The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2016, ••, ••–••. With 7 figures.

Nuclear DNA based species delineations of *Coccus* scale insects in symbiosis with plants and ants, and the role of plant epicuticular wax in structuring associations

SWEE-PECK QUEK[1*,†], SHOUHEI UEDA[2,3†], PENNY J. GULLAN[4], TAKUMASA KONDO[5], MITSURU HATTORI[2,6], TAKAO ITIOKA[7], KAORI MURASE[8] and TAKAO ITINO[2,9]

[1]*Department of Integrative Biology, University of California, Berkeley, CA, 94720, USA*
[2]*Department of Biology, Faculty of Science, Shinshu University, 3-1-1 Asahi, Matsumoto, Nagano, 390-8621, Japan*
[3]*Graduate School of Life and Environmental Science, Osaka Prefecture University, 1-1 Gakuen-cho, Nakaku, Sakai, Osaka, 599-8531, Japan*
[4]*Division of Evolution, Ecology & Genetics, Research School of Biology, The Australian National University, Acton, Canberra, ACT, 2601, Australia*
[5]*Corporación Colombiana de Investigación Agropecuaria (CORPOICA), Centro de Investigación Palmira, Calle 23, Carrera 37, Continuo al Penal, Palmira, Valle, Colombia*
[6]*Graduate School of Fisheries and Environmental Sciences, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki, 852-8521, Japan*
[7]*Graduate School of Human and Environmental Studies, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, 606-8501, Japan*
[8]*Graduate School of Natural Sciences, and Research Center for Biological Diversity, Nagoya City University, 1 Yamanohata, Mizuho-cho, Mizuho-ku, Nagoya, Aichi, 467-8501, Japan*
[9]*Institute of Mountain Science, Shinshu University, 3-1-1 Asahi, Matsumoto, Nagano, 390-8621, Japan*

Received 11 June 2016; revised 26 August 2016; accepted for publication 26 August 2016

We undertook phylogenetic analysis of nuclear DNA to elucidate species boundaries in the symbiotic *Coccus* scale insects associated with mutualistic *Crematogaster* ants and *Macaranga* plants occurring in the ever-wet forests of Southeast Asia. The coccid specimens clustered into ten lineages, each corresponding to a morphospecies assignment. The lineage identified as *C. secretus* was separated from the Main Clade by an outgroup. We also examined all pairwise associations among the three symbiont guilds to understand how patterns of association were structured. The analyses revealed that each ant, plant or coccid operational (taxonomic) unit often associated with multiple O(T)Us of each of the other two guilds. However, where testing was feasible, a 'preference' for one or sometimes two partner O(T)Us of each guild was often detected. Mutual 'preferences' or 'avoidances' were relatively common among the symbionts, and no conflicts of interest were apparent. The network of preferred partners among all three guilds showed compartmentalization structured by the presence/absence of plant epicuticular wax, suggesting that this feature plays a fundamental role in how the symbionts select partners that best serve their needs. To a lesser degree, the network was also structured by whether the host plant stems were ant-excavated or hollowed naturally. © 2016 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2016, **00**, 000–000.

KEYWORDS: ant–plant mutualism – Borneo – *Crematogaster borneensis* group – host specificity – Malay Peninsula – myrmecophyte – plant–insect interactions – rain forest – Sundaland.

*Corresponding author. E-mail: queksp@gmail.com
[†]These authors contributed equally to this work.

INTRODUCTION

The ever-wet forests of Southeast Asia, spanning Borneo, parts of the Malay Peninsula, and Sumatra, harbour a myrmecophyte system comprising *Macaranga* plants and *Crematogaster* ants which depend specifically on each other for survival (Fiala *et al.*, 1989). A third member of this system, *Coccus* hemipterans (Fig. 1), is essential to the survival of the ants (Heckroth, Fiala & Maschwitz, 2001; Handa & Itioka, 2011), and therefore of the plants as well, yet research on them has lagged far behind that of the ants and plants. Only five studies have been published to date (Heckroth *et al.*, 1998, 2001; Ueda *et al.*, 2008, 2010; Handa *et al.*, 2012) since the first taxonomic descriptions in 1921 by Morrison and 1952 by Takahashi, compared to well over 50 for the ants and plants combined.

In providing domatia in hollow stems, lipid- and protein-rich food bodies (Heil *et al.*, 1998), sugar-rich phloem sap via the conduit of *Coccus* scale insects (hereafter referred to as coccids), and in some cases extra-floral nectar, the plants provide all resources needed by the ants. The plants and ants also supply all resources needed by the coccids by providing phloem sap as food, and ant-protected domatia as refuge. Heckroth *et al.* (2001) noted that coccids themselves were never consumed by the ants even under starvation conditions, and that adult ants consumed coccid honeydew while the food bodies were harvested for the ant brood. In the greenhouse, with the absence of herbivores and competing vines, usually ant-inhabited *Macaranga* are able to grow and thrive without ants (Fiala & Maschwitz, 1992). In their native habitat, without the protection of their specific resident ants, the plants are eaten by insects and smothered by vines, leading to near-certain death unless recolonized quickly by those ants (Fiala *et al.*, 1989; Itioka *et al.*, 2000; Heil *et al.*, 2001). The myrmecophytes (referring to the host plants) occur in three lineages of *Macaranga* (Davies, 2001; Davies *et al.*, 2001; but see Blattner *et al.*, 2001): sections *Pachystemon*, *Pruinosae*, and *Winklerianae*, all of which house *Crematogaster* ants. This paper focuses primarily on the symbiosis involving sections *Pachystemon* and *Pruinosae*, which predominantly host ants in the *Crematogaster borneensis* group (*sensu* Feldhaar, Maschwitz & Fiala, 2016). Section *Winklerianae* was excluded because we lack collections of their inhabitant coccids and ants, and the latter are not part of the *C. borneensis* group.

Figure 1. Worker(s) of *Crematogaster borneensis* group tending *Coccus* sp. (possibly *C. penangensis*) in opened stems of *Macaranga bancana*. Photos by T. Komatsu.

Based on several studies looking at ant (morpho) species (Fiala *et al.*, 1999; Feldhaar *et al.*, 2003a, 2016), mitochondrial (mt) DNA (Feldhaar *et al.*, 2003a; Quek *et al.*, 2004, 2007), and nuclear DNA microsatellite genotyping (Feldhaar, Gadau & Fiala, 2010; Ueda *et al.*, 2015) we know that the associations between *Macaranga* species groups and *Crematogaster* (morpho)species and DNA lineages show a fair-to-high degree of specificity. This specificity has been attributed in part to the presence/absence of epicuticular wax crystals secreted on the terminal stem sections and leaves of some *Macaranga* species (Federle *et al.*, 1997; Federle, Rohrseitz & Hölldobler, 2000; Quek *et al.*, 2004), conferring a glaucous (powdery, bluish-white) appearance.

Among the *Coccus* scale insects associated with *Macaranga*, seven species have been named (Morrison, 1921; Takahashi, 1952). Much of what is known about their patterns of association with the ants and plants derives from Heckroth *et al.* (1998). Their survey of the hollow internodes of 19 *Macaranga* species (843 coccid specimens) retrieved 22 morphospecies with varying degrees of abundance and specificity to *Macaranga* species/clades and/or ant morphospecies. Further, their surveys indicated an obligate association between the coccids and ant-

inhabited *Macaranga*, noting that most of the *Coccus* morphospecies were associated only with *Macaranga*, although three morphospecies were also found elsewhere (to varying degrees).

The *Coccus* scale insects of *Macaranga* are taxonomically challenging even for coccid taxonomists. Under field conditions, identification is virtually impossible partly due to their small size and because species boundaries have yet to be worked out, with many species unnamed. Phylogenetic analysis based on mtDNA has so far proven unhelpful in sorting them into morphologically coherent groups (Ueda *et al.*, 2008; P. J. Gullan, unpubl. data).

Here, we investigate whether phylogenetic analysis of nuclear DNA from *Macaranga*-inhabiting *Coccus* specimens can delimit morphologically coherent groups that might suggest where species boundaries lie. We also analyse all pairwise associations between the ants, plants and coccids collected in this study, present a synthesis of the patterns of association, focusing on the role of plant epicuticular wax in structuring those patterns, and speculate on the underlying mechanisms by which those patterns arise.

MATERIAL AND METHODS

Sampling

One hundred and 51 adult female *Coccus* specimens were sequenced from collections made from 145 trees representing 19 *Macaranga* species (25 *Macaranga* species in sections *Pachystemon* and *Pruinosae* are known to harbour ants in the *Crematogaster borneensis* group) from 12 locations: six in Borneo and six in Malaya (for this paper, Malaya includes Singapore, peninsular Malaysia, and Bintan island, Indonesia). Collections from Sumatra are absent. Most of the coccids reported here were used in Ueda *et al.* (2008, 2010). Sampling of coccids was haphazard (as opposed to random, which denotes a statistical process designed to achieve an unbiased sample reflecting the diversity of the population at large). When different (to the naked eye) coccid forms were encountered, all forms were collected. For outgroups, we selected two *Coccus* species (*C. hesperidum* and *C. viridis*) and a member of the genus *Pulvinaria*, which belongs to the same subfamily (Coccinae) as *Coccus* (Hodgson, 1994). *C. hesperidum* is a close relative of the *Coccus* species from *Macaranga* (Lin *et al.*, 2013). Supporting Information (Table S1) shows the collection localities, GenBank accession numbers, host plant species, and mtDNA lineage identities of the attendant *Crematogaster* ant symbionts (as used by Quek *et al.*, 2007) associated with the *Coccus* specimens used here (see also file *Coccus List Table S1.xlsx*, doi: 10.5061/dryad.6q1q6). The mtDNA lineage identity of the coccids as published by Ueda *et al.* (2010) is also provided in Supporting Information (Table S1). Supporting Information (Table S2) provides a tally of the number of: (1) specimens, (2) collecting localities and (3) coccid species associated with each host plant species. A map of coccid collection localities can be found in Quek *et al.* (2007, Fig. 3). Collecting permission was obtained from: (1) Sabah Parks, (2) Sarawak Forest Department, and (3) National Parks Board, Singapore.

DNA extraction, voucher specimens, polymerase chain reaction and sequencing

DNA was extracted from single 95%-ethanol-preserved adult female coccids using a 'salting-out' protocol (Sunnucks & Hales, 1996). Exoskeletons were left intact and slide-mounted using a modification of the method of Williams & Granara de Willink (1992). Most specimens were cleared by placing them in cold 10% KOH overnight and then gently heating to 40 °C for a few hours before expressing the body contents in water to which a drop of detergent was added. Cuticles were then stained for several hours in acidified alcohol containing a few drops of acid fuchsin solution, prior to dehydration in a series of alcohol baths, and then transferred through three xylene baths prior to mounting in Canada balsam on microscope slides. Each specimen was assigned to species or morphospecies by P. J. Gullan and/or T. Kondo based on cuticular morphology examined under a compound microscope. These voucher specimens will be deposited in the Australian National Insect Collection (ANIC), CSIRO, Canberra, when work on their morphological taxonomy is completed and published.

Two nuclear genes, *elongation factor 1* α (*EF-1*α) and *wingless* (*WG*) were amplified by polymerase chain reaction (PCR) using TaKaRa Ex *Taq* polymerase (TaKaRa Bio, Shiga, Japan). *WG* sequences were amplified using the primers WG1 and WGR0 reported in Brower & Desalle (1998) and Braby, Vila & Pierce (2006), respectively. *EF-1*α sequences were amplified using the following primers designed by S. Ueda. (1) Coc-efs-3: 5′-TAA AGC CGA CGG TAA ATG CCT-3′ (2) Coc-efs-4: 5′-CAG GAT GTG TAC AAA ATT GGT-3′ (3) Coc-efa-3: 5′-ACA CTT CAT CCA TTC GAT TGG GA-3′ and (4) Coc-efa-5: 5′-TAC CTG AGC GGT GAA GTC AGC-3′. The PCR temperature profile used for *EF-1*α and *WG* was: 30 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 60 s. After amplification the PCR products were purified with ExoSap-IT (USB, Cleveland, OH, USA). Both directions were sequenced (BigDye Terminator Cycle Sequencing Kit, electrophoresis on an ABI 3130

Genetic Analyzer, and editing and alignment using SeqScape v. 2.5, all from Applied Biosystems, Lifetechnologies.com). Insertions and deletions (indels) were removed.

Phylogenetic analyses and morphospecies assignments

The *EF-1α* and *WG* segments combined yielded 888 base pairs for phylogenetic inference after the removal of indels (see Supporting Information, Table S3, for breakdown and for DNA character statistics). Homogeneity of base frequency was tested for each data partition (*EF-1α* exon 1, *EF-1α* intron 1, *EF-1α* intron 2, and *WG*) using a chi-squared-test in KAKUSAN 4 (Tanabe, 2007). The chi-squared-test did not reject homogeneity of nucleotide frequency in all pairs of sequences ($P = 1$ for each of the four data partitions; Supporting Information, Table S3). We also assessed the degree of substitution saturation by plotting transition and transversion rates against genetic distance for each of the data partitions using DAMBE (Xia & Xie, 2001). Substitution saturation was not detected for any of the data partitions ($P < 0.001$; Supporting Information, Fig. S1). Maximum Likelihood (ML) phylogenetic analysis was done on the *EF-1α* + *WG* sequences combined. The *EF-1α* segment comprises two introns and an exon, and the *WG* segment comprises one exon. The exons were each partitioned into 1st, 2nd and 3rd codon positions, and each intron was assigned to a separate partition, totaling eight data partitions. The best-fitting substitution model for each partition (Supporting Information, Table S4) was selected using Bayesian Information Criterion 5 (BIC5) in the KAKUSAN 4 software package (Tanabe, 2007). The eight data partitions were analyzed simultaneously using TREEFINDER version Oct. 2008 (Jobb, von Haeseler & Strimmer, 2004) with ML clade support provided by 1000 bootstrap replicates. Parsimony bootstrap support and Bayesian posterior probabilities (BPP) also were obtained for the nodes recovered in the ML analysis. Using PAUP* 4.0b10 (Swofford, 2002), parsimony bootstrap support was assessed with 1000 bootstrap replicates, using heuristic searches with tree bisection and reconnection branch swapping, and 100 random addition replicates per bootstrap replicate. BPP were obtained in MrBayes version 3.1.2 (Huelsenbeck & Ronquist, 2001). The Bayesian analysis was run for 5 million generations, sampling every 1000 generations, using the default run settings in which two independent analyses are performed, each with one cold and three heated chains. We plotted the log-likelihood for each sampling point against generation time to identify the stationarity phase and discarded the initial 1000 trees obtained in the pre-stationarity phase as burn-in.

Patterns of association among symbionts

We tested for biased association between all pairs of symbionts in both directions, yielding six sets of tests as follows: (1) ants to coccids, (2) coccids to ants, (3) ants to plants, (4) plants to ants, (5) coccids to plants, and (6) plants to coccids. For the coccids, we used as operational taxonomic units (OTUs) the ten species or morphospecies resulting from the phylogenetic and morphological analyses (see Results). For the ant OTUs we used the mitochondrial DNA lineages (matrilines) from Quek *et al.* (2007; 17 of them). For the plants, we used the *supra*-specific groups described in Figure 2 (in black boxes or grey ellipses). The plant groupings were designated to test whether (1) phylogenetic relatedness, (2) presence/absence of epicuticular wax, and (3) naturally hollowing stems vs. ant-excavated stems (i.e., PAC vs. PRU, Fig. 2) have a structuring effect on patterns of associations. The only plant group in Figure 2 that is known for certain to be non-monophyletic is gPAC, a paraphyletic grade. As for the others, monophyly is not certain, but plausible, given our incomplete understanding of phylogenetic relationships among myrmecophytic *Macaranga* species. Thus for the host plants, operational unit, or OU is a more appropriate term than OTU, and we use 'O(T)U' when referring to symbionts from any or all of the three guilds (where a guild is the ants, plants, or coccids, each taken as a whole). See figure 3 in Davies (2001) for some traits that distinguish these plant groups from one another. For the ants, 24 host plant species from sections *Pachystemon* and *Pruinosae* are represented, out of 25 species known to host *Crematogaster borneensis group* ants, and for the coccids, 19 host plant species are represented (see Fig. 2). For the analysis of association between ants and plants, 40 samples (see Supporting Information, Table S5) in addition to those used in Quek *et al.* (2007) are included here. A list of all ant samples used (with relevant collection information) is found in the file *Ant List.xlsx* (doi: 10.5061/dryad.6q1q6). For each O(T)U, Bornean samples were analyzed separately from samples from Malaya (or Malaya + Sumatra).

Exact multinomial tests were done between each symbiont O(T)U and all the O(T)Us in each of the other guilds that it could theoretically encounter to determine if the observed proportions deviated from the expected proportions. The expected proportions were computed from the pool of partners co-existing with the symbiont O(T)U in question (i.e., present in the sites where the symbiont was found and therefore could, in theory, could be encountered by it). Because exact multinomial tests do not reveal which particular pairings are responsible for low P values, we also subjected each pairwise combination to exact binomial tests to identify any pairs which deviated in their observed proportions. As an example,

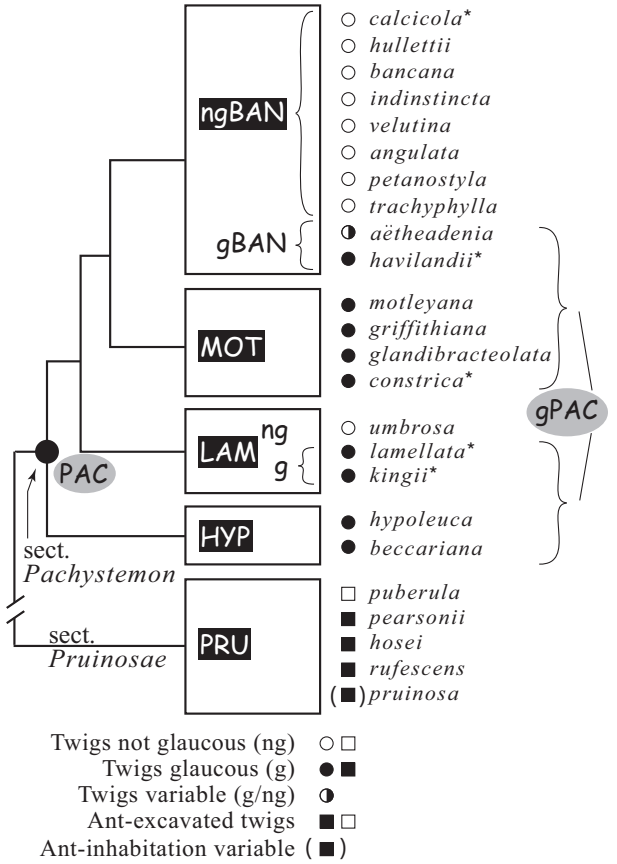

Figure 2. Simplified phylogeny of myrmecophytic *Macaranga* species (simplified from Davies, 2001) harbouring ants from the *Crematogaster borneensis* group (*sensu* Feldhaar *et al.*, 2016). The groups labelled in white font and black boxes are the lower level operational units (OUs) erected for the purposes of this study (BAN, *bancana* group; MOT, *motleyana* group; LAM, *lamellata* group; HYP, *hypoleuca* group; PRU, *Macaranga* section *Pruinosae*). The two groups labelled in black font and grey ellipses are the higher level OUs (also erected for the purposes of this study), which comprise multiple lower level units; PAC = *Macaranga* section *Pachystemon*, which contains all the lower level OUs except PRU; gPAC = the glaucous members of section *Pachystemon*, containing HYP, MOT and some of LAM). g = glaucous and ng = not glaucous (glaucous twigs are covered with epicuticular wax crystals). In this study, all specimens of *M. aëtheadenia* sampled had glaucous stems thus that species is included in the gPAC group. *M. velutiniflora* is the only species not represented in this study and has been omitted. With the exception of the five species marked with an asterisk (*), all species shown are represented in the analyses of associations between ants and plants. Asterisked species are not represented in the coccid sampling and analyses of associations between coccids and plants.

to test if ant lineage A showed a biased association with the plant group MOT, we compared the proportion of ant A samples that were found with MOT group members (i.e., observed proportion = 11/13 in Supporting Information, Table S6, first data line), with the proportion of total sampled plants made up by MOT group members in all the locations where ant A was sampled (i.e., expected proportion = 27/103). We then used the exact binomial two-tailed test to reject (or not reject) the null hypothesis that the two proportions were not statistically different. The coccid exhibited a positive bias ('preference') if it was observed to associate with the ant at a proportion that was significantly greater than that ant's proportional availability (the expected proportion), and a negative bias ('avoidance') if that proportion was significantly less. For this paper, the terms 'preference' and 'avoidance' (and their verb forms) refer strictly to *patterns* of associations that appear biased in comparison to the expected proportions in which partner O(T)Us are available, and are not to be taken in the literal sense. An Excel file showing how the values for the exact bi- and multinomial tests were calculated is provided in the file *Ant Plant Coccus counts for binomial tests.xlsx* (doi: 10.5061/dryad.6q1q6).

Where multiple coccid morphospecies were found with the same ant colony, each coccid morphospecies was treated separately such that the comparison was done separately for the coccid as both chooser and 'choosee' (object).

The exact binomial and exact multinomial tests were done in R, using the EMT (Exact Multinomial Test) package by Uwe Menzel for the latter. In total, 495 exact binomial tests were done (43 coccid-to-ant, 35 ant-to-coccid, 79 coccid-to-plant, 81 plant-to-coccid, 137 ant-to-plant, and 120 plant-to-ant). There were several ant–plant and coccid–plant pairings for which the other insect partner (coccid and ant, respectively) was not collected or not sequenced (i.e., not identified to OTU level; while host plants can be identified to species in the field most of the time, ants must be sequenced to be assigned to their mtDNA lineage, and not all ant colonies were collected with their coccids; coccids cannot be identified in the field and must be prepared, mounted and examined under a microscope by trained eyes). Several O(T)Us were not tested due to their small sample sizes ($N \leq 3$) or due to the absence of options, as in: (1) the case of ant lineage L (which occurs only where a single plant OU is available), and (2) coccids in the Malay Peninsula (by happenstance, these coccids were only collected during the sampling of one ant lineage, K, but were not collected when other ant lineages were sampled).

RESULTS

DNA, MORPHOLOGY, AND PHYLOGENETIC CLUSTERS

The 152 ingroup specimens were assigned to ten morphospecies, five of which have been formally

described (Morrison, 1921; Takahashi, 1952; *Coccus caviramicolus, C. macarangae*, *C. penangensis*, *C. secretus* and *C. tumuliferus*). Phylogenetic analyses of nuclear (nr) DNA also yielded ten lineages (C1 through C10 in Fig. 3), eight of which were well supported (ML bootstrap values ranging from 75% to 100%), while the remaining two (the sister pair C8 and C9) comprised one specimen each. C8 and C9 were assigned to separate morphospecies because they were morphologically distinct and genetically well separated (see branch lengths in Fig. 3). The ten lineages were each well supported (where $N > 1$)

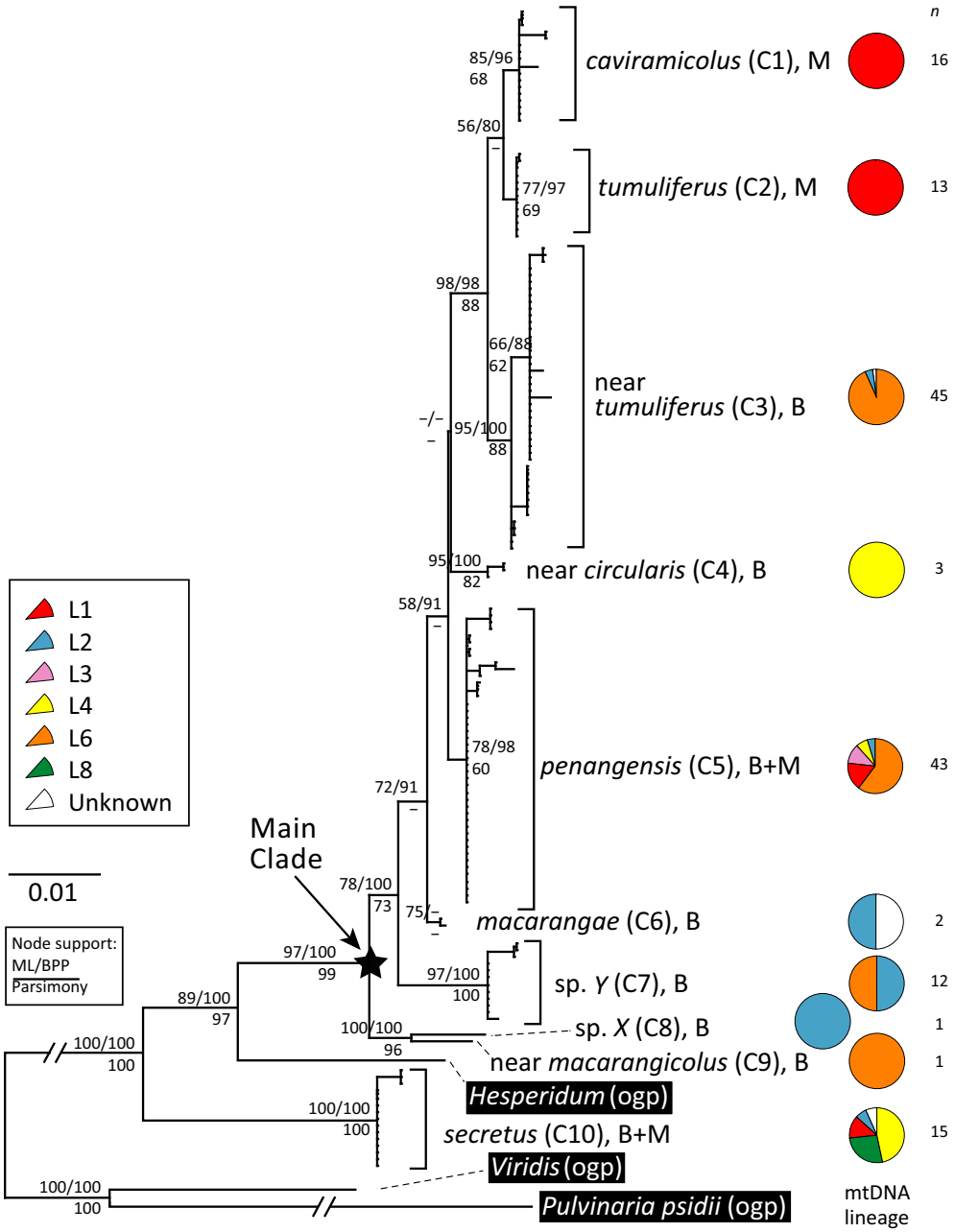

Figure 3. Maximum likelihood phylogenetic reconstruction of 151 ingroup *Coccus* specimens using DNA segments from the nuclear genes *EF-1α* and *WG*. Outgroups are blackboxed. Dash in node support means that the node was not recovered in the Parsimony or Bayesian analyses. B, Borneo; M, Malaya (for this paper: Singapore, Peninsular Malaysia and offshore islands, and Bintan Island, Indonesia); sampling for *Coccus* did not include Sumatra. Pie charts show the mitochondrial DNA lineage composition from Ueda *et al.* (2010).

and congruent with morphospecies assignments, but support for the branching relationships among these lineages was variable: ML bootstrap did not support one of the nodes in Figure 3, and support for other nodes ranged from 56% to 97%. The lineages also varied in size, ranging from single individuals in C8 and C9 to 45 individuals in C3 (sample sizes listed in Fig. 3). The majority of *Macaranga*-associated coccids formed a monophyletic group comprising nine lineages, C1–9. The specimens identified as *C. secretus* (C10), however, formed a single clade separated from the rest by the outgroup *C. hesperidum*. The nrDNA-based phylogeny thus indicates that symbiotic *Coccus* species from *Macaranga* are not monophyletic. Supporting Information (Table S7) shows the correspondence between the (morpho)species reported by Heckroth *et al.* (1998) and those in the present study.

Patterns of association among symbionts

Exact binomial tests for biased associations are presented in Supporting Information (Table S6) and a graphical representation of partner use, based on Supporting Information (Table S6), is presented in Figures 4–6.

As the objective of this study is to identify trends in the patterns of association, we also include and discuss findings where $0.05 < P < 0.09$ so as to include the plethora of suggestive but meaningful patterns typically ignored under the conventionally set limit of 0.05. For each of the six sets of tests (ants to coccids, coccids to ants, ants to plants, plants to ants, coccids to plants, and plants to coccids), noteworthy patterns in the results are provided in Supporting Information (Notes, under S0). The broad picture that emerged was that, in general: (1) most O(T)Us within a given guild associate with multiple O(T)Us of the other two guilds, but exhibit positive bias for only one or two O(T)U(s) in each guild, with the exception of those O(T)Us for which sampling was very limited ($N \leq 3$), or where options were not available, such as coccid–ant pairings in Malaya (Fig. 4), and ant lineage L (Fig. 6); (2) however, several O(T)Us were found to associate exclusively with single partner O(T)Us despite the availability of others (these are listed in Supporting Information, Notes, under S0, based on Figures 4–6); (3) coccid OTUs showed non-overlapping preferences towards plant OUs and towards ant OTUs; (4) the same was true for the plants (towards coccids and ants); (5) however, many instances of overlapping preferences were seen among the ant OTUs towards plants as well as towards coccids (see Supporting Information, Notes, under S0); (6) conflicts of interest were not detected – all instances of biased associations were met in the reverse direction with either the same bias or no bias (e.g., '+' with '+', or '+' alone, but never '+' with '–' in Table 1, which lists all pairwise associations showing P values < 0.09).

The network of preferred partners shows compartmentalization

We constructed a network showing all the instances of positive bias detected in the exact binomial tests. The O(T)Us in the network grouped into three compartments (I, II and III in Fig. 7). Ant E and plant LAM make up compartment III. Compartment I contains only non-glaucous hosts from the BAN group (in section *Pachystemon*) and their preferred or preferring symbionts. The host plants in compartment II are glaucous plants from section *Pachystemon* as well as host plants from section *Pruinosae* (in Fig. 7, the latter includes only Bornean members, which comprise three glaucous and one non-glaucous species), thus compartment II mostly contains glaucous host plants and their preferred/preferring associates. Ant-excavated vs. naturally hollowing stems (PRU vs. PAC) did not appear to have a clear-cut partitioning effect on the network, but some degree of partitioning was present (plant PRU, ant G1 and coccid C7 form a loop that is bidirectional in two sectors and unidirectional in one, but PRU and G1 were linked to other OTUs).

DISCUSSION

In this study, we have shown that the *Coccus* associates of myrmecophytic *Macaranga* are not monophyletic, and comprise at least ten morphologically distinguishable species that concur with phylogenetic analyses of nuclear DNA. *C. secretus* emerged as separate from the remaining nine which form a monophyletic group. Analysis of the patterns of association among all three guilds show that preferences and avoidances are common and no conflicts of interest are apparent. Furthermore, a network of preference patterns among all three guilds reveals compartmentalization according to the presence or absence of epicuticular wax on the *Macaranga* host plants. To a lesser degree, ant-excavated stems also impose some structure on preference patterns.

The following caveats should be noted: (1) The geographic distributions noted in Figure 3 are largely dictated by sampling extent and intensity. In this study, *Coccus macarangicolus* ($N = 1$), *C. caviramicolus, C. tumuliferus, C.* near *tumuliferus* and *C. macarangae* each appear to be restricted to either Malaya or Borneo (Sumatra was not sampled). However, the more extensive study by Heckroth *et al.*

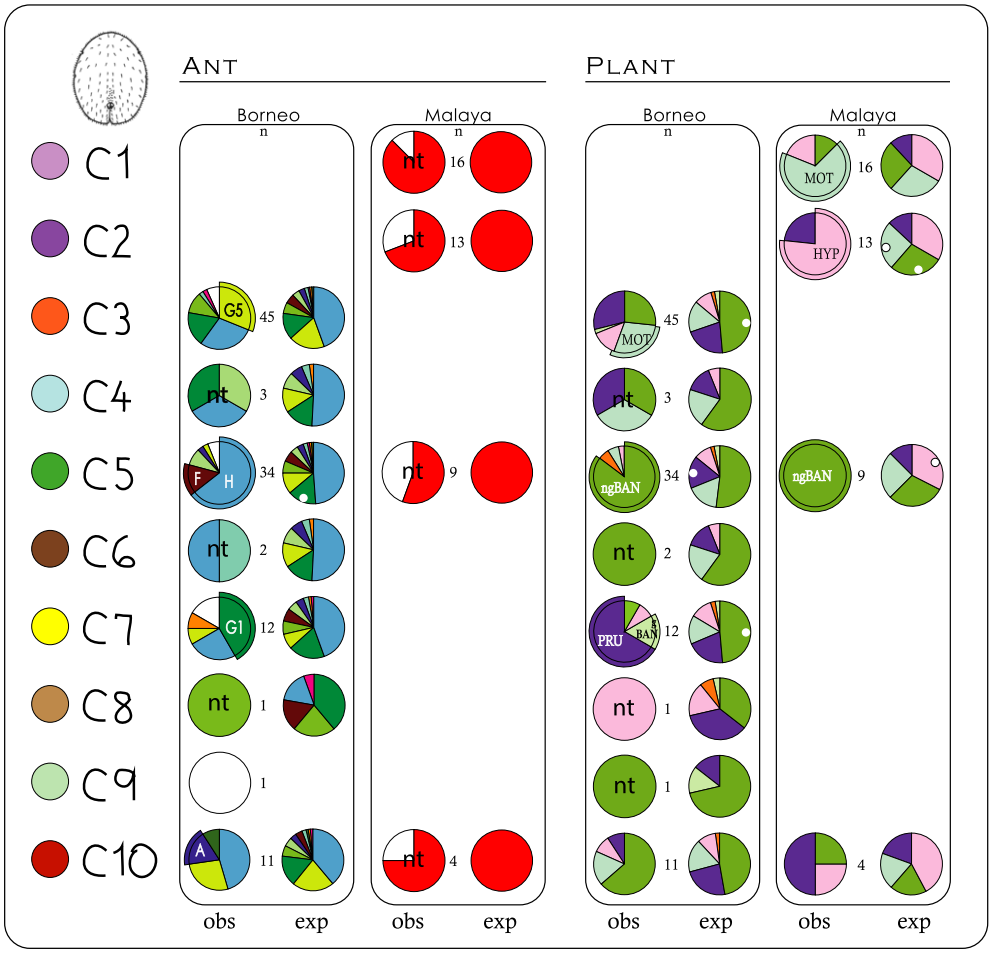

Figure 4. Coccid association patterns with ants and plants (data from Supporting Information, Table S6). Left column in each box shows the observed proportion ('obs') of each partner O(T)U, and right columns show the expected proportion ('exp'). Enlarged pie sectors in the 'obs' columns are associates which are favoured disproportionately relative to their availability (i.e., 'preferred'), as suggested by exact binomial tests ($P < 0.09$). 'Avoided' partners are indicated by white circles in the 'exp' column pies. 'nt' denotes 'not tested', and white pie sectors indicate unknown ant associates. Numbers between pie charts show sample sizes for the left (obs) column. Note that, by happenstance, Malayan coccids were collected during the sampling of ant K only. Coccids are colour-coded (in the extreme left of figure) to serve as keys for the pie sectors in Figures 5 and 6. Likewise, colour keys for the pie sectors in this figure are shown in Figure 5 (plants) and Figure 6 (ants). Plant OUs PAC and gPAC were not represented because they comprise multiple lower level OUs (see Fig. 2) and thus cannot be inserted into the pie charts. PAC and PRU are complements of each other, thus a coccid that prefers one in the binomial tests, by definition, avoids the other with identical *P*-values (see Supporting Information, Table S6).

(1998) noted that *C. tumuliferus*, *C. macarangae*, *C. caviramicolus,* and *C. macarangicolus* occurred in both regions, although taxonomic issues may have confounded previous identifications (P. J. Gullan and T. Kondo, unpubl. data). (2) Because of the large number of binomial tests done, a proportion of *P* values will suggest a bias (preference or avoidance) where none is present, by chance alone. At the same time, some of these tests will also fail to detect real preferences/avoidances due to: (1) insufficient sample size, i.e. lack of power, and/or (2) an ant or coccid OTU preferentially inhabiting an area where its preferred partner is dominant and/or its avoided partner is rare (and thus the similarity between the observed and expected proportions will fail to reject the null hypothesis). Therefore, we did not perform mathematical corrections for the *P* values in Supporting Information (Table S6). In addition to the questionable utility of *P*-value corrections in general (see Feise, 2002; Rothman, 1990), there is as yet no published method for correcting *P* values in tests which use count data where all the comparisons are part of

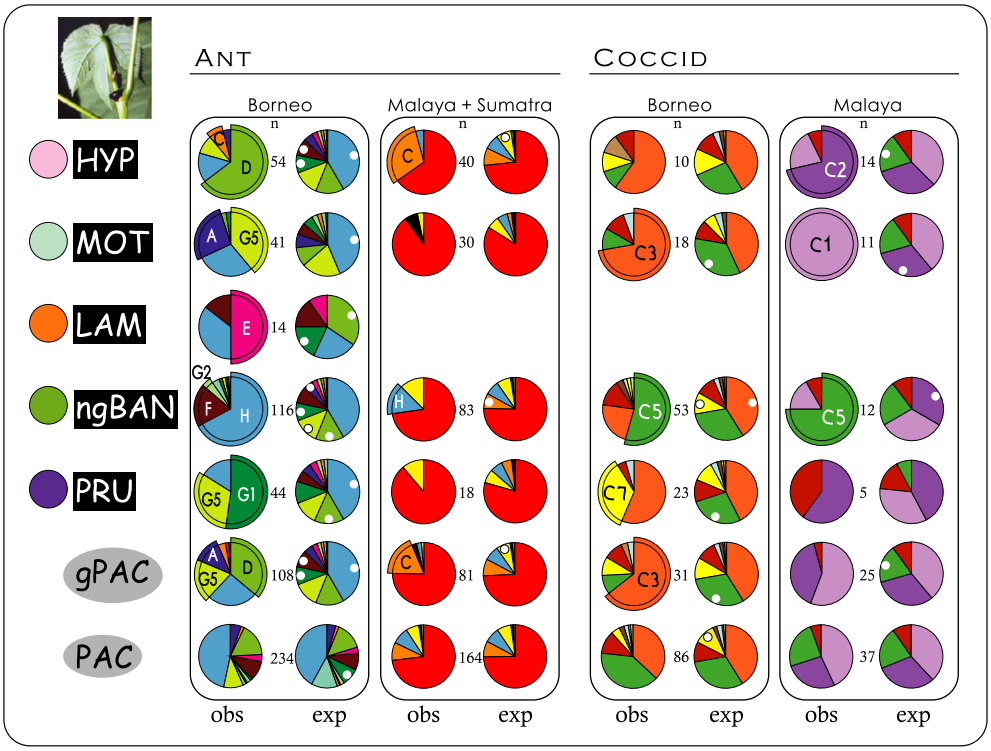

Figure 5. Plant association patterns with ants and coccids (data from Supporting Information, Table S6). Left column in each box shows the observed proportion ('obs') of each partner OTU, and right columns show the expected proportion ('exp'). Enlarged pie sectors in the 'obs' columns are associates which are favoured disproportionately relative to their availability (i.e., 'preferred'), as suggested by exact binomial tests ($P < 0.09$). 'Avoided' partners are indicated by white circles in the 'exp' column pies. Numbers between pie charts show sample sizes for the left (obs) column. Plant O(T)Us are colour-coded (in the extreme left of figure) to serve as keys for the pie sectors in Figures 4 and 6. Likewise, colour keys for the pie sectors in this figure are shown in Figure 4 (coccids) and Figure 6 (ants).

a finite and planned/prescribed pool of comparisons. Thus readers should bear in mind that the test results on which Figures 4–7 and Table 1 are based (i.e., Supporting Information, Table S6) will harbour both Type I and Type II errors.

ANT MATRILINES VS. SPECIES – A CAVEAT

The ant OTUs are matrilines (based on maternally-inherited mitochondrial DNA lineages) and may not represent species. Incomplete lineage sorting and hybridization are known issues when attempting to infer species boundaries in young clades using a single non-segregating locus. Feldhaar *et al.* (2016) recently assigned *Crematogaster borneensis* group into eight species, including five newly described species, based on the morphology of female sexuals. Those species concurred more with DNA analyses based on microsatellite loci and the nuclear gene *EF-1α* than with mitochondrial DNA (Feldhaar, Fiala & Gadau, 2004; Feldhaar *et al.*, 2010, 2016). However, a study by Ueda *et al.* (2015) based on a single locality showed that microsatellite variation patterns concurred with that of mitochondrial DNA for all five of the matrilines present at that location. We were unable to unequivocally match all the ant matrilines with the species described by Feldhaar *et al.* because not all the *Macaranga* branches we sampled for ants contained female sexuals (for further information, see Supporting Information, Notes, under S1). However, based on Feldhaar *et al.* (2016) and unpublished data (Supporting Information, Notes, under S2), it appears that some of the older lineages in the mtDNA phylogeny by Quek *et al.* (2007), namely A and B, but also C + D together, are reliable species proxies. As for the others, further investigations beyond the scope of this study are necessary to uncover the relative contributions of hybridization vs. incomplete lineage sorting in distorting species boundaries in the mtDNA phylogeny.

Nevertheless, there are good reasons for using mtDNA lineages as taxonomic units for these ants, as long as the above caveats are taken. Mitochondrial DNA lineages can be determined without ambiguity

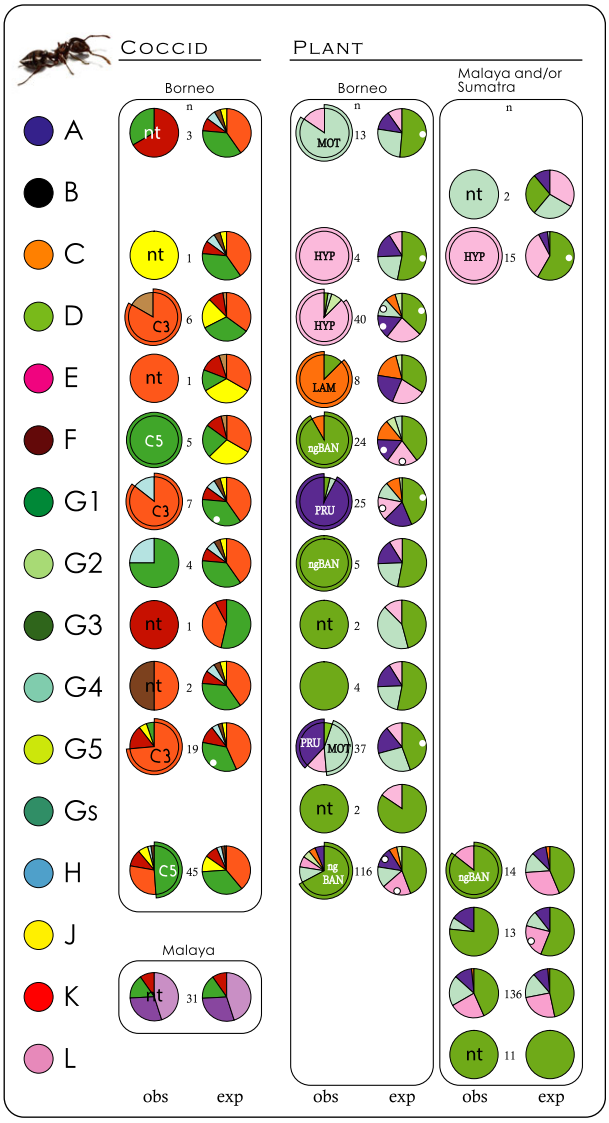

Figure 6. Ant association patterns with coccids and plants (data from Supporting Information, Table S6). Left column in each box shows the observed proportion ('obs') of each partner O(T)U, and right columns show the expected proportion ('exp'). Enlarged pie sectors in the 'obs' columns are associates which are favoured disproportionately relative to their availability (i.e., 'preferred'), as suggested by exact binomial tests ($P < 0.09$). 'Avoided' partners are indicated by white circles in the 'exp' column pies. 'nt' denotes 'not tested'. Numbers between pie charts show sample sizes for the left (obs) column. Ant lineages B, Gs, J and L were not collected with coccids. Ants are colour-coded (in the extreme left of figure) to serve as keys for the pie sectors in Figures 4 and 5. Likewise, colour keys for the pie sectors in this figure are shown in Figure 4 (coccids) and Figure 5 (plants). Plant OUs PAC and gPAC were not represented because they comprise multiple lower level OUs (see Fig. 2) and thus cannot be inserted into the pie charts. PAC and PRU are complements of each other, thus a coccid that prefers one in the binomial tests, by definition, avoids the other with identical *P*-values (see Supporting Information, Table S6).

where molecular facilities exist regardless of whether or not female reproductives are sampled, allowing for larger samples and thus greater testing power. In contrast, species boundaries are fluid and hypothetical, and species identity can only be determined by well trained experts, and only when female reproductives have been sampled. Mitochondrial markers also have greater resolving power as demonstrated in these ants (17 matrilines vs. 8 morphospecies), revealing recent gene flow corridors or barriers that are not readily apparent in morphology.

Macaranga myrmecophytes harbour two clades of coccids – a generalist species and a clade comprising specialist morphospecies

The *Coccus* specimens from *Macaranga* segregated into two clades separated by the outgroup *C. hesperidum*, a worldwide agricultural pest (Gill, 1988; Williams & Watson, 1990). Although *C. hesperidum* appears as sister to the Main Clade (denoted by a star in Fig. 3), Lin *et al.* (2013) showed that the sister to *C. hesperidum* was not a *Macaranga* coccid. The outlying clade contained all the specimens identified as *C. secretus* and the Main Clade contained the remaining specimens. Heckroth *et al.* (1998) noted that although *C. secretus* mainly associated with *Macaranga*, it also was found in the domatia of other ant–plant systems, particularly those involving *Cladomyrma* (see also Moog *et al.*, 2005) and other *Crematogaster* ants, and occasionally found in association with additional ant genera. We found that *C. secretus* in most cases makes up a small proportion of the coccid partners of ant or plant O(T)Us (see Figs 4–6).

In the Main Clade, the coccid (morpho)species that were abundant enough for testing showed biased associations towards particular ant or plant O(T)Us, and vice versa. Our study did not include collections (ants or coccids) from *Macaranga winkleri*, a Bornean myrmecophyte in a separate section of *Macaranga* harbouring obligate and mutually specific ants unrelated to those in this study (Fiala *et al.*, 1999). Heckroth *et al.* (1998) noted that a coccid morphospecies, labelled C41 in their study, was found mostly on *M. winkleri* and rarely on other host plant species. *M. winkleri* additionally harboured five other morphospecies of coccid (some of which are in the present study). Morphologically, Heckroth's C41 is close to *C. penangensis* (C5; P. J. Gullan, pers. observ.) and thus it is likely to be part of the Main

Table 1. Mutual preference or avoidance between symbionts

Mutual?	Coccid	+/–	Ant	*P* values (B)	*P* values (M)	Ant	+/–	Coccid	*P* values (B)	*P* values (M)
Y	C3	+	G5	*	n.a.	G5	+	C3	*	n.a.
	C3		D	X	n.a.	D	+	C3	*	n.a.
	C3		G1	X	n.a.	G1	+	C3	*	n.a.
Y	C5	+	F	*	n.a.	F	+	C5	***	n.a.
(Y)	C5	+	H	*	n.a.	H	+	C5	#	n.a.
(Y)	C5	–	G1	*	n.a.	G1	–	C5	#	n.a.
	C5		G5	X	n.a.	G5	–	C5	**	n.a.
	C7	+	G1	*	n.a.	G1		C7	X	n.a.
	C10	+	A	#	n.a.	A		C10	n.t.	n.a.

Mutual?	Coccid	+/–	Plant	*P* values (B)	*P* values (M)	Plant	+/–	Coccid	*P* values (B)	*P* values (M)
Y	C1	+	MOT	n.a.	**	MOT	+	C1	n.a.	****
	C1	+	gPAC	n.a.	*	gPAC		C1	n.a.	X
Y	C2	+	HYP	n.a.	**	HYP	+	C2	n.a.	**
Y	C2	–	ngBAN	n.a.	*	ngBAN	–	C2	n.a.	*
Y	C2	–	MOT	n.a.	*	MOT	–	C2	n.a.	*
Y	C3	+	MOT	*	n.a.	MOT	+	C3	*	n.a.
Y	C3	+	gPAC	*	n.a.	gPAC	+	C3	*	n.a.
Y	C3	–	ngBAN	**	n.a.	ngBAN	–	C3	**	n.a.
Y	C5	+	ngBAN	****	****	ngBAN	+	C5	***	***
(Y)	C5	–	HYP	X	*	HYP	–	C5	X	####
Y	C5	–	PRU	**	X	PRU	–	C5	***	X
Y	C5	–	gPAC	**	***	gPAC	–	C5	*	**
	C5		MOT	X	X	MOT	–	C5	*	X
	C5	+	PAC	**	X	PAC		C5	X	X
Y	C7	+	PRU	***	n.a.	PRU	+	C7	**	n.a.
Y	C7	–	ngBAN	**	n.a.	ngBAN	–	C7	*	n.a.
(Y)	C7	–	PAC	***	n.a.	PAC	–	C7	#	n.a.

Mutual?	Ant	+/–	Plant	*P* values (B)	*P* values (M+S)	Plant	+/–	Ant	*P* values (B)	*P* values (M+S)
Y	A	+	MOT	****	n.a.	MOT	+	A	****	n.a.
Y	A	+	gPAC	****	n.a.	gPAC	+	A	**	n.a.
Y	A	–	ngBAN	****	n.a.	ngBAN	–	A	**	n.a.
Y	Cms	+	HYP	n.a.	****	HYP	+	Cms	n.a.	****
Y	Cms	+	gPAC	n.a.	****	gPAC	+	Cms	n.a.	**
Y	Cms	–	ngBAN	n.a.	****	ngBAN	–	Cms	n.a.	**
Y	Cb	+	HYP	****	n.a.	HYP	+	Cb	**	n.a.
(Y)	Cb	+	gPAC	**	n.a.	gPAC	+	Cb	###	n.a.
	Cb	–	ngBAN	*	n.a.	ngBAN		Cb	X	n.a.
Y	D	+	HYP	****	n.a.	HYP	+	D	****	n.a.
Y	D	+	gPAC	****	n.a.	gPAC	+	D	****	n.a.
Y	D	–	ngBAN	****	n.a.	ngBAN	–	D	****	n.a.
Y	D	–	PRU	**	n.a.	PRU	–	D	**	n.a.
	D	+	PAC	**	n.a.	PAC		D	X	n.a.
	D	–	MOT	####	n.a.	MOT		D	X	n.a.
	D		LAM	X	n.a.	LAM	–	D	*	n.a.
Y	E	+	LAM	****	n.a.	LAM	+	E	***	n.a.
Y	F	–	HYP	**	n.a.	HYP	–	F	*	n.a.
Y	F	–	gPAC	****	n.a.	gPAC	–	F	***	n.a.
Y	F	+	ngBAN	****	n.a.	ngBAN	+	F	***	n.a.
	F	+	PAC	*	n.a.	PAC		F	X	n.a.
	F	–	PRU	*	n.a.	PRU		F	X	n.a.
Y	G1	–	HYP	*	n.a.	HYP	–	G1	*	n.a.
Y	G1	–	ngBAN	****	n.a.	ngBAN	–	G1	***	n.a.
Y	G1	+	PRU	****	n.a.	PRU	+	G1	****	n.a.

Table 1. *Continued*

Mutual?	Ant	+/–	Plant	*P* values (B)	*P* values (M+S)	Plant	+/–	Ant	*P* values (B)	*P* values (M+S)
Y	G1	–	PAC	****	n.a.	PAC	–	G1	****	n.a.
Y	G1	–	gPAC	**	n.a.	gPAC	–	G1	**	n.a.
	G1		LAM	X	n.a.	LAM	–	G1	####	n.a.
(Y)	G2	+	ngBAN	##	n.a.	ngBAN	+	G2	#	n.a.
Y	G5	+	MOT	*	n.a.	MOT	+	G5	**	n.a.
(Y)	G5	+	gPAC	*	n.a.	gPAC	+	G5	##	n.a.
Y	G5	–	ngBAN	****	n.a.	ngBAN	–	G5	****	n.a.
Y	G5	+	PRU	**	n.a.	PRU	+	G5	***	n.a.
	G5	–	PAC	**	n.a.	PAC		G5	X	n.a.
Y	H	–	HYP	***	X	HYP	–	H	****	X
Y	H	–	gPAC	**	*	gPAC	–	H	***	X
Y	H	+	ngBAN	****	**	ngBAN	+	H	****	*
Y	H	–	PRU	*	X	PRU	–	H	***	X
	H	+	PAC	*	X	PAC		H	X	X
	H		MOT	X	X	MOT	–	H	####	X
Y	J	–	HYP	n.a.	*	HYP	–	J	n.a.	*

Preferences (+) and avoidances (–) between symbiont O(T)Us as determined by exact binomial tests (from Supporting Information, Table S4). Mutual preference or avoidance is evident when the symbionts have the same sign (+ or –) in both directions, and this is indicated by 'Y' or '(Y)' in the column first column. The left and right sections of each table (demarcated by a vertical line) are *vice versa* comparisons between O(T)Us (e.g., coccid to ant, vs. ant to coccid). A blank entry under the header '+/–' means no preference/avoidance was found. Y = yes, and (Y) is for the case when one or both directions has $0.05 < P < 0.09$. *P*-values are shown separately for Borneo and Malaya or Malaya + Sumatra, i.e., *P* (B) and *P* (M) or *P* (M + S). *P*-value codes (non-directional): ####$P < 0.09$; ###$P < 0.08$; ##$P < 0.07$; #$P < 0.06$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; n.a., not applicable; X: $P > 0.1$, and n.t., not tested (due to small sample size). For this paper, Malaya encompasses Singapore, Peninsular Malaysia and its offshore islands, and Bintan Island, Indonesia, and Sumatra encompasses Sumatra and Lingga Island. *Coccus* OTUs follow the species/morphospecies in Figure 3. *Crematogaster* ant OTUs are mtDNA lineages following Quek *et al.* (2007); *Macaranga* plant OUs follow Figure 2. Note that PRU and PAC are complements of each other, such that preference for one means avoidance of the other. Note also that MOT, HYP, and glaucous members of LAM are part of 'gPAC' (see Fig. 2). See Supporting Information (Table S6) for O(T)Us excluded from preference analyses (due to small sample sizes, or absence of options).

Clade. If so, the Main Clade of coccids spans two obligate tripartite systems: (1) the myrmecophytes in *Macaranga* sections *Pachystemon* and *Pruinosae* which host *Crematogaster borneensis* group ants, and (2) the myrmecophyte *M. winkleri* which hosts an unrelated *Crematogaster* ant. Heckroth *et al.* (1998) noted that other coccid (morpho)species (namely *C. macarangae* and a morphospecies labelled C214), while mainly associated with myrmecophytic *Macaranga* were also found in association with other plants and ants, albeit much more rarely than *C. secretus*. The differing patterns of association between *C. secretus* and the Main Clade coccids (described above) suggest that the Main Clade may be more obligately associated with *Macaranga* myrmecophytes whereas *C. secretus* may be a more facultative or opportunistic symbiont.

Whereas the earlier branching relationships in Figure 3 are very well supported (supporting two non-sister clades of *Macaranga*-associated coccids), the earlier branching relationships in the mtDNA trees reported in Ueda *et al.* (2008: fig. 2; Ueda *et al.*, 2010: fig. 1a) are not well supported and thus reveal nothing about the monophyly of the mtDNA sequences of the *Macaranga*-associated coccids (but see further discussion in the following section). This is not the case, however, for figure 1b in Ueda *et al.* (2010) using an exemplar-based tree with longer mtDNA sequences compared to figure 1a (see also next section).

The nuclear DNA clusters correlate with *Coccus* morphospecies boundaries

The *Coccus* samples clustered into ten well supported (where $N > 1$) nuclear DNA lineages that concurred with morphospecies assignments, comprising five described species and another five to be formally described (by P. J. Gullan and T. Kondo, unpubl. data). The congruence between morphospecies and nrDNA lineages contrasts with the mitochondrial DNA phylogeny published previously (Ueda *et al.*, 2008, 2010) in which most of the morphospecies, including the outlying species *C. secretus*, were

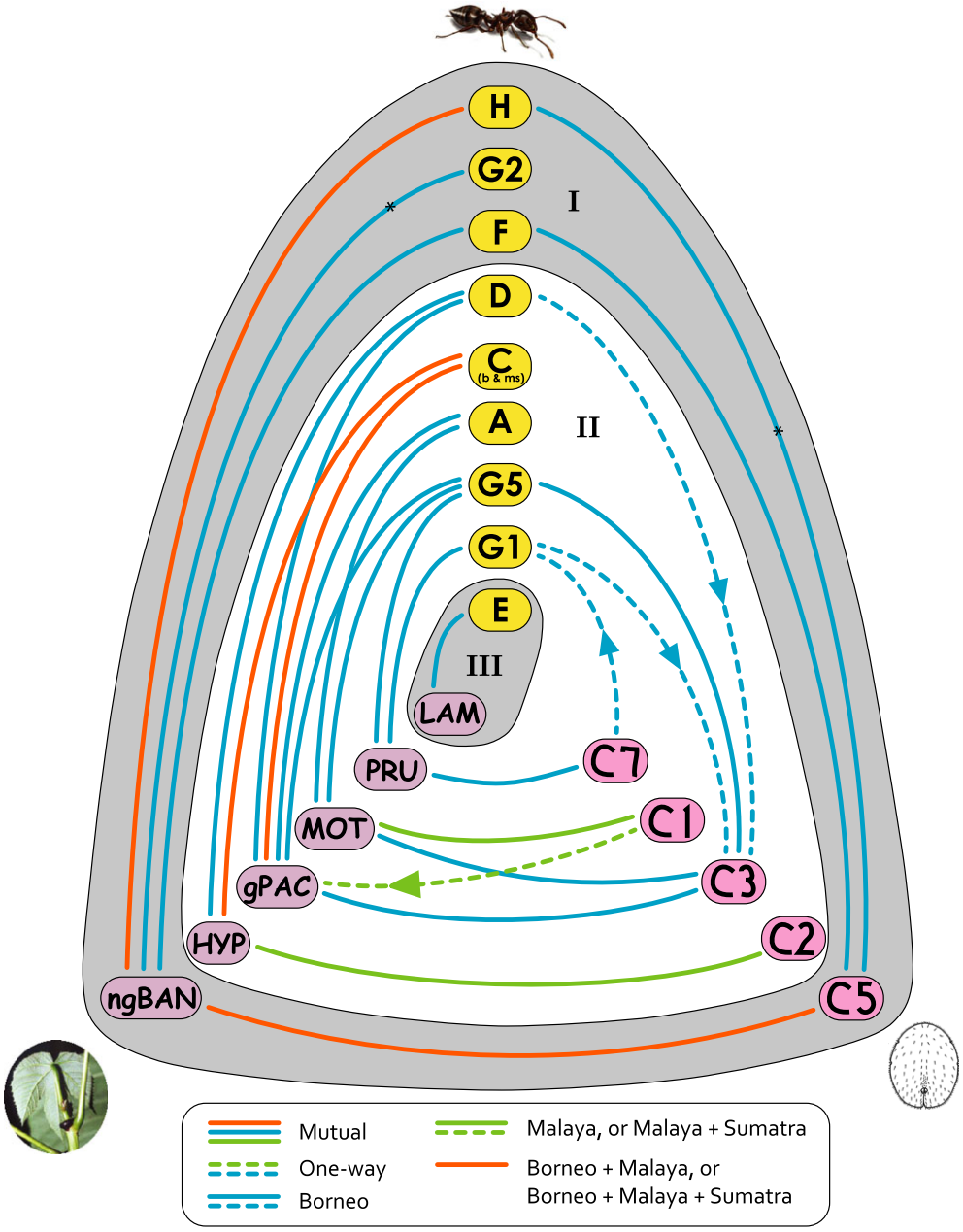

Figure 7. Network of 'preferred' associates among ants, plants and Main Clade coccids (data from Table 1). Solid lines represent mutual/bidirectional preference and dotted lines represent unidirectional preference where the arrow points from chooser to choosee. All preferences having $P < 0.09$ in exact binomial tests are shown. Asterisks (*) indicate $0.5 < P < 0.09$ for one or both directions. The network separates into three compartments, I (grey outer), II (white) and III (grey inner). Compartment I holds non-glaucous host plants from section *Pachystemon* and their preferred/preferring symbionts, and compartment III contains host plants (glaucous and non-glaucous) in the LAM clade and ant E. Compartment II contains glaucous host plants in section *Pachystemon* and their preferred/preferring associates as well as Bornean PRU clade plants and their preferred/preferring associates (three of the four PRU clade plants in this compartment are glaucous). Note that in most cases, the preferred partner is not the sole partner; see Figures 4–6 for the complete set of partners for each ant, coccid, and plant O(T)U. C10 (*C. secretus*) is not part of the Main Clade and has been omitted (the only bias in association involving this species, as either chooser or choosee, was a preference for ant A).

scattered across multiple mtDNA lineages (unpublished morphological data from P. J. Gullan). Such a pattern is consistent with a copying of the mtDNA locus into the nuclear genome prior to the last common ancestor of the Main Clade and *C. secretus* (Sunnucks & Hales, 1996; Bensasson *et al.*, 2001). If

true, the resulting phylogeny would likely contain a mix of the nuclear copies and *bona fide* mtDNA sequences and would not reflect the true cladogenetic history of the mitochondrial genome. The presence of nuclear copies of mitochondrial DNA in the *Coccus* datasets of Ueda *et al.* (2008) would also explain why no biased associations (either preference or avoidance by the scales) with ants or plants was detected, in stark contrast to the results of this study. Interestingly, the 20-exemplar mtDNA tree based on a longer stretch of mtDNA in Ueda *et al.* (2010: 1021 base pairs in dataset 2 *vs*. 521 base pairs in dataset 1 in fig. 1b vs. a, respectively) also shows *Coccus hesperidum* breaking up the monophyly of the *Macaranga* coccids as seen here, suggesting that NUMTs (nuclear copies of mitochondrial DNA) may be less of a problem (or possibly not at all) in that particular dataset (dataset 2). If NUMTs are indeed present in the data of Ueda *et al.* (2008, 2010, particularly dataset 1), then our understanding of the age and biogeographic history of the scales based on that dataset must be re-evaluated. In addition to, or possibly instead of NUMTs, hybridization or incomplete lineages sorting may contribute to the lack of congruence with morphology in those studies.

Coccids and plants, but not ants, show non-overlapping partner preference

Most OTUs within a given guild associate with multiple OTUs of each of the other two guilds but show preference towards only one or two of them (similarly noted by Heckroth *et al.*, 1998). This suggests a complex history of adaptation and even co-adaptation to specific partner traits albeit under unpredictable ecological settings that favour flexibility in partner use. A few O(T)Us, however, appear to be monogamous in their partnerships despite the availability of multiple options. In general, exact binomial tests of the coccids' and plants' associations with each of their partner guilds revealed non-overlapping preferences as a rule. Where the coccid OTUs overlapped in their preferences (C1 and C3 both preferred MOT, Fig. 4), they were geographically separated (C1 from Malaya, C3 from Borneo). On the other hand, tests of the ants' associations with each of their partner guilds revealed overlapping preferences as a rule. A likely reason is that the 17 mitochondrial lineages (from Quek *et al.*, 2007) represent too fine a division (resulting in too many ant OTUs) and do not reflect taxa that operate as differentiated units in their selection of partners (cf. eight species by Feldhaar *et al.*, 2016; see also the caveat regarding matrilines vs. species and the next section).

Reciprocal preferences/avoidances and compartmentalized preference networks defined by the presence/absence of epicuticular wax

Many cases of preference or avoidance were reciprocal (Table 1). Furthermore, the network of preferred partners (Fig. 7) among all three guilds showed compartmentalization attributable to the absence or presence of epicuticular wax (compartments I and II, respectively). We have known that this feature determines which ants can gain access to a given plant (see Introduction), but the finding that all pairwise preferences, sometimes in both directions, can be grouped by the presence/absence of this trait was unexpected.

Ant-excavated vs. naturally hollowing stems did not impose a clear-cut partition on the network. However, of the five *Macaranga* species with ant-excavated stems (i.e., PRU clade), four of them are glaucous, thus the two traits are confounded to some extent. Nevertheless, some degree of partitioning was evident (plant PRU, ant G1, and coccid C7). Feldhaar *et al.* (2016: 203) noted that the ants that specialize on this group of plants are the largest among the *Macaranga* ants, presumably because stems requiring excavation favour larger ants. The compartment comprising ant E and plant LAM may indicate evolutionary specialization by this ant matriline to plant traits in LAM yet to be identified.

A multitude of mutualism-related traits are correlated with the presence/absence of epicuticular wax

Federle *et al.* (1997) showed that epicuticular wax crystals prevent non-mutualistic ants from gaining access to the plants, and that phytoecious ants (ants obligately inhabiting specialized live plant domatia) from non-glaucous *Macaranga* species showed impaired ability to climb glaucous species. Federle *et al.* (2000) also showed that the phytoecious ants from glaucous species were inferior to those from non-glaucous species in their ability to cling to a smooth surface under centrifugal acceleration. Several other studies have further demonstrated differences in mutualism-related traits in both the ants and plants, including (comparing glaucous vs. non-glaucous *Macaranga* species, or comparing ants from glaucous vs. non-glaucous species): (1) exposed vs. hidden food bodies (Federle & Rheindt, 2005), (2) absence vs. presence of prostomata (a region of thinned translucent domatium wall that facilitates the creation of openings; Federle *et al.*, 2001), (3) high vs. low levels of leaf toughness and chemical defences against herbivory (Nomura, Itioka & Itino, 2000), (4) low vs. high intensity of vine-pruning

(Federle, Maschwitz & Hölldobler, 2002), and (5) less vs. more aggressive defensive behavior (Itioka *et al.*, 2000). Feldhaar *et al.* (2003b) also reported a suite of differences between two groups of ant morphospecies: the group which overwhelmingly inhabited glaucous species had smaller queens and workers, lower density of workers, produced alates at smaller colony sizes, and only colonized saplings, whereas the group inhabiting a mixture of glaucous and non-glaucous species showed the converse pattern and was able to colonize saplings and branches of adult trees. Quek *et al.* (2004) also showed a pattern of coevolutionary diversification in the ant mtDNA and plant phylogenies, where the older lineages of ants live almost exclusively in glaucous *Macaranga* species, which make up the older lineages of *Macaranga* (Blattner *et al.*, 2001; Davies *et al.*, 2001). Thus the presence/absence of epicuticular wax crystals appears to be the first order filter by which ants and plants select partners that best serve their needs.

That the presence/absence of epicuticular wax was found to partition the three-way network of preferred partners despite the ants being defined by matriline rather than by species suggests that the broad-scale patterns we have detected in this study are robust (despite the issues underscored in the caveats). On the other hand, considering that female alates are the ones who select *Macaranga* hosts, and females (workers) determine which coccid species to accept or reject, it is possible that the patterns of partner preference based on matrilines are 'tighter' than those based on nuclear DNA lineages or species. Interesting insights could certainly be gained from a comparison of partner preferences between the two taxonomic modes. Nevertheless, we now know that host stem traits (glaucous vs. not glaucous, and ant-excavated vs. naturally hollow) are important predictors of ant identity, whether that identity is defined by matriline, morphology (i.e., species), or nuclear DNA lineage (see Feldhaar *et al.*, 2003a, 2004, 2010, 2016; Quek *et al.*, 2007; Ueda *et al.*, 2015, and Supporting Information, Notes under S2).

Ants may be the architects behind the preference patterns

We know that female alates (winged reproductives) of the *Crematogaster borneensis* group can chemically distinguish among *Macaranga* species when selecting host plants to found new colonies (Inui *et al.*, 2001; Jürgens *et al.*, 2006; see also Edwards *et al.*, 2006; Grangier *et al.*, 2009 for other ant–plant systems). Unlike female ant alates which actively disperse and select their host plants, the coccid crawlers (first nymphal stage during which they disperse) are subject to the vagaries of wind dispersal (Handa *et al.*, 2012), consistent with observations by Heckroth *et al.* (2001) of first- and second-instar nymphs aggregating on *Macaranga* shoot tips (presumably in anticipation of catching a wind draft). Handa *et al.* observed that, upon landing on an ant-inhabited *Macaranga*, the crawlers of symbiotic coccid species are either carried into the domatia by ants or find their own way to a domatium hole, whereas non-symbiotic species were always thrown off the plant by the ants. Heckroth *et al.* (2001) further noted in an experiment that ants from *M. hypoleuca* (an intensely glaucous species) had the capacity to distinguish between symbiotic coccid species, accepting *C. tumuliferus* (C2 and possibly C3, belonging to the glaucous compartments in white in Fig. 7) but throwing *C. penangensis* (C5, a member of the non-glaucous compartment, in grey) off the plant. We can reasonably assume that most/all of the ant OTUs show some selectivity towards coccid species, and the associations observed between certain coccids and plants could simply be a by-product of the ants exerting their preferences for coccid and plants. The same could also be true of the apparent preferences shown by the coccids or plants towards the ants – such patterns may simply reflect the ants' effectiveness at commandeering their preferred plant and coccid resources. The ants may also be selecting coccids based on the interaction between coccid and plant. For example, they may select the coccid species best able to evade or tolerate plant defensive compounds (if present), extract phloem sap and produce honey dew with the ants' optimal nutritional profile from that particular host plant species.

We also know that ant inhabitants may change as host plants mature – for example, one ant species often secondarily colonizes mature trees, while another specializes on saplings (Feldhaar *et al.*, 2016). Thus some of the coccids may have been selected by previous rather than current ant inhabitants.

Coccids and plants may also shape the association patterns

We cannot, however, assume that the plants or coccids are merely objects of the ants' manipulations. Fiala *et al.* (1999) noted that specificity between *Crematogaster* and *Macaranga* increases with the age of the plant, and thus after initial host selection by ant foundresses, ecological or physiological sorting processes continue to select for better-matched partnerships. The association patterns between coccids and plants may have nothing to do with the ants in contrast to the speculations of the previous section, but reflect varying abilities among the coccids in tolerating plant chemical defences, since the coccids are technically plant pests. Coccid crawlers may be able

to distinguish between favourable and unfavourable ant or plant species, and attempt to re-disperse if they find themselves in company of the 'wrong' plants or ants after alighting (however, mature coccids are sessile and can only leave/enter domatia when transported by ants, according to Heckroth *et al.*, 2001). Such abilities would be advantageous given that the ants are known to selectively remove coccids from their host plants (as previously noted).

More likely, the observed patterns of associations among ants, coccids and plants are shaped by a variety of factors (each contributing to varying degrees depending on ecological conditions), and possibly by chance as well.

After more than 30 years of intensive research on *Macaranga* and their ants, we have a reasonable grasp of the ecological and/or physiological differences among the plants and among the ants, and what may have driven their respective diversifications. While we may have made some advances in the taxonomy of the *Macaranga* coccids herein, many species remain poorly (and un-) documented, preventing inferences that might lead to hypotheses about how their association with *Macaranga* or *Crematogaster* evolved. Much more work remains to be done before we can understand how this group originated and diversified. In the face of unrelenting destruction of their habitat, this will be an endeavour that is as urgent as it is challenging.

ACKNOWLEDGEMENTS

We sincerely thank K. Inamori, S. Finnegan and Y. Sato for analytical support; S. J. Davies; Y. Inui, C. Handa, T. Okubo, U. Shimizu-kaya, T. Komatsu and Y. Kataoka for sampling support; C. R. Marshall and P. Roopnarine for critical comments, C. Looy for logistic support, S. Usami and S. Matsuda for technical support, and L. Yoon for graphic advice. We are especially grateful to three anonymous reviewers whose comments facilitated substantial improvement of the manuscript. We thank Sabah Parks, National Parks Board (Singapore) and Sarawak Forest Department for collecting permission. This research was funded by the Japan Ministry of Education, Culture, Sports, Science and Technology (18370010 and 22255001 to TI, 21870012 and 23770018 to SU).

REFERENCES

Bensasson D, Zhang D-X, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution* **16:** 314–321.

Blattner FR, Weising K, Bänfer G, Maschwitz U, Fiala B. 2001. Molecular analysis of phylogenetic relationships among myrmecophytic *Macaranga* species (Euphorbiaceae). *Molecular Phylogenetics and Evolution* **19:** 331–344.

Braby MF, Vila R, Pierce NE. 2006. Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zoological Journal of the Linnean Society* **147:** 239–275.

Brower AVZ, Desalle R. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Molecular Biology* **7:** 73–82.

Davies SJ. 2001. Systematics of *Macaranga* sects. *Pachystemon* and *Pruinosae* (Euphorbiaceae). *Harvard Papers in Botany* **6:** 371–448.

Davies SJ, Lum SKY, Chan R, Wang LK. 2001. Evolution of myrmecophytism in Western Malesian *Macaranga* (Euphorbiaceae). *Evolution* **55:** 1542–1559.

Edwards DP, Yu DW, Hassall M, Sutherland WJ. 2006. Assembling a mutualism: ant symbionts locate their host plants by detecting volatile chemicals. *Insectes Sociaux* **53:** 172–176.

Federle W, Rheindt FE. 2005. *Macaranga* ant–plants hide food from intruders: correlation of food presentation and presence of wax barriers analyzed using phylogenetically independent contrasts. *Biological Journal of the Linnean Society* **84:** 177–193.

Federle W, Maschwitz U, Fiala B, Riederer M, Hölldobler B. 1997. Slippery ant–plants and skillful climbers: selection and protection of specific ant partners by epicuticular wax blooms in *Macaranga* (Euphorbiaceae). *Oecologia* **112:** 217–224.

Federle W, Rohrseitz K, Hölldobler B. 2000. Attachment forces of ants measured with a centrifuge: better 'wax-runners' have a poorer attachment to a smooth surface. *Journal of Experimental Biology* **203:** 505–512.

Federle W, Fiala B, Zizka G, Maschwitz U. 2001. Incident daylight as orientation cue for hole-boring ants: prostomata in *Macaranga* ant plants. *Insectes Sociaux* **48:** 165–177.

Federle W, Maschwitz U, Hölldobler B. 2002. Pruning of host plant neighbours as defence against enemy ant invasions: *Crematogaster* ant partners of *Macaranga* protected by 'wax barriers' prune less than their congeners. *Oecologia* **132:** 264–270.

Feise RJ. 2002. Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology* **2:** 8.

Feldhaar H, Fiala B, Gadau J, Mohamed M, Maschwitz U. 2003a. Molecular phylogeny of *Crematogaster* subgenus *Decacrema* ants (Hymenoptera: Formicidae) and the colonization of *Macaranga* (Euphorbiaceae) trees. *Molecular Phylogenetics and Evolution* **27:** 441–452.

Feldhaar H, Fiala B, Hashim RB, Maschwitz U. 2003b. Patterns of the *Crematogaster-Macaranga* association: the ant partner makes the difference. *Insectes Sociaux* **50:** 9–19.

Feldhaar H, Fiala B, Gadau J. 2004. Characterization of microsatellite markers for plant ants of the genus

Crematogaster subgenus Decacrema. *Molecular Ecology Notes* **4:** 409–411.

Feldhaar H, Gadau J, Fiala B. 2010. Speciation in obligately plant-associated *Crematogaster* ants: host distribution rather than adaption towards specific hosts drives the process. In: Glaubrecht M, Schneider H, eds. *Evolution in action: case studies in adaptive radiation, speciation and the origin of biodiversity*. New York: Springer Verlag Berlin Heidelberg, 193–213.

Feldhaar H, Maschwitz U, Fiala B. 2016. Taxonomic revision of the obligate plant-ants of the genus *Crematogaster* Lund (Hymenoptera: Formicidae: Myrmicinae), associated with *Macaranga* Thouars (Euphorbiaceae) on Borneo and the Malay Peninsula. *Sociobiology* **63:** 651–681.

Fiala B, Maschwitz U. 1992. Food bodies in the genus *Macaranga* and their significance for the evolution of myrmecophytism. *Botanical Journal of the Linnean Society* **110:** 61–75.

Fiala B, Maschwitz U, Pong TY, Helbig AJ. 1989. Studies of a South East Asian ant–plant association: protection of *Macaranga* trees by *Crematogaster borneensis*. *Oecologia* **79:** 463–470.

Fiala B, Jakob A, Maschwitz U, Linsenmair KE. 1999. Diversity, evolutionary specialization and geographic distribution of a mutualistic ant–plant complex: *Macaranga* and *Crematogaster* in South East Asia. *Biological Journal of the Linnean Society* **66:** 305–331.

Gill RJ. 1988. *The scale insects of California: part 1. The soft scales (Homoptera: Coccoidea: Coccidae)*. Sacramento, CA: California Department of Food & Agriculture.

Grangier J, Dejean A, Malé P-JG, Solano P-J, Orivel J. 2009. Mechanisms driving the specificity of a myrmecophyte–ant association. *Biological Journal of the Linnean Society* **97:** 90–97.

Handa C, Itioka T. 2011. Effects of symbiotic coccid on the plant-ant colony growth in the myrmecophyte *Macaranga bancana*. *Tropics* **19:** 139–144.

Handa C, Ueda S, Tanaka H, Itino T, Itioka T. 2012. How do scale insects settle into the nests of plant-ants on *Macaranga* myrmecophytes? Dispersal by wind and selection by plant-ants. *Sociobiology* **59:** 1–12.

Heckroth H-P, Fiala B, Gullan PJ, Idris AH, Maschwitz U. 1998. The soft scale (Coccidae) associates of Malaysian ant–plants. *Journal of Tropical Ecology* **14:** 427–443.

Heckroth H-P, Fiala B, Maschwitz U. 2001. Integration of scale insects (Hemiptera: Coccidae) in the South-East Asian ant–plant (*Crematogaster* (Formicidae) – *Macaranga* (Euphorbiaceae)) system. *Entomologica, Bari* **33:** 287–295.

Heil M, Fiala B, Kaiser W, Linsenmair KE. 1998. Chemical contents of *Macaranga* food bodies: adaptations to their role in ant attraction and nutrition. *Functional Ecology* **12:** 117–122.

Heil M, Fiala B, Maschwitz U, Linsenmair KE. 2001. On benefits of indirect defence: short- and long-term studies of antiherbivore protection via mutualistic ants. *Oecologia* **126:** 395–403.

Hodgson CJ. 1994. *The scale insect family Coccidae: an identification manual to genera*. Wallingford, UK: CAB International.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Inui Y, Itioka T, Murase K, Yamaoka R, Itino T. 2001. Chemical recognition of partner plant species by foundress ant queens in *Macaranga-Crematogaster* myrmecophytism. *Journal of Chemical Ecology* **27:** 2029–2040.

Itioka T, Nomura M, Inui Y, Itino T, Inoue T. 2000. Difference in intensity of ant defense among three species of *Macaranga* myrmecophytes in a southeast Asian dipterocarp forest. *Biotropica* **32:** 318–326.

Jobb G, von Haeseler A, Strimmer K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology* **4:** 18.

Jürgens A, Feldhaar H, Feldmeyer B, Fiala B. 2006. Chemical composition of leaf volatiles in *Macaranga* species (Euphorbiaceae) and their potential role as olfactory cues in host-localization of foundress queens of specific ant partners. *Biochemical Systematics and Ecology* **34:** 97–113.

Lin Y-P, Kondo T, Gullan P, Cook LG. 2013. Delimiting genera of scale insects: molecular and morphological evidence for synonymising *Taiwansaisettia* Tao, Wong and Chang with Coccus Linnaeus (Hemiptera: Coccoidea: Coccidae). *Systematic Entomology* **38:** 249–264.

Moog J, Saw LG, Hashim R, Maschwitz U. 2005. The triple alliance: how a plant ant, living in an ant–plant, acquires the third partner, a scale insect. *Insectes Sociaux* **52:** 169–176.

Morrison H. 1921. Some nondiaspine Coccidae from the Malay Peninsula, with descriptions of apparently new species. *Philippine Journal of Science* **18:** 637–677.

Nomura M, Itioka T, Itino T. 2000. Variations in abiotic defense within myrmecophytic and non-myrmecophytic species of *Macaranga* in a Bornean dipterocarp forest. *Ecological Research* **15:** 1–11.

Quek S-P, Davies SJ, Itino T, Pierce NE. 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.

Quek S-P, Davies SJ, Ashton PS, Itino T, Pierce NE. 2007. The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. *Molecular Ecology* **16:** 2045–2062.

Quek S-P, Ueda S, Gullan PJ, Kondo T, Hattori M, Itioka T, Murase K, Itino T. 2016. Data from: Nuclear-DNA-based species delineations of *Coccus* scale insects in symbiosis with plants and ants, and the role of plant epicuticular wax in structuring associations. *Dryad Digital Repository*. doi: 10.5061/dryad.6q1q6.

Rothman KJ. 1990. No adjustments are needed for multiple comparisons. *Epidemiology* **1:** 43–46.

Sunnucks P, Hales DF. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in

aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution* **13:** 510–524.

Swofford DL. 2002. *PAUP*: phylogenetic analysis using parsimony. Ver. 4.0b10*. Sunderland, MA: Sinauer Associates.

Takahashi R. 1952. Some species of nondiaspine scale insects from the Malay Peninsula. *Insecta Matsumurana* **18:** 9–17.

Tanabe AS. 2007. KAKUSAN: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. *Molecular Ecology Notes* **7:** 962–964.

Ueda S, Quek S-P, Itioka T, Inamori K, Sato Y, Murase K, Itino T. 2008. An ancient tripartite symbiosis of plants, ants and scale insects. *Proceedings of the Royal Society B* **275:** 2319–2326.

Ueda S, Quek S-P, Itioka T, Murase K, Itino T. 2010. Phylogeography of the *Coccus* scale insects inhabiting myrmecophytic *Macaranga* plants in Southeast Asia. *Population Ecology* **52:** 137–146.

Ueda S, Nagano Y, Kataoka Y, Komatsu T, Itioka T, Shimizu-kaya U, Inui Y, Itino T. 2015. Congruence of microsatellite and mitochondrial DNA Variation in acrobat ants (*Crematogaster* Subgenus *Decacrema*, Formicidae: Myrmicinae) inhabiting *Macaranga* (Euphorbiaceae) myrmecophytes. *PLoS ONE* **10:** e0116602.

Williams DJ, Granara de Willink MC. 1992. *Mealybugs of Central and South America*. Wallingford: CAB International.

Williams DJ, Watson GW. 1990. *The scale insects of the tropical South Pacific region. Part 3: the soft scales (Coccidae) and other families*. Wallingford: CAB International.

Xia X, Xie Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity* **92:** 371–373.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Saturation plots for the datasets.
Table S1. List of *Coccus* samples.
Table S2. Tally of number of specimens, localities and coccid species for each *Macaranga* host plant species.
Table S3. Character statistics for DNA data sets.
Table S4. Substitution models selected by BIC5.
Table S5. List of additional 40 *Macaranga*-associated *Crematogaster* samples used in this study.
Table S6. Tests for preference or avoidance between all pairwise combinations of O(T)Us among guilds, using exact multinomial and exact binomial tests.
Table S7. Species/morphospecies reported in Heckroth *et al.* (1998) compared with those in the present study.
Notes. Supporting information cited in RESULTS and DISCUSSION.

SHARED DATA

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.6q1q6 (Quek *et al.*, 2016).