



A few steps forward and no steps back: long-distance dispersal patterns in small carpenter bees suggest major barriers to back-dispersal

Sandra Rehan^{1*} and Michael Schwarz²

¹Department of Biological Sciences, University of New Hampshire, 46 College Road, Durham, NH 03824, USA, ²School of Biology, Flinders University of South Australia, GPO Box 2100, Adelaide, SA 5000, Australia

ABSTRACT

Aim Inequality between rates of forward and backward dispersal is one mechanism that will produce regional differences in biotic composition. Any such differences will affect ecosystem assembly. Thus, it is important to understand why unequal dispersal rates might occur. Recent studies have indicated strong differences between forward and backward dispersal in two major bee groups, Hylaeinae and Allodapini. Here, we use a third bee group, tribe Ceratinini, to determine whether this might be a more general phenomenon.

Location Worldwide.

Methods Ceratinini is globally distributed (excepting Antarctica) and contains the single genus *Ceratina* and 21 subgenera, most of which are endemic to continents. We examined the phylogeny of Ceratinini using sequence data from three genes and 99 species, and inferred ancestral distributions using Bayesian and maximum likelihood approaches.

Results We show that the tribe Ceratinini had an African origin in the early Eocene. This was followed by four dispersals out of Africa in the Eocene and early Oligocene and three further dispersals between the Miocene and the present. There has been no back-dispersal of Ceratinini to Africa. The lack of back-dispersal cannot be explained by species abundance in non-African regions, and physical barriers to back-dispersal seem unlikely.

Main conclusions Our results are similar to findings for two other major bee groups and indicate that back-dispersal is very rare. It is unlikely that physical barriers can explain these patterns. However, dispersal patterns might be explained by assembly rules such as niche pre-emption or by reduced opportunities for escaping into enemy-free space. At a continental level, it is likely that the composition of bee fauna is more strongly influenced by very early dispersal events rather than by later physical barriers to dispersal.

Keywords

Ceratina, historical biogeography, Hymenoptera, long-distance dispersal, molecular phylogeny, Xylocopinae.

*Correspondence: Sandra Rehan, Department of Biological Sciences, University of New Hampshire, 46 College Road, Durham, NH 03824, USA.
E-mail: sandra.rehan@unh.edu

INTRODUCTION

Bees comprise a key element in angiosperm communities because of their role in pollination. Understanding how regional differences in bee diversity have arisen is therefore important for understanding how wider angiosperm-based communities have assembled and evolved. Dispersal of bee groups across regional scales and the influence of that dispersal on local ecosystems are key factors in community

assembly. The recent application of molecular phylogenetics to major bee groups is transforming our understanding of the ways in which long-distance dispersal (LDD) of bees has shaped regional fauna.

Kayaalp *et al.* (2013) showed that a globally distributed hylaeine bee genus, *Hylaeus* (family Colletidae), originated in Australia during the Oligocene and is now highly diverse and abundant there. *Hylaeus* has dispersed twice to New Zealand, where it remains a minor element, but a single additional

dispersal out of Australia led to its distribution on all continents except Antarctica. Despite this remarkable capacity of *Hylaeus* for LDD, there is no evidence for a second dispersal out of Australia (other than to New Zealand) and no evidence for its back-dispersal into Australia.

A similar paucity in LDD with lack of back-dispersal is evident in the allodapine bee genus *Braunsapis*. This genus originated in Africa and dispersed only once into Asia (Michener, 2007). *Braunsapis* then dispersed further into northern Australia, where it has become abundant and diverse (Reyes, 1991), but there is no evidence of its back-dispersal from Australia to Asia or from Asia to Africa (Fuller *et al.*, 2005). This pattern of limited forward dispersal and lack of back-dispersal is also thought to have occurred in the bee tribe Ctenoplectrini (Schaefer & Renner, 2008). Similarly, the highly diverse halictine bee genus *Lasioglossum* showed a single dispersal from Laurasia into Australia, where it has become highly diverse and a major element of the bee fauna; back-dispersal of *Lasioglossum* appears to be restricted to the subgenus *Homalictus*, and elements of that group have not moved further north than Thailand and southern India (Michener, 1979; Danforth *et al.*, 2004). Limited LDD is also thought to have occurred in the major tropical bee tribe Meliponini (Rasmussen & Cameron, 2010), while multiple forward and back-dispersals have occurred in tribe Bombini (Hines, 2008) and family Colletidae (Almeida *et al.*, 2012).

A scarcity of LDD and back-dispersal events could simply reflect the limited probability of these events. If a single forward dispersal is unlikely, then the probability of two such events occurring is very low, and the chance of back-dispersal might be less than that of forward dispersal if substantial time is required for colonizing populations to reach sizable numbers or diversity. Geophysical factors may also lead to asymmetry in the probability of forward and back-dispersal. For example, for species in which dispersal is aided by wind or ocean currents, the West Wind Drift might make movement from western to eastern landmasses much more likely than movement from east to west. However, successful LDD events might also be influenced by ecological processes that arise from the presence of related taxa in the target region. We now briefly consider these processes.

In a key review paper, Silvertown (2004) argued that the floral composition of the Canary and Micronesian islands suggested that early dispersals into these regions reduced the probability of successful colonization by subsequent dispersals of related taxa via 'niche pre-emption'. The key notion here is that members of a taxonomic group that have evolved adaptations to their local region might be able to exploit similar unoccupied niches when dispersing to a new region, but subsequent immigrants might be excluded from those now-colonized niches. Other factors may also make second or later dispersals less likely to succeed than earlier ones. For instance, early dispersing species may arrive in enemy-free space, relaxing initial fitness constraints. The new region may not contain parasites, predators or pathogens adapted to the new colonizer, which would enable it to

expand in the new region and to evolve defences as enemies evolve adaptations to the new lineage.

Understanding the origins of current regional patterns of bee diversity requires an understanding of factors that may have influenced past bee distributions. These factors include the ages of major bee groups, tectonic relationships that could affect opportunities for dispersal or lead to divergence via vicariance, and climatic changes. The ability to explore these factors has increased with the development of analytical techniques based on molecular phylogenetics that allow the ages of phylogenetic divergence to be estimated (e.g. Shapiro *et al.*, 2011; Drummond *et al.*, 2012).

Discerning patterns that involve rare dispersal events is difficult precisely because of the rarity of such events and requires broad comparative studies that capture a wide range of data. Our previous study used molecular phylogenetics to explore the age and historical biogeography of the bee tribe Ceratinini, and inferred an African origin in the Eocene (Rehan *et al.*, 2010). However, that study suffered from poor representation of many subgenera, especially those from the New World where only 11 species were included. Here, we provide DNA sequence data that deepen the geographical representation of Ceratinini, and we use this extended taxon set to explore the biogeographical history of this tribe. We examine whether long-distance and back-dispersal events have been rare for Ceratinini in particular, and for bees more generally.

MATERIALS AND METHODS

Taxonomy and choice of taxa

The tribe Ceratinini comprises a single genus, *Ceratina*, and approximately 200 described species (Michener, 2007). Taxon names used here are based on the subgeneric classification of Michener (2007). Our study included 99 *Ceratina* species representing 16 of 21 described subgenera and covering all ecozones of ceratinine diversity: Afrotropical (32 species), Malagasy (four species), Oriental (29 species), Palaearctic (15 species), Nearctic (six species) and Neotropical (13 species). Importantly, this study improved on the taxon sampling of Rehan *et al.* (2010) by including two additional subgenera, the monospecific *Protopithitis aereola* and a new undescribed subgenus. We included 10 additional New World species and 22 newly sequenced Old World species. The five missing subgenera are largely monospecific and endemic to single ecoregions (Michener, 2007). The choice of species for inclusion in our study was based on the availability of sequence data and our attempt to cover as many subgenera as possible whilst not excluding any species that were available to us, regardless of their subgeneric assignment. However, conspecific specimens were not included.

Voucher specimens are housed in the collections of M.P. Schwarz at Flinders University of South Australia. In addition to the 99 *Ceratina* species, we included 25 species from two tribes of the Xylocopinae: Allodapini (12 species) and

Xylocopini (seven species), as well as two ctenoplectrine and four corbiculate bees, to allow fossil calibration points and to help root the ingroup. The allodapines were included because this tribe is the extant sister group to Ceratinini (Sakagami & Michener, 1987; Roig-Alsina & Michener, 1993; Engel, 2001; Rehan *et al.*, 2013) and is thus likely to be most appropriate for rooting the ceratinine clade. The split between Ceratinini and Allodapini also provided a minimum-age calibration point because there is support for a sister relationship between allodapines and the Baltic amber fossil tribe Boreallodapini with Ceratinini as the next most basal clade (Engel, 2000). We included the seven *Xylocopa* species to root the subfamily Xylocopinae (Rehan *et al.*, 2012). The four corbiculate and two ctenoplectrine bees provided another age calibration point between the xylocopines and apines (Cardinal *et al.*, 2010).

DNA sequence acquisition

The molecular dataset comprised 2487 base pairs (bp) of aligned nucleotide characters from one nuclear gene, the 773-bp F2 paralogue of elongation factor-1 alpha (EF-1 α F2) and two mitochondrial genes: the 1285-bp cytochrome oxidase *c* subunit one (COI), and the 429-bp cytochrome *b* (*cyt b*). Tissue sampling, DNA extraction, polymerase chain reaction (PCR) amplification, sequencing, and alignment protocols followed Rehan *et al.* (2010). GenBank accession numbers and distribution information for all species are provided in Table S1 in Appendix S1 of the Supporting Information.

Phylogenetic analyses

We used Bayesian techniques implemented in BEAST 1.6.2 (Drummond & Rambaut, 2007) to recover the Ceratinini phylogeny from ultrametric trees where branch length was proportional to divergence time. We partitioned sequence data into four groups: codon positions 1 and 2 for the nuclear gene EF-1 α and for the two mitochondrial genes combined, and codon position 3 for these same genes. We unlinked substitution models and nucleotide composition for each partition, but we linked clock and topology models. On the basis of earlier hierarchical log-likelihood (lnL) and Akaike's information criterion (AIC) analyses in MODELTEST 3.06 (Posada & Crandall, 1998), a GTR+I+ Γ model was used for each partition, with all parameters subsequently estimated in BEAST. We used an uncorrelated log-normal relaxed clock model, a Yule process of speciation, and a randomly generated starting tree. Fifty million iterations were run with sampling every 5000 iterations, and a burn-in of 10 million iterations was determined by examining parameter values (including lnL) as a function of iteration number using TRACER 1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). Analyses were run three times to check for convergent model outcomes.

No reliable internal fossil calibration points were available for Ceratinini; therefore, we used an external calibration

point consisting of a normally distributed prior for the root node with a mean of 107 Ma and a standard deviation of 10 Myr, based on molecular phylogenetic studies by Cardinal *et al.* (2010) and Cardinal & Danforth (2011), which covered all major apid subfamilies. This estimate of 107 Ma was older than the approximate age (95 Ma) for the same divergence by Cardinal & Danforth (2013), but 95% highest posterior density (HPD) limits in the latter paper indicated overlap between the two estimates.

Biogeographical analyses

We aimed to identify likely ancestral regions for key nodes in the Ceratinini phylogeny, as well as likely dispersal and vicariance events. This required extant distributions to be defined in a way that reflected both existing distributions and biogeographically meaningful regions. We defined six relevant areas: Afrotropical, Malagasy, Oriental, Palaearctic, Nearctic and Neotropical regions (Fig. 1). These areas largely coincide with ecozones, which are the broadest biogeographical divisions of the Earth's land surface and are based on the distribution of terrestrial organisms. Patterns of plant and animal distribution were shaped by plate tectonic processes, which have redistributed the world's landmasses over geological time. Madagascar has traditionally been included in the Afrotropical ecozone, but the separation of the Madagascar–India landmass from continental Africa beginning about 160 Ma, and the rifting of India from Madagascar approximately 85 Ma, led to a unique endemic Malagasy biota with many recent bee colonization events (Rabinowitz *et al.*, 1983; Fuller *et al.*, 2005; Yoder & Nowak, 2006; Eardley *et al.*, 2009; Rehan *et al.*, 2010; Chenoweth, 2011).

We used three methods to infer the historical biogeography of the Ceratinini: a dispersal–vicariance–cladogenesis (DEC) maximum likelihood approach implemented in LAGRANGE version 2 (Ree & Smith, 2008); and a Bayesian binary MCMC [BBM, Markov chain Monte Carlo (MCMC)] technique and a statistical dispersal–vicariance analysis (S-DIVA) approach, both implemented in RASP version 2.0 beta (Yu *et al.*, 2011). The BBM method in RASP is modified from source code of MRBAYES 3.1.2 (Ronquist & Huelsenbeck, 2003) and the S-DIVA analysis is modified from Ronquist's (2001) DIVA 1.2 software. We utilized the six geographical regions outlined above for all analyses. Only one species of *Ceratina* (*C. australensis*) is known in the Australian region, and *C. australensis* may be conspecific with the Indian *C. propinqua* (Michener, 1965); as a result, we treated the Indo-Australian region as part of the Oriental region in the analyses. Of the 99 *Ceratina* species assessed, five have ranges that cover two of our geographical regions (one species occurs in the Nearctic and Neotropics, and four occur in the Oriental and Palaearctic), so we set two regions as the maximum that could be occupied by ancestral nodes for BBM and S-DIVA.

For the LAGRANGE analysis, we precluded certain dispersals and regions that could be simultaneously occupied by an ancestral species. We permitted dispersal only between

adjacent regions (i.e. between the Palaearctic and Nearctic, the Oriental and Nearctic, the Oriental and Afrotropical, and the Malagasy and Afrotropical regions). We did not set penalties for permitted dispersals, and dispersal direction was always symmetric. We did not vary permissible ranges or dispersals over time.

We used the LAGRANGE configurator (<http://www.reelab.net/lagrange/configurator/index>) to enter our consensus Bayesian chronogram and matrix of current species ranges and then set the assumptions for ranges and dispersal outlined above. The resulting Python script was run on a Macintosh OS X 10.7 platform.

For the RASP analyses, we coded the outgroup (represented by the allodapine *Macrogalea zanzibarica*) as having a 'wide' distribution, meaning that it could occur in any of the regions occupied by our ingroup, but we also explored setting the root node to an African distribution to reflect the origin of the allodapines (Chenoweth & Schwarz, 2011). The MCMC analyses were run for 100,000 generations, with sampling every 100 iterations and a burn-in of 500. Although this was a small run size compared to the BEAST analysis, the BBM analysis explored a single model parameter (geographical distribution) with six discrete states, whereas the BEAST analysis contained dozens of model parameters, including topology and branch length, many of which were continuous. We ran the analyses three times to check for convergent outcomes.

Some joint ancestral distributions were precluded from our S-DIVA analyses on the basis of historical geography; we did not permit ancestral species to occupy both Africa or Madagascar and the New World because of the large physical separation between these regions throughout the time frame of our phylogeny. One problem with DIVA, and hence with S-DIVA, is that it treats vicariance as a default optimization (Yuan *et al.*, 2005) and can indicate a vicariance event over a dispersal event, even when vicariance is precluded by historical geographical relationships (Schwarz *et al.*, 2006). Therefore, we considered the possibility that inferred vicariance might actually be dispersal when examining our S-DIVA results.

Timing of LDD events and lack of back-dispersal

Our biogeographical analyses indicated that there were five dispersal events from Africa to the Oriental, New World and Palaearctic regions, and two from Africa to Madagascar. Most of these events occurred close to the Ceratinini root. In order to examine whether LDD events were biased towards the root, we recovered dates (node ages) of all branching events that involved the African lineage and divided these into two groups: one in which both branches had African distributions, and another in which one branch led directly to a clade with a non-African distribution. We compared these two sets of node ages using logistic regression with node type as the categorical dependent variable and node age as the independent variable, and queried whether divergence time predicted nodes leading to LDD events. One complication with this approach is that for LDD events that led to a

single terminal species, dispersal could have occurred any time between the stem node and the present, so stem age may overestimate the actual time of dispersal. This complication arose for both Malagasy and Palaearctic dispersals involving *Ceratina minutula*. To remove these complications, we conducted further analyses using only non-African clades with multiple terminals and with crown rather than stem ages. Crown ages provided a minimum dispersal date and so may have underestimated the age of a dispersal event.

Comparing the relative expected probability of forward dispersal from Africa and of back-dispersal into Africa is problematic because differences in species abundance in source regions should affect the probability of LDD. One way to compare these probabilities, albeit with caveats, is to use a null hypothesis that assumes equal probability of any speciation event giving rise to a daughter lineage that produces an LDD event involving Africa and a non-African region, regardless of whether the parent clade is African. The frequency of forward dispersal (from Africa) and back-dispersal can then be contrasted using chi-square tests.

RESULTS

Phylogenetic relationships

BEAST analysis recovered *Neoceratina* as a monophyletic (posterior probability, PP = 1.0) clade, and its sister-clade relationship to the remaining ceratinines was also strongly supported (PP = 1.0) with the remaining group recovered as monophyletic with high support (PP = 0.95) (Fig. 1, Appendix S1: Fig. S1). Of the two daughter clades arising from this latter node, monophyly was strongly supported (PP = 0.96) for the clade containing *Pithitis* (an undescribed African subgenus referred to as 'new subgenus'), the Oriental clade including *Ceratinidia* and *Lioceratina*, and the New World subgenera. There was moderate support (PP = 0.86) for monophyly of the clade comprising *Euceratina*, *Malgatina*, and a variety of African and Malagasy taxa in the subgenera *Megaceratina*, *Ceratina s. s.*, *Hirashima*, *Simioceratina* and *Cenoceratina*. Our results also suggested that *Ceratina s. s.* may be paraphyletic with respect to *Megaceratina* and *Copoceratina*.

The clade containing *Pithitis* and *Protolithitis* was well supported (PP = 1.0) as were the African *Pithitis* (PP = 1.0) and Oriental *Pithitis* (PP = 1.0), which were paraphyletic to the African subgenus *Protolithitis*. The African + Malagasy subgenus *Hirashima* was recovered as monophyletic (PP = 1.0) with two Malagasy species. The Palaearctic subgenus *Euceratina* was rendered polyphyletic with the Malagasy subgenus *Malgatina* (PP = 1.0). The subgenera *Lioceratina* and *Ceratinidia*, which were mostly Oriental but which included some Palaearctic species, were recovered as monophyletic (PP = 1.0) and were strongly supported as sister clades (PP = 1.0) to the monophyletic (PP = 0.95) New World subgenera *Zadontomerus*, *Calloceratina* and *Ceratinula* (Fig. 1, Appendix S1: Fig. S1).

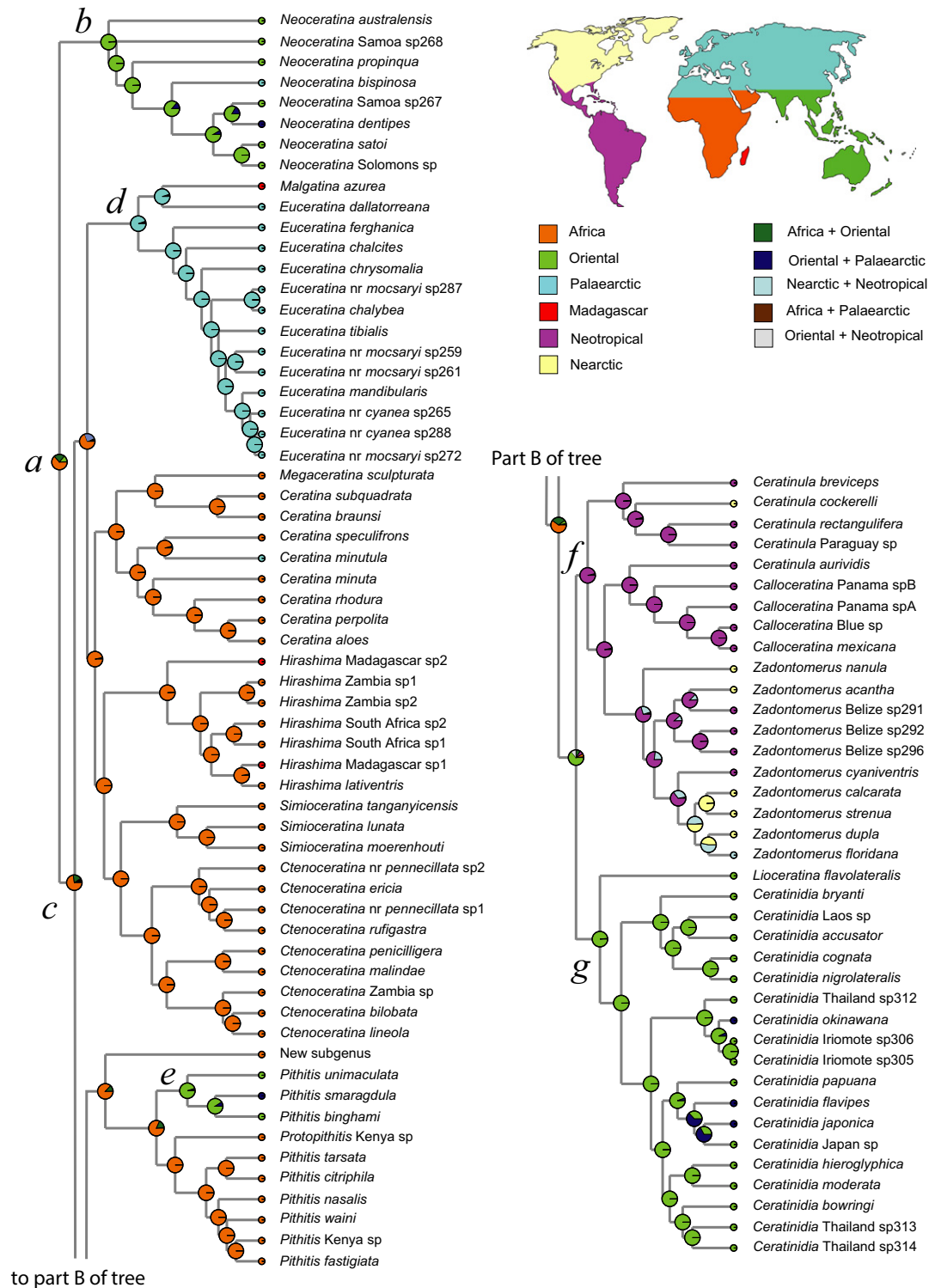


Figure 1 Maximum credibility chronogram of small carpenter bees from BEAST analysis with inferred ancestral distributions from a Bayesian binary MCMC [BBM, Markov chain Monte Carlo (MCMC)] analysis. The relative probabilities of inferred ancestral distributions are indicated by the relative ‘slices’ in the pie charts for internal nodes; ancestral distributions were permitted to simultaneously occupy two regions. Regions are colour coded in the top right-hand side of the figure. Posterior probability values for the maximum credibility chronogram are given in Fig. S1 in Appendix S1. Point-estimates and their 95% highest posterior density limits for the ages of six key clades: (a) Ceratinini, (b) *Neoceratina*, (c) all *Ceratina* except *Neoceratina*, (d) *Euceratina* and *Malgatina*, (e) Oriental *Pithitis*, (f) all New World taxa, and (g) *Lioceratina* and *Ceratinidia* are provided in Table 1.

Divergence dates

The inferred crown age of the Ceratinini was approximately 55.6 Ma, and the 95% HPD interval for our estimate was 41.5–72 Ma (Table 1). The crown age for the predominantly Palaearctic subgenus *Euceratina* was approximately 33.8 Ma. The crown age of the New World subgenera was dated at approximately 40 Ma, and the predominantly Oriental clade comprising *Lioceratina* and *Ceratinidia* had an estimated crown age of approximately 36.6 Ma (Table 1, Appendix S1: Figs S2 & S3).

Ceratinini biogeography

Our RASP BBM analysis treated the outgroup (Allodapini) as having a wide distribution (Fig. 1) or an ancestral African distribution (Appendix S1: Fig. S4) and assessed the relative probability of ancestral distributions for each *Ceratina* species. These two analyses were largely in agreement, except for a higher probability of an African ancestral range for the Ceratinini root node in the latter analysis. The S-DIVA analysis revealed an African or joint African–Oriental origin of the Ceratinini (Appendix S1: Fig. S5).

Whilst LAGRANGE analyses do not enable ready interpretation of ancestral nodes per se, they enable inference of probable ancestral ranges from nodes (Appendix S1: Table S2). Our LAGRANGE and RASP analyses were largely in agreement (Appendix S1: Fig. S6); possible discrepancies occurred in the root nodes of *Hirashima* and the New World clades, in that the more-likely distributions failed to show lnL values that were twice those of the next most likely distribution.

Long-distance dispersal and lack of back-dispersal

LDD events from Africa tended to be more frequent closer to the root (Fig. 2). When stem age was used for all LDDs out of Africa, logistic regression showed a significant effect of age (eight dispersals, Wald = 4.36, $P = 0.037$). We also found a significant effect when we restricted the analyses to

Table 1 Point-estimates and their 95% highest posterior density limits (in brackets) for the ages (in Ma) of six key small carpenter bee clades indicated by letters in Fig. 1.

Clade	Stem age (Ma)	Crown age (Ma)
Node a (Ceratinini)	79.4 (60.0, 100.8)	55.6 (41.5, 72.0)
Node b (<i>Neoceratina</i>)	55.6 (41.5, 72.0)	42.0 (25.1, 57.8)
Node c (all <i>Ceratina</i> except <i>Neoceratina</i>)	55.6 (41.5, 72.0)	51.4 (37.9, 65.4)
Node d (<i>Euceratina</i> + <i>Malgatina</i>)	47.9 (34.2, 61.2)	33.8 (22.8, 46.0)
Node e (Oriental <i>Pithitis</i>)	29.0 (19.5, 40.3)	20.4 (10.9, 30.8)
Node f (all New World taxa)	43.3 (31.2, 56.2)	40.0 (28.1, 52.0)
Node g (<i>Lioceratina</i> and <i>Ceratinidia</i>)	43.3 (31.2, 56.2)	36.6 (25.9, 49.0)

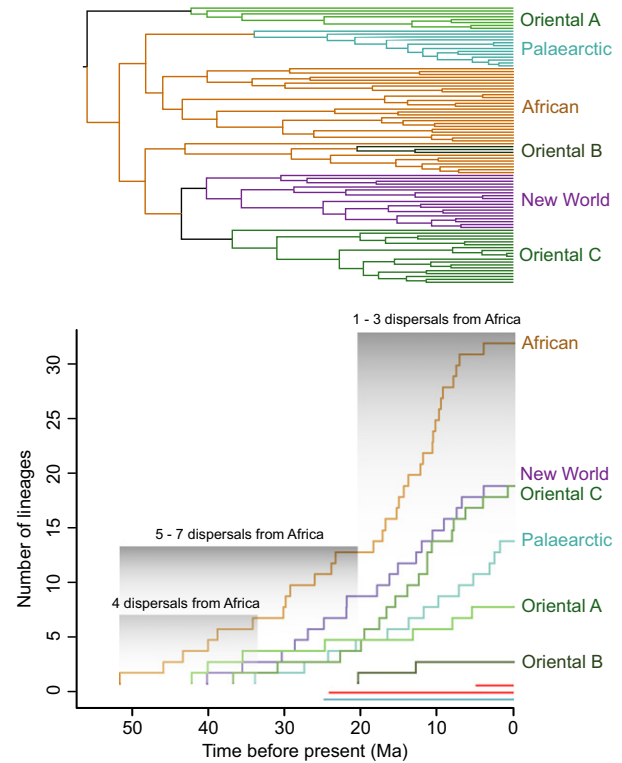


Figure 2 Summary of ‘out-of-Africa’ dispersal events of small carpenter bees. The top panel is condensed from the maximum credibility chronogram of Fig. 1 with species names omitted and Malagasy species excluded, and with regional distributions colour coded. The bottom panel shows the corresponding lineages-through-time (LTT) plots for each regional clade (excluding Madagascar), with the three Oriental clades indicated separately. The two single inferred dispersals into Madagascar and the single inferred dispersal into the Palaearctic are indicated by red and blue lines, respectively, at the bottom of the LTT plots; these are indicated by continuous lines because crown ages, the starting point of each line, also represent stem ages for each of these single-species clades and dispersal times cannot be inferred. Grey-shaded boxes summarize time frames for inferred dispersals from Africa, indicating a clustering of these dispersals closer to the root, despite accelerating lineage numbers in the African clade after approximately 20 Ma.

dispersing clades that subsequently speciated in their new regions (four dispersals, Wald = 4.97, $P = 0.026$), but the effect was non-significant when crown ages were used for the latter group (Wald = 3.31, $P = 0.069$).

There was no back-dispersal to Africa (Fig. 1). We used chi-square tests to assess the significance of this pattern by comparing the numbers of nodes that did or did not show LDD events between Africa and other regions. We excluded the New World clade from this analysis because we considered it to be unlikely that dispersal could occur from that region directly to Africa. When all out-of-Africa events were included, the chi-square analyses showed a significant difference in expected forward and backward dispersals ($\chi^2_1 = 10.39$, $P = 0.001$); there was also a significant difference when dispersals to Madagascar were excluded ($\chi^2_1 = 7.57$, $P = 0.006$).

DISCUSSION

This study provides the most comprehensive phylogeographical assessment to date of a large and diverse bee clade that has dispersed to and colonized all continents (except Antarctica) – as well as Madagascar and some Pacific islands. The larger taxon set compared to that of our previous study (Rehan *et al.*, 2010) provides greater confidence in estimating the number of long-distance dispersal events. When the BBM and S-DIVA analyses are combined, and when the bias of S-DIVA towards inferring vicariance over dispersal despite biogeographical realities is considered, the following events appear likely.

First, the ancestral ceratinine node was in Africa. The inferred crown age of the Ceratinini was nearly 9 Myr older than the crown age (47 Ma) estimated in our study of the Xylocopinae as a whole (Rehan *et al.*, 2012), but the 95% HPD interval for our estimate (41.5–72 Ma) included that earlier point estimate. Our analyses did not completely exclude a joint African/Oriental distribution for this node. However, its age was in the early to middle Eocene when the Tethys Sea was still substantial, so a joint African/Oriental distribution would have required an enormous geographical range that included what are now large parts of Europe and the Middle East, despite water barriers between Africa and those regions. A joint ancestral distribution thus seems unlikely. However, one scenario that could give rise to an apparent joint African/Asian ancestral distribution would be an early dispersal from Africa to the then-close Indian plate in the early Eocene (Fig. 3a), with this lineage then being ‘ferried’ to Asia as India moved rapidly northwards, ultimately giving rise to *Neoceratina*. Such a scenario was recently suggested for metriorrhynchine beetles (Sklenarova *et al.*, 2013), which have poor long-distance dispersal abilities but show a key divergence between African and Asian clades in the late

Palaeocene. However, dispersal distances from Africa to the Indian plate would have been very substantial (Fig. 3a).

Second, there were multiple dispersal events out of Africa. There were two LDDs from Africa to the Orient, one in the mid-Eocene that gave rise to *Lioceratina* + *Ceratinidia*, and another in the Oligocene that led to the Oriental *Pithitis* clade. Moreover, there were two dispersals from Africa to the Palaeartic, one giving rise to *Euceratina* in the late Eocene or Oligocene, and the other leading to *Ceratina minutula* and possibly to other *Ceratina s. s.* species not covered by our samples. In addition, three LDDs into Madagascar occurred, each of which comprised a single tip clade such that lower time frames for these events cannot be bracketed by crown ages. One of these Malagasy lineages, *Malgatina*, was nested within the predominantly Palaeartic subgenus *Euceratina*, but a dispersal via Africa followed by subsequent extinction of intermediate African lineages cannot be excluded.

Third, there were several dispersal events from the Oriental region to the Palaeartic, but most of these arose from apparent range extensions, which might have been associated with the extensive contiguity between those regions. Finally, there was a single dispersal into the New World from the same root clade that dispersed in the Oriental region, giving rise to a clade comprising *Ceratinula*, *Colloceratina* and *Zadontomerus*.

Because our results indicated multiple dispersals out of Africa, we considered whether these events might have occurred randomly with respect to time and whether the lack of back-dispersal to Africa could be explained simply as a function of rare event frequency.

Dispersal out of Africa

Our results suggest that the first dispersal out of Africa was to the Oriental region in the early to middle Eocene, shortly after the crown age of the Ceratinini, which resulted in the *Neoceratina* clade. At that time, the Tethys Sea separated Africa from the Orient, but the Indian plate was juxtaposed between India and what is now Southeast Asia, so this early dispersal could have involved over-water movement to the Indian plate, which later collided with the Eurasian plate.

The next dispersal out of Africa was inferred to involve a joint Oriental/Palaeartic distribution, rapidly followed by radiation of the largely Oriental *Lioceratina/Ceratinidia* clade and a single dispersal into the New World, which also resulted in a major radiation into Neotropical and Nearctic regions. These LDD events were recovered as having middle-Eocene time frames. After these events, a subsequent dispersal in the Oligocene from Africa to the Palaeartic resulted in *Euceratina*, and a Miocene dispersal from Africa to the Orient resulted in the Asian clade *Pithitis*.

The occurrence of just three ceratinine dispersals from Africa to the Orient was surprising given the near-continuous coastline that linked these regions after closure of the Tethys Sea; the analyses suggested that these dispersals were statistically biased towards older times. It is remarkable that

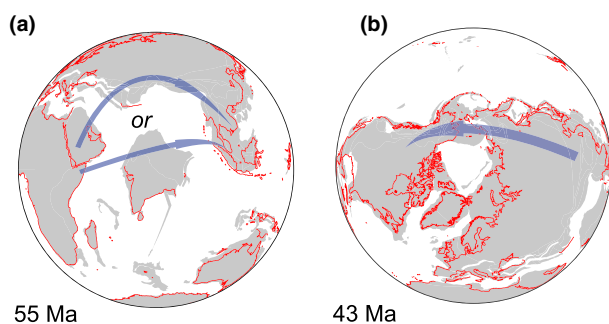


Figure 3 Reconstructed tectonic plates for geological periods corresponding to (a) the first inferred dispersal of small carpenter bees from Africa to the Oriental region in the early Eocene, and (b) the only dispersal event into the New World in the mid-Eocene. Plate fragments are indicated by grey shading, and current shorelines are indicated by red lines. Arrows indicate inferred dispersal directions from the RASP analyses. Plate reconstructions and current shorelines were obtained using the ODSN webpage (<http://www.odsn.de/odsn/services/paleomap/paleomap.html>).

there have been no apparent dispersals from Africa to Asia since the early Miocene, despite the rapid accumulation of African lineages since that time. Allodapine bees show even more restricted dispersal from Africa to Asia, with only one such dispersal in the early Miocene, despite a tribal crown age equivalent to that of the Ceratinini (Rehan *et al.*, 2012) and very similar levels of species abundance in Africa (Eardley & Urban, 2010). Like the ceratinines, there has been no back-dispersal of allodapines into Africa despite their large radiation in the Orient (Fuller *et al.*, 2005). Our results also indicated only a single direct dispersal from Africa to the Palaearctic some 27 million years ago, which again was remarkable because of the lack of substantial water barriers since closure of the Tethys Sea.

Two dispersals were indicated from Africa to Madagascar, and a third was possible if *Malgatina* represents a dispersal event from the Palaearctic to Madagascar via Africa. This third event could have occurred at any time over the last 25 Myr since *Malgatina* is a monotypic subgenus for which a stem age, but not crown age, can be estimated. The same situation occurred for the two Malagasy *Hirashima* species, with stem ages of approximately 20 Ma and 5 Ma.

A single dispersal into the New World

Another major biogeographical result from our analyses concerns the origin of the New World ceratinines. We found a single successful dispersal into the New World, approximately 43 Ma (Fig. 3b); as with the limited dispersal out of Africa, this was surprising because a substantial intermittent land bridge linked the eastern Palaearctic to the New World via Beringia. Furthermore, this dispersal occurred during a period when the global climate was substantially cooler than it was in the Eocene and late Oligocene but still warmer than that from the mid-Miocene onwards (Pearson *et al.*, 2007). Therefore, it is difficult to argue that this single dispersal was possible because of unusually favourable climatic conditions at that time.

Whilst our inferred Neotropical distribution for the ancestral New World ceratinines might seem counterintuitive given a Beringia dispersal event, glaciation cycles beginning in the Pliocene may explain this. The Pleistocene glaciations covered the central and eastern Nearctic from 3 Ma to 12 ka, leaving few refugia in the western Nearctic (Swenson & Howard, 2005). As ice sheets retreated, the eastern Nearctic could have been re-occupied by *Zadontomerus* species from southern Neotropical refugia.

Few steps forward and no steps back: a role for niche pre-emption?

Our combined results indicated that a small number of LDD events occurred, even though *Ceratina* is now common on all continents except Antarctica. The analyses further showed that dispersal back to Africa was either very rare (and not captured in our sampling) or was absent. This mismatch in

forward and back-dispersal for Africa was statistically significant. Although the single LDD event prevented a statistical comparison of dispersal into and out of the New World, the findings were remarkable given the species abundance in that region (Michener, 2007). Plate-tectonic and climatic factors did not suggest that the dearth of subsequent re-dispersal and back-dispersal was a result of physical barriers; indeed, physical barriers between the major regions in our study were thought to have declined since the initial dispersal events.

The limited forward and lack of back-dispersal in ceratinines is similar to patterns in other taxa that have led to suggestions that other assembly rules constrain the composition of fauna and flora. One important possibility here is niche pre-emption, which has been proposed as a factor in the assemblage of island biota (e.g. Carine *et al.*, 2004; Silvertown *et al.*, 2005; Cardillo *et al.*, 2008; Cooper *et al.*, 2008). Niche pre-emption could arise if early dispersing lineages radiate into unoccupied habitats, making it difficult for later immigrants from the same source clade to compete for those niches. However, niche availability could extend beyond simple resource exploitation. Early dispersers may arrive into enemy-free space that lacks parasites and predators adapted to these species, so that the 'new' niche comprises a mixture of both resources and escape from enemies. Adaptation of these early dispersers to their new habitat, and the eventual adaptation of local biota to exploit these early colonizers, might increase the challenges faced by subsequent arrivals.

The next step will be to assess alternative factors and the role of niche in community assembly (e.g. Weiher *et al.*, 2011; Wiens, 2011). A challenge arises if LDD events are intrinsically rare, in that discriminating between stochastic effects and the presence of assembly rules is problematic where very small sample sizes limit the ability to draw inferences. However, it is important to distinguish between these alternatives because each has the potential to explain major differences in regional species composition, but in very different ways. Assembly rules that include factors such as niche pre-emption and escape into enemy-free space predict that ecosystem structure is strongly influenced by early dispersal events rather than by physical barriers to dispersal per se (Cavender-Bares *et al.*, 2009; Wiens, 2012). Our data support this contention.

The importance of understanding the factors behind observed dispersal patterns has significance beyond interpreting the biogeographical history of particular taxa. Bees are a major component of pollination systems, so factors that influence the assemblage of regional bee fauna are likely to have a substantial influence on pollination ecology and ecosystem composition (Roulston & Goodell, 2011). Our findings of limited forward dispersal and the absence of back-dispersal are very similar to findings for allodapines (Schwarz *et al.*, 2006; Chenoweth & Schwarz, 2011). Our data are also similar to those for hylaeine bees, which are highly diverse in Australia and for which a single dispersal out of Australia ultimately resulted in worldwide distribution but did not lead to back-dispersal (Kayaalp *et al.*, 2013). In contrast, there is strong evidence for multiple interchanges between

Old and New World bumble bee lineages (Hines, 2008) and for multiple interchanges involving Australian and South American colletid lineages (Almeida *et al.*, 2012). Understanding the factors that permit or constrain successful dispersal is important to understanding the evolution of pollinator systems, and therefore angiosperm communities, at very broad regional scales.

ACKNOWLEDGEMENTS

Funding was provided by the University of New Hampshire to S.M.R. and by Discovery grants from the Australian Research Council to M.P.S. Special thanks to Laurence Packer, Lincoln Best, Nicholas de Silva and Fritz Gusenleitner for specimens.

REFERENCES

- Almeida, E.A.B., Pie, M.R., Brady, S.G. & Danforth, B.N. (2012) Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. *Journal of Biogeography*, **39**, 526–544.
- Cardillo, M., Gittleman, J.L. & Purvis, A. (2008) Global patterns in the phylogenetic structure of island mammal assemblages. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1549–1556.
- Cardinal, S. & Danforth, B.N. (2011) The antiquity and evolutionary history of social behavior in bees. *PLoS ONE*, **6**, e21086.
- Cardinal, S. & Danforth, B.N. (2013) Bees diversified in the age of eudicots. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122686.
- Cardinal, S., Straka, J. & Danforth, B.N. (2010) Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. *Proceedings of the National Academy of Sciences USA*, **107**, 16207–16211.
- Carine, M.A., Russell, S.J., Santos-Guerra, A. & Fransisco-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.
- Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.
- Chenoweth, L.B. (2011) *The evolution and diversification of the allodapine bees*. PhD Thesis, Flinders University, Adelaide, South Australia.
- Chenoweth, L.B. & Schwarz, M.P. (2011) Insights into the biogeographical origins and speciation history of the exoneurine allodapine bees of Australia (Hymenoptera, Apidae). *Journal of Biogeography*, **38**, 1471–1483.
- Cooper, N., Rodríguez, J. & Purvis, A. (2008) A common tendency for phylogenetic overdispersion in mammalian assemblages. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 2031–2037.
- Danforth, B.N., Brady, S.G., Sipes, S.D. & Person, A. (2004) Single-copy nuclear genes recover Cretaceous-age divergences in bees. *Systematic Biology*, **53**, 309–326.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Eardley, C.D. & Urban, R. (2010) Catalogue of Afrotropical bees (Hymenoptera: Apoidea: Apiformes). *Zootaxa*, **2455**, 1–548.
- Eardley, C.D., Ginkungu, M. & Schwarz, M.P. (2009) Bee conservation in Sub-Saharan Africa and Madagascar: diversity, status and threats. *Apidologie*, **40**, 355–366.
- Engel, M.S. (2000) A new interpretation of the oldest fossil bee (Hymenoptera: Apidae). *American Museum Novitates*, **3296**, 1–11.
- Engel, M.S. (2001) A monograph of the Baltic amber bees and evolution of the Apoidea (Hymenoptera). *Bulletin of the American Museum of Natural History*, **259**, 1–192.
- Fuller, S., Schwarz, M.P. & Tierney, S.M. (2005) Phylogenetics of the allodapine bee genus *Braunsapis*: historical biogeography and long-range dispersal over water. *Journal of Biogeography*, **32**, 2135–2144.
- Hines, H.M. (2008) Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). *Systematic Biology*, **57**, 58–75.
- Kayaalp, P., Schwarz, M.P. & Stevens, M.I. (2013) Rapid diversification in Australia and two dispersals out of Australia in the globally distributed bee genus *Hylaeus*. *Molecular Phylogenetics and Evolution*, **66**, 668–678.
- Michener, C.D. (1965) A classification of the bees of the Australian and South Pacific regions. *Bulletin of the American Museum of Natural History*, **130**, 1–362.
- Michener, C.D. (1979) Biogeography of the bees. *Annals of the Missouri Botanical Garden*, **66**, 277–347.
- Michener, C.D. (2007) *The bees of the world*, 2nd edn. Johns Hopkins University Press, Baltimore, MD.
- Pearson, P.N., van Dongen, B.E., Nicholas, C.J., Pancost, R.D., Schouten, S., Singano, J.M. & Wade, B.S. (2007) Stable warm tropical climate through the Eocene Epoch. *Geology*, **35**, 211–214.
- Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Rabinowitz, P.D., Coffin, M.F. & Falvey, D. (1983) The separation of Madagascar and Africa. *Science*, **220**, 67–69.
- Rasmussen, C. & Cameron, S.A. (2010) Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. *Biological Journal of the Linnean Society*, **99**, 206–232.
- Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
- Rehan, S.M., Chapman, T.W., Craigie, A.I., Richards, M.H., Cooper, S.J.B. & Schwarz, M.P. (2010) Molecular phylog-

- eny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. *Molecular Phylogenetics and Evolution*, **55**, 1042–1054.
- Rehan, S.M., Leys, R. & Schwarz, M.P. (2012) A mid-cretaceous origin of sociality in xylocopine bees with only two origins of true worker castes indicates severe barriers to eusociality. *PLoS ONE*, **7**, e34690.
- Rehan, S.M., Leys, R. & Schwarz, M.P. (2013) First evidence for a massive extinction event affecting bees close to the K-T boundary. *PLoS ONE*, **8**, e76683.
- Reyes, S.G. (1991) Revision of the bee genus *Braunsapis* in the Oriental Region (Apoidea: Xylocopinae: Allodapini). *University of Kansas Science Bulletin*, **54**, 179–207.
- Roig-Alsina, A. & Michener, C.D. (1993) Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). *University of Kansas Science Bulletin*, **55**, 123–162.
- Ronquist, F. (2001) *DIVA version 1.2*. Computer program and manual. Available at <http://www.ebc.uu.se/systzoo/research/diva/diva.html>.
- Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inferences under mixed model. *Bioinformatics*, **19**, 1572–1574.
- Roulston, T.H. & Goodell, K. (2011) The role of resources and risks in regulating wild bee populations. *Annual Review of Entomology*, **56**, 293–312.
- Sakagami, S.F. & Michener, C.D. (1987) Tribes of the Xylocopinae and origin of the Apidae. *Annals of the Entomological Society of America*, **80**, 439–450.
- Schaefer, H. & Renner, S. (2008) Phylogeny of the oil bee tribe Ctenoplectrini (Hymenoptera: Anthophila) based on mitochondrial and nuclear data: evidence for early Eocene divergence and repeated out-of-Africa dispersal. *Molecular Phylogenetics and Evolution*, **47**, 799–811.
- Schwarz, M.P., Fuller, S., Tierney, S.M. & Cooper, S.J.B. (2006) Molecular phylogenetics of the exoneurine allodapine bees reveal an ancient and puzzling divergence from Africa to Australia. *Systematic Biology*, **55**, 31–45.
- Shapiro, B., Ho, S.Y.W., Drummond, A.J., Suchard, M.A., Pybus, O.G. & Rambaut, A. (2011) A Bayesian phylogenetic method to estimate unknown sequence ages. *Molecular Biology and Evolution*, **28**, 879–887.
- Silvertown, J. (2004) The ghost of competition past in the phylogeny of island endemic plants. *Journal of Ecology*, **92**, 168–173.
- Silvertown, J., Francisco-Ortega, J. & Carine, M. (2005) The monophyly of island radiations: an evaluation of niche pre-emption and some alternative explanations. *Journal of Ecology*, **93**, 653–657.
- Sklenarova, K., Chesters, D. & Bocak, L. (2013) Phylogeography of poorly dispersing net-winged beetles: a role of drifting India in the origin of Afrotropical and Oriental fauna. *PLoS ONE*, **8**, e67957.
- Swenson, N.G. & Howard, D.J. (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.
- Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T. & Bentivenga, S. (2011) Advances, challenges and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2403–2413.
- Wiens, J.J. (2011) The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2336–2350.
- Wiens, J.J. (2012) Why biogeography matters: historical biogeography vs. phylogeography and community phylogenetics for inferring ecological and evolutionary processes. *Frontiers in Biogeography*, **4**, 128–135.
- Yoder, A.D. & Nowak, M.D. (2006) Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 405–431.
- Yu, Y., Harris, A.J. & He, X.J. (2011) *RASP: Reconstruct Ancestral State in Phylogenies v.2.0.1.0, beta 1, build 110304*. Available at: <http://mnh.scu.edu.cn/soft/blog/RASP>.
- Yuan, Y.M., Wohlhauser, S., Möller, M., Klackenberg, J., Callmander, M.W. & Küpfer, P. (2005) Phylogeny and biogeography of *Exacum* (Gentianaceae): a disjunctive distribution in the Indian Ocean basin resulted from long distance dispersal and extensive radiation. *Systematic Biology*, **54**, 21–34.

SUPPORTING INFORMATION

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

Appendix S1 Additional tables (Tables S1 & S2) and figures (Figs S1–S6).

BIOSKETCHES

Sandra Rehan studies social evolution and genetics. Her research on the origin of social behaviour combines behavioural ecology, comparative genomics and molecular phylogenetics to answer questions about the ecological constraints and genetic underpinnings selecting for social behaviour.

Michael Schwarz focuses on evolution, combining molecular phylogenetics, behaviour, biogeography, diversity and ecology. His research uses bees in Australia, Africa, Madagascar, the Pacific and Asia as a model system.

Editor: Peter Pearman