wallacea are a collage of continental fragments, volcanic island arcs related to subduction, and possibly also accreted volcanic plateaus from mantle hotspots (Zahirovic *et al.*, 2016). Their complex tectonic evolution is evident in the network of suture zones (Zahirovic *et al.*, 2016), indicating the consumption of ancient ocean gateways during the main phase of equatorial ocean gateway closure since Australia's northward drift from Antarctica (Whittaker, J. M., Williams, S. E., and Müller, 2013). The islands between the Bird's Head Peninsula in north-western New Guinea and the east/southeast arms of Sulawesi represent the 'Sula Spur', which is a continental promontory attached to the Australian continent. The arrival of these blocks in the near-equatorial region by 20 million years ago marks the final closure of the Indonesian oceanic gateway, and the onset of the main phase of biological exchange between Australia and Asia. Most of the concentration and uplift of islands in Wallacea has likely occurred subsequent to this closure (in the last 20 million years) (Zahirovic *et al.*, 2016). The Lesser Sunda Islands are thought to have become emergent in the last ~10 million years, due to the interplay of Australia-Sundaland collision and the opening of the Banda Sea (Hinschberger *et al.*, 2005). Parts of Maluku and Sulawesi are also built on older Australian continental crust, forming the Sula Spur (Audley-Charles *et al.*, 1979), but have experienced extensive uplift and reconfiguration in the last 20 million years.

Phylogenetic supermatrix assembly

We compiled a species-level supermatrix of genetic data for the Columbidae plus selected outgroups using sequence data downloaded from NCBI (https://www.ncbi.nlm.nih.gov/). This used specified exemplar reference gene sequences drawn from the key Columbiformes phylogenetic references, with BLASTn+organism NCBI database searches. The emphasis was on widely sampled loci used in published phylogenetic studies with a reasonable proportion of taxa and major lineages. The key references are provided in the Supplementary File, and genes and GenBank accessions listed in Table S1. Species-level taxonomy followed the International Ornithology Council World Bird List v. 9.2 https://www.worldbirdnames.org/updates/. Recent phylogenomic analyses of the neoaves (Jarvis *et al.*, 2014; Prum *et al.*, 2015; Reddy *et al.*, 2017) indicated the Pteroclidae and Mesitornithidae as outgroups to the Columbidae. As there is patchy gene sequence for these taxa, we pooled data for each lineage to create composite family representatives, along with a composite Cuculiformes taxon as a further outgroup. Genetic data was aligned with MAFFT (v. 7.245) (Katoh & Standley, 2013) using the local-pair (L-INS-i) algorithm, alignments assembled into a custom Microsoft Excel database, and nomenclature rationalized to IOC9.2 (with the help of cross-referencing via Wikipedia using common names). Gene trees were inferred by IQ-TREE (Nguyen *et al.*, 2015) ultrafast bootstrap consensus (Hoang *et al.*, 2018), using models of sequence evolution identified by ModelFinder as implemented in IQTREE (Kalyaanamoorthy *et al.*, 2017). These trees were then scanned for non-monophyletic genera and species (using the custom script GTREER5), and the database updated by excluding aberrant accessions or in some cases revising nomenclature. Where necessary sequence sets were then realigned (as above).

Some long genes are routinely sequenced in fragments (e.g. RAG-1, COI), so in order to maximize data for the COI gene we also used a consensus method where the alignment was reduced down to a single consensus sequence per species, based on the most common base per site (with ties scored as ambiguous). This in effect picks the most commonly sequenced sub-lineages, and is a simple way to combine data and discount aberrant sequences (wrong loci/taxa etc). These consensus alignments were then subjected to the same procedure of gene tree and genera monophyly scans as above. We also used mitogenome data as follows. As across the relevant taxa gene order is preserved, we first aligned the entire mitogenomes then excised the set of commonly used genes and added the sequences to their respective alignments. The remainder (referred to as mtg-block) was kept as a separate alignment, after deleting the non-coding D-loop region. Concatenated supermatrix sequence data then used the best (longest accepted) single exemplar sequence per gene per species. These gene alignments were then compacted by removing regions with little or no data (<10% taxa per gene) and ambiguous alignment regions (via GTREER5, ALISCORE v2.2; Misof & Misof 2009).

Two versions of supermatrix - with and without the mtg-block - were analysed; the latter to avoid distorting

4

the result due to biased mitogenome sampling (missing from several key groups) and nucleotide saturation effects on relative divergence (especially for deeper outgroup lineages). Final analyses used the supermatrix without the mtg-block (as the six genes add enough well-sampled mtDNA, and empirically results were very similar). This final supermatrix comprised 247 out of 344 recognised pigeon species (72% complete) including sections of four nuclear and six mitochondrial gene loci, amounting to 11,100 sites 39% data-complete; 1,125,420 defined bases in 1,262 sequences (including 63 COI consensus) from 1527 accessions out of a total database of 3,639 accessions (with 37 rejected). Of 49 IOC9.2 Columbidae genera only three (all monotypic) were missing (*Starnoenas* , *Microgoura* and *Cryptophaps*).

Phylogenetic inference

Phylogenetic resolution was first explored using IQ-TREE ultra-fast bootstrap consensus with multi-partition sequence evolution model selected via ModelFinder (see Table S2). This was done for both with and without mtg-block datasets, and a cut down version of nuclear data only for 94 taxa with >2 nuclear genes. To further investigate the consistency of support among genes for primary lineages, we employed ASTRAL consensus species tree from gene trees approach (rather than site support methods due to the patchwork supermatrix data). This used a cut down dataset of 59 taxa for 8 genes.

Bayesian relaxed-clock analyses were performed in BEAST v2.4 (Drummond*et al.* , 2006; Bouckaert *et al.* , 2014), with a lognormal clock and Yule speciation prior (each with 1/X parameterization), and using a seven partition sequence evolution model optimized by ModelFinder (nuclear gene codon positions, intron+12S, mitochondrial codon positions; Table S2). Dating calibrations comprised of two Columbidae fossil constraints and two secondary outgroup priors.*Rupephaps,* a ptilinopine fruit pigeon from New Zealand, was applied as exponential (offset=22.0, mean=3.0) on the stem subtending*Hemiphaga* , *Lopholaimus* and *Gymnophaps* (Worthy*et al.* , 2009). *Primophaps* , a phabine pigeon from northern Australia, was applied as exponential (offset=19.0, mean=3.0) on the main Australian radiation of terrestrial seed-eating pigeons (stem subtending *Leucosarcia* to *Phaps*) (Worthy, 2012). Two normal secondary priors, based on Prum et al. (2015), were placed on the root (mean=63.0, sd=4.3) and Columbimorphs (mean=58.0, sd=4.8). Preliminary tests indicated a birth-death speciation prior was not required above the simpler Yule prior.

We ran two independent Bayesian MCMC analyses of 50 million steps sampling every 5,000, with a burnin of 20% that returned all parameters with ESS >500 (Rambaut *et al.* , 2014), with near identical maximum clade credibility (MCC) consensus tree topology, node support, and dating (see electronic supplementary material for the complete BEAST2 xml file). Runs were then pooled. For the purposes of biogeographic analyses, in addition to a MCC tree, a subset of 100 trees were drawn from the BEAST posterior sample to reasonably represent the topological and branch length variation of the total posterior. Outgroup taxa were then removed (Phytools R package; Revell 2012).

Defining biogeographic regions

To define a practical biogeographical categorization of regions, we balanced capturing the complexity of our focal regions (i.e. the islands between Australia and continental south-east Asia), with limiting biogeographic model complexity (i.e. minimising states for which there is insufficient data to estimate parameters). We also structured regions to a) allow comparisons between the major continents and islands with long histories of isolation from continental landmasses, and b) to allow some investigation of biogeographic patterns within the islands of the geologically complex IAA.

To begin with, we used 21 sub-regions or areas to describe insular and continental region pigeon diversity (Table S3). Geographic range data was initially summarised from the Handbook of the Birds of the World (Baptista *et al.* , 1997) and then cross-referenced and updated based on two online databases (https://www.iucnredlist.org/ and https://www.worldbirdnames.org/new/). Grouping of areas was initially explored using turnover of phylogenetic diversity (phylobeta: Graham & Fine 2008; Leprieur et al. 2012). Phylogenetic beta-diversity was estimated using Simpson's phylobeta index (pβsim) using the BEAST MCC tree (after removing outgroups), and compiled into a pairwise pβsim distance matrix (using custom script GTREER5) (see Dryad repository). Using this distance matrix we visualised differences between regions

5

using both hierarchical clustering (average = UPGMA) and Multi Dimensional Scaling (MDS) 2-D plot (hclust and cmdscale, 'stats' R package; R Core Team 2016)). Five taxa that occur in >8 sub-regions were removed from estimation of pβsim to limit emphasis on recent connections at the expense of underlying older endemicity, which is the focus of this study.

We then assigned taxa to broad regional categories (hereafter referred to as regions) based on modifying the pre-existing historical regionalizations (e.g. Jønsson et al. 2010, and also including the IOC9.2 breeding range zones) with information from both phylogenetic turnover (Fig. S1) and geological models for the history and formation of islands around the Indo-west Pacific (see above and Zahirovic et al. 2016). This resulted in an eight region scheme: 1) New World – North, South and Middle America: 2) – Old World (Europe, Africa and mainland Asia combined); 3) – Philippines; 4) Wallacea – Sulawesi and the Lesser Sundas; 5) West Melanesia - New Guinea plus the islands of Maluku; 6) Indian Ocean Islands – especially the Mascarenes and Madagascar; 7) Australia (not including New Guinea) and; 8) Pacific Ocean – the islands of the south-west Pacific including East Melanesia, New Caledonia and New Zealand. Assignment of phylogenetically sampled species to these eight regions is listed in Table S4. We note that our configuration of Wallacea does not include Maluku. This is because our phylogenetic turnover analyses indicate that much of Maluku is more closely allied to New Guinea (Fig. S1A). Geological (Hill & Hall 2003) and biological (Oliver *et al.* , 2022) data also indicate a close relationship between islands of Maluku and New Guinea. We separated Australia from West Melanesia in our analyses because: a) phylobeta analyses indicate the dominant pigeon faunas of the two regions are relatively discrete and b) geological models indicate that the New Guinea has been separated from the main Australian landmass by at least a shallow sea for most the late Cenozoic (see geological context above). Species that occur across regions were scored as such and allowed to have multiple states in analyses.

To investigate the role of islands in pigeon diversification and dispersal, we collated regions into broader categories of islands (a composite of the Wallacea, West Melanesia, Philippines and Pacific Ocean regions) versus continents (Americas, Old World and Australia). This was done by post hoc combining states and state changes inferred in the full eight region analyses into summary results (see below).

Historical biogeography modelling

We estimated ancestral states across the pigeon radiation using two different approaches: 1) Discrete-state Markov rate-based Maximum Likelihood (Mk-ML) model (Pagel, 1999; Huelsenbeck, Nielsen, & Bollback, 2003), as developed from O'Hara et al. (2019); 2) a suite of Bayesian biogeographic models as implemented in BioGeoBears (Matzke, 2013).

Markov Maximum Likelihood ancestral states inference

We inferred ancestral states using the rayDISC function in the R package corHMM (Beaulieu, O'Meara, & Donoghue, 2013), with node.states="marginal", state.recon="subsequently" and root.p="maddfitz". This algorithm allows multistate tip input, and infers marginal likelihoods (probabilities) for each state at each node including tips, the latter being in effect an implied state of common ancestor of the multistate tip (Felsenstein 2004, p. 255). Options for interpreting this post-analysis tip state marginal are to i) use these values or ii) to recode as equal probabilities to partially account for terminal dispersals within the tip lineage, or iii) explicitly count all implied terminal dispersals. In our eight-region scheme 86% of species are endemic to a single region, hence the moderate proportion (14%) of multistate species is a tolerable level for this type of ancestral state method to accommodate.

In order to select an optimal ML model structure, we applied an approximation intended to achieve a practical fair evaluation of what can be a prohibitively large number of possible rate models (Huelsenbeck *et al.* , 2003; Felsenstein, 2004). Using the single consensus tree, an average of ER, SYM and ARD model rates (in order to buffer against potentially miss-specified extreme values) were rank ordered into a serially increasing number of rate category models from the simplest single rate (ER), evaluated by BIC (Table S5). We selected a four-rate model structure representing a balance of complexity and support, which was then used for all subsequent analyses (Table S6).

6

Rather than produce a single best ancestral state result, we focussed on broad biogeographical inferences summarizing both model ancestral state and phylogenetic uncertainty. To do this we applied stochastic node state mapping using 200 re-samplings of node state marginal likelihoods, to each of 100 BEAST tree samples. Randomly sampling node states according to marginal likelihoods selects an explicit node state – multiple resampling then fairly represents the marginal likelihood (Huelsenbeck *et al.* , 2003; Revell, 2012; O'Hara *et al.* , 2019). Doing this for a set of tree samples then includes phylogenetic uncertainty in the whole inference.

Transition events are summarized according to parent to daughter node (including tip) states; cases where the state remains the same are referred to as endemic cladogenesis. These results can be divided up by divergence age into time-bins, assigned by daughter node age. For counting state lineages, branches spanning a time point can be assigned to closest parent or daughter node state. These procedures for summarizing an explicitly resolved set of ancestral states are straightforward (Revell, 2012) and can be applied to any ancestral state method including DEC models (Bribiesca-Contreras *et al.* , 2019).

This entire procedure can be repeated on multiple trees and the entire set of transition events and state lineages per time-bin combined into a final matrix integrating the tree and model variation (Bribiesca-Contreras *et al.* , 2019). Such summary results of events per time-bin are best interpreted as a summary of the relative frequency or probability density of such events rather than an explicit count. Results can further be integrated by post-hoc combining several individual states and changes into larger summary categories, such as continental versus islands. By summarizing the results across the whole marginal likelihood the method accounts for multiple solutions and hence in effect to some degree accounts for multistate solutions, particularly for combined region summaries.

Analyses were executed, and summarized into 2 million year time-bin profile plots, in R using a custom script ASSMR2. For phylogenetic visualizations onto the single MCC consensus tree, we averaged marginal likelihoods across all 100 sample trees plus consensus tree, for all nodes in common with the consensus tree (identical taxon bipartitions) and the linked parent node. Nodes (hence branches) with maximum state probability <67% were then marked as indecisive. Except for inclusion of the parent node, these are often used procedures (Matzke, 2013; O'Hara *et al.* , 2019).

BioGeobears

The probability of ancestral ranges was also estimated using BioGeoBEARS (Matzke, 2013) in R (R Core Team, 2016). Pigeon species were assigned to the same eight regions as the Markov Maximum likelihood inference. The maximum range size was set to eight regions as some species are widespread across all regions. Other settings were left as default. We used the standard six different biogeographic models tested by BioGeoBears: BAYAREA, which does not allow dispersal at cladogenesis; DIVA, which allows vicariance [i.e. if parent lineage occurs in (x,y), then one daughter lineage can occur in only (x) whereas the other can occur in only (y)] but disallows subset speciation, in which one daughter lineage inherits a subset of the distribution of the parent [i.e. if parent lineage occurs in (x,y), then one daughter lineage can occur in only (x) or only (y) whereas the other can occur in (x,y)]; DEC, which allows both vicariance and subset speciation; and BAYAREA+J, DIVA+J and DEC+J, which are equivalent to the previous three models but also allow jump dispersal through founder effects at cladogenesis(Matzke 2014). The BioGeoBEARS analysis was run using two separate phylogenetic inputs: 1) using the BEAST MCC tree and 2) a subset of 100 BEAST trees. Results of the BioGeoBEARS run across multiple trees were averaged. Repeated analyses were run to account for possible different topologies and poorly supported nodes in the phylogeny. The best-fit model for the pigeon data was DEC+J, followed by DIVALIKE+J (LnL -444.3 and -455.1, respectively; Table S7).

Results

Our Columbidae supermatrix phylogeny (Figs. 2A, S2) most closely resembles that of Lapiedra et al. (2021), which of published columbiform trees has the most data. The tree was well resolved, with the majority of nodes significantly supported (84% of nodes had ultrafast bootstrap (BS) > 95%, and 73% of nodes had posterior probability (PP) > 0.95). However, the phylogeny shows a striking feature of relatively low resolution among the basal primary lineages. This is evident in both ML and time-calibrated Bayesian

7

inferences (Fig. S2), and also when using just nuclear data alone, or in the low consistency among individual gene trees (Fig S3). This is exacerbated by the long stem divergence to the closest living outgroups. Including the extra mtg-block does not make any material change to topology. The root (and implied American) origins for pigeons are therefore weakly supported, despite considerable data (all of these nodes are informed by at least 9/10 genes) and high level of resolution elsewhere in tree. These basal nodes are associated with relatively short internodes (BS and PP with Pearson product moment correlation coefficient to branch length >0.5), and the shape of the overall pigeon phylogeny describes a long stem followed by rapid early radiation.

Results on general patterns of ancestral state estimation are consistent and complimentary across the Mk-ML (Figs. 3, S4) and DEC+J(Figs. S5–6) methods. Summary and discussion of patterns of inference will focus on the Mk-ML results.

All models and analyses support lineage accumulation (Fig. 2B) and endemic cladogensis (Fig. 2C) within the islands of the IAA through the Oligo-Miocene (Fig. S7). Endemic cladogensis summaries of combined islands versus combined continental regions (Figs. 3A–C) further indicate that the island systems have been continuously generating in-situ diversity since this time, including a peak relative to continents following a rapid accumulation of lineages in the early Oligocene. Three major and strongly supported clades (Ptilinopini, Raphiniae and Phabini) are each largely endemic or centred on eastern IAA, and group together with weak support into a large insular supra-clade in Bayesian analyses (Fig. 2A). The Cuckoo Doves (*Macropygia* and allied genera within the Colombini) also diversified on IAA islands (Fig. S4), but are much younger.

West Melanesia (comprising of New Guinea and Maluku) is the most frequently inferred insular state throughout the Oligo-Miocene, and is further the most frequent regional state for the entire Columbiformes through the Oligocene (Fig. 2B). The Rhaphinae and Ptilinopini reconstruct to West Melanesia, while the ancestral state for the Phabines is split across Wallacea, the Pacific and West Melanesia (Fig. 2). Three endemic pigeon radiations on the Philippines are inferred to be of Miocene age and largely derived from other insular lineages (Fig. 2). The Pacific and Indian oceans show no endemic clades older than the Miocene.

Insular, Old World and Australian pigeon faunas are relatively discrete, however, we inferred a number of exchanges between continental and insular regions. Upstream colonisations are more frequent than downstream colonisation up until the late Miocene (Figure 3B). In the Oligocene many reconstructions infer an initial upstream shift into the old world (Eurasia and Africa) at the base of weakly supported clade spanning *Treron* -*Turtur* (Fig. 2A). An alternative hypothetical reconstruction has the *Treron* and *Oena* / *Turtur* ancestral lineages independently colonising the Old World, contributing a separate mid-Miocene peak in inferred upstream shifts. In the early Miocene the mainly terrestrial Phabines colonised Australia from islands (Fig. 3D). From the mid-Miocene onwards at least ten lineages within the Ptilinopini are inferred to have colonised Australia or Sundaland independently (Fig. 2A). A majority of these dispersals are inferred in widespread and vagile species or species complexes that occur across both continental and insular regions.

In contrast, there is a negligible signal of downstream colonisation until the late Miocene (Fig. 3C), when the ancestor of the Cuckoo-Doves (*Macropygia* and allied genera) shifted from Continental to Insular regions (Fig. 2A); whether this colonisation occurred from the Old or New World is ambiguous. Through the late Plio-Pleistocene the overall number of inferred colonisation events increases, and downstream events begin to outnumber upstream events, with many shifts linked to widespread species or species complexes that occur across both continental and insular regions in genera such as *Macropygia* ,*Ptilonopus* , *Streptopelia* and *Treron* .

Discussion

Upstream colonisation from the Indo-Australian Archipelago

While there is strong evidence that upstream colonisation from islands to continents does occur, it has until recently been considered an anomalous trajectory (Bellemain & Ricklefs, 2008). The islands of the IAA appear to be the source area for at least one global radiation of birds (Jonsson *et al.* , 2010). Recent

macroevolutionary analyses of Columbiformes have also emphasised the importance of islands in their early evolutionary history, but did not focus on spatial or temporal patterns of dispersal. Conversely, other data indicate that the biota of continental Asia serves as a strong filter on upstream colonisation by either blocking successful colonisation completely (Oliver *et al.*, 2018b; White *et al.*, 2021) or by strongly filtering taxa by ecology (Letsch *et al.*, 2020).

Limitations of our phylogeny (70% sampling) and poor resolution at basal nodes) and historical reconstruction methods mean that we are able to give an accurate summary of trends, but not a detailed reconstructions of all events. Nonetheless, in support of Lapiedra (et al. 2021) we found that in the extant radiation of Columbiformes upstream colonisations from the IAA dominated downstream colonisations through much of the Oligo-Miocene. The importance of islands in the IAA is particularly striking given their comparatively small areal extent when compared to nearby continents (Fig. 2). The regional geography and ecological context of the earliest inferred upstream shifts into the Old World are relatively unclear, with the alternative hypotheses being one or two shifts from islands into the broader "Old World" in a weakly supported clade spanning *Treron* - *Turtur* . The Oligo-Miocene timing and direction of these colonisations is potentially comparable with upstream colonisation of the Oscine passerines (Jonsson *et al.*, 2010; Moyle *et al.*, 2016), although the outcomes in terms of diversification differ starkly (less than 40 species versus over 5000). A more strongly resolved phylogenetic tree, and ideally fossils, would be needed to shed more light on the trajectory of potential early upstream shifts in these lineages.

The geographical, temporal and ecological context of more recent inferred upstream colonisations in Columbiformes is clearer. The specialist fruit-eating arboreal pigeons (fruit doves) of the Ptilinopini have colonised nearby continents on at least ten separate occasions (and probably more if missing taxa such as *Ptilinopus alligator* and intraspecific populations are added to the phylogeny). Ecological niche shifts to arboreality in the ancestor of the Ptilinopini underpinned extensive radiation in the IAA and Pacific (˜100 species), and potentially also predisposed them to repeatedly colonise nearby continents (Lapiedra *et al.*, 2021). In contrast, speciation within continental regions appears to have been very limited. Indeed, around half the inferred upstream shifts involve species level taxa that occur across islands and continents. Furthermore most Ptilinopini in continental areas (and especially Asia) remain closely associated with coastal habitats and islands (e.g. *Ducula aenea* , *Ptilinopus jambu, P. melanospilus*), or other habitats with often less biodiverse communities such as montane areas (e.g.*Ptilinopus porphyreus* , *Ducula badia*) (Baptista *et al.*, 1997). This conforms with a hypothesis that upstream colonisation is challenging (White *et al.*, 2021), and even when frequent, ecological filtering may limit new arrivals to comparatively species poor environments on 'mainlands' (Mayr & Diamond, 2001).

Upstream colonisation by terrestrial-feeding Raphinini and Phabines has been even more limited, again suggesting ecological filtering at continental margins. Raphinini are entirely restricted to islands, despite their evident history of dispersing across vast distances into the Indian and Pacific oceans. The one striking outlier is provided by the Phabines, which are inferred to have colonised Australia from islands in the early Miocene and subsequently radiated across the continent. The relative success of this upstream colonisation could have been mediated by major climatic changes (aridification) and the concomitant opening up of niches and expansion of key food sources for terrestrial taxa, especially grasses (Toon *et al.*, 2015). Like the Raphinini, the Phabines have been largely unable to colonise continental Asia. The only exception is in the genus *Geopelia* , which originated in Australia, but has likely colonised the expanding savannahs of the Sundaland only recently during the Plio-Pleistocene. Thus, our data support the hypothesis that islands have been a net source area for continents (Lapiedra *et al.*, 2021), while simultaneously indicating that continents (and especially the lowland rainforests of Asia) present a formidable ecological filter to upstream colonisation (White *et al.*, 2021).

Mid-Cenozoic insular radiation in the IAA

Temporal and spatial patterns of insular diversification within the IAA are poorly understood, and have at times been contentious (Jonsson*et al.*, 2010; Moyle *et al.*, 2016). A key challenge is that while geological reconstructions suggest the position of key geological terranes, it is often unclear if these were above or below

9

water (Cao *et al.* , 2017). A further issue for biogeographic analysis is that the contemporary proximity of geological features differs greatly from paleogeographic configurations. Differing geological, biogeographic and cultural delineations of key regions have potentially further confounded analyses and comparison, most importantly New Guinea is not always considered part of Melanesia (Mayr & Diamond, 2001) and sometimes even lumped together with Australia as a single landmass of Sahul.

Our analyses emphasise that islands in the east of the IAA, here considered to be part of the broader Melanesian region, have been a hotspot of pigeon evolution and speciation since the Oligocene (Fig. S2, S7). Deep (mid-Cenozoic) insular origins have now been inferred for several radiations centred on Melanesia, strongly indicating that key terrane complexes (Vitiaz Arc and Sepik Arc) and/or the proto-Papuan region have likely been shaping diversity across the IAA since the mid-Cenozoic (Aggerbeck *et al.* , 2014; Oliver *et al.* , 2018a; Bank *et al.* , 2021) (Table S8). Wallacea to the west also appears to be a potentially important source of upstream colonists in pigeons, although this smaller region shows younger and weaker signals of lineage accumulation and endemic cladogenesis (Fig 2B,C) and geological evidence suggest more recent arrival and uplift of key geological features (Zahirovic *et al.* , 2016). To the north-west the Philippine pigeon fauna (Fig. S7B) appears to be relatively young and derived when compared to that of Melanesia or Wallacea. However, in light of the extreme geological dynamism of the region and the mobility of pigeons, we suggest that further discussion of the extent to which Oligo-Miocene insular diversification of pigeons may be linked to specific geological features within Melanesia - such as the Caroline, Sepik or Melanesian Arcs - remains speculative.

While the importance of Melanesian islands in the early diversification of pigeons is strongly supported, the ultimate source from which this region was colonised is obscured by the long stem lineage, the apparent rapid radiation at the base of the Columbiformes and the sparse fossil record. This contrasts with most other old insular radiations in the IAA, for which biogeographic analyses typically point clearly to either Australian or Asian origins (Table S8). One intriguing potential scenario for pigeons is South America; with westward dispersal to island arcs, leading to populations on quasi-continental fragments (Melanesia, Australia/Zealandia and Philippines) and later dispersal to Old World. This is speculative, but we suggest is as consistent with the current information on distribution, diversity, phylogeny, dispersal ability, biogeography and fossil record of pigeons as any other inference. Other Pacific lineages also show evidence of deep 'out-of-South America' origins (Malone, Reynoso, & Buckley, 2017), showing that eastward migration across the Pacific is possible, if not common.

The rapid early radiation of pigeons around the mid-Cenozoic also mirrors the estimated timing of initial diversification in two lizard radiations similarly centred on island arcs in the IAA (Oliver *et al.* , 2018a; Slavenko *et al.* , 2022). This concordance may reflect the timeframe when island arcs were especially conducive to biotic colonisation and diversification – for example the putative Philippine-Caroline-Vitiaz arcs (Zahirovic *et al.* , 2016). An alternative or complementary explanation is a shared legacy of global climatic shifts linked to changes in southern ocean circulation patterns (Zachos *et al.* , 2001). Analyses of plant (Nge *et al.* , 2020) and reptile (Oliver & Hugall, 2017) lineage diversification patterns in Australia, patterns of lineage diversity in mammal faunas globally (Stadler, 2011), and substantial turnover in marine and terrestrial fossil records (McGowran *et al.* , 2004; Sun *et al.* , 2014) all suggest the early Oligocene as a time of profound biotic turnover.

Challenging the cohesiveness of "Sahul"

Growing evidence for the biogeographic significance of terranes in the IAA also emphasises that considering Australian and New Guinea together as 'Sahul' may often mask two distinct mid Cenozoic centres of diversification with independent histories of isolation. The first likely consisted of tropical and mountainous island arcs and/or the Proto-Papuan Archipelago to the north. Many taxa with poor overwater dispersal abilities were likely absent (e.g. most terrestrial mammals) – potentially setting the stage for novel patterns of ecological release and diversification in pigeons (Lapiedra *et al.* , 2021), and other lineages that were able colonise (Aggerbeck *et al.* , 2014; Oliver, Skipwith, & Lee, 2014; Tallowin *et al.* , 2020; Roycroft *et al.* , 2022). The second potential centre of diversification is the main subaerial portion of the Australian continent to the

south, which was likely more arid and more temperate. Historical differentiation across these regions may be reflected in the contemporary diversity of Australian pigeons. On the one hand, the Australian phabine radiation appears be associated with the second region, and is centred on temperate forests, woodlands and deserts. In contrast Australia's tropical rainforests are dominated by lineages of Ptilinopini that are largely nested within insular radiations, suggesting a recently assembled fauna dominated by upstream colonists.

This growing evidence that a cohesive "Sahul" may be a relatively recent geological feature has important implications for biogeographic analyses. First, analyses of "Sunda-Sahul biotic" exchange should account for the probability that for much of the last 35 million years or more, the contemporary geological features that comprise "Sahul" were probably spread across two environmentally and geographically discrete regions separated by sea barriers. This would explain the often striking differentiation between the Australian and Melanesian biotas, especially in taxa with tropical or Asian origins (Joyce *et al.* , 2021; Oliver et al. 2022). Second, considering Melanesia, and especially New Guinea, as separate from Australia may also set the stage for a more holistic understanding of the longer-term role that island-to-continent dispersal has played in the assembly of the Australian biota, and especially the Australian rainforest biota. As a case in point, pigeons are key dispersers of rainforest fruit in Australia's rainforests (Crome, 1975). Yet here we provide evidence that the Australian fruit-specialist pigeon fauna is relatively young and derived from Melanesia (Fig. 3D). If fruit-specialist pigeons are indeed recent colonists, this may have important implications for understanding the assembly and evolutionary dynamics of the Australian rainforest flora, especially the recent influx of many plant species from the north (Sniderman & Jordan, 2011).

Data accessibility. All customs scripts (GTREER5.sh, ASSMR2.R), the supermatrix data and trees, and geological data will be available in The Dryad depository (to be completed).

References

Aggerbeck M, Fjeldsa J, Christidis L, Fabre PH & Jonsson KA .2014 . Resolving deep lineage divergences in core corvoid passerine birds supports a proto-Papuan island origin. *Molecular Phylogenetics and Evolution* **70** : 272–285.

Audley-Charles M, Carter DJ, Barber A, Norvick M & Tjokrosapoetro S . 1979 . Reinterpretation of the geology of Seram: implications for the Banda Arcs and northern Australia.*Journal of the Geological Society* **136** : 547–566.

Bank S, Cumming RT & Bradler S . 2021 . A tree of leaves: Phylogeny and historical biogeography of the leaf insects (Phasmatodea: Phylliidae). *Communications Biology* : 932.

Baptista LF, Collar NJ, Haffer JH, Horblit HM, de Juana E, Payne RB, Rowley I, Trail PW & Turner DA . 1997 . *Handbook of the birds of the world. Volume 4: Sandgrouse to Cuckoos.* (J del Hoyo, A Elliot, and J Sargatal, Eds.). Barcelona: Lynx Editions.

Beaulieu JM, O'Meara BC & Donoghue MJ . 2013 . Identifying hidden rate changes in the evolution of a binary morphological character: The evolution of plant habit in campanulid angiosperms. *Systematic Biology* **62** : 725–737.

Bellemain E & Ricklefs RE . 2008 . Are islands the end of the colonization road? *Trends in Ecology and Evolution***23** : 461–468.

Bocek M & Bocak L . **2019** . The origins and dispersal history of the trichaline net-winged beetles in Southeast Asia, Wallacea, New Guinea and Australia. *Zoological Journal of the Linnean Society* **185** : 1079–1094.

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A & Drummond AJ . **2014** . BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology* **10** .

Bribiesca-Contreras G, Verbruggen H, Hugall AF & O'Hara TD .**2019** . Global biogeographic structuring of tropical shallow-water brittle stars. *Journal of Biogeography* **46** : 1287–1299.

Cao W, Zahirovic S, Flament N, Williams S, Golonka J & Dietmar Muller R . **2017** . Improving global paleogeography since the late Paleozoic using paleobiology. *Biogeosciences* **14** : 5425–5439.

Carine MA, Russell SJ, Santos-Guerra A & Francisco-Ortega J .**2004** . Relationships of the Macaronesian and mediterranean floras: Molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in Convolvulus (Convolvulaceae).*American Journal of Botany* **91** : 1070–1085.

Cibois A, Thibault JC, Bonillo C, Filardi CE, Watling D & Pasquet E . **2014** . Phylogeny and biogeography of the fruit doves (Aves: Columbidae). *Molecular Phylogenetics and Evolution***70** : 442–453.

Cibois A, Thibault JC, Bonillo C, Filardi CE & Pasquet E .**2017** . Phylogeny and biogeography of the imperial pigeons (Aves: Columbidae) in the Pacific Ocean. *Molecular Phylogenetics and Evolution* **110** : 19–26.

Crome FHJ . **1975** . The Ecology of Fruit Pigeons in Tropical Northern Queensland. *Australian Wildlife Research***2** : 155–185.

Crottini A, Madsen O, Poux C, Strauss A, Vieites DR & Vences M .**2012** . Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K – T boundary in Madagascar.*Proceedings of the National Academy of Sciences USA* **109** : 5358–5363.

Deng J, Yang X, Zhang ZF & Santosh M . **2015** . Early Cretaceous arc volcanic suite in Cebu Island, Central Philippines and its implications on paleo-Pacific plate subduction: Constraints from geochemistry, zircon U-Pb geochronology and Lu-Hf isotopes.*Lithos* **230** : 166–179.

Drummond AJ, Ho SYW, Phillips MJ & Rambaut A . **2006** . Relaxed phylogenetics and dating with confidence. *PLoS Biology***4** : 699–710.

Economo EP, Klimov P, Sarnat EM, Guenard B, Weiser MD, Lecroq B & Knowles LL . **2015** . Global phylogenetic structure of the hyperdiverse ant genus Pheidole reveals the repeated evolution of macroecological patterns. *Proc. R. Soc. Lond. B* **282** : 20141416.

Felsenstein J . **2004** . *Inferring phylogenies* . Sunderland, MA: Sinauer Associates.

Filardi CE & Moyle RG . **2005** . Single origin of a pan-Pacific bird group and upstream colonization of Australasia.*Nature* **438** : 216–9.

Golonka J, Krobicki M, Pajak J, Van Giang N & Zuchiewicz W .**2006** . *Global Plate Tectonics and Paleogeography of Southeast Asia* . Krakow: Faculty of Geology, Geophysics and Environmental Protection, AGH University of Science and Technology, Arkadia.

Graham CH & Fine PV. 2008 . Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters* **11** : 1265–1277.

Haq BU, Hardenbol J & Vail PR . **1987** . Chronology of fluctuating sea levels since the Triassic. *Science* **235** : 1156–1167.

Harrington L, Zahirovic S, Flament N & Muller RD .**2017** . The role of deep Earth dynamics in driving the flooding and emergence of New Guinea since the Jurassic. *Earth and Planetary Science Letters* **479** : 273–283.

Hill KC & Hall R . **2003** . Mesozoic-Cenozoic evolution of Australia's New Guinea margin in a west Pacific context. In:*Geological Society of Australia Special Publication* ., 259–283.

Hinschberger F, Malod J, Rehault J, Villeneuve M, Royer J & Burhanuddin S . **2005** . Late Cenozoic geodynamic evolution of eastern Indonesia. *Tectonophysics* **404** : 91–118.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ & Vinh LS .**2018** . UFBoot2: Improving the ultrafast bootstrap approximation.*Molecular Biology and Evolution* **35** : 518–22.

Huelsenbeck JP, Nielsen R & Bollback JP . **2003** . Stochastic Mapping of Morphological Characters. *Systematic Biology* **52** : 131–158.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, Suh A, Weber CC, Fonseca RR da, Li J, Zhang F, Li H, Zhou L, Narula N, Liu L, Ganapathy G, Boussau B, Bayzid MS, Zavidovych V, Subramanian S, Gabaldon T, Capella-Gutierrez S, Huerta-Cepas J, Rekepalli B, Munch K, Schierup M, Lindow B, Warren WC, Ray D, Green RE, Bruford M, Zhan X, Dixon A, Li S, Li N, Huang Y, Derryberry EP, Bertelsen MF, Sheldon F, Brumfield RT, Mello C, Lovell P V., Wirthlin M, Samaniego JA, Velazquez AMV, Alfaro-Nunez A, Campos PF, Sicheritz-Ponten T, Pas A, Bailey T, Scofield P, Bunce M, Lambert D, Zhou Q, Perelman P, Driskell AC, Ruby G, Shapiro B, Xiong Z, Zeng Y, Liu S, Li Z, Liu B, Wu K, Xiao J, Yinqi X, Zheng Q, Zhang Y, Yang H, Wang J, Smeds L, Rheindt FE, Braun M, Fjeldsa J, Orlando L, Barker K, Jonsson KA, Johnson W, Koepfli KP, O'Brien S, Haussler D, Ryder OA, Rahbek C, Willerslev E, Graves GR, Glenn TC, McCormack J, Burt D, Ellegren H, Alstrom P, Edwards S, Stamatakis A, Mindell DP, Cracraft J, Braun EL, Warnow T, Jun W, Gilbert MTP & Zhang G . **2014** . Whole Genome Analyses Resolve the Early Branches in the Tree of Life of Modern Birds.*Science* **346** : 1320–1331.

Jonsson KA, Fabre P henri, Ricklefs RE & Fjeldsa J .**2010** . Major global radiation of corvoid birds originated in the proto-Papuan archipelago. *Proceedings of the National Academy of Sciences USA* **108** : 2328–2333.

Joyce EM, Thiele KR, Slik JWF & Crayn DM . **2021** . Plants will cross the lines: Climate and available land mass are the major determinants of phytogeographical patterns in the Sunda-Sahul Convergence Zone. *Biological Journal of the Linnean Society***132** : 374–387.

Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A & Jermiin L. 2017 . ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **14** : 587–589.

Katoh K & Standley DM . **2013** . MAFFT multiple sequence alignment software version 7: Improvements in performance and usability.*Molecular Biology and Evolution* **30** : 772–780.

Lapiedra O, Sayol F, Garcia-Porta J & Sol D . **2021** . Niche shifts after island colonization spurred adaptive diversification and speciation in a cosmopolitan bird clade. *Proceedings of the Royal Society B* .

Leprieur F, Albouy C, de Bortoli J, Cowman PF, Bellwood DR & Mouillot D . **2012** . Quantifying phylogenetic beta diversity: Distinguishing between 'true' turnover of lineages and phylogenetic diversity gradients. *PLoS ONE* **7** .

Letsch H, Balke M, Toussaint EFA, Narakusumo RP, Fiedler K & Riedel A . **2020** . Transgressing Wallace's Line brings hyperdiverse weevils down to earth. *Ecography* **43** : 1329–1340.

Low T . **2014** . *Where song began: Australia's birds and how they changed the world.* Melbourne: Penguin.

MacArthur RM & Wilson E. 1967 . *The Theory of Island Biogeography. Monographs in Population Biology. vol. 1* . Princeton, NJ: Princeton University Press.

Mahoney L, Mclaren S, Hill K, Kohn B, Gallagher K & Norvick M .2019 . Late Cretaceous to Oligocene burial and collision in western Papua New Guinea: Indications from low-temperature thermochronology and thermal modelling. *Tectonophysics* **752** : 81–112.

Malone CL, Reynoso VH & Buckley L . 2017 . Never judge an iguana by its spines: Systematics of the Yucatan spiny tailed iguana, Ctenosaura defensor (Cope, 1866). *Molecular Phylogenetics and Evolution* **115** : 27–39.

Matthews TJ, Rigal F, Triantis KA & Whittaker RJ .2019 . A global model of island species–area relationships.*Proceedings of the National Academy of Sciences of the United States of America* **116** : 12337–12342.

Matzke NJ . 2013 . Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing.*Frontiers of Biogeography* **5.4** : 242–248.

Mayr E & Diamond JM . 2001 . *The Birds of Northern Melanesia: Speciation, Dispersal, Biogeography* . Oxford: Oxford University Press.

McGowran B, Holdgate GR, Li Q & Gallagher SJ . 2004 . Cenozoic stratigraphic succession in southeastern Australia.*Australian Journal of Earth Sciences* **51** : 459–496.

Merdith AS, Williams SE, Collins AS, Tetley MG, Mulder JA, Blades ML, Young A, Armistead SE, Cannon J, Zahirovic S & Muller RD .2021 . Extending full-plate tectonic models into deep time: Linking the Neoproterozoic and the Phanerozoic. *Earth-Science Reviews* **214** .

Misof B & Misof K . 2009 . A Monte Carlo Approach Successfully Identifies Randomness in Multiple Sequence Alignments: A More Objective Means of Data Exclusion. *Systematic Biology* **58** : 21–34.

Moyle RG, Oliveros CH, Andersen MJ, Hosner PA, Benz BW, Manthey JD, Travers SL, Brown RM & Faircloth BC . 2016 . Tectonic collision and uplift of Wallacea triggered the global songbird radiation. *Nature Communications* **7** : 1–7.

Nge FJ, Biffin E, Thiele KR & Waycott M . 2020 . Extinction pulse at Eocene-Oligocene boundary drives diversification dynamics of two Australian temperate floras. *Proceedings of the Royal Society B: Biological Sciences* **287** .

Nguyen LT, Schmidt HA, Von Haeseler A & Minh BQ .2015 . IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32** : 268–274.

Nicholson KE, Glor RE, Kolbe JJ, Larson A, Hedges SB & Losos JB . 2005 . Mainland colonization by island lizards.*Journal of Biogeography* **32** : 929–938.

Norvick MS . 2003 . *New Paleographic Maps of the Northern Margins of the Australian Plate: Updated Report* .

O'Hara TD, Hugall AF, Woolley SNC, Bribiesca-Contreras G & Bax NJ . 2019 . Contrasting processes drive ophiuroid phylodiversity across shallow and deep seafloors. *Nature* **565** : 636–639.

Oliver PM, Brown RM, Kraus F, Rittmeyer E, Travers SL & Siler CD . 2018a . Lizards of the lost arcs: mid-Cenozoic diversification, persistence and ecological marginalization in the West Pacific. *Proceedings of the Royal Society B* **285** : 20171760.

Oliver PM, Blom MPK, Cogger HG, Fisher RN, Richmond JQ & Woinarski JCZ . 2018b . Insular biogeographic origins and high phylogenetic distinctiveness for a recently depleted lizard fauna from

Christmas Island, Australia. *Biology Letters* **14** : 20170696.

Oliver PM, Bower D, Mcdonald PJ, Kraus F, Luedtke J, Neam K, Hobin K, Chauvenet ALM, Allison A, Arida E, Clulow S, Gunther R, Nagombi E, Tjaturadi B, Travers SL & Richards SJ . **2022** . Melanesia holds the world's most diverse and intact insular amphibian fauna.*Communications Biology* **In press** .

Oliver PM & Hugall AF . **2017** . Phylogenetic evidence for mid-Cenozoic turnover of a diverse continental biota. *Nature Ecology & Evolution* **1** : 1896.

Oliver PM, Skipwith P & Lee MSY . **2014** . Crossing the line: increasing body size in a trans-Wallacean lizard radiation.*Biology letters* **10** : 20140479.

Pagel M . **1999** . The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* **48** : 612–622.

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Moriarty Lemmon E & Lemmon AR . **2015** . A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing.*Nature* **526** : 569–573.

R Core Team . **2016** . (2013). , Vienna, Austria. *R: A language and environment for statistical computing.*

Rambaut A, Suchard MA, Xie D & Drummond AJ . **2014** . Tracer v1.6.

Reddy S, Kimball RT, Pandey A, Hosner PA, Braun MJ, Hackett SJ, Han KL, Harshman J, Huddleston CJ, Kingston S, Marks BD, Miglia KJ, Moore WS, Sheldon FH, Witt CC, Yuri T & Braun EL . **2017** . Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Systematic Biology* **66** : 857–879.

Revell LJ . **2012** . phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3** : 217–223.

Roycroft E, Fabre PH, MacDonald AJ, Moritz C, Moussalli A & Rowe KC . **2022** . New Guinea uplift opens ecological opportunity across a continent. *Current Biology* : 1–10.

Schellart W. & Spakman W . **2015** . Australian plate motion and topography linked to fossil New Guinea slab below Lake Eyre.*Earth and Planetary Science Letters* **421** : 107–116.

Slavenko A, Tamar K, Tallowin OJS, Kraus F, Allison A, Carranza S & Meiri S . **2022** . Revision of the montane New Guinean skink genus Lobulia (Squamata: Scincidae), with the description of four new genera and nine new species. *Zoological Journal of the Linnean Society* **195** : 220–278.

Sniderman JMK & Jordan GJ . **2011** . Extent and timing of floristic exchange between Australian and Asian rain forests.*Journal of Biogeography* **38** : 1445–1455.

Stadler T . **2011** . Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences* **108** : 6187–6192.

Steadman DW . **2006** . *Extinction and Biogeography of Tropical Pacific Birds* . University of Chicago Press.

Steadman DW & Takano OM . **2020** . A new genus and species of pigeon (Aves, Columbidae) from the Kingdom of Tonga, with an evaluation of hindlimb osteology of columbids from Oceania.*Zootaxa* **4810** : 401–420.

Sun J, Ni X, Bi S, Wu W, Ye J, Meng J & Windley BF .**2014** . Synchronous turnover of flora, fauna, and climate at the Eocene–Oligocene Boundary in Asia. *Scientific Reports***4** : 7463.

Tallowin OJS, Meiri S, Donnellan SC, Richards SJ, Austin CC & Oliver PM . **2020** . The other side of the Sahulian coin: Biogeography and evolution of Melanesian forest dragons (Agamidae).*Biological*

Journal of the Linnean Society **129** .

Toon A, Crisp MD, Gamage H, Mant J, Morris DC, Schmidt S & Cook LG . **2015** . Key innovation or adaptive change? A test of leaf traits using Triodiinae in Australia. *Scientific Reports* **5** : 12398.

White AE, Dey KK, Stephens M & Price TD . **2021** . Dispersal syndromes drive the formation of biogeographical regions, illustrated by the case of Wallace's Line. *Global Ecology and Biogeography* **30** : 685–696.

Whittaker, J. M., Williams, S. E., and Muller RD .**2013** . Revised tectonic evolution of the Eastern Indian Ocean.*Geochemistry, Geophysics, Geosystems* **14** : 1891–1909.

Whittaker RJ, Fernandez-Palacios JM, Matthews TJ, Borregaard MK & Triantis KA . **2017** . Island biogeography: Taking the long view of natures labor atories. *Science* **357** .

Worthy TH, Hand SJ, Worthy JP, Tennyson AJD & Scofield RP .**2009** . A large fruit pigeon (Collimbidae) from the early Miocene of New Zealand. *Auk* **126** : 649–656.

Worthy TH . **2012** . A phabine pigeon (Aves:Columbidae) from Oligo-Miocene Australia. *Emu* **112** : 23–31.

Zachos J, Pagani M, Sloan L, Thomas E & Billups K .**2001** . Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science (New York, N.Y.)* **292** : 686–693.

Zahirovic S, Matthews KJ, Flament N, Muller RD, Hill KC, Seton M & Gurnis M . **2016** . Tectonic evolution and deep mantle structure of the eastern Tethys since the latest Jurassic.*Earth-Science Reviews* **162** : 293–337.
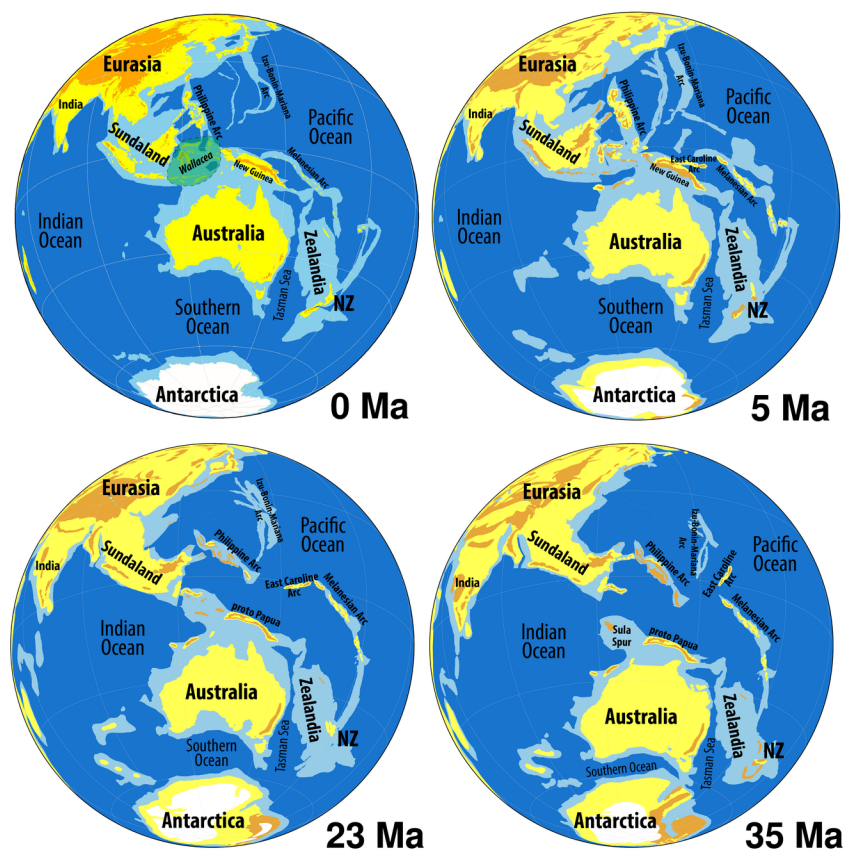
Figures.

Figure 1. Tectonic reconstructions for the Indo-Australian Archipelago and South-west Pacific region from Zahirovic et al. (2016). Orange demarcates putative mountain building, yellow subaerial land, light blue shallow sea (flooded continental, shelf, and volcanic plateau crust) and dark blue deep sea. Note the composite nature and changing configuration of Wallacea, the greatly varying position of key ribbon arc terranes, and the consistent inference of an isolated proto-Papuan Archipelago across all time slices.
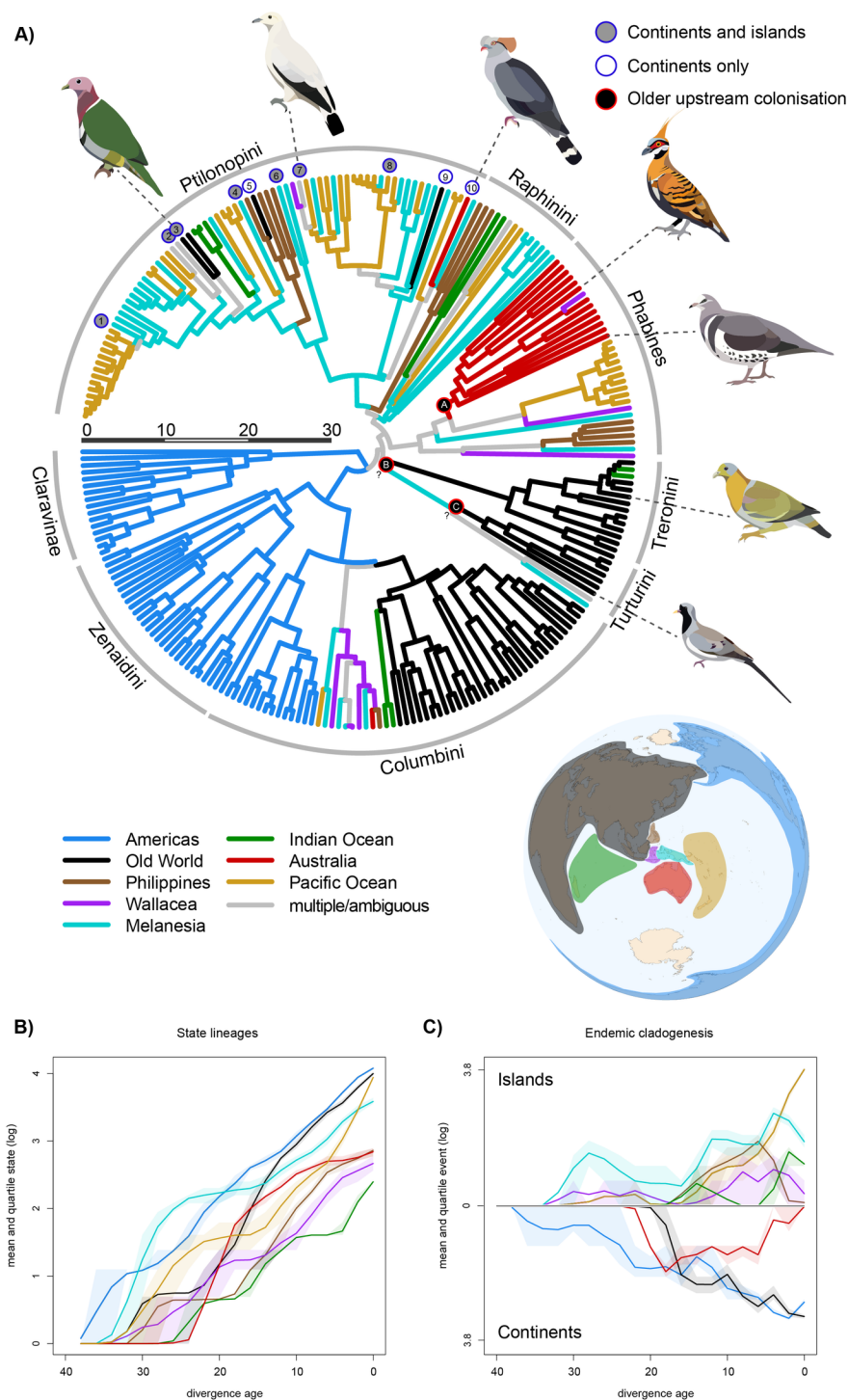
Figure 2. A) Phylogeny and ancestral state estimations for the global radiation of pigeons based on eight postdoc biogeographic regions. Contemporary taxa that occur on continents (C) or islands and continents (IC) within the fruit specialist Ptilinopini are numbered 1)*P. reginae* (IC), 2) *P. superbus* (IC), 3) *P. porphyreus* (IC), 4) *P. melanospilus* (IC), 5) *P. jambu*(C), 6) *P. magnificus* (IC), 7) *Ducula bicolor /luctuosa* complex (IC), 8) *D. aenea* complex (IC), 9) *Ducula badia* (C) and 10) *Lopholaimus antarticus* (C). Nodes

subtending potential older upstream colonisations in other clades are indicated with asterisks. Figure S4 shows the full tree with tip names and ancestral state reconstructions; B) lineage accumulation plots for the eight biogeographic regions; C) endemic cladogenesis plots for eight postdoc biogeographic regions. All biogeographic model result plots show median values, with 50% confidence intervals indicated by colour-shading.
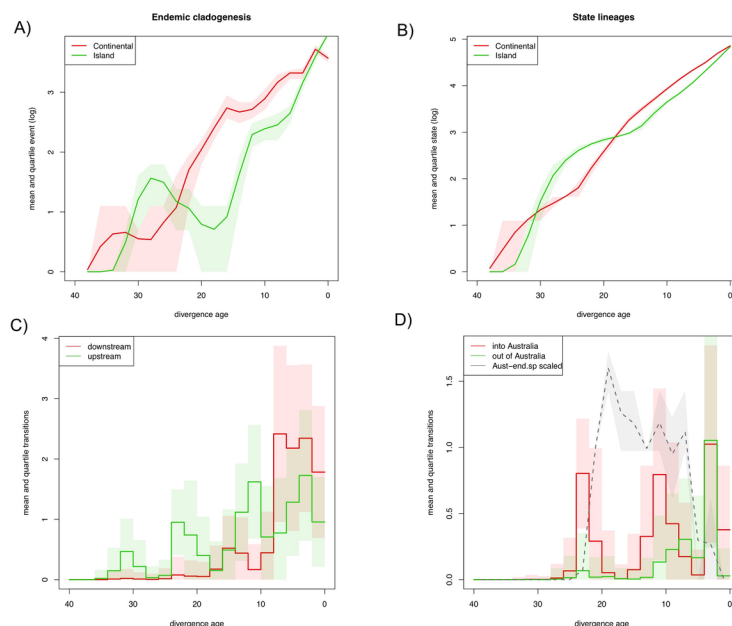

Figure 3. Columbidae continental (Australia, Old World, Americas) versus insular (Indian Ocean, Wallacea, West Melanesia, Pacific Ocean, Philippines) biogeographic model summaries. A) insular (green) versus continental (red) endemic cladogensis; B) continental versus insular lineage accumulation; C) inferred upstream (insular to continental) versus downstream (continent alto insular) dispersals; and D) colonisations into (all from islands) and out of Australia, juxtaposed against estimates of endemic cladogensis in Australia (dotted line). Dispersal event and geographic state mean and variation drawn from 200 node marginal re-samplings each of 100 posterior sample trees, summarized into 20 two-million year time bins. All biogeographic model result plots show median values, with 50% confidence intervals indicated by colour-shading.

Hosted file

`Oliver_2022_supp_submitted.docx` available at https://authorea.com/users/513832/articles/589874-oligo-miocene-colonisation-and-radiation-within-south-west-pacific-arc-terranes-underpinned-repeated-upstream-continental-colonisations-in-pigeons-columbiformes