PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

New Guinea highland origin of a widespread arthropod supertramp

Michael Balke, Ignacio Ribera, Lars Hendrich, Michael A. Miller, Katayo Sagata, Aloysius Posman, Alfried P. Vogler and Rudolf Meier

Proc. R. Soc. B published online 1 April 2009
doi: 10.1098/rspb.2009.0015

PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

FirstCite® e-publishing

Proc. R. Soc. B
doi:10.1098/rspb.2009.0015
Published online

New Guinea highland origin of a widespread arthropod supertramp

Michael Balke[1,*], Ignacio Ribera[2,3], Lars Hendrich[4], Michael A. Miller[4], Katayo Sagata[5], Aloysius Posman[6], Alfried P. Vogler[7,8] and Rudolf Meier[1]

[1]*Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore*
[2]*Museo Nacional de Ciencias Naturales (CSIC), 28006 Madrid, Spain*
[3]*Institute of Evolutionary Biology (CSIC-UPF), 08003 Barcelona, Spain*
[4]*Zoological State Collection, 81247 Munich, Germany*
[5]*Papua New Guinea Institute for Biological Research, Pacific Estate, Goroka 44, Papua New Guinea*
[6]*Papua New Guinea Binatang Research Center, Madang 51, Papua New Guinea*
[7]*Department of Entomology, The Natural History Museum, London SW7 5BD, UK*
[8]*Division of Biology, Silwood Park Campus, Imperial College London, Ascot SL5 7PY, UK*

The biologically and geologically extremely diverse archipelagos of Wallacea, Australasia and Oceania have long stimulated ecologists and evolutionary biologists. Yet, few molecular phylogenetic analyses of the terrestrial fauna have been carried out to understand the evolutionary patterns. We use dense taxon and character sampling of more than 7000 bp DNA sequence data for a group of diving beetles ranging from the Holarctic throughout Asia to as far east as French Polynesia. We here show that an ecologically diverse, common and widespread (Portugal to New Zealand) arthropod supertramp species originated in the highlands of New Guinea, *ca* 6.0–2.7 Myr ago. The approximately 25 closely related species are narrow endemics in Australasia/Oceania. The ancestor of this clade colonized that region from Eurasia *ca* 9–7 Myr ago. Our finding contradicts the widely held view of local endemism as an evolutionary dead end, as we find multiple peripatric speciation events within the Pleistocene and complex colonization patterns between the Oriental and Australian zoogeographic regions, including the recolonization of Eurasia, jumping across Wallace's line and colonization of continental Australia out of New Guinea. Our study strongly highlights the importance of dispersal over water gaps in shaping biogeographic patterns.

Keywords: molecular phylogeny; evolution of Oceanian fauna; origin of widespread species; *Rhantus* colonization of Wallacea; Wallace's line

1. INTRODUCTION

Island radiations are thought to undergo evolutionarily short 'taxon cycles' of diversification and rapid demise, before being superseded by different lineages of colonizers (Bellemain & Ricklefs 2008). The archipelagos of Wallacea (eastern Indonesia), Melanesia (including New Guinea) and Oceania have long served as a natural laboratory to study the evolutionary dynamics of such colonizations and biological radiations (Wallace 1859; Wilson 1961; Diamond 1974; Mayr & Diamond 2001; Novotny *et al.* 2007; Trewick & Cowie 2008). Yet, the faunal origins and mechanisms responsible for the region's diversification as well as their contribution to global diversity remain poorly understood. Several conferences have addressed the region's highly complex biogeography, often linking clade diversification and biogeographic patterns to plate tectonic activity (Keast & Miller 1996; Hall & Holloway 1998; Metcalfe *et al.* 2001; SAGE 2009).

*Author and present address for correspondence: Zoological State Collection, Münchhausenstrasse 21, 81247 Munich, Germany (michael_balke@yahoo.de).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2009.0015 or via http://rspb.royalsocietypublishing.org.

Received 5 January 2009
Accepted 12 March 2009

However, modern phylogenetic studies, especially at a wider geographical scale, are scarce. While zoologists find with Wallace's line a sharp biogeographic boundary in the centre of Indonesia, the situation is different for plants. Botanists classify the entire Malayan archipelago, the Philippines and New Guinea as one phytogeographic region, Malesia (van Steenis 1979), but for some plants, e.g. palms, Wallace's line seems to be an important boundary, too (e.g. Dransfield 1981). Genera such as *Cyrtandra* (Gesneriaceae) colonized Oceania in a single event out of southeast Asia via Malesia, radiating extensively to produce numerous island endemics (Cronk *et al.* 2005; Clark *et al.* 2008). Similarly, drynaroid ferns possibly originated in southeast Asia and West Malesia, and then dispersed across Indonesia and the Philippines to New Guinea (Schneider *et al.* 2008). Molecular phylogenetics of sandalwoods suggests an Australian origin followed by complex long-distance dispersal patterns throughout the Pacific as far west as Juan Fernandez (Harbaugh & Baldwin 2007).

Phylogenetic research on insects suggests that rare dispersal events out of Eurasia followed by vicariance on island chains played a prominent role in the evolution of the Wallacean and Melanesian faunas (Boer & Duffels 1996). Molecular phylogenetic studies on passerine birds

suggest a more complex scenario. They reveal colonization of Australasia/Oceania out of Eurasia followed by a colonization sequence in the opposite direction in recent geological time, as well as repeated transgressions of Wallace's line (Filardi & Moyle 2005; Jønsson & Fjeldså 2006). The evolution of megadiverse drosophilid clades in Hawaii and their unexpected subsequent colonization of continental areas and remote islands suggested by O'Grady & DeSalle (2008) is evidence that, for arthropods, even more complex biogeographic scenarios are possible. Polynesian spiders in the family Thomisidae have probably colonized Oceania twice: once from Eurasia and once from the New World (Garb & Gillespie 2006). Extended taxon and DNA character sampling thus unveiled more complex scenarios than that were apparent from morphological data, most notably the colonization of larger islands from remote, smaller ones (Filardi & Moyle 2005, see also Gillespie *et al.* 2008).

However, detailed molecular phylogenetic analyses in the region remain rare for arthropods (Garb & Gillespie 2006) and particularly insects (Austin *et al.* 2004). The few existing studies suggest, for example, unidirectional colonization of New Guinea and Melanesia out of Eurasia via Wallacea (Leys *et al.* 2002; Kodandaramaiah & Wahlberg 2007), or out of Australia (Balke *et al.* 2007*a*; Braby & Pierce 2007) followed by recent local radiations. Comparative studies on terrestrial species with ranges spanning the entire, vast region between southeast Asia and Oceania have not been undertaken to date.

We use molecular phylogenetic analyses to study the evolutionary history of aquatic beetles in the genus *Rhantus* to document the complexities of faunal evolution in these diverse archipelagos. The approximately 45 known species of the monophyletic *Rhantus suturalis* group are commonly encountered in the wetlands of the Holarctic, Oriental and Australian regions. These medium-sized, approximately 1 cm long predatory beetles are also abundant in southeast Asian, Wallacean and Melanesian highland habitats, but avoid tropical lowlands (Balke 2001). Most of the approximately 25 species of Australasia/Oceania are nevertheless narrow endemics on islands or single mountain chains with at least 12 endemic species in New Guinea alone (Balke 2001). However, one species, *R. suturalis*, is extremely widespread, ranging from New Zealand to the Iberian Peninsula. This species is a first colonizer of emerging lentic habitats where it is frequently replaced by other species during habitat succession. It is surprisingly the only species of the *R. suturalis* group found in numerous remote Wallacean island highland habitats (Balke 2001).

Based on the study of 28 species of the *R. suturalis* group and the comprehensive geographical sampling of *R. suturalis* itself, we aim to reconstruct the faunal evolution of this extraordinarily diverse and geologically complex part of the world.

2. MATERIAL AND METHODS

(a) *Taxon sampling and DNA sequencing*

We focused on the *R. suturalis* species group (Balke 2001; Balke *et al.* 2007*b*) and included 18 out of the 22 known species from Wallacea, Melanesia, Australia and Oceania, and two undescribed species from that region. Specimens for the widespread *R. suturalis* were obtained from its entire range (Europe, Central Asia, India, Japan, Wallacea, Melanesia, Australia and New Zealand; see the electronic supplementary material, table S1); while this sampling spans a large geographical area, with localities up to 18 000 km apart, we have not yet managed to acquire larger series from the individual localities for a population genetics analysis. The *suturalis* group otherwise occurs in the Holarctic and Oriental regions, from where we included eight species. To place the species of the *R. suturalis* group in a wider phylogenetic context, we included most other known Colymbetini species east of mainland Asia, i.e. *Rhantus* of the *pacificus* group and *Carabdytes upin* as outgroups (Balke *et al.* 2007*b*).

We sequenced more than 7 kb from the nuclear histone 3 (322 bp), wingless (352 bp), elongation factor 1-α (1092 bp, and approximately 180 bp intron sequence used for a subset of species), and 18S rRNA genes (602 aligned bp), as well as from the mitochondrial cytochrome *c* oxidase subunit 1–tRNA^{Leu}–cytochrome *c* oxidase subunit 2 (*cox1–cox2*; 2084 aligned bp), cytochrome *b*–tRNA^{Ser}–NADH dehydrogenase subunit 1 (*cob–nad1*) (1564 aligned bp), 16S rRNA–tRNA^{Leu}–NADH dehydrogenase subunit 1 (*16S rRNA–nad1*) (751 aligned bp) and 12S rRNA (384 aligned bp). Primers are listed in the electronic supplementary material, table S2. PCR protocols followed Cho *et al.* (1995) and Balke *et al.* (2004) for *EF1α*.

(b) *Alignment and data matrices*

The protein-coding genes were length invariant while length variation of the tRNAs and rRNAs amounted to a maximum of 11 bp length difference between pairs of sequences. In addition, *Rhantus pseudopacificus* had an autapomorphic 16 bp insertion in *rrnL*. We used the multiple alignment software MUSCLE with default settings to align the sequences (Edgar 2004). In order to explore whether indel placement influences topology, we carried out separate analyses excluding all nucleotide positions of the matrix containing indels.

Of the aligned 7151 characters in the combined analyses, 5288 characters were invariant and 1381 were parsimoniously informative, including 590 informative characters in the *cox1–cox2* fragment, 353 characters in *cob–nad1*, 156 characters in *EF1α*, 93 characters in *16S rRNA–nad1*, 67 characters in *H3*, 62 characters in *wng*, 40 characters in *12S rRNA* and 21 characters in *18S rRNA*.

We analysed nuclear DNA separately, using a dataset reduced to 20 individuals to avoid large amounts of missing data particularly for the long *EF1α* fragment. Here, we included the *EF1α* intron for most species except the highly deviating outgroup *Dytiscus dimidiatus*. Intron sequence length was 167–177 bp (aligned 180 bp; longest sequences in *Rhantus poellerbauerae*, *Rhantus montheithi* and *R. pseudopacificus*). The length of the aligned nDNA dataset, including the intron, was 2548 bp.

(c) *Phylogenetic analysis and molecular dating*

Phylogenetic relationships were inferred using parsimony (PAUP*, Swofford 2002) and partitioned Bayesian (MRBAYES v. 3.1, Huelsenbeck & Ronquist 2001) analyses. MRMODELTEST v. 2.2.1 (Nylander 2004) was used in conjunction with PAUP* for finding the best models as suggested by the hierarchical likelihood ratio tests for the combined dataset as well as for the partitions. Bayesian analyses were run with the data partitioned according to gene fragments. The fragments *cox1–cox2*, *cob–nad1* and *16S rRNA–nad1* were treated as three separate partitions. For the Bayesian analyses, we used the default priors starting

with random trees and ran two separate runs of three heated and one cold Markov chains for 4 000 000 generations, sampled every 500 generations. Posterior probabilities (PP) were calculated after discarding the first 5000 trees before likelihood values became stationary. Parsimony heuristic searches used 1000 replicates of random sequence addition with indels treated as a fifth character state in PAUP*. Bootstrap support (Felsenstein 1985) was obtained from 100 pseudoreplicates with 10 random sequence addition replicates each.

Maximum-likelihood analysis was conducted with GARLI (Zwickl 2006) under the GTR+I+Γ model using the default options, terminating the search after the analyses ran for 10 000 generations without significant increase in the likelihood of the topology. The bootstrap support of the likelihood trees was based on 250 replications.

We analysed the reduced nDNA dataset with GARLI and bootstrap support is based on 250 replications as above. Marker performance and interaction were explored by calculating partitioned Bremer support values (PBS; Baker & DeSalle 1997) for the tree topology shown in figure 1, with indels treated as fifth characters, and the full species dataset except for the *EF1α* intron. Constrained trees were generated with TreeRot v. 3 (Sorenson & Franzosa 2007) and partitioned for mtDNA, *H3*, *wng*, *EF1α* exons and *18S rRNA*. The parsimony analysis was run in PAUP*. To normalize the summed PBS values for each partition, we divided its PBS by the number of characters in the partition.

We estimated divergence times of splits using the Bayesian relaxed phylogenetic approach implemented in Beast v. 1.4.7 (Drummond & Rambaut 2007), using only the mitochondrial sequence and constraining the nine nodes marked in figure 1, which were present in all trees sampled from the chain for dating. The model was set to GTR+I+Γ with four rate categories. The uncorrelated lognormal relaxed molecular clock model was used to estimate substitution rates, with a Yule process of speciation as the tree prior. We ran two separate analyses for a combined 10^7 generations and used Tracer v. 1.4 to determine convergence of the two chains, measure the effective sample size of each parameter and calculate the mean and 95 per cent highest posterior density interval for divergence times (Rambaut & Drummond 2004).

We set the age of the node between *R. intermedius* (Tahiti) and *R. schereri* (Bora Bora) +*R. debilis* (Tahiti) to 1 Myr, assuming the origin of the two Tahitian species *R. intermedius* and *R. debilis* on that large island, and dispersal-mediated allopatry (*sensu* Clark *et al.* 2008) of *R. schereri* by dispersal to Bora Bora out of Tahiti. One Myr is the approximate age of Tahiti island (0.4–1.2 Myr, Guillou *et al.* 2005; Neall & Trewick 2008), and obtained a mean rate of 0.019 substitutions per site per Myr, with a 95% confidence interval between 0.011 and 0.028. This rate is in agreement with the 0.0115 substitutions per site per Myr widely assumed for insect mtDNA (Brower 1994) that we here also used for alternative age estimations.

(d) *Biogeography*

We used dispersal–vicariance analysis as implemented in the software Diva (see Ronquist 1996) under default settings, i.e. allowing widespread ancestral species (maxtrees = number of recognized areas), to infer ancestral ranges at the nodes of interest. Biogeographic analysis was also conducted by parsimony character mapping, using area as an unordered, multistate character traced on the topology of figure 1 with MacClade (Maddison & Maddison 1998). For this analysis, we defined the following biogeographic areas (character states) that are generally coherent in their overall faunal composition: Oriental region, Holarctic, Australia (incl. New Zealand and New Caledonia), Sundaland/Wallacea (the transitional zone between Oriental region and Australia), New Guinea and Eastern Melanesia+Polynesia.

We also tested how alternative area codings impact our conclusions. By redrawing area boundaries, this might change the conclusions about ancestral and derived area and hence dispersal routes in the Diva analysis. The alternatives tested and implications are outlined under results, illustrated in the electronic supplementary material and additional constrained analyses are described in the electronic supplementary material.

3. RESULTS

(a) *Molecular phylogenetics*

The combined analysis of 10 gene fragments (six mitochondrial, 4783 bp; and four nuclear, 2368 bp (excluding the *EF1α* intron), with a total of 7151 aligned base pairs, table 1) resulted in a strongly supported and well-resolved topology, at both the species and population level (figure 1). This exact topology was supported by parsimony as well as Bayesian phylogenetic methods in which model parameters were estimated independently for each gene. The protein-coding mitochondrial markers provided the highest number of informative characters, while the nuclear *18S rRNA* had the lowest contribution (table 1). Partitioned Bremer support analysis contrasting the node support of the combined mtDNA markers, *H3*, *wng*, *EF1α* exons and *18S rRNA* shows that mtDNA has the highest summed PBS (414.66), while all nDNA markers combined had 113.36 PBS. Normalized against the number of characters in each partition, mtDNA still contributed nearly twice the PBS of nDNA (0.086 versus 0.047) (see the electronic supplementary material, table S4). The normalized PBS of *H3* and *wng*, however, was higher than that of the mtDNA partition (0.194 and 0.116 versus 0.086, respectively). *18S rRNA* had a poor overall performance, conflicting with the other genes (see the electronic supplementary material, figure S2*F*) in the PBS analysis (*note*: *18S* analysed separately does not reveal a supported topology in conflict with the combined topology, as GARLI bootstrapping collapses all nodes). In the PBS analysis, *R. suturalis* was supported with positive PBS values for mtDNA, and the nuclear *H3* and *EF1α* exons; both the northern and southern clades of *R. suturalis* had positive PBS values for mtDNA and the nuclear *H3*. The lack of support from *EF1α* is most likely in this analysis due to the large amount of missing data, and, for the northern clade, the inclusion of *R. pederzanii*, which is clearly driven by the mtDNA genes (see §4).

Only one of nine backbone nodes differed between Bayesian and maximum-likelihood analysis of the unpartitioned dataset (see §2 and the electronic supplementary material), moving the Oriental clade above node 6 and the Melanesian clade above node 7 (see the electronic supplementary material, figure S1*A*). The topology remained robust to the removal of all sites with indels from the aligned dataset. The backbone of the tree and all crucial nodes discussed below were highly supported with

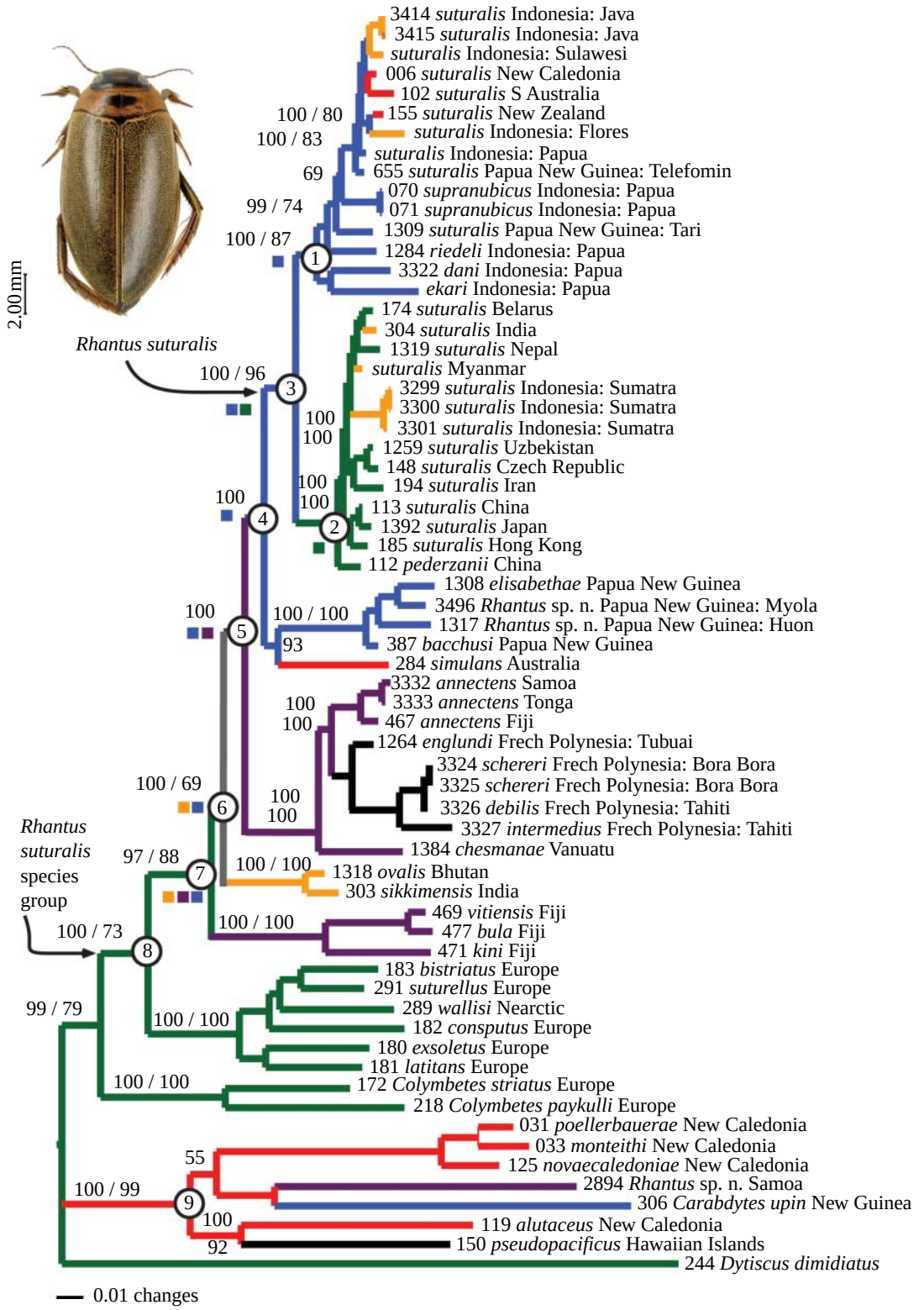

Figure 1. Phylogram of Bayesian analysis with data partitioned according to genetic loci. Numbers in circles refer to nodes discussed in the text and numbers above branches are Bayesian PP (×100)/ML bootstrap values obtained from GARLI, indicated for all crucial nodes (only one number = Bayesian pp). Branch colours denote distribution areas: orange, Oriental region; blue, New Guinea; red, Australia, New Zealand and New Caledonia; purple, Melanesia; black, Oceania (Polynesia); green, Holarctic. Extents of branch coloration derived from parsimony optimization with MacClade. Colour squares at the nodes denote ancestral ranges inferred with Diva, using the same colour code as for the branches.

Bayesian PP > 0.95, as well as high maximum-likelihood and parsimony bootstrap support values. Parsimony analysis in PAUP* yielded 24 trees of 6102 steps (gaps = fifth base, CI = 0.42, RI = 0.60, see the electronic supplementary material for analyses of different partitions and indel treatment).

The *R. suturalis* group was monophyletic (figure 1, node 8, PP 1.0) and *R. suturalis* was placed within a clade of narrowly endemic New Guinean and Australo-Oceanian species (figure 1, node 1). Sampling throughout its geographical range revealed two separate Asian/Palaearctic and Australian/Melanesian lineages, although

Table 1. Marker performance and tree statistics. (Parsimony analysis using PAUP*, CI, consistency index; RI, retention index; inf, parsimony informative.)

	no. of taxa	no. of characters	variable characters	inf characters	inf characters/ no. of characters	tree length	no. of trees	CI	RI
cox1–cox2	64	2083	133	595	0.28	3132	60	0.34	0.58
cob–nad1	53	1565	114	367	0.23	1509	>5000	0.43	0.59
h3	50	322	10	52	0.16	166	>5000	0.54	0.79
Wng	43	352	42	62	0.18	187	1241	0.67	0.81
EF1α	20	1092	51	152	0.14	283	288	0.79	0.82
12S	41	384	28	40	0.10	126	826	0.62	0.75
16S rRNA–nad1	45	751	89	93	0.12	352	>1000	0.61	0.70
18S rRNA	52	602	18	8	0.01	39	>500	0.82	0.90
combined	64	7151	479	1387	0.19	6102	24	0.41	0.60

sequence divergence within each of these was low even at great geographical distances. In the barcoding (www.barcodinglife.org) gene *cox1*, haplotypes from the Czech Republic and Sumatra, which are separated by 9500 km, differ by only 2.42 per cent, and those from Belarus and India (5000 km) by 0.42 per cent, i.e. well within the 2–3% intraspecific divergence widely seen in insects. The southern lineage of *R. suturalis* was paraphyletic with respect to three species restricted to minute ranges in the New Guinean highlands, and the northern lineage included one species found only in the southeast Chinese highlands. These subordinated endemics in the northern and southern clades took basal positions as the sister to the other clade members in the Bayesian and maximum-likelihood analysis, but using parsimony, *R. pederzanii* from China was the sister of *R. suturalis* from Myanmar, and *R. suturalis* from Tari in Papua New Guinea was the sister of the other individuals of the southern clade. Hence, *R. suturalis* gave rise to local endemics that are clearly distinct in male genitalia and external traits.

Separate analysis of only the nuclear DNA data (see the electronic supplementary material, figure S2*A*) including 180 bp *EF1α* intron sequence recovered the northern and southern clades and confirmed the nested position of *R. ekari* and *R. supranubicus* within *R. suturalis*, but *R. pederzanii* was recovered outside *R. suturalis* and not assigned to the northern clade above node 2. Even the two *EF1α* exons and the *EF1α* intron alone each support a northern and southern clade within *R. suturalis*, and the placement of *R. ekari* and *R. supranubicus* in the southern clade, but place *R. pederzanii* outside *R. suturalis* (see the electronic supplementary material, figures S2*B*,*C*). The southern and northern clades of *R. suturalis* can be diagnosed by four fixed character changes in the *EF1α* exon plus intron sequence that would result in their recognition as two separate species under some species concepts (Cracraft 1983; Wheeler & Meier 2000). Analyses of wng, h3 and 18S separately (see the electronic supplementary material figures S2*D*–*F*) reveal little resolution and collapsed bootstrap topologies (not shown), but importantly not conflicting, supported topologies with respect to figure 1.

A group of New Guinean species and the southwest Australian endemic *R. simulans* (node 4) was basal to *R. suturalis*, followed by a clade of Melanesian + Polynesian (node 5), a clade of two Oriental (node 6) and a clade of Fijian species (node 7). Holarctic members of the *suturalis* group, and the Holarctic genus *Colymbetes* were sister to all of these clades. Several other *Rhantus* species endemic to Oceanian islands, and the New Guinean *C. upin* were placed in a strongly supported clade sister to the remaining Colymbetini (node 9, PP 1.0, Balke *et al.* 2007*b*).

(b) *Faunal evolution and biogeography*

Biogeographic reconstruction using a dispersal–vicariance method as well as character optimization places the origin of *R. suturalis* or its ancestor in the New Guinean highlands (asterisk in figure 2), followed by a complex pattern of large range expansion throughout the Old World and secondary back migrations into the Indonesian islands from both Eurasian and Australian region (figure 2, nodes B–E). The Australasian/Oceanian region was colonized by the *R. suturalis* group from Eurasia (figure 1, nodes 6 and 7: according to parsimony character mapping, or only figure 1, node 7: dispersal–vicariance analysis).

Rhantus suturalis occurs sympatrically or even syntopically with narrowly endemic species in eastern New Guinea. *Rhantus suturalis* appears to be ecologically more generalist, and more abundant in newly occurring or man-made habitats such as irrigation ditches and watering holes in open land, while the narrowly endemic species such as the co-occurring *R. elisabethae* and *R. bacchusi* seem to prefer more primary forest water holes. Detailed ecological studies remain desirable however. The western part of New Guinea's central mountain chain interestingly only supports endemic species such as *R. supranubicus*, *R. dani*, *R. ekari* and *R. riedeli*. These endemics are nested within *R. suturalis* (clade at node 1) and were possibly peripheral isolates undergoing morphological change with no subsequent contact with the 'parental' *R. suturalis*. *Rhantus dani* and *R. ekari* are nevertheless morphologically highly similar to *R. suturalis*, differing slightly in male genital curvature (Balke 2001).

The two main clades of *R. suturalis* (figure 1, node 3) separated *ca* 2.7 or 4.3 Myr ago, using either the split of the Tahitian species or the insect mtDNA molecular clock of Brower (1994) to date the ultrametric tree derived from the BEAST analysis (4.2–1.4 Myr 95% confidence interval with the Tahitian dating or 6.7–2.3 Myr using Brower's rate), in the Late to Early Pliocene. Eurasia was recolonized from New Guinea by the ancestors of one of

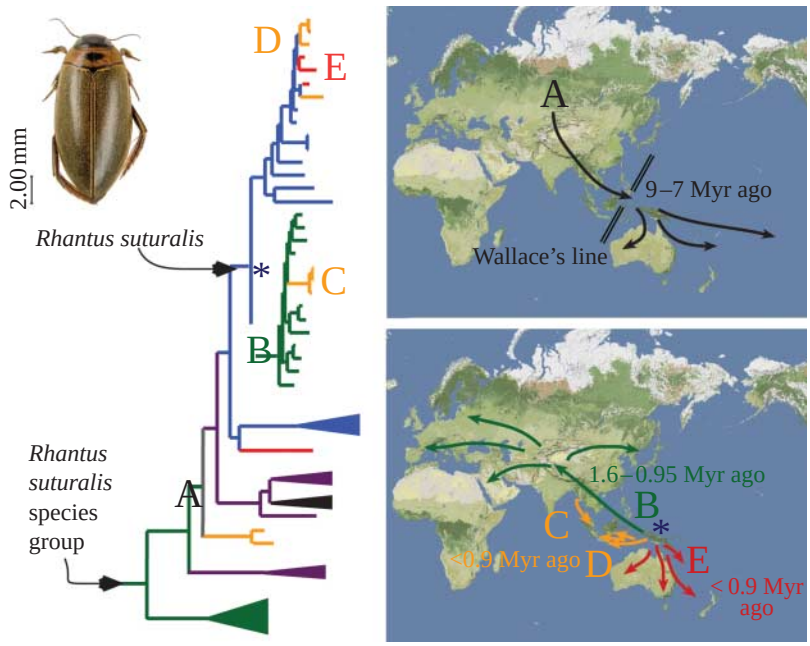

Figure 2. Generalized phylogram from figure 1 depicting major colonization events (A–E). Branch colours: orange, Oriental region; blue, New Guinea; red, Australia, New Zealand and New Caledonia; purple, Melanesia; black, Oceania (Polynesia); green, Holarctic; origin of *R. suturalis* in the New Guinea highlands (asterisk) and its range expansion towards the north (green arrows) (B); and colonization of Wallacea, Sumatra and the Australian region (orange and red arrows) (C, D, E).

these lineages, i.e. *Rhantus* crossed Wallace's line again, and the populations in the northern clade of *R. suturalis* (figure 2) spread over the entire Palaearctic region including north Africa and wide stretches of the Oriental region, diversifying there in the past *ca* 1.6 Myr or 950 000 years (1.5–0.4 Myr or 2.4–0.9 Myr) (node 2). However, this colonization left no faunal trace in Wallacea, the region between New Guinea/Melanesia and Eurasia. The only *Rhantus* known from there, the widespread *R. suturalis*, entered only in the Late Pleistocene, out of New Guinea again, to colonize Flores, Sulawesi and Java less than 900 000 years ago (1.2–0.3 Myr ago or 1.5–0.6 Myr ago; figure 2). *Rhantus suturalis* has recently colonized Sumatra from Eurasia, also less than 900 000 years ago (1.1–0.3 Myr ago or 1.7–0.6 Myr ago; figure 2). This colonization scenario is based on the topology within the northern and southern clades of *R. suturalis*, which is driven by mtDNA. However, the assignment of a Javanese specimen to the southern clade (figure 1, node 1), and a Sumatran specimen to the northern clade (figure 1, node 2) in the separate analysis of the *EF1α* intron (see the electronic supplementary material, figure S2*B*), shows that the directionality described here is supported by nDNA, too.

The Melanesian and Pacific Colymbetini faunal diversity encompasses an older, possibly Early Miocene component (figure 1, node 9) that is widely distributed from New Guinea, via the Solomons, Fiji, Samoa to Hawaii, with high levels of species endemism.

Finally, we constrained topologies where all *R. suturalis* from east and all specimens from west of Wallace's line group together, and a topology where all *Rhantus* of the suturalis group from east/west of Wallace's line have their own clade to enforce Wallace's line as a powerful dispersal boundary for these insects. Both these topologies were statistically significantly worse than our preferred topology as suggested by a Shimodaira–Hasegawa test (see also the electronic supplementary material).

(c) *Alternative distributional area coding*

The following alternative area delimitations were tested.

(i) New Zealand and New Caledonia were coded as separate areas, not part of the Australian region, as both are at least in part old landmasses with comparably high degrees of local endemism (e.g. Sanmartin & Ronquist 2004). This concerns only nodes subordinated within the clade at node 1. Thus, the overall biogeographic scenario does not change.

(ii) New Guinea, Solomon Islands, Vanuatu and Fiji were merged into Melanesia together with New Caledonia. This more conservative coding for the southern areas merely suggests a larger ancestral area and reveals a scenario where the Palaearctic and Oriental regions were colonized out of Melanesia, rather than specifically suggesting New Guinea highlands as the origin of *R. suturalis*.

(iii) Only the central mountain chain (=ocean floor uplift) of New Guinea was coded as 'New Guinea' without the localities on the Huon peninsula (*Rhantus* sp.n. Huon) and the Papuan Peninsula (*Rhantus* sp.n. Myola), which are parts of a former oceanic Melanesian island arc (e.g. Polhemus & Polhemus 1998) and therefore were assigned to Melanesia. This separation of New Guinean localities seems however inappropriate, as these different geological elements were already accreted to each other (Hall & Holloway 1998) before *Rhantus* began to diversify in the region (this paper). This coding scheme (see the electronic supplementary material figure, S1*B*) does not affect the preferred dispersal–vicariance scenario, as New Guinea remains the ancestral area of *R. suturalis*. Using parsimony reconstruction, this character coding introduces ambiguity as to whether *R. suturalis* originated in New Guinea or Melanesia.

(iv) Vanuatu was assigned to New Guinea instead of Melanesia, following Sanmartin & Ronquist (2004, see the electronic supplementary material, figure S1*C*). Under this coding scheme, the lineage leading to *R. suturalis* is inferred to be confined to New Guinea at an even earlier stage in the phylogeny (at nodes 5/6, rather than node 4 under the original coding; see the electronic supplementary material, figure S2).

4. DISCUSSION

Recent molecular studies in various lineages of animals have shown the great importance of dispersal in biogeography (Sanmartin *et al.* 2008), which has long been accepted for the flora. Plants, probably because of the dispersal ability of seeds, are less affected by geographical barriers and it is now believed that many southern phytogeographic patterns were shaped by recent dispersal (see Sanmartin & Ronquist 2004 for a comprehensive multi-taxon analysis; Harbaugh & Baldwin 2007, Pacific Sandalwoods; Keppel *et al.* 2008, southwest Pacific cycads). For the fauna, simple colonization scenarios have been replaced by more complex ones (Bellemain & Ricklefs 2008), particularly since recently diverged, morphologically conserved species can be included. Oscine passerine birds originated in the Australian/Melanesian region, and subsequently dispersed into southeast Asia after Australia entered a northern position close enough for successful dispersion, *ca* 30–15 Myr ago (Barker *et al.* 2004; Jønsson & Fjeldså 2006). A few species subsequently dispersed back to the Australian region. Thus, Melanesia + Australia not only received the northern biota as had been generally thought, but also served as a diversity pump in the opposite direction (Jønsson & Fjeldså 2006). Monarch flycatchers have dispersed from Eurasia into Melanesia and Oceania recently, diversifying over the past *ca* 2 Myr into numerous insular endemics. Out of this area of high endemism, some species dispersed to the larger island of New Guinea in the past 1 Myr (Filardi & Moyle 2005).

Here, we show the emergence of a very widespread and ecologically versatile 'supertramp' (Diamond 1974), *R. suturalis*, from a clade of Melanesian endemics that originated in the remote New Guinea highlands (figure 1, node 3). Samples of *R. suturalis* from throughout its geographical range revealed numerous genotypes representing two main clades that correspond to the Asian/Palaearctic and Australian/Melanesian populations, respectively. Both clades contain narrow endemics restricted to small areas in the New Guinean highlands, some of them morphologically still highly similar to *R. suturalis* (*R. ekari*, *R. dani*, *R. supranubicus*, *R. riedeli*) and southwest China (*R. pederzanii*), respectively (figure 1). Therefore, *R. suturalis* also gave rise to local endemics distinguishable by male genitalia, and in some species, external traits (Balke 2001). These differences possibly result in reproductive isolation when in contact with its dispersive ancestor. Extensive fieldwork in New Guinea has never uncovered the sympatric occurrence of *R. suturalis* with one of its nested species. However, *R. suturalis* frequently occurs even syntopically with distantly related species in the clade containing *R. bacchusi* and *R. elisabethae* (above node 4). While nDNA data support the inclusion of species such as *R. supranubicus* and *R. riedeli* (see the electronic supplementary material figures S2*A*,*B*) in the southern clade of *R. suturalis*, the inclusion of the Chinese *R. pederzanii* in the northern clade (figure 1) is clearly driven by mtDNA data alone and not recovered in any analysis of nDNA data (see the electronic supplementary material, figure S2). Although *R. suturalis* and *R. pederzanii* are morphologically very different species with respect to body size, genital structures and shape of the male claws, this might be an interesting case of introgression calling for analysis of additional, more variable genetic markers and especially different populations of *R. pederzanii*.

The traditional taxonomic scenario for the Australasian diving beetles of the *R. suturalis* group was in line with the literature on other arthropods (Gressitt 1956; Boer & Duffels 1996; Balke 2001; Kodandaramaiah & Wahlberg 2007) and older literature on birds. It suggested unidirectional colonization out of Eurasia, across Wallace's line, to enter Melanesia and the Australian–Polynesian region with subsequent diversification into numerous endemic species. Their subsequent diversification produced approximately 25 narrow endemics, which are rather abundant in some places and an important part of Melanesian/Oceanian freshwater ecosystems (figure 2). *Rhantus suturalis* is the only species of this genus in most of Australia, and, strikingly, in the vast Malay archipelago between Asia and New Guinea/Australia. Islands such as Sulawesi, Java, Borneo, Lombok, Flores and Timor all feature extensive highlands habitats yet only support this one species. Our results suggest that this region was colonized only during the past few hundred thousand years, and more extensive specimen sampling for population genetic analyses is required to understand whether the lack of morphological differentiation is due to a comparatively recent colonization of the region or continued gene flow despite the highly patchy insular range.

The surprising morphological similarity among populations of *R. suturalis* (Balke 2001) and lack of regional ecological specialization support a conservative recognition of only one paraphyletic, morphologically well-defined species. However, the presence of fixed character changes in the *EF1α*, although only few, would support the recognition of two phylogenetic species *sensu* Wheeler & Platnick (see debate in Wheeler & Meier 2000).

Rhantus suturalis occurs sympatrically or even syntopically with narrowly endemic species in eastern New Guinea. *Rhantus suturalis* appears to be ecologically more generalist, and more abundant in newly occurring or man-made habitats such as irrigation ditches and watering holes in open land, while the narrowly endemic species such as the co-occurring *R. elisabethae* and *R. bacchusi* seem to prefer more primary forest water holes. Detailed ecological studies remain desirable however. The western part of New Guinea's central mountain chain interestingly supports only endemic species such as *R. supranubicus*, *R. dani*, *R. ekari* and *R. riedeli*. These endemics are nested within *R. suturalis* (clade at node 1) and were possibly peripheral isolates undergoing morphological change with no subsequent contact with the parental *R. suturalis*. *R. dani* and *R. ekari* are nevertheless morphologically highly similar to *R. suturalis*, differing slightly in male genital curvature (Balke 2001).

Rhantus suturalis inhabits saline desert pools, steppe lakes, peat bogs, high-altitude puddles, stream pools, and even cattle troughs and other man-made waters throughout its range. The northern and southern clades are now close to making contact—our localities from Sumatra and Java are only 1000 km apart, but there are several potential highland habitats in between. Sampling these areas might offer a rare opportunity to better understand the evolutionary dynamics of the diversification or reinforcement process.

5. CONCLUSIONS

The complex biogeographic pattern revealed here involves multiple transgressions of Wallace's line, one of the most widely known biogeographic boundaries of the world (van Oosterzee 1997). Although considered isolated by distance from the northern and western biota, New Guinea instead contributed a supertramp species with a vast geographical range and large ecological amplitude. As islands are increasingly recognized as a 'pump' of species diversity rather than an evolutionary sink (Filardi & Moyle 2005; Garb & Gillespie 2006), *R. suturalis* must be among the most extreme cases of an island emigrant founding new continental lineages. Bellemain & Ricklefs (2008) argued that 'more diverse continental communities are difficult to invade. Accordingly, one might expect 'reverse colonists' to occupy ecologically peripheral or specialized positions in continental communities, perhaps with little potential for further diversification.' However, *R. suturalis*'s generalist habitat associations that include high-altitude peat swamps, steppe lakes and marshlands might have favoured the invasion of diverse continental areas. Here, phylogenetic data reveal the opposite pattern, and the initial ecological diversity of the widespread supertramp *R. suturalis* may have triggered further isolation of lineages and possibly speciation.

This work was funded by the German Science Foundation, DFG (BA2152/1-1, 1-2, 3-1, 3-2, 4-1, 6-1), the Ministry of Education in Singapore (R-154-000-270-112), the Spanish MICINN (CGL2007-61665/BOS) and the US National Science Foundation (DEB-0451971 to R. Gillespie). We also received generous support from the UK Darwin Initiative and the Linnean Society of London. We thank R. Sites, A. Ramsdale, M. Jäch, T. Inoda, H. Shaverdo, K. Mailautoka, A. Kinibel and E. Claridge for their valuable specimens. David Lohman and Gwynne Lim kindly read the manuscript. M.B. thanks Ditta Amran for her support in the field and A. Riedel for digital photographs of beetles. Voucher specimens will be deposited in the national insect collections of Papua New Guinea and Fiji. We thank the PNG Department of Environment and Conservation for long-term support.

REFERENCES

Austin, A. D. *et al.* 2004 Insects 'Down Under' diversity, endemism and evolution of the Australian insect fauna: examples from select orders. *Aust. J. Entomol.* **43**, 216–234. (doi:10.1111/j.1326-6756.2004.00448.x)

Baker, R. H. & DeSalle, R. 1997 Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst. Biol.* **46**, 654–673. (doi:10.2307/2413499)

Balke, M. 2001 Biogeography and classification of New Guinea Colymbetini (Coleoptera: Dytiscidae). *Invertebr. Taxon.* **15**, 259–275. (doi:10.1071/IT98008)

Balke, M., Ribera, I. & Vogler, A. P. 2004 MtDNA phylogeny and biogeography of Copelatinae, a highly diverse group of tropical diving beetles (Dytiscidae). *Mol. Phylogenet. Evol.* **32**, 866–880. (doi:10.1016/j.ympev.2004.03.014)

Balke, M., Pons, J., Ribera, I., Sagata, K. & Vogler, A. P. 2007*a* Infrequent and unidirectional colonization of hyperdiverse *Papuadytes* diving beetles in New Caledonia and New Guinea. *Mol. Phylogenet. Evol.* **42**, 505–516. (doi:10.1016/j.ympev.2006.07.019)

Balke, M., Wewalka, G., Alarie, Y. & Ribera, L. 2007*b* Molecular phylogeny of Pacific island Colymbetinae: radiation of New Caledonian and Fijian species (Coleoptera, Dytiscidae). *Zool. Scr.* **36**, 173–200. (doi:10.1111/j.1463-6409.2006.00265.x)

Barker, F. K., Cibois, A., Schikler, P., Feinstein, J. & Cracraft, J. 2004 Phylogeny and diversification of the largest avian radiation. *Proc. Natl Acad. Sci. USA* **101**, 11 040–11 045. (doi:10.1073/pnas.0401892101)

Bellemain, E. & Ricklefs, R. 2008 Are islands the end of the colonisation road? *Trends Ecol. Evol.* **23**, 461–468. (doi:10.1016/j.tree.2008.05.001)

Boer, A. J. & Duffels, J. P. 1996 Historical biogeography of the cicadas of Wallacea, New Guinea and the West Pacific: a geotectonic exploration. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **124**, 153–177. (doi:10.1016/0031-0182(96)00007-7)

Braby, M. F. & Pierce, N. E. 2007 Systematics, biogeography and diversification of the Indo-Australian genus *Delias* Hübner (Lepidoptera: Pieridae): phylogenetic evidence supports an 'out-of-Australia' origin. *Syst. Entomol.* **32**, 2–25. (doi:10.1111/j.1365-3113.2006.00349.x)

Brower, A. V. Z. 1994 Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl Acad. Sci. USA* **91**, 6491–6495. (doi:10.1073/pnas.91.14.6491)

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. & Zhao, S. 1995 A highly conserved nuclear gene for low-level phylogenetics: elongation factor 1-alpha recovers morphology-based tree for heliothine moths. *Mol. Biol. Evol.* **12**, 650–656.

Clark, J. R., Ree, R. H., Alfaro, M. E., King, M. G., Wagner, W. L. & Roalson, E. H. 2008 A comparative study in ancestral range reconstruction methods: retracing the uncertain histories of insular lineages. *Syst. Biol.* **57**, 693–707. (doi:10.1080/10635150802426473)

Cracraft, J. 1983 Species concepts and speciation analysis. *Curr. Ornithol.* **1**, 159–187.

Cronk, Q. C. B., Kiehn, M., Wagner, W. L. & Smith, J. F. 2005 Evolution of *Cyrtandra* (Gesneriaceae) in the Pacific Ocean: the origin of a supertramp clade. *Am. J. Bot.* **92**, 1017–1024. (doi:10.3732/ajb.92.6.1017)

Diamond, J. M. 1974 Colonisation of exploded volcanic islands by birds: the supertramp strategy. *Science* **184**, 803–806. (doi:10.1126/science.184.4138.803)

Dransfield, J. 1981 Palms and Wallace line. In *Wallace's line and plate tectonics* (ed. T. C. Whitmore), pp. 43–56. Oxford, UK: Clarendon.

Drummond, A. J. & Rambaut, A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)

Edgar, R. C. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (doi:10.1093/nar/gkh340)

Felsenstein, J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791. (doi:10.2307/2408678)

Filardi, C. E. & Moyle, R. G. 2005 Single origin of a pan-Pacific bird group and upstream colonisation of Australasia. *Nature* **438**, 216–219. (doi:10.1038/nature04057)

Garb, J. E. & Gillespie, R. G. 2006 Island hopping across the central Pacific mitochondrial DNA detects sequential colonisation of the Austral Islands by crab spiders (Araneae: Thomisidae). *J. Biogeogr.* **33**, 201–220. (doi:10.1111/j.1365-2699.2005.01398.x)

Gillespie, R., Claridge, E. M. & Goodacre, S. L. 2008 Biogeography of French Polynesia: diversification within and between a series of hotspot archipelagoes. *Phil. Trans. R. Soc. B* **363**, 3347–3361. (doi:10.1098/rstb.2008.0124)

Gressitt, J. L. 1956 Some distribution patterns of Pacific Island faunae. *Syst. Zool.* **5**, 11–33. (doi:10.2307/2411649)

Guillou, H., Maury, R. C., Blais, S., Cotten, J., Legendre, C., Guille, G. & Caroff, M. 2005 Age progression along the society hotspot chain (French Polynesia) based on new unspiked K–Ar ages. *Bull. Soc. Géol. Fr.* **176**, 135–150. (doi:10.2113/176.2.135)

Hall, R. & Holloway, J. D. (eds) 1998 The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In *Biogeography and geological evolution of SE Asia*, pp. 99–124 Leiden, The Netherlands: Backhuys.

Harbaugh, D. T. & Baldwin, B. G. 2007 Phylogeny and biogeography of the sandalwoods (Santalum, Santalaceae): repeated dispersals throughout the Pacific. *Am. J. Bot.* **94**, 1030–1042. (doi:10.3732/ajb.94.6.1028)

Huelsenbeck, J. P. & Ronquist, F. 2001 MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. (doi:10.1093/bioinformatics/17.8.754)

Jønsson, K. A. & Fjeldså, J. 2006 A phylogenetic supertree of oscine passerine birds (Aves: Passeri). *Zool. Scr.* **35**, 149–186. (doi:10.1111/j.1463-6409.2006.00221.x)

Keast, A. & Miller, S. E. (eds) 1996 *The origin and evolution of Pacific Island Biotas, New Guinea to Eastern Polynesia: patterns and processes*. Leiden, The Netherlands: Backhuys.

Keppel, G., Hodgskiss, P. & Plunkett, G. M. 2008 Cycads in the insular Pacific: dispersal or vicariance? *J. Biogeogr.* **35**, 1004–1015. (doi:10.1111/j.1365-2699.2007.01869.x)

Kodandaramaiah, U. & Wahlberg, N. 2007 Out-of-Africa origin and dispersal mediated diversification of the butterfly genus *Junonia* (Nymphalidae: Nymphalinae). *J. Evol. Biol.* **20**, 2181–2191. (doi:10.1111/j.1420-9101.2007.01425.x)

Leys, R., Cooper, S. J. B. & Schwarz, M. P. 2002 Molecular phylogeny and historical biogeography of the large carpenter bees, genus *Xylocopa* (Hymenoptera: Apidae). *Biol. J. Linn. Soc.* 77, 249–266. (doi:10.1046/j.1095-8312.2002.00108.x)

Maddison, W. P. & Maddison, D. R. 1998 MacClade, *analysis of phylogeny and character evolution, v. 3.08*. Sunderland, MA: Sinauer.

Mayr, E. & Diamond, J. 2001 *The birds of northern Melanesia*, p. 492. Oxford, UK: Oxford University Press.

Metcalfe, I., Smith, J. M. B., Morwood, M. & Davidson, I. D. 2001 *Faunal and floral migrations and evolution in SE Asia*. Leiden, The Netherlands: Balkema.

Neall, V. E. & Trewick, S. A. 2008 The age and origin of the Pacific islands: a geological overview. *Phil. Trans. R. Soc. B* **363**, 3293–3308. (doi:10.1098/rstb.2008.0119)

Novotny, V. *et al.* 2007 Low beta diversity of herbivorous insects in tropical forests. *Nature* **448**, 692–695. (doi:10.1038/nature06021)

Nylander, J. A. A. 2004 MrModeltest *v2*. Uppsala, Sweden: Evolutionary Biology Centre, Uppsala University.

O'Grady, P. M. & DeSalle, R. 2008 Out of Hawaii: the origin and biogeographic history of the genus *Scaptomyza* (Diptera, Drosophilidae). *Biol. Lett.* **4**, 195–199. (doi:10.1098/rsbl.2007.0575)

Polhemus, D. A. & Polhemus, J. T. 1998 Assembling New Guinea: 40 million years of island arc accretion as indicated by the distributions of aquatic Hemiptera (Insecta). In *Biogeography and geological evolution of SE Asia* (eds R. Hall & J. D. Holloway), pp. 327–340. Leiden, The Netherlands: Backhuys.

Rambaut, A. & Drummond, A. J. 2004. Tracer V1.4. See http://beast.bio.ed.ac.uk/Tracer.

Ronquist, F. 1996 Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* **46**, 195–203. (doi:10.2307/2413643)

SAGE 2009 *Southeast Asian Gateway Evolution, an international Conference held at Royal Holloway*. London, UK: University of London. (www.sage2009.rhul.ac.uk/index.html)

Sanmartín, I. & Ronquist, F. 2004 Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* **53**, 216–243. (doi:10.1080/10635150490423430)

Sanmartín, I., van der Mark, P. & Ronquist, F. 2008 Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *J. Biogeogr.* **35**, 428–449. (doi:10.1111/j.1365-2699.2008.01885.x)

Schneider, H., Kreier, H.-P., Hovenkamp, P. & Janssen, T. 2008 Phylogenetic relationships of the fern genus *Christiopteris* shed new light onto the classification and biogeography of drynarioid ferns. *Bot. J. Linn. Soc.* **57**, 645–656. (doi:10.1111/j.1095-8339.2008.00825.x)

Sorenson, M. D. & Franzosa, E. A. 2007 TreeRot, *v. 3*. Boston, MA: Boston University.

Swofford, D. L. 2002 *PAUP*. phylogenetic analysis using parsimony (*and other methods), v. 4.0b*. Sunderland, MA: Sinauer.

Trewick, S. A. & Cowie, R. H. 2008 Evolution on Pacific islands: Darwin's legacy. *Phil. Trans. R. Soc. B* **363**, 3289–3291. (doi:10.1098/rstb.2008.0127)

van Oosterzee, P. 1997 *Where worlds collide. The Wallace Line*. Ithaca, NY: Cornell University Press.

van Steenis, G. G. J. 1979 Plant-geography of East Malesia. *Bot. J. Linn. Soc.* **79**, 97–178. (doi:10.1111/j.1095-8339.1979.tb01511.x)

Wallace, A. R. 1859 On the zoological geography of the Malay Archipelago (communicated by Charles Darwin to the LSL meeting of 3 November 1859). *J. Proc. Linn. Soc. Lond.* **4**, 172–184. (doi:10.1111/j.1096-3642.1860.tb00090.x)

Wheeler, Q. B. & Meier, R. (eds) 2000 *Species concepts and phylogenetic theory. A debate*. New York, NY: Columbia University Press.

Wilson, E. O. 1961 The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* **95**, 169–193. (doi:10.1086/282174)

Zwickl, D. J. 2006 Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, The University of Texas, Austin, TX.