

Molecular phylogeny of Indo-Pacific carpenter ants (Hymenoptera: Formicidae, *Camponotus*) reveals waves of dispersal and colonization from diverse source areas

Ronald M. Clouse^{a,b,*}, Milan Janda^{c,d}, Benjamin Blanchard^{e,†}, Prashant Sharma^b, Benjamin D. Hoffmann^f, Alan N. Andersen^f, Jesse E. Czekanski-Moir^{g,§}, Paul Krushelnycky^h, Christian Rabeling^{d,†}, Edward O. Wilson^d, Evan P. Economoⁱ, Eli M. Sarnat^j, David M. General^k, Gary D. Alpert^d and Ward C. Wheeler^b

^aDepartment of Bioinformatics and Genomics, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC, 28223, USA; ^bDivision of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th St., New York City, NY, 10024, USA; ^cBiology Centre, Czech Academy of Sciences, Branisovska 31, 370 05, Ceske Budejovice, Czech Republic; ^dMuseum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA, 02138, USA; ^eDepartment of Ecology and Evolutionary Biology, University of Michigan, 2019 Kraus Natural Science Building, Ann Arbor, MI, 48109, USA; ^fCSIRO Land and Water Flagship, Darwin, NT, 0822, Australia; ^gDepartment of Biology, Program in Ecology and Evolutionary Biology, University of Oklahoma, Norman, OK, 73019, USA; ^hDepartment of Plant and Environmental Protection Sciences, University of Hawai'i at Mānoa, 3050 Maile Way, Room 310, Honolulu, HI, 96822, USA; ⁱOkinawa Institute of Science and Technology, 1919-1 Tancha, Onna-son, Kunigami gun, Okinawa, Japan, 904-0495; ^jDepartment of Entomology, University of Illinois, 320 Morrill Hall, 505 South Goodwin Avenue, Urbana, IL, 61801, USA; ^kUniversity of the Philippines at Los Baños, Museum of Natural History, Los Baños, 4031, Laguna, Philippines; [†]Present address: Committee on Evolutionary Biology, University of Chicago, 1025 E. 57th Street, Culver Hall 402, Chicago, IL, 60637, USA; [§]Present address: State University of New York, College of Environmental Science and Forestry, One Forestry Drive, 246 Illick Hall, Syracuse, NY, 13210, USA; [†]Present address: Department of Biology, University of Rochester, River Campus Box 270211, Rochester, NY, 14627, USA

Accepted 2 September 2014

Abstract

Ants that resemble *Camponotus maculatus* (Fabricius, 1782) present an opportunity to test the hypothesis that the origin of the Pacific island fauna was primarily New Guinea, the Philippines, and the Indo-Malay archipelago (collectively known as Malesia). We sequenced two mitochondrial and four nuclear markers from 146 specimens from Pacific islands, Australia, and Malesia. We also added 211 specimens representing a larger worldwide sample and performed a series of phylogenetic analyses and ancestral area reconstructions. Results indicate that the Pacific members of this group comprise several robust clades that have distinctly different biogeographical histories, and they suggest an important role for Australia as a source of Pacific colonizations. Malesian areas were recovered mostly in derived positions, and one lineage appears to be Neotropical. Phylogenetic hypotheses indicate that the orange, pan-Pacific form commonly identified as *C. chloroticus* Emery 1897 actually consists of two distantly related lineages. Also, the lineage on Hawai'i, which has been called *C. variegatus* (Smith, 1858), appears to be closely related to *C. tortuganus* Emery, 1895 in Florida and other lineages in the New World. In Micronesia and Polynesia the *C. chloroticus*-like species support predictions of the taxon-cycle hypothesis and could be candidates for human-mediated dispersal.

© The Willi Hennig Society 2014.

The relationship between the biotas of Pacific islands and those of nearby large landmasses is a long-standing line of inquiry among biogeographers. A funda-

mental outstanding question is whether Pacific islands were primarily colonized via New Guinea, the Philippines, and the Indo-Malay Archipelago (“Malesia”) (Gressitt, 1974; Holloway, 1998; Keppel et al., 2009), from presumably Southeast Asian ancestors, or have a more complex and diverse history. Like so many bio-

*Corresponding author:
E-mail address: ronaldmclouse@gmail.com

geographical puzzles in the Indo-Pacific, Alfred Russell Wallace made the first significant contribution with his recognition of Eurasian and Australian biotas on either side of the Lombok and Makassar Straits (Wallace, 1859). Various modifications to Wallace's Line followed (Huxley, 1868; Simpson, 1977), but Mayr (1944) first articulated the argument that the divide depended on dispersal ability and ecological requirements and only applied to animals, perhaps only vertebrates. In contrast, tropical-adapted plants in Southeast Asia, being generally good dispersers, readily colonized across Malesia, being outcompeted by Australian colonizers only in drier areas.

Gressitt (1974) proposed that the New Guinean, Melanesian, Micronesian, and even Polynesian insect faunas were fundamentally Eurasian ("Oriental") in origin (areas are labelled in Fig. 1). Insects are good at crossing oceanic barriers, and their colonization from Southeast Asia is presumably because of abiotic affinities (i.e. the shared tropical environment of the Pacific and Southeast Asia). However, recent molecular studies suggest that Pacific island histories are more complex than a simple colonization from Malesia (Garb and Gillespie, 2006; Keppel et al., 2009; Gillespie et al., 2012), although it is still the main source of the region's biota. Australia (Harbaugh et al., 2009) and the New World (Baldwin and Wagner, 2010; Sharma and Giribet, 2011) now appear to be more significant sources of biodiversity for the South Pacific region than previously believed, islands have been recognized as candidate sources for widespread lineages (instead of merely being recipients of mainland dispersers) (Filardi and Moyle, 2005; O'Grady and Desalle, 2008; Balke et al., 2009), and colonizations apparently do not always proceed in a stepping-stone fashion (Filardi and Moyle, 2005; Keppel et al., 2009). However, the basis for any general statement on Pacific island colo-

nization—especially for animals and including remote island archipelagos—is lacking.

In this context we conducted a molecular phylogenetic analysis of Indo-Pacific carpenter ants in the diverse group of *Camponotus maculatus*-like species. Begun as an expanding number of subspecies under the West African species *Camponotus maculatus* (Fabricius, 1782), this group is now a highly diverse but taxonomically unstable constellation of lineages found mostly in the Indo-Pacific and Africa (Donisthorpe, 1915; McArthur and Leys, 2006). Members of the group in Australia were recently analysed using mitochondrial sequence data (McArthur and Leys, 2006), inferring a lack of support for a close relationship between Australian and African lineages, raising *C. humilior* Forel, 1902 to species level, and identifying *C. crozieri* McArthur and Leys, 2006 as a novel species (McArthur and Leys, 2006). Still, based on morphology and considering the full array of African and Indo-Pacific species (McArthur and Leys, 2006, had only South African samples in their study), *C. maculatus*-like species appear to have extensive Old World taxonomic affinities. Thus, our starting hypothesis was that ancestry of Pacific lineages can generally be traced back to Malesia, with lineages on islands in the remote Pacific being derived from those to the West, and those in Malesia being derived from those in Asia and Africa. We also hypothesized that northern Australian lineages were derived from New Guinea through the Torres Strait region, or perhaps from eastern Indonesia directly through north-western Australia. Finally, following Gressitt (1982) and basic tenets of island biogeography relating island remoteness and size to colonization (MacArthur and Wilson, 1963; Gillespie and Roderick, 2002), we hypothesized that different species on remote islands were sister taxa resulting from single colonization events.

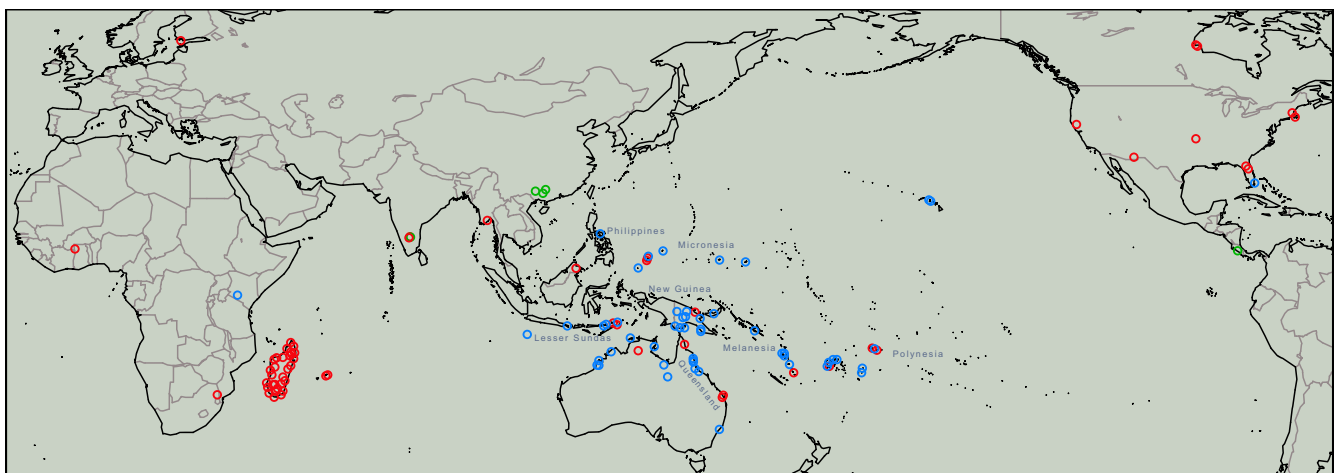


Fig. 1. Localities of all specimens used in this study. Specimens with red circles had only COI sequences available, and those with blue circles had other markers available. Green circles indicate specimens used in follow-up analyses (see supplementary material).

Methods

Specimen collection and selection

We assembled an extensive range of tissues from *C. maculatus*-like specimens via new field collections and sampling older, mounted specimens (Fig. 1). We included *C. chloroticus* Emery, 1897, *C. eperiamorum* Clouse 2007, *C. irritans kubaryi* Mayr, 1876, *C. humilior*, the *C. dorycus* (Smith 1860) group, and a wide variety of other *C. maculatus*-like morphotypes. We also downloaded sequences for all *maculatus*-like specimens found on GenBank and BOLD (Ratnasingham and Hebert, 2007), and, given the instability of *C. maculatus* taxonomy and its dubious monophyly, we included a selection of all *Camponotus* subgenera with publicly available data, especially those with more than one genetic marker available.

Of the 357 specimens used for this study, we collected and sequenced 119 ourselves. In addition we sequenced 27 specimens from older museum collections. Sequences for 211 terminals were gathered from public databases (BOLD, GenBank), of which 56 were previously published. The sources, museum and author research collections where specimens are deposited, and data coverage are provided in supplementary Table S1. Images of specimens representing different terminals collected by the authors will be available online at the following public databases: www.newguineants.org, www.antweb.org, and www.antwiki.org.

We aimed to conduct our analyses with minimal assumptions about the relationships among different *Camponotus* lineages, including the monophyly of *C. maculatus*-like species. Some *Camponotus* species in the region, such as certain Micronesian ones described by Clouse (2007a), were not included both because they are extremely rare, known from only a few historical specimens, and also because they present a number of characters at considerable variance from those generally associated with *C. maculatus*. However, we did include the widespread Pacific species *C. reticulatus* Roger, 1863 and the supposedly convergently evolved Australian species *C. claripes* Mayr, 1876. After preliminary analyses, some outgroups, such as *C. tortuganus* Emery, 1895, placed among *C. maculatus* specimens, and thus we added more *Camponotus* terminals for which we could find public data, including representatives of all *Camponotus* subgenera. For further outgroups, we used *Odontomachus simillimus* Smith, 1958 and *Strumigenys* specimens (for which we had fresh tissues and success at sequencing most genetic markers), as well as publicly available data from other formicine genera. *O. simillimus* was used to root the phylogeny, because we were able to obtain from it the complete gene sampling that we utilized for the ingroup.

DNA sequencing

Total DNA was extracted using Qiagen's DNEasy® tissue kit (Valencia, CA, USA). We then amplified the mitochondrial markers cytochrome *c* oxidase I (COI) and cytochrome b (CytB), and the nuclear markers carbomoyl-phosphate synthase II (CAD), elongation factor-1 α (EF-1 α), long wavelength rhodopsin (LR), and wingless (Wg). Primer sequences, fragment lengths, and temperature protocols are provided in Table 1. The touchdown temperature profile “TD6” is from Chenail et al. (2010). Fifty of the publicly available COI sequences extended as much as 831 bases past our reverse primer, and the COI sequences from the *C. maculatus*-group study of McArthur and Leys (2006) had only the end part of this longer fragment. Sequence data were inspected and edited using Sequencher™ (Gene Codes Corp., Ann Arbor, MI, USA), SeaView v. 4 (Gouy et al., 2010), and BioEdit (Hall, 2007).

Phylogenetic reconstructions

We performed three independent phylogenetic analyses on three different selections of terminals. Terminal sets had 337 members (95.6% of all terminals) with at least one mitochondrial marker (“any-one”; many of which were obtained from public databases with only COI sequenced), 164 members (46.5%) with at least two markers, one of which was mitochondrial (“any-two”), and 132 members (37.4%) with at least three markers (“any-three”). For the “any-one” terminals, the COI sequence needed to be nearly complete in order for the data to be included in the alignment. The goal of making three data sets was to explore the tradeoff between increased amounts of data versus increased numbers of terminals. In the latter case, we would expect resampling supports to be low, but we would also be able to see the placement of interesting terminals that have only COI available (such as *C. floridanus*). Our main concern was avoiding the inclusion of terminals that had only a small amount of nuclear data, as those did not provide as much variation as mitochondrial data, and such terminals would probably place randomly and also lower resampling support.

Most sequences with only publicly available COI sequences had 658 bp that overlapped with the fragment we generated, sometimes with a maximum of 30 bases missing from either end of the DNA sequence. Thirty terminals had an additional 831 bases of COI, which overlapped with the 12 terminals included from the McArthur and Leys (2006) *C. maculatus*-group study, which utilized only this latter 831-bp fragment of COI and CytB. This latter fragment of COI was analysed as a separate partition in

Table 1

Primers (forward above, reverse below) and protocols used for sequencing DNA in this study, with approximate marker size, intron size, indel number, and relevant references

Forward primer	Protocol	Aligned length (bp)	Intron aligned length (bp)	Indels	Citation
Cytochrome <i>c</i> oxidase subunit I (COI)					
LCO1490	GGTCAACAAATCATAAAGATATTGG	46*	658	n/a	0
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA				
Cytochrome <i>b</i> (CytB)					
Cb1-fw	TATGTACTACCATGAGGACAAATATC	CytB-ant [†]	433	n/a	0
CB2	ATTACACCTCCTAATTTATTAGGAAT				
Carbamoyl-phosphate synthase II (CAD)					
CD892F	GGYACCGGRCGTTGYTAYATGAC	TD53 [‡]	769	308	17–50
CD1491R	GCCGCARTTNAGRGCRGTYTGYCC				
Elongation factor 1-alpha (EF-1 α)					
F1-1424-fw	GCGCCKGCGGCTCTCACCACCGAGG	EF1a-ant [§]	359	n/a	0
F1-1829-rev	GGAAGCCTCGACGCACATMGG				
Long wavelength rhodopsin (LR)					
LR143F	GACAAAGTKCCACCRGARATGCT	TD6 [¶]	588	123	20–50
LR639ER	YTTACCGRITCCATCCRAACA				
Wingless (Wg)					
578F	TGCACNGTGAARACYTGCTGGATGCG	TD53 [‡]	409	49	0–18
1032R	ACYTCGCAGCACCARTGGAA				

*[94 °C (1:00 min)], 35 cycles of [94 °C (0:30), 46 °C (0:30), 72 °C (1:30)], [72 °C (10:00)].

[†][94 °C (1:00 min)], 35 cycles of [94 °C (1:00), 50 °C (1:00), 72 °C (1:00)], [72 °C (10:00)].

[‡][94 °C (1:00 min)], [94 °C (1:00), 63–53 °C (1:00), 72 °C (1:00)], four cycles of [94 °C (1:00), 53 °C (1:00), 72 °C (1:00)], 25 cycles of [94 °C (0:30), 58 °C (0:45), 72 °C (0:45)], [72 °C (3:00)].

[§][94 °C (3:00 min)], [94 °C (0:30), 70–62 °C (0:45), 72 °C (1:30)], 25 cycles of [94 °C (0:30), 61 °C (0:30), 72 °C (1:00)], [72 °C (7:00)].

[¶][94 °C (1:00 min)], [94 °C (1:00), 58–45 °C (1:00), 72 °C (1:00)], 25 cycles of [94 °C (0:30), 58 °C (0:45), 72 °C (0:45)], [72 °C (3:00)].

tree searches. One specimen, BOLD ID ANIND001-11, identified as *C. compressus* (once a subspecies of *C. maculatus*) from India, had only the 658-bp fragment, but it was missing 92 bases at the end and was extremely unstable in preliminary analyses; being recovered in nearly any possible position under different search strategies and resampling calculations, it was therefore excluded from the main analyses. It was the only terminal excluded based on instability and a lack of data. One specimen with an even shorter fragment of the 658-bp COI fragment, a specimen of *C. aurosus* from Madagascar, was not particularly unstable and was included alongside many other Madagascar specimens, including another also identified as *C. aurosus*.

We aligned the sequences in MAFFT (Katoh et al., 2002) under default settings. Aligned marker and intron lengths, as well as indel numbers are provided in Table 1. We performed tree searches under the maximum-likelihood criterion in RAxML (Stamatakis et al., 2008) on the CIPRES (Miller et al., 2010) computing cluster, using GTR and a partitioned model. The same alignment was used to conduct tree searches using Bayesian inference, using MrBayes v. 3.1.2

(Ronquist and Huelsenbeck, 2003) with a unique model of sequence evolution with corrections for a discrete gamma distribution and a proportion of invariant sites specified for each partition, as selected in Modeltest v. 3.7 (Posada and Crandall, 1998) under the Akaike information criterion. Default priors were used starting with random trees, and four runs, each with three hot Markov chains and one cold Markov chain, were performed until the average deviation of split frequencies reached < 0.01 (2–107 generations). After burnin samples were discarded, sampled trees were combined in a single majority consensus topology, and the percentage of nodes was taken as posterior probability. Finally, we searched under dynamic homology and parsimony in POY 5 (Varón et al., 2009), using an equal transformation cost scheme (gaps, transversions, and transitions all costing 1). Initial search strategies for the any-two and any-three terminal sets were determined automatically by the program using the “search(max_time DD:HH:MM)” command and various other cost transformation schemes (combinations of gaps costing 1–4 times the cost of transversions, transversions costing 1–4 times transitions, and gap-openings costing 3 with

gap-extensions 1). For the any-one terminals, starting trees of initial searches were obtained from RAxML searches and random builds in POY, as Wagner building under dynamic homology with so much missing data taxed memory resources; these trees were then read into POY and swapped using SPR and TBR under different cost schemes. Finally, the best trees from all preliminary searches were read into POY and put through a final round of swapping under equal costs, and bootstrap support was calculated using 100 pseudoreplicates, aligned data (“static_approx”), and the best tree as the starting tree.

Ancestral area reconstructions

We used the Bayesian ancestral area reconstruction program implemented in RASP (Yu et al., 2011), coding each terminal for its region as indicated in Table S1. RASP is designed specifically for ancestral area reconstructions, giving different costs to inferred biogeographical events, such as vicariance and dispersal, and it is derived from the programs DIVA (Ronquist, 1997) and S-DIVA (Yu et al., 2010). Ancestral areas were not constrained, and a probability was allowed for each coded region. Because the number of areas that can be input to the program is limited, we combined Fiji and Polynesia, Micronesia and Palau, New Guinea and the Solomons, and the Holarctic and Japan for reconstructions on the any-one and any-two phylogenies. Vanuatu was coded separately in all reconstructions. These combinations had no impact on results, as they consisted of regions whose specimens were recovered as closely related in the final trees.

COI branch length optimizations

We explored the feasibility of dating diversification events in *C. maculatus*-like lineages by generating an ultrametric tree with 95% confidence intervals for node heights. Using only the COI partition, we took the tree for the any-two terminal set recovered under maximum-likelihood and made it ultrametric with optimized branch lengths in the program BEAST (Drummond and Rambaut, 2007), as implemented on the CIPRES computing cluster (Miller et al., 2010). Eight terminals that did not have COI in the original terminal set were removed from the tree before the analysis. Setting the mutation rate to 1, with a lognormal relaxed clock, we ran four chains for 5 million generations with a GTR + Γ model and empirical base frequencies. All runs stabilized with nearly identical likelihoods (log likelihood range < 1), so the maximum clade credibility tree was derived from the run with the highest log likelihood effective sample size, using a burnin of 500 000 generations.

Results

Phylogenetic hypotheses

From the tree searches on the any-two and any-three terminal sets we recovered trees that had five main clades holding most of the terminals of interest (Fig. 2, clades I–V, support values in Table 2). Using the any-one terminal set, Clade II, containing African and Malagasy lineages, became mixed with the large number of Old World terminals added from public databases (Fig. S1). The remaining four clades were generally recovered and supported using the any-one terminal set (clades III and IV under parsimony being exceptions) and can be characterized as follows: Clade I contains *C. tortuganus* from southern Florida and *C. variegatus* (Smith, 1858) from Hawai'i (pictured in Fig. S2), as well as members of the *C. dorycus* (Smith, 1860) group (which includes species in New Guinea and *C. dorycus confusus* Emery, 1887 from Australia); Clade III is a relatively small clade containing species from the Lesser Sundas and the Philippines; Clade IV contains some darker lineages from Micronesia, including *C. eperiamorum*, a terminal from Christmas Island and other lineages from New Guinea, and then a large clade containing all the specimens from Fiji, Samoa, Tonga, and some Vanuatu and New Guinean terminals that have been called *C. chloroticus*; and, finally, Clade V contains the bulk of Australian, New Guinean, Vanuatuan, and Micronesian specimens, including lineages previously identified as *C. novaehollandiae* Mayr, 1870, *C. crozieri*, *C. humilior*, and *C. chloroticus*. Clades III and IV tended to be recovered as sister groups, and in the any-one analyses, when two *C. floridanus* (Buckley, 1866) COI sequences were included, that species and Clade I were recovered as monophyletic in maximum-likelihood and Bayesian searches.

The main clades, especially IV and V, were recovered under a variety of conditions. They were stable under different cost schemes in POY, the any-two and any-three phylogenies when gaps and transversions cost 2 (instead of 1, like transitions) being identical to those from equal costs, for example; the main difference with the any-one phylogeny when gap and transversion costs were higher was that Clade II was recovered as a paraphyletic grade at the base of Clade IV. The clades, especially I, IV, and V, were usually recovered under separate partitions, and all five clades were recovered when using just the nuclear marker Wg (separate phylogenies for Wg, LR, EF-1 α , and CytB are given in Fig. S3). *Camponotus dolendus* Forel, 1892, which had two markers and was recovered in Clade II in the likelihood any-two phylogeny, was unstable, placing near or sister to Clade III under Bayesian and parsimony, and among a large group of

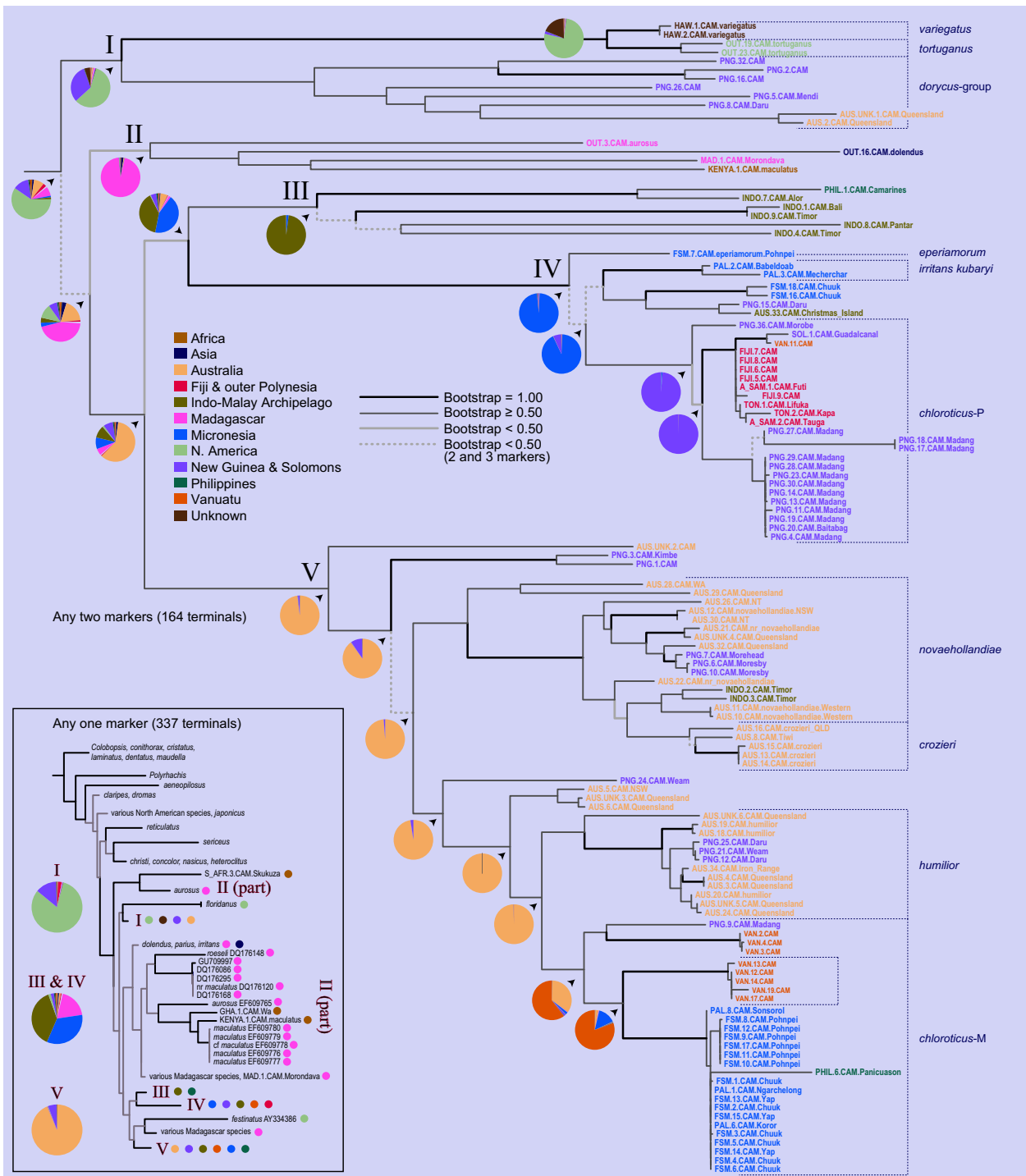


Fig. 2. Maximum-likelihood phylogram of *Camponotus maculatus*-like species in the Pacific, using 164 terminals that had at least two of the seven markers sequenced (outgroups not shown). Branches reflect bootstrap supports as calculated in RAxML: black for 100%, dark grey for 50%–99%, light grey for less than 50%, and dashed light grey for < 50% in the analysis using terminals with any three sequenced markers. Pie charts at nodes show probabilities of different ancestral area reconstructions, as calculated in a Bayesian framework in RASP. Current taxonomic designations are noted to the right of certain Pacific clades, with the *chloroticus*-like forms distinguished by a regional name for Micronesia and Polynesia. A clade of Vanuatu specimens recovered inside of *chloroticus*-M but containing large black forms (such as VAN.17 and VAN.19) is identified by a second bracket inside *chloroticus*-M. The inset shows a simplified phylogeny recovered under maximum-likelihood using 337 terminals with any one marker sequenced. Most clades are collapsed with geographical designations shown as dots to the right, and reconstructions for major Pacific clades are shown to the left. The clade where specimens identified as *C. maculatus* have been recovered is expanded.

Table 2

Bootstrap support values, as a percentage, for each of the major clades (I–V) and each terminal set (any-one etc.), using maximum-likelihood (ML), parsimony (P), and Bayesian (B) analyses

Clade	Any one			Any two			Any three		
	ML	P	B	ML	P	B	ML	P	B
I	69	66	88	82	81	99	80	84	99
II	–	–	–	47	–	–	57	92	97
III	70	59	95	85	–	99	87	84	100
IV	98	97	100	100	100	100	100	100	100
V	89	53	100	98	91	100	99	96	100

mostly Malagasy terminals in the any-one phylogeny. Under parsimony *C. floridanus* also placed inside this group, away from a sister relationship with Clade I, as found from likelihood and Bayesian analyses.

We found that *C. maculatus*-like species in the Pacific were composed of several distinct lineages that originated via diverse colonization pathways. What has been considered a single species, *C. chloroticus*, was inferred as two species in our analysis, one predominately occurring in Micronesia, the other mostly in Polynesia, with overlapping ranges in Melanesia. Ancestral area reconstructions pointed to Australia and even North America as likely sources for different clades, and deep origins for the group remain unresolved.

Ancestral area reconstructions

Ancestral area reconstructions supported the following origins for the five main clades in the any-two phylogeny recovered under maximum-likelihood (Fig. 2): I—North America (58%, New Guinea 31%), II—Madagascar (95%), III—Indonesia (98%), IV—Micronesia (98%), and V—Australia (95%). New Guinea was reconstructed as the ancestral locality for species in Polynesia (excluding Hawai'i; Clade IV), but before that the ancestral location was probably Micronesia and before that possibly Australia. For the large clade of Australian, New Guinean, Vanuatuan, Micronesian, Indonesian and Philippine species (Clade V), the ancestral locality was reconstructed as Australia. Other than the Polynesian colonization within Clade IV, New Guinea was reconstructed as an endpoint for several lineages, and Queensland ancestors appear to have diversified in all directions, including into Malesia. Ancestral area reconstructions for Clades III+IV and V in the any-two phylogeny (Fig. 2) were interpreted on maps as historical pathways (Figs 3 and 4, respectively). These reconstructions suggest that Clade IV has most recently moved south from Micronesia into New Guinea and Polynesia (Fig. 3), a result driven by the consistent recovery of Micronesian species as the early lineages in Clade IV, and Polynesian ones more

derived (Fig. 2). The large number of lineages in Clade V appear to have originated in Queensland, which was the source of multiple colonizations of Melanesia, one of which subsequently colonized Micronesia (Fig. 4).

Branch length optimizations and dating

Optimizing COI branch lengths on the any-two phylogeny using a variable clock in BEAST produced an ultrametric tree with $\ln L = -15\,566.291$ (Fig. S4). Confidence intervals were large, however, often half as long as branch lengths themselves. Combined with ambiguities over the proper mitochondrial mutation rate, we found dating attempts unconvincing. Divergence rates for ants have been estimated to range from about 2%/Ma (Quek et al., 2004; Steiner et al., 2006; Goropashnaya et al., 2007) to 3%/Ma (Solomon et al., 2008; Pennings et al., 2011), and even as high as 5%/Ma (Leppanen et al., 2011). Taking into account the 95% confidence interval on the initial diversification of *C. maculatus*-like lineages in the Pacific (which has a per-site mutation probability in the COI tree of 0.2976–0.4535), as well as the range of mitochondrial mutation rates suggested for ants (half the diversification rate, 0.01–0.025), we would estimate the diversification to have started approximately 12–45 Ma. An alternative approach was to use the emergence of the Micronesian island of Pohnpei and its endemic species, *C. eperiamorum*, as a rough calibration point, but even this date, as well as that of other Micronesian islands, is difficult to ascertain. Volcanic activity continued on Pohnpei from 0.9 to 8.7 Ma (Hafiz Ur et al., 2013), and when the island was truly above water and quiet enough to allow the permanent establishment of a flora and fauna is not clear.

Discussion

Our results do not support the hypothesis that *C. maculatus*-like species colonized the Pacific directly from Malesia (Indonesia, the Philippines, and New Guinea). Instead of being recovered near the base of our phylogenies, specimens from Indonesia and the Philippines were recovered either together in Clade III, which had little effect on the group's deepest ancestral area reconstruction, or in highly derived positions in clades IV and V. New Guinean specimens were recovered in a variety of positions, often derived. The colonization of Polynesia and its reconstruction back to New Guinea is the clearest west-to-east colonization we recovered, and the Polynesian species appears to be the result of a recent colonization from the Solomons or Vanuatu. However, it is difficult to draw from this any support

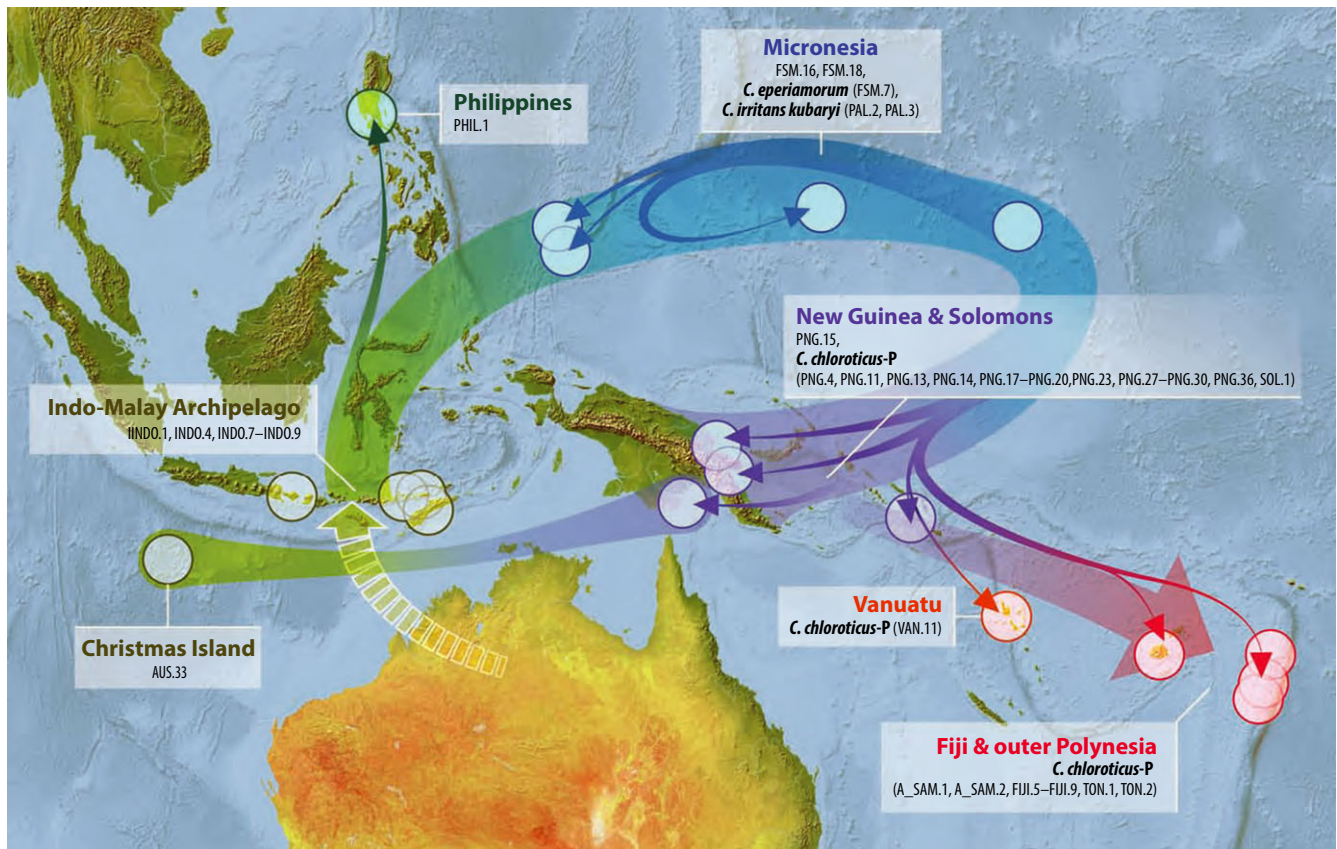


Fig. 3. Map showing origin and possible dispersal route for Clade III+IV, based on the tree recovered under maximum-likelihood and the Bayesian ancestral reconstruction model implemented in RASP (Fig. 2). The reconstruction of Australia as the origin is ambiguous, as is the location of the ancestor of AUS.33 on Christmas Island and PNG.15 in Daru.

for a general out-of-Malesia colonization route in Polynesia, as in Clade V the species that extends throughout Micronesia and into the Philippines appears to be the result of a dispersal out of Vanuatu, which itself originated in Queensland.

Morphology and taxonomy

We have noted species determinations where they are the clearest, and we found that what has been understood as *C. chloroticus*—a yellow–orange lineage found throughout the Pacific—is actually polyphyletic; the proper application of this name is the subject of a separate study, and here we simply distinguish the Polynesian and Micronesian lineages as “*chloroticus*-P” and “-M.” The important aspect of *C. maculatus*-group morphology and taxonomy for our test of colonization histories in the Pacific is that they frequently give a misleading measure of diversity. Also, the molecular data show the presence of multiple distantly related clades on remote islands, which does not support Gressitt’s (1982) hypothesis that diversity in such places is probably the result of a single colonization event. For example, in Micronesia, endemic species on

Pohnpei, Chuuk, and Palau that could easily be interpreted from their morphologies as local variations or speciations out of *C. chloroticus*-M (Clade V) are actually members of a distantly related lineage (Clade IV) that was on the islands perhaps millions of years before; and in Vanuatu, “*C. chloroticus*” is actually comprised of members of clades IV and V (Fig. 5). It appears that *C. maculatus*-like species in the Pacific, in addition to being a collection of long-separated lineages that have followed odd routes around the islands, are also capable of rapidly generating new morphologies, both divergent and convergent, and thus easily confounding taxonomic studies, as has been recently noted in the ant *Pheidole megacephala* (Wills et al., 2014) and an Indo-Pacific bird species (Irestedt et al., 2013). Outside of the Pacific, and indeed outside of the *C. maculatus*-group, the genus as a whole is morphologically confounding, with few of our clades conforming to subgeneric designations. Indeed, similar to the results of Brady et al. (2006), even the genus *Camponotus* as a whole was not recovered as monophyletic, the subgenus *Colobopsis* falling outside *Polyrhachis* and the remaining *Camponotus* in the any-one analyses.

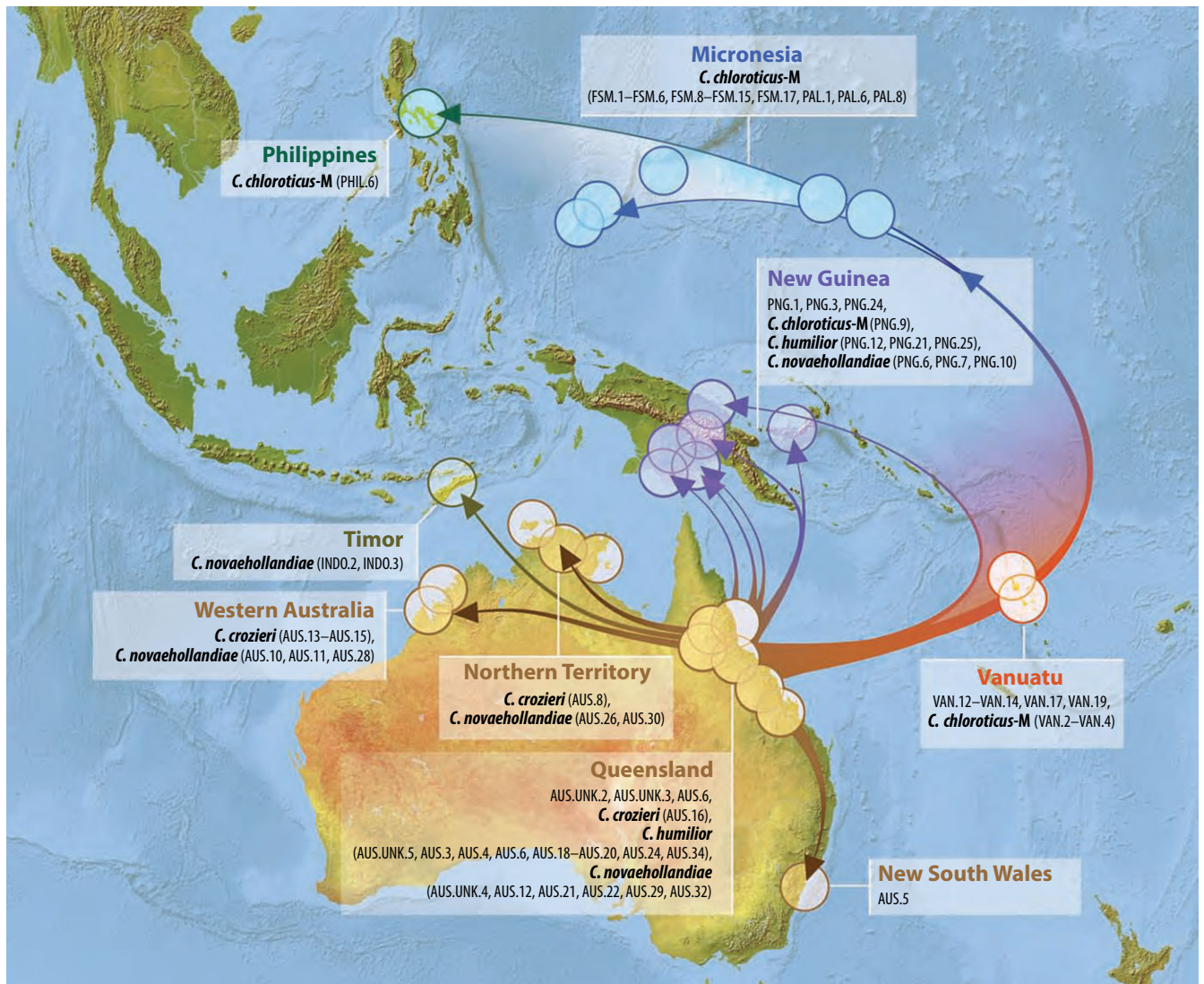


Fig. 4. Map showing origin and possible dispersal route for Clade V, based on the tree recovered under maximum-likelihood and the Bayesian ancestral reconstruction model implemented in RASP (Fig. 2).

In general we do find support for most *maculatus*-like lineages in the Pacific originating from a common ancestor, as they are usually recovered as more closely related to each other than to other Pacific groups, like the widespread Pacific species *C. reticulatus* or the convergent *C. claripes* and related species in southern Australia. However, we also find that a close connection between the Pacific lineages and *C. maculatus* and its allies in Africa or even Asia is difficult to support when using a large sample of species.

Deep origins

Most of the *C. maculatus*-like species in the Pacific (Clade V and perhaps clades III and IV) appear to have derived from the Australian wet tropics, but they are not closely related to the endemic Gondwanan

fauna there, which includes *C. aeneopilosus* Mayr, 1862, *C. dromas* Santschi 1919, and *C. claripes*. Presumably the Queensland ancestor originated in Southeast Asia, but the signal of that connection is clouded by back-tracking and a lack of Asian material. Relationships among the Pacific lineages and Asian and African species in our best trees are unstable, with short branch lengths and low resampling support.

The close relationship between *C. tortuganus* and *C. variegatus* (Fig. S2) is one of the most stable and supported relationships in our study but not readily explained. Counter to a hybridization explanation (which might confuse phylogenetic reconstruction), the relationship was recovered using separate nuclear and mitochondrial partitions (with high support from RAxML, 100% with Wg, 99% with LR, 83% with EF-1 α , and 100% with CytB; Fig. S3), and the



Fig. 5. Selection of specimens from Vanuatu, illustrating morphological variation and convergence in the *maculatus*-group.

sequences are different enough between the species that one does not appear to be an unusual colour morph of the other. We also tested for long branch attraction by removing one species or the other and rerunning tree searches in RAxML; we still recovered Clade I with *C. tortuganus* or *C. variegatus*, with high support (70–85%), regardless of the terminal set, and *C. floridanus* was still recovered as sister to Clade I in the any-one analyses. *C. tortuganus* was originally described in the *maculatus*-group, and it remains in the subgenus *Tanaemyrmex*, but *C. floridanus*, which placed close to *C. tortuganus* and Clade I in any-one analyses, is in the subgenus *Myrmothrix*.

The origins of Pacific *maculatus*-group lineages prior to Australia will probably only be resolved by analysing a global sample of fresh specimens from which

the full complement of molecular markers can be obtained, as well as additional collections from the Pacific. To this end we did a follow-up analysis using CytB sequences reported for a collection from southern China (including one identified as *C. variegatus*) (Pang et al., 2009), COI sequences of three unidentified species from Costa Rica, and COI from an Indian specimen identified as the former *maculatus*-group species *C. compressus* Fabricius, 1787 (excluded earlier due to a lack of data and instability). The resulting placements suggest that Clade I is indeed part of a large Neotropical lineage (Figs S5 and S6), that Clade III is part of a larger Southeast Asian lineage, and that species determinations based on morphology, including that of *C. variegatus*, are easily misled by convergence. We also recovered *C. compressus* from India as sister

to Clade V (with high bootstrap support), a promising indication of the utility of Asian samples in clarifying the origin of Clade V prior to Australia.

We found that the sister species of what is called *C. variegatus* on Hawai'i is *C. tortuganus* in Florida, and that this pair probably has a close relationship with some of the other New World species in our sample. This raises several questions about this species which, given our focus on Indo-Pacific and Old World species, we are unable to resolve here. Is the Hawaiian *Camponotus* the same species as those identified as *C. variegatus* in Sri Lanka, Singapore, and China (Pang et al., 2009)? We have no way of knowing without examining and sequencing those specimens ourselves. In addition, is it an ancient inhabitant of Hawai'i or a more recent arrival of an unknown species from the Americas, perhaps even mediated by human activity? Again, we are unable to determine that here without more specimens from the New World. If the Hawaiian *Camponotus* is a native Hawaiian ant, that would contradict the long-held belief that Hawai'i has no native ants (Zimmerman, 1948; Wilson and Taylor, 1967; Wilson, 1996). This possibility may not be as radical as it sounds, for the hypothesis that remote islands in the eastern Pacific may have completely lacked ants before human contact has recently been challenged by subfossil data, reports of which are under preparation (N. Porch, pers. comm.).

We remain open to the hypothesis that hybridization has contributed to the lack of deep phylogenetic resolution and of distinct morphological characters that are useful for taxonomic purposes. Using a manageable subset of our data to develop a parsimony network analysis, it was found that certain species were best explained as the direct results of reticulate events (the joining of lineages through time) in the most parsimonious network (Kannan and Wheeler, 2014), and we recommend expanding this line of investigation with *Camponotus* going forward.

Drivers of colonization

Ecology, reproductive biology, and human-mediated transfer are important sources of complexity in colonization patterns (Foucaud et al., 2010; Cerdá et al., 2011; Rabeling and Kronauer, 2013; Waters et al., 2013). We note that *C. chloroticus*-M and *C. chloroticus*-P show the pattern of favouring marginal habitats (Clouse, 2007b; Sarnat and Economo, 2012) while also (i) covering large ranges and (ii) being two of the more recent lineages in our phylogeny. This is consistent with Wilson's hypothesis (Wilson, 1959, 1961) of the historical taxon cycle, which postulates that lineages first evolve morphologies and behaviour that allow them to expand their ranges from restricted local areas. The large range eventually divides as popula-

tions adapt to local conditions, become specialists, and speciate. They tend to be found in interior habitats as local endemics, especially during times of human colonization, when coastal regions undergo extensive disturbance and homogenization among islands. This then sets the stage for the cycle to repeat. This can perhaps be seen in Clade IV on Pohnpei Island, where *C. eperiamorum*, an older colonist than *C. chloroticus*-M, is now restricted to upland, native forest and is an island endemic of only moderate abundance. For most other *maculatus*-like lineages in the region, detailed habitat and range data have not been compiled, but our phylogeny offers a first step to identifying these lineages and using them to understand the processes of historical colonization in the Pacific. One aspect of colonization emphasized in the original taxon-cycle hypothesis, as well as the later theory of island biogeography (MacArthur and Wilson, 1963), was the role of large landmasses as sources of island faunas, but here we find that not necessarily to be the case, as has been seen recently in other groups (Filardi and Moyle, 2005; O'Grady and Desalle, 2008; Balke et al., 2009).

The taxon cycle has been tested in Melanesia using ants in the genus *Pheidole* (Economo and Sarnat, 2012), where a comprehensive data set has been built over many years. Key predictions of the hypothesis were supported, such as the connectivity of species in lowland habitats. In addition, the impact of marginal habitat being converted to human-disturbed zones was presented as a new area of inquiry, perhaps setting the stage for a global taxon cycle driven by the modern wave of human-associated exotics. Perhaps the *chloroticus*-like lineages in Micronesia and Polynesia are early examples of this, having been spread primarily by ancient human travel. Their preference for low-elevation and coastal habitats puts them in closer contact with people, whose voyages are based on cultural connections and would more easily explain the odd combination of wide but constrained distributions among specific groups of oceanic islands. Nonetheless, other than *C. variegatus*, none of the *maculatus*-like species in the Pacific has been hypothesized to be exotic, and our phylogeny does not suggest the presence of any pan-tropical species in the group.

Acknowledgements

We thank the following funders for supporting this work: DARPA (W911NF-05-1-0271), Marie Curie Fellowship (PIOFGA2009-25448), Czech Science Foundation (P505/12/2467), Czech Ministry of Education Grant (6007665801), US National Science Foundation (DEB 0515678), Putnam Expedition Grants (Museum of Comparative Zoology), the Harvard Society of Fellows, the Australian Studies Committee of Harvard,

and the American Museum of Natural History's Research Experience for Undergraduates program (NSF). We are grateful to Stefan Cover at the Museum of Comparative Zoology for assistance with sampling mounted specimens. Simon Robson at James Cook University generously made available specimens from the late Ross H. Crozier's collection, the localities for which were determined with the assistance of Ellen Schluens and Corrie Moreau. We are grateful to P. Riha for the graphic design of maps, to M. Borovanska and E. Youngerman for specimens and data management, to the staff of the New Guinea Binatang Research Center for field assistance, to V. Novotny, S. E. Miller and N. Pierce for assistance with Papua New Guinea (PNG) projects, to the PNG Department of Environment and Conservation for assistance with research permits, to the Department of Environmental Protection and Conservation of the Republic of Vanuatu for assistance with research permits, and to Miss Donna Kalfatak and the citizens of Penaoru village for assistance with field work. Research in Micronesia was made possible through the generous assistance of the Chuuk Visitor's Bureau, the College of Micronesia, The Conservation Society of Pohnpei, Yap State Department of Resources and Development, the Belau National Museum, and the University of Guam. Special field assistance on the island of Tol (in Chuuk) was kindly provided by the Techuo family, and on Pohnpei Island by Amos Eperiam and Nixon Daniel. Specimen loans from Palau were arranged by Alan Olsen. Daniel Janies and the Department of Bioinformatics and Genomics at UNC Charlotte provided important computing and imaging support. Brian Fisher helped us locate sequences of important Malagasy and African specimens. Anupam Kumar and Sameer Siddiqi investigated hybridization and prepared data for network analysis. We thank the three anonymous reviewers for their helpful suggestions on earlier versions of this manuscript.

References

- Abouheif, E., Wray, G.A., 2002. Evolution of the gene network underlying wing polyphenism in ants. *Science* 297, 249–252.
- Baldwin, B.G., Wagner, W.L., 2010. Hawaiian angiosperm radiations of North American origin. *Ann. Bot.* 105, 849–879.
- Balke, M., Ribera, I., Hendrich, L., Miller, M.A., Sagata, K., Posman, A., Vogler, A.P., Meier, R., 2009. New Guinea highland origin of a widespread arthropod supertramp. *Proc. Biol. Sci.* 276, 2359–2367.
- Brady, S.G., Schultz, T.R., Fisher, B.L., Ward, P.S., 2006. Evaluating alternative hypotheses for the early evolution and diversification of ants. *Proc. Natl Acad. Sci.* 103 (48), 18172–18177.
- Cerdá, X., Angulo, E., Caut, S., Courchamp, F., 2011. Ant community structure on a small Pacific island: only one native species living with the invaders. *Biol. Invasions* 14, 323–339.
- Chenuil, A., Hoareau, T.B., Egea, E., Penant, G., Rocher, C., Aurelle, D., Mokhtar-Jamai, K., Bishop, J.D.D., Boissin, E., Diaz, A., Krakau, M., Luttikhuisen, P.C., Patti, F.P., Blavet, N., Mousset, S., 2010. An efficient method to find potentially universal population genetic markers, applied to metazoans. *BMC Evol. Biol.* 10, 276.
- Clouse, R.M., 2007a. New ants (Hymenoptera: Formicidae) from Micronesia. *Zootaxa* 1475, 1–19.
- Clouse, R.M., 2007b. The ants of Micronesia (Hymenoptera: Formicidae). *Micronesica* 39, 171–295.
- Donisthorpe, H., 1915. The type of *Camponotus* (*Myrmoturba*) *maculatus*. *Entomol. Rec. J. Var.* 27, 221–222.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Economu, E.P., Sarnat, E.M., 2012. Revisiting the ants of Melanesia and the taxon cycle: historical and human-mediated invasions of a tropical archipelago. *Am. Nat.* 180, E1–E16.
- Filardi, C.E., Moyle, R.G., 2005. Single origin of a pan-Pacific bird group and upstream colonization of Australasia. *Nature* 438, 216–219.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R.C., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotech.* 3, 294–299.
- Foucaud, J., Orivel, J., Loiseau, A., Delabie, J.H.C., Jourdan, H., Konghouleux, D., Vonshak, M., Tindo, M., Mercier, J.-L., Fresneau, D., Mikissa, J.-B., McGlynn, T., Mikheyev, A.S., Oettler, J., Estoup, A., 2010. Worldwide invasion by the little fire ant: routes of introduction and eco-evolutionary pathways. *Evol. Appl.* 3, 363–374.
- Garb, J.E., Gillespie, R.G., 2006. Island hopping across the central Pacific: mitochondrial DNA detects sequential colonization of the Austral Islands by crab spiders (Araneae: Thomisidae). *J. Biogeogr.* 33, 201–220.
- Gillespie, R.G., Roderick, G.K., 2002. Arthropods on islands: colonization, speciation, and conservation. *Annu. Rev. Entomol.* 47, 595–632.
- Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R., Roderick, G.K., 2012. Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol. Evol.* 27, 47–56.
- Goropashnaya, A.V., Fedorov, V.B., Seifert, B., Pamilo, P., 2007. Phylogeography and population structure in the ant *Formica exsecta* (Hymenoptera, Formicidae) across Eurasia as reflected by mitochondrial DNA variation and microsatellites. *Ann. Zool. Fenn.* 44, 462–474.
- Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27, 221–224.
- Gressitt, J.L., 1974. Insect biogeography. *Annu. Rev. Entomol.* 19, 293–321.
- Gressitt, J.L., 1982. Pacific-Asian biogeography with examples from the Coleoptera. *Entomol. Generalis* 8, 1–11.
- Hafiz Ur, R., Hideo, N., Kei, K., 2013. Geological origin of the volcanic islands of the Caroline group in the Federated States of Micronesia, Western Pacific. *South Pacific Stud.* 33, 101–118.
- Hall, T., 2007. BioEdit, Ver. 7.0.9. Ibis Biosciences, An Abbott Company. Available at: <http://www.mbio.ncsu.edu/bioedit/page2.html>.
- Harbaugh, D.T., Wagner, W.L., Allan, G.J., Zimmer, E.A., 2009. The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *J. Biogeogr.* 36, 230–241.
- Holloway, J.D., 1998. Geological signal and dispersal noise in two contrasting insect groups in the Indo-Australian tropics: R-mode analysis of pattern in Lepidoptera and cicadas. In Hall, R., Holloway, J.D. (eds), *Geological Signal and Dispersal Noise in Two Contrasting Insect Groups in the Indo-Australian Tropics: R-Mode Analysis of Pattern in Lepidoptera and Cicadas*. Backhuys Publishers, Leiden, The Netherlands, pp. 291–314.
- Huxley, T.H., 1868. On the classification and distribution of the Alectoromorphae and Heteromorphae. *J. Zool.* 1868, 294–319.

- Irestedt, M., Fabre, P.H., Batalha-Filho, H., Jonsson, K.A., Roselaar, C.S., Sangster, G., Ericson, P.G., 2013. The spatio-temporal colonization and diversification across the Indo-Pacific by a 'great speciator' (Aves, *Erythropitta erythrogaster*). *Proc. Biol. Sci.* 280 (1759), 20130309.
- Jermiin, L.S., Crozier, R.H., 1994. The cytochrome b region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: sequence divergence in Hymenoptera may be associated with nucleotide content. *J. Mol. Evol.* 38, 282–294.
- Kannan, L., Wheeler, W.C., 2014. Exactly computing the parsimony scores on phylogenetic networks using dynamic programming. *J. Comput. Biol.* 21 (4), 303–319.
- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Keppel, G., Lowe, A.J., Possingham, H.P., 2009. Changing perspectives on the biogeography of the tropical South Pacific: influences of dispersal, vicariance and extinction. *J. Biogeogr.* 36, 1035–1054.
- Leppanen, J., Vepsalainen, K., Savolainen, R., 2011. Phylogeography of the ant *Myrmica rubra* and its inquiline social parasite. *Ecol. Evol.* 1, 46–62.
- MacArthur, R.H., Wilson, E.O., 1963. An equilibrium theory of insular zoogeography. *Evolution* 17, 373–387.
- Mayr, E., 1944. Wallace's line in the light of recent zoogeographic studies. *Q. Rev. Biol.* 19, 1–44.
- McArthur, A.J., Leys, R., 2006. A morphological and molecular study of some species in the *Camponotus maculatus* group (Hymenoptera: Formicidae) in Australia and Africa, with a description of a new Australian species. *Myrmecol. Nachr.* 8, 99–110.
- Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" In Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, Louisiana. pp 1–8.
- O'Grady, P., Desalle, R., 2008. Out of Hawaii: the origin and biogeography of the genus *Scaptomyza* (Diptera: Drosophilidae). *Biol. Lett.* 4, 195–199.
- Pang, G.-F., Qin, F., Zhou, S.-Y., Qin, X.-M., 2009. Molecular phylogenetic research on some species of *Camponotus* (Hymenoptera: Formicidae). *J. Guangxi Normal Univ.* 27, 62–66.
- Pennings, P.S., Achenbach, A., Foitzik, S., 2011. Similar evolutionary potentials in an obligate ant parasite and its two host species. *J. Evol. Biol.* 24, 871–886.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Quek, S.P., Davies, S.J., Itino, T., Pierce, N.E., 2004. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). *Evolution*, 58, 554–570.
- Rabeling, C., Kronauer, D.J.C., 2013. Evolution of thelytokous parthenogenesis in eusocial Hymenoptera. *Annu. Rev. Entomol.* 58, 273–292.
- Ratnasingham, S., Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Mol. Ecol. Notes* 7, 355–364.
- Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.*, 46 (1), 195–203.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sarnat, E., Economo, E.P., 2012. The ants of Fiji. *Univ. Calif. Publ. Entomol.* 132, 1–398.
- Schultz, T.R., Brady, S.G., 2008. Major evolutionary transitions in ant agriculture. *Proc. Natl Acad. Sci. USA* 105 (14), 5435–5440.
- Sharma, P.P., Giribet, G., 2011. The evolutionary and biogeographic history of the armoured harvestmen - *Laniatores* phylogeny based on ten molecular markers, with the description of two new families of Opiliones (Arachnida). *Invertebr. Syst.* 25, 106–142.
- Simpson, G.G., 1977. Too many lines; the limits of the Oriental and Australian zoogeographic regions. *Proc. Am. Philos. Soc.* 121, 107–120.
- Solomon, S.E., Mauricio Bacci, J., Joaquim Martins, J., Vinha, G.G., Mueller, U.G., 2008. Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS ONE* 3, e2738.
- Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.* 57, 758–771.
- Steiner, F., Schlick-Steiner, B., Konrad, H., Linksvayer, T., Quek, S., Christian, E., Stauffer, C., Buschinger, A., 2006. Phylogeny and evolutionary history of queen polymorphic *Myrmecina* ants (Hymenoptera: Formicidae). *Eur. J. Entomol.* 103, 619–626.
- Varón, A., Sy Vinh, L., Wheeler, W.C., 2009. POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics* 26, 72–85.
- Wallace, A.R., 1859. Letter from Mr. Wallace concerning the geographical distribution of birds. *Ibis* 1, 449–454.
- Ward, P.S., Downie, D.A., 2005. The ant subfamily Pseudomyrmecinae (Hymenoptera: Formicidae): phylogeny and evolution of big-eyed arboreal ants. *Syst. Entomol.* 30, 310–335.
- Ward, P.S., Brady, R.H., Fisher, B.L., Schultz, T.R., 2010. Phylogeny and biogeography of dolichoderine ants: effects of data partitioning and relict taxa on historical inference. *Syst. Biol.* 59 (3), 342–362.
- Waters, J.M., Fraser, C.I., Hewitt, G.M., 2013. Founder takes all: density-dependent processes structure biodiversity. *Trends Ecol. Evol.* 28, 78–85.
- Wills, B.D., Moreau, C.S., Wray, B.D., Hoffmann, B.D., Suarez, A.V., 2014. Body size variation and caste ratios in geographically distinct populations of the invasive big-headed ant, *Pheidole megalcephala*. *Biol. J. Linn. Soc.* 113, 423–438.
- Wilson, E.O., 1959. Adaptive shift and dispersal in a tropical ant fauna. *Evolution* 13, 122–144.
- Wilson, E.O., 1961. The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* 95, 169–193.
- Wilson, E.O., 1996. Hawaii: a world without social insects. *Bishop Mus. Occas. Pap.* 45, 3–7.
- Wilson, E.O., Taylor, R., 1967. The ants of Polynesia (Hymenoptera: Formicidae). *Pac. Insects Mon.* 14, 1–109.
- Yu, Y., Harris, A.J., He, X.-J., 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. *Mol. Phylogenet. Evol.* 56 (2), 848–850.
- Yu, Y., Harris, A.J., He, X.-J., 2011. RASP (Reconstruct Ancestral State in Phylogenies), Ver. 2.0 beta. Available at <http://mnh.scu.edu.cn/soft/blog/RASP>.
- Zimmerman, E.C., 1948. *Insects of Hawaii, Volume 1: Introduction*. University of Hawaii Press, Honolulu.

Supporting Information

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

Fig. S1. All phylogenetic hypotheses generated in this study. Trees in the same row have the same selection of terminals (any one marker (at least one mitochondrial), any two markers, and any three markers). Trees in the same column were analysed under the same optimality criterion, alignment, and cost scheme. The left column contains trees analysed in RAxML under maximum-likelihood and a partitioned model, in the middle column are trees analysed in MrBayes using the same alignment as in RAxML, and in the

right column are trees analysed in POY using dynamic homology for length-variable markers and a cost scheme in which transitions, transversions, and gaps cost 1. The top middle tree (B) has taxonomic points of interest noted in red.

Fig. S2. Specimens of *C. tortuganus* from Florida and *C. variegatus* from Hawai'i sequenced in this study. Scale bar = 1.0 mm for frontal views, 1.5 mm for lateral views.

Fig. S3. Phylogenetic hypotheses recovered under maximum-likelihood using single markers and all terminals that had each of those markers sequenced (except COI, which is the main driver of the any-one phylogenies), outgroups not shown, and *C. maculatus*-group clades (I–V) noted and coloured. The clade containing the Hawaiian and Floridian specimens is shown with a star and bootstrap values.

Fig. S4. Phylogeny of terminals with any two markers recovered under maximum-likelihood, then with COI branch lengths optimized in BEAST. Error bars represent 95% confidence intervals.

Scale bar below represents the per-site mutation probability.

Fig. S5. Detailed view of any-one (A), any-two (B), and any-three (C) analyses including specimens from southern China (orange terminal names) and Costa Rica (purple terminal names). Clades IV and V have been collapsed for readability, and likelihood bootstrap support values for all nodes are shown. Two specimens identified as *C. variegatus* are highlighted with arrows: one from Hawai'i (black arrows) and the other from China (orange arrows).

Fig. S6. Hypothetical Neotropical origin and Pacific dispersal route for Clade I.

Table S1. All terminals used in phylogenetic analyses, with basic taxonomic information, original collection codes, locality information, GenBank numbers, and participating phylogenetic analyses noted. Taxonomic information is provided for publicly available sequences and known island endemics and is based upon AntCat (antcat.org).