Syst. Biol. 50(2):194–214, 2001

Multiple Molecular Data Sets Suggest Independent Origins of Highly Eusocial Behavior in Bees (Hymenoptera:Apinae)

SYDNEY A. CAMERON[1,2] AND PATRICK MARDULYN[3]

Department of Entomology, University of Arkansas, Fayetteville, Arkansas 72701, USA; E-mail: scameron@comp.uark.edu

Abstract.—Different views of the pattern of social evolution among the highly eusocial bees have arisen as a result of discordance between past molecular and morphology-based phylogenies. Here we present new data and taxa for four molecular data sets and reassess the morphological characters available to date. We show there is no significant character incongruence between four molecular data sets (two nuclear and two mitochondrial), but highly significant character incongruence leads to topological incongruence between the molecular and morphological data. We investigate the effects of using different outgroup combinations to root the estimated tree. We also consider various ways in which biases in the sequence data could be misleading, using several maximum likelihood models, LogDet corrections, and spectral analyses. Ultimately, we concede there is strong discordance between the molecular and morphological data partitions and appropriately apply the conditional combination approach in this case. We also find two equally well supported placements of the root for the molecular trees, one supported by 16S and 28S sequences, the other supported by cytochrome *b* and opsin. The strength of the evidence leads us to accept two equally well supported hypotheses based on analyses of the molecular data sets. These are the most rigorously supported hypotheses of corbiculate bee relationships at this time, and frame our argument that highly eusocial behavior within the corbiculate bees evolved twice independently. [Apines; combined phylogenetic analysis; corbiculate bees; Hymenoptera; insects; molecular phylogeny.]

The four tribes of corbiculate bees, named for the specialized pollen-carrying structure on the hindleg known as the corbicula, exhibit novel features distinguishing them from the rest of the bees (Apoidea). The four tribes include the major groups of social bees, honey bees (Apini) and stingless bees (Meliponini), which form some of the most elaborate societies on earth; the intermediately social bumblebees (Bombini); and the mostly solitary, occasionally communal orchid bees (Euglossini) (Michener, 1974; Garófalo et al., 1998). Corbiculate bees are often the dominant pollinators within both natural and agro-ecosystems throughout temperate and tropical regions of the world. These bees offer the entire range of social organizations for comparative study and provide a rich source for investigating patterns of social evolution.

From an evolutionary standpoint, the form of highly eusocial behavior found in Apini and Meliponini is unique among bees. No other bees exhibit the extreme obligate division of labor between queen and worker, in which queens are exclusively responsible for egg-laying and can do little else, while the many sterile workers perform the remaining tasks of colony maintenance. Along a conceptual social spectrum, eusociality in Bombini appears to be intermediate between the elaborate obligate form acquired by Apini/Meliponini and the more plastic eusocial systems characteristic of some of the phylogenetically distant sweat bees (Halictidae) (Packer, 1990; Eickwort et al., 1996). Halictid-like plasticity is not found in the corbiculate bees, and the form of eusociality exhibited by halictids is uniquely derived within that family and not homologous to that of the corbiculate bees.

Knowledge of the historical pattern of evolution among the corbiculate bees is implicit to understanding the evolution of their diverse social behavior. Currently, two main phylogenetic hypotheses are competing: Either the two highly eusocial tribes are sister taxa that arose from a common highly eusocial ancestor (single origin), or the highly eusocial tribes arose independently along different lineages (dual origin). Most morphology-based systematists, most notably Michener, have classified Apini and Meliponini as a distinct monophyletic group (Michener, 1944, 1974, 1990; Prentice, 1991;

[1]Author for correspondence.

[2]Current address: Department of Entomology, University of Illinois, 505 S. Goodwin Ave., Urbana, Illinois 61801, USA. E-mail: scameron@life.uiuc.edu

[3]Current address: Unit of Evolutionary Genetics (CP 300), Free University of Brussels (ULB), 6041 Gosselies, Belgium.

Roig-Alsina and Michener, 1993; Schultz et al., 1999), thus inferring a single transition from primitive eusociality to highly eusocial behavior. We refer to this as the traditional view. Occasionally, other investigations of morphology have called into question the single-origin hypothesis (Winston and Michener, 1977; Kimsey, 1984; Plant and Paulus, 1987). However, the strongest challenge to the traditional view has come from independent investigations of molecular data (Cameron, 1993, and references therein; Koulianos et al., 1999; Mardulyn and Cameron, 1999), which consistently reject the hypothesis that Apini and Meliponini belong to the same clade. Instead, each molecular data set to date shows a strong relationship between Bombini and Meliponini, thus suggesting two independent origins of highly eusocial behavior.

The accumulation of multiple molecular data sets contradicting the results of morphological analyses has focused attention on the need for a unified phylogenetic framework for the corbiculate bees. Recently, Chavarría and Carpenter (1994) voiced criticism against analyzing the morphological and molecular data sets independently, arguing for simultaneous analysis of multiple data sets (Kluge, 1989; Nixon and Carpenter, 1996; DeSalle and Brower, 1997, and references therein). Their combined parsimony analysis of the only consequential molecular data set available at the time (Cameron, 1993), together with several partially overlapping morphological data sets, supported the traditional view of the monophyly of Apini and Meliponini. However, the approach of combining strongly discordant data in a global parsimony analysis, without consideration of heterogeneous patterns of character change, can be ineffective (Larson, 1994; Lockhart and Cameron, 2001) or lead to erroneous results (Huelsenbeck et al., 1996; Cunningham, 1997). In addition, a much larger body of molecular data is now available for study than was available at the time of their reanalysis.

The purpose of the present study is to examine the basis of the discordance between the molecular and morphology trees. We include in our analysis sequences from two nuclear genes, the major opsin (LW *Rh*; Mardulyn and Cameron, 1999) and 28S (D2 and D3 expansion region, new data reported here). We have also sequenced additional taxa to supplement previously applied mtDNA data sets, 16S (Cameron, 1993) and cytochrome *b* (cyt b) (Koulianos et al., 1999).

One problem with previous studies comparing different data sets is that the exemplar species examined for each corbiculate tribe have differed widely among studies. We therefore chose a set of exemplars that maximizes the overlap among all the molecular studies and the morphological investigation of Roig-Alsina and Michener (1993), which consolidated the largest number of new and previously used characters to date. Five outgroup taxa were selected from four other monophyletic tribes within the Apidae (Roig-Alsina and Michener, 1993): *Melissodes* (Eucerini), *Centris* (Centridini), *Anthophora* and *Habropoda* (Anthophorini), and *Xylocopa* (Xylocopini).

To compare the DNA and morphological data, we constructed a new data matrix of all available morphological characters (taken from the literature) for these taxa. The ultimate solution to this problem of incongruence should inform not only the pattern of highly eusocial evolution in the bees but also some of the underlying causes of incongruence among data from different sources.

Materials and Methods

Taxa Examined

The use of exemplars to represent the corbiculate tribes is justified on the basis that each tribe has been recognized as a monophyletic group by several independent studies of morphology and DNA (summarized in Michener, 1990, and references therein; Cameron, 1993). However, because taxon sampling is a critical aspect in phylogenetic studies (Hillis, 1998; Graybeal, 1998; Poe, 1998), we sampled across a greater diversity of tribal genera and species than in prior molecular or recent morphological analyses of the corbiculate bees. For the Apini (a monogeneric tribe) we sampled from three of the six described species. For Euglossini we sampled from four of the five genera. For Bombini (another monogeneric tribe) we sampled from 3 of the 35 described subgenera. For Meliponini we sampled from 4 of the 21 genera. Outgroups were selected to represent a spectrum of phylogenetic affinity to the corbiculate clade as proposed by Roig-Alsina and Michener (1993), *Centris*

TABLE 1. Taxa examined.

Subfamily/tribe	Species	Collection site	Collector[b]
Apinae			
Apini (6)[a]	*Apis mellifera*	Arkansas, USA	SAC
	Apis nigrocincta	Sulawesi	GWO
	Apis dorsata	India	SAC
Bombini (239)	*Bombus pennsylvanicus*	Arkansas, USA	SAC
	Bombus avinoviellus	India	SAC
	Bombus terrestris	Great Britain	HS
Meliponini (430)	*Trigona hypogea*	Brazil	SAC
	Scaptotrigona depilis	Brazil	SAC
	Tetragona dorsalis	Panama	DWR
	Lestrimelitta limao	Panama	DWR
	Melipona compressipes	Brazil	SAC
Euglossini (174)	*Eufriesea caerulescens*	Mexico	SAC
	Euglossa imperialis	Panama	RSH
	Exaerete frontalis	Panama	RSH
	Eulaema meriana	Costa Rica	SAC
Outgroups			
Apinae			
Eucerini	*Melissodes rustica*	Arkansas, USA	SAC
Centridini	*Centris inermis*	Costa Rica	SAC
Anthophorini	*Anthophora pacifica*	California, USA	RWT
	Habropoda depressa	California, USA	RWT
Xylocopinae			
Xylocopini	*Xylocopa virginica*	Missouri, USA	SAC

[a]Numbers in parentheses next to the four corbiculate bee tribal names are the approximate numbers of species in those tribes.
[b]SAC, Sydney Cameron; GWO, Gard Otis; HS, Horst Schwarz; DWR, David Roubik; RSH, Regula Schmid-Hempel; RWT, Robbin Thorp.

being considered the closest relative and *Xylocopa* the most distant. Table 1 indicates the taxonomic affinity, collection site, and collector for each species examined. Voucher specimens for all taxa used in this investigation are deposited in the Illinois Natural History Survey of the University of Illinois, Urbana-Champaign.

Morphology

All morphological characters and character states used in this study (Appendix 1) were taken from Roig-Alsina and Michener's (1993) analysis of the long-tongued bees. Only characters informative within the ingroup as defined here (Apini + Meliponini + Bombini + Euglossini), or supporting the monophyly of the ingroup relative to the outgroup, or supporting a relationship between one or more of the outgroups to the ingroup are included from that study. The 95 characters include both adult (70) and larval (25) attributes. Several characters were taken directly or reformulated from earlier studies: Winston and Michener (1977); McGinley (1981); Kimsey (1984); Michener and Brooks (1984); Schönitzer (1986); Plant and Paulus (1987); Michener (1990); and Prentice (1991). Chavarría and Carpenter (1994) also recoded some characters from these studies, but we have used the Roig-Alsina and Michener codings, which are more detailed and were examined across a greater number of taxa, including outgroups. Appendix 2 shows the matrix of morphological character states used in the analyses.

Polymerase Chain Reaction (PCR) and DNA Sequencing

Genomic DNA was extracted from fresh, frozen (−80°C), and ethanol-preserved tissue from the thorax, abdomen, or legs of each bee. Tissue was ground in sodium dodecyl sulfate homogenization buffer, incubated for 1–2 h with proteinase K at 60°C, followed by four phenol/chloroform extractions, ethanol precipitation, and resuspension in TE buffer (10 mM Tris, 1 mM EDTA). For each species we obtained sequences of the following: (1) a fragment of the opsin LW*Rh* gene ~700 bp long, including 502 bp of coding sequence corresponding to nucleotide positions 421–922 of the *Apis mellifera* sequence published by Chang et al. (1996), which we obtained by using primers given in Mardulyn and Cameron (1999); (2) a 586-bp fragment of the cyt b

mitochondrial gene, by using the primers 5′-CGT TTA ATT CAY ATA AAT GG-3′ (Koulianos et al., 1999) and the CB-N-11367 5′-ATT ACA CCT CCT AAT TTA TTA GGA AT-3′ (Simon et al., 1994); (3) a fragment of the ribosomal 28S gene ~680 bp long, including the D2 and D3 expansion regions (Hancock et al., 1988), by using the primers 5′-AAG AGA GAG TTC AAG AGT ACG TG-3′ (modified from Belshaw and Quicke, 1997) and 5′-TAG TTC ACC ATC TTT CGG GTC CC-3′ (Mardulyn and Whitfield, 1999); (4) a fragment of the ribosomal 16S mitochondrial gene ~530 bp long, by using the primers 16SWb 5′-CAC CTG TTT ATC AAA AAC AT-3′ (Dowton and Austin, 1994) and 874–16SlR (Cameron et al., 1992). A subset of the 16S and cyt b sequences were taken from Cameron (1993) and Koulianos et al. (1999), respectively. PCR conditions included an initial denaturation step of 30 sec at 94°C, followed by 35 cycles of a 60-sec denaturation at 94°C, 60-sec annealing at 55–60°C (opsin, 28S) or 50–55°C (cyt b, 16S), and 60-sec extension at 72°C, with a final extension step of 2 min at 72°C. PCR products were purified by using the Wizard® PCR Prep DNA Purification System (Promega). Sequencing was conducted with an ABI 377 automated sequencer, using the PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit FS (Perkin-Elmer) according to manufacturer's specifications. Both strands were sequenced for all taxa.

DNA Sequences

New sequences were deposited in GenBank under accession numbers AF181577–AF181618 and aligned sequences can be downloaded from the *Systematic Biology* website or from TreeBase (Sanderson et al., 1994). The cyt b and opsin sequences were aligned manually with the program SeqPup version 0.6 (Gilbert, 1996). The 28S and 16S sequences were aligned based on the criterion of maximum parsimony by using the program Malign 2.5 (command "build", randorderns 10, align swap, contig) (Wheeler and Gladstein, 1995), specifying a gap cost/change cost ratio of 1.5, 2, 4, and 6. CLUSTAL X (Thompson et al., 1997) was also used to align 28S and 16S (gap opening of 10, gap extension of 0.05), for comparison with Malign. With CLUSTAL, the ingroup sequences were first aligned separately, then the outgroup sequences were added and aligned to the ingroup sequences (profile alignment in CLUSTAL). To explore the effects of different alignment parameters, modified 16S and 28S data sets were constructed, excluding variable-length regions that are difficult to align unambiguously. Such regions were identified by comparing the different alignments obtained for each data set under the different alignment methods described above.

Base frequencies and uncorrected pairwise nucleotide sequence divergences were calculated for each gene fragment by using the computer program PAUP*4.0 (test versions 4.0 d64 [provided by D. L. Swofford] and beta version b1a and b2a [Swofford, 1998]). A chi-square test of homogeneity of base frequencies across taxa was performed with PAUP*4.0. MacClade 3.07 (Maddison and Maddison, 1992) was used to estimate the frequency distribution of observed number of substitutional changes per character for each gene and for each codon position of the protein-coding genes. Nucleotide sequences of the cyt b and opsin fragments were translated into amino acid sequences by using MacClade.

Phylogenetic Analyses

Tribal relationships of the corbiculate bees were inferred from parsimony analyses, implemented in PAUP*4.0. All parsimony analyses utilized branch-and-bound searches, unless otherwise indicated. Parsimony procedures for 28S and 16S sequences were applied to both the reduced (variable-length regions excluded) and complete data sets (Malign alignment, gap cost/change cost ratio-4 for 28S, 1.5 for 16S). Maximum parsimony trees were first estimated separately for each of the four nucleotide data sets and the morphological data set. Gaps were treated as missing characters. Unrooted parsimony analyses were conducted first, using the ingroup taxa only. To root the trees, parsimony analyses were conducted after addition of five outgroup taxa (see Table 1). This two-step procedure is justified by the fact that rooting the corbiculate bees tree is problematic (see Results). Conducting unrooted analyses first allows us to separate the problem of estimating the phylogenetic relationships among the four corbiculate bee tribes from the problem of inferring the position of the root. Additional analyses included linking

data sets into the following partitions: (a) protein-coding genes, (b) ribosomal DNA, (c) all DNA data, (d) DNA and morphology data. Before assembling multiple data sets into new data partitions, we thought it important to estimate the homogeneity among the original and the new partitions. Incongruence length difference (ILD) tests (Farris et al., 1994) were implemented in PAUP* with invariant characters removed (Cunningham, 1997). Multiple pairwise tests were performed on each of the original molecular partitions (16S, cyt b, opsin, 28S) and on new partitions (a) versus (b) and (c) versus morphology.

A weighted parsimony analysis was performed on the opsin LW*Rh* data set because of the demonstrated saturation of transitions (TIs) at third position sites in this gene within Apinae (Mardulyn and Cameron, 1999). Following the rationale of Mardulyn and Cameron, we weighted TIs in third positions four times lower than other substitutions in this analysis. Possible saturation of different substitution types was also investigated for the other gene fragments by plotting the number of TIs against the number of transversions (TVs) for all possible pairs of taxa.

Bootstrap analyses implemented in PAUP* (1,000 replicates, heuristic search, simple addition sequence, TBR swapping) and decay indices (also known as Bremer support values; Bremer, 1988) implemented in TreeRot 1 (Sorenson, 1996) were performed to provide measures of relative support for each node estimated in the above analyses.

We applied maximum likelihood (ML) analysis to each molecular data set, which explicitly allows us to correct for multiple substitutions at a given site and for rate heterogeneity among sites. Two ML models were used, depending on the class of sequence data. The General Time Reversible (GTR) (Yang, 1994) model was applied to the AT-rich 16S and cyt b sequences (see Whitfield and Cameron, 1998); the HKY85 model (Hasegawa et al., 1985; see Yang, 1993) was applied to the 28S and opsin sequences. The TI/TV ratio, gamma shape parameter, and proportion of invariable sites were estimated directly from the data during the ML analysis, when possible, or were estimated before the ML analysis by using the maximum parsimony (MP) tree as a first reasonable estimate of the phylogeny (Swofford et al., 1996). The proportion of invariable sites was also estimated with outgroups excluded to minimize the influence of high amounts of change contributed by the outgroups. Base frequencies were set to their empirical values. ML searches were implemented in PAUP* as heuristic searches, as-is addition sequence, and TBR branch swapping.

Incorrect phylogenetic estimations can occur when patterns of substitutions vary across the tree (nonstationarity), resulting in differences in base composition among lineages (Hasegawa and Hashimoto, 1993; Lockhart et al., 1994) or differences in the distribution of variable sites (covariotide/covarion pattern of changes; Lockhart et al., 1996, 1998). Nonstationarity can lead to phylogenetic errors as follows. If, by chance, two distantly related taxa share a more similar distribution of invariable sites than they share with more closely related taxa, they will tend to attract one another erroneously in phylogenetic analysis. In very much the same way, two taxa that have a similar base composition will tend to attract each other, even if they are distantly related. Given that two different rootings of the ingroup were inferred (described below) with the four molecular data sets, we tested whether the phylogenetic position of either root could have been incorrectly estimated as a result of nonstationarity. Each molecular data set was analyzed separately with each outgroup, one at a time. Using an approach described by Steel et al. (2000), we constructed PAUP* matrices of LogDet distances (correcting for differences in base composition among taxa; Lockhart et al., 1994), after removing different proportions of the sites (0%, 30%, 50%, and 60%) that were assumed to be invariable. Estimates of the proportion of invariable sites were also calculated by using both ML and the capture–recapture method (implemented with SplitsTree 2.4; Huson, 1998) described in Steel et al. (2000). To test whether the support or conflict for the position of the root of the tree was stable to the removal of an increasing number of invariant sites, we obtained a distance Hadamard spectrum (Penny et al., 1996; implemented with Spectrum 2.0 [Charleston and Page, 1997]) for each LogDet distance matrix. We also used parametric bootstrapping (Huelsenbeck, 1997) to test whether the rooting conflict observed among the different data sets could be the result of the so-called "long-branch attraction"

artifact (Felsenstein, 1978; Hendy and Penny, 1989). For that test, simulated data sets were generated by using SequenceGenerator version 1.1 (Rambaut and Grassly, 1997).

Lastly, we used Templeton's test (Templeton, 1983) to determine whether nodes that conflict between two trees, estimated from different data partitions, were significantly incompatible (see examples in Cameron, 1993; Cunningham, 1997; Larson, 1994). This test allows exploration of the goodness of fit of alternative trees to a given data partition. For example, the morphology data can be optimized onto a tree topology estimated from the molecular data (alternative tree) and tested to determine whether the alternative topology is significantly less optimal for the morphology data than is the maximum parsimony tree estimated from the morphology data (optimal tree). Parsimony analysis of the data partitions optimized onto alternative topologies was implemented in PAUP* (branch-and-bound) to obtain the number of extra steps per character required by the alternative topology. Statistical assessment of the number of extra steps required by the alternative topology relative to the optimal tree was made by using the Wilcoxon sign rank test (Wardlaw, 1985). Failure to detect a significant difference ($P < 0.05$) between the number of steps for the optimal and alternative trees would suggest, following the above example, that the molecular tree is nearly optimal for the morphology data. Hence the conflict between the data sets for a given topology would be insignificant.

RESULTS

Data Characteristics

Ninety-five phylogenetically informative morphological characters (70 adult and 25 larval) were identified and taken from the Roig-Alsina and Michener (1993) data set (Appendices 1 and 2). We grouped the characters into eight categories: head (14 characters), mouthpart (15), mesosomal (11), leg (15), wing (6), metasomal (5), male genitalic (4), and larval (25). Fifty-one of these characters were informative within the ingroup; the remainder were included to clarify outgroup relationships. Several of the characters informative within the ingroup are coded as presence/absence (e.g., characters 29, 55, 63, 67) and, therefore, have the potential to group taxa based on the convergent loss of a feature. Although we recognize this as a potential problem, there is no strong a priori rationale for excluding these characters.

With the introduction of gaps, the complete 16S and 28S data sets used for analyses contain 534 and 738 aligned sites, respectively. Of these sites, 185 were parsimony-informative sites for 16S, and 168 were parsimony-informative sites for 28S. For some analyses, 38 nucleotides were excluded from the 28S data set and 196 nucleotides from the 16S data set, corresponding to ambiguous regions of alignment (see Materials and Methods). The cyt b data set contains 586 aligned sites (239 parsimony-informative sites), and the opsin data set contains, after removing the two introns, 502 aligned sites (144 parsimony-informative sites). These two protein-coding genes do not include any indels. The average base frequencies for each gene fragment, and for the different codon positions of the two protein-coding genes, are given in Table 2. As expected, a strong AT bias is observed in the two mitochondrial genes, being particularly high at the third position sites for cyt b. The chi-square test of homogeneity of base frequencies across taxa resulted in significant P-values in only one case: for third position sites in the cyt b fragment ($P = 0.001$).

The occurrence of saturation of TIs relative to TVs in third positions was inferred for the gene fragments by plotting the number of TIs against the number of TVs. Third position TIs of opsin showed a pattern suggestive of saturation (Mardulyn and Cameron, 1999). However, no clear pattern of saturation was observed for the other three genes (results not shown). Because of the strong

TABLE 2. Average base frequencies for each molecular data set.

	Codon position	Base			
		A	C	G	T
Opsin					
	1st	32.6	13.7	25.7	28.0
	2nd	22.8	21.6	19.4	36.2
	3rd	21.2	30.3	22.2	26.4
Cyt b					
	1st	33.3	13.0	15.8	38.0
	2nd	23.3	19.1	12.0	45.7
	3rd	43.1	3.5	2.9	50.5
28S		19.9	28.7	30.0	21.4
16S		40.6	7.3	13.2	39.0

AT bias in the two mitochondrial genes, the majority of substitutions are TVs. In the case of the cyt b fragment, most of the changes have occurred at third positions. Although saturation of substitutions was not obvious at those sites, we nonetheless conducted parsimony analyses of the amino acid sequences for the cyt b fragment, thereby considering only nonsynonymous substitutions.

Morphological Analyses

A single most-parsimonious tree was obtained from analysis of the entire set of morphological characters (Fig. 1a). Separate analysis of the adult characters resulted in a single most-parsimonious-tree (Fig. 1b), and 20 equally parsimonious trees were obtained from analysis of the larval characters (strict consensus shown in Fig. 1c). The morphology-based trees a and b of Figure 1 show the same ingroup topology for the four tribes, (((Apini + Meliponini) + Bombini) + Euglossini), whereas tree c is mostly unresolved. All morphological characters combined, adult characters alone and larval characters alone show strong support (bootstrap value 92–99%; decay index 3–7) for a single highly eusocial clade (Apini + Meliponini). The larval characters do not resolve relationships among the remaining taxa (Fig. 1c). The Apini + Meliponini relationship is unaffected by the exclusion of potentially convergent (29, 55, 63, and 67) or highly homoplastic (60) characters supporting that clade.

Unrooted Molecular Analyses

The strict consensus of the most-parsimonious trees obtained from unrooted analysis of each of the four DNA data sets is shown in Figure 2. The potential monophyly of each of the four apine tribes is well supported by all four genes, with one exception: the tribe Bombini is paraphyletic for the 16S tree. More importantly, all four genes strongly support, with high bootstrap values, the same split among the four apine tribes: (Apini + Euglossini)(Bombini + Meliponini) (Fig. 2). The parsimony analysis performed on the amino acid sequences of cyt b resulted in the same tribal topology, with a bootstrap value of 99% (tree not shown). Note that regardless of where these trees might be rooted, it is

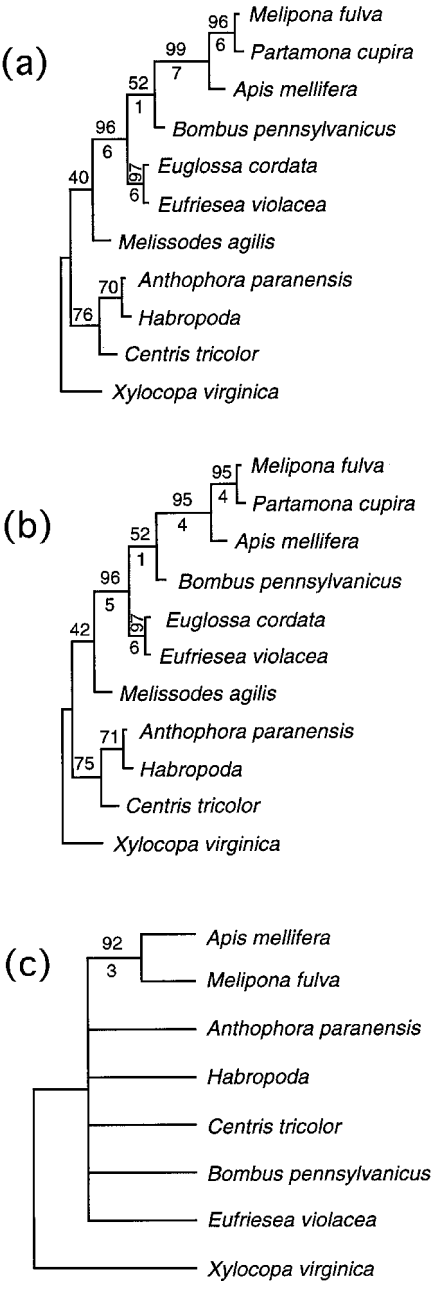

FIGURE 1. Rooted maximum parsimony trees (MPT) for six ingroup and five outgroup taxa of corbiculate bees estimated from morphological characters. Exhaustive searches were performed for all analyses. All characters were unordered. Numbers above branches are bootstrap values; numbers below are decay indices. (a) MPT estimated by using both adult and larval characters. Use of MP resulted in a single tree, having 190 steps; CI = 0.60, RI = 0.66. (b) MPT estimated by using adult characters only resulted in a single tree of 142 steps; CI = 0.61, RI = 0.68. (c) MPT estimated from larval characters only, yielding a strict consensus tree (48 steps) of 20 equally parsimonious trees.

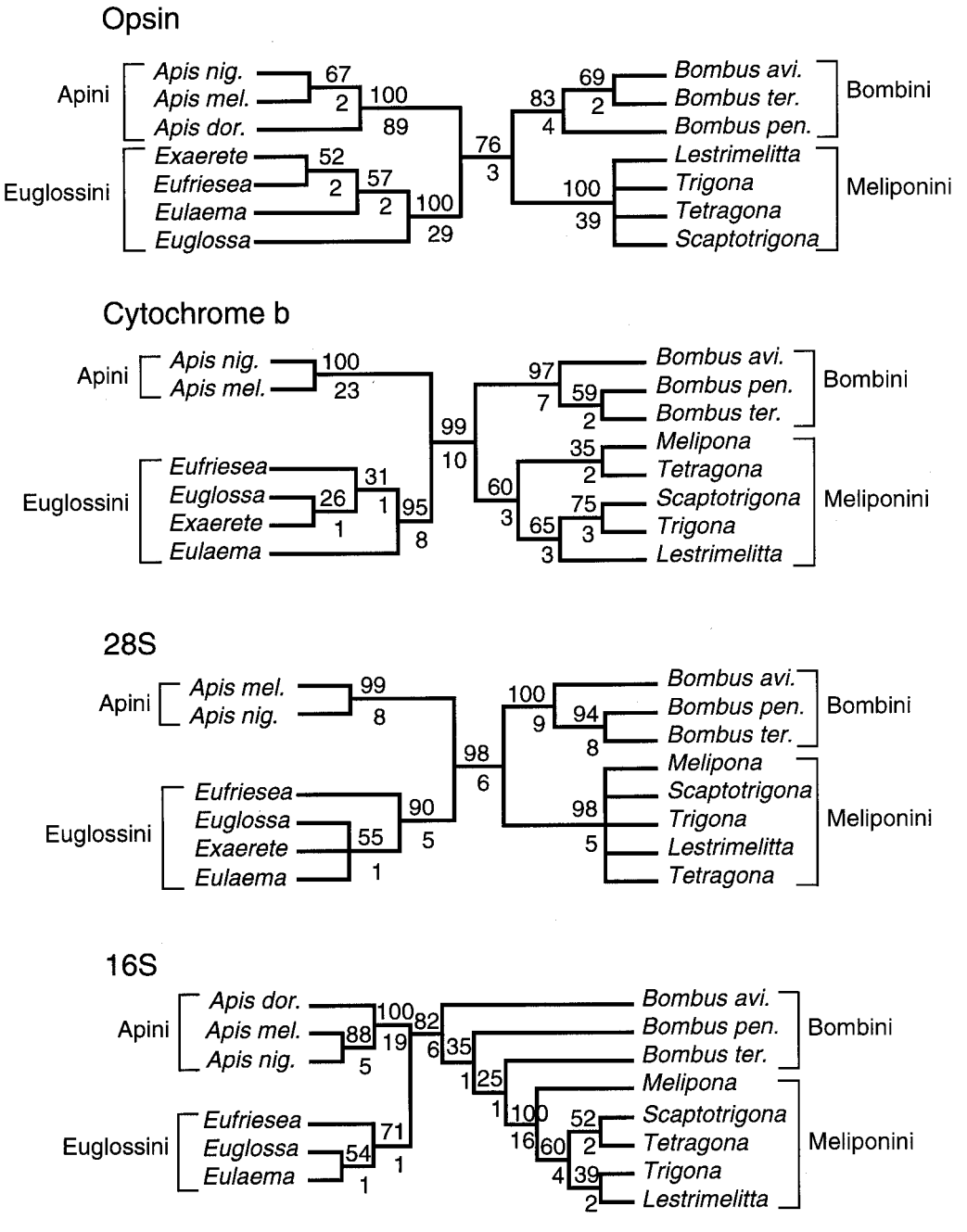

FIGURE 2. Most-parsimonious unrooted tree topologies (strict consensus) for only the 14 ingroup taxa of corbiculate bees, estimated from each of four gene fragments: opsin (weighted parsimony analysis, strict consensus of 3 MP trees, tree length = 604, CI = 0.79, RI = 0.86), cyt b (1 tree, tree length = 647, CI = 0.51, RI = 0.47), 28S (15 MP trees, tree length = 581, CI = 0.63, RI = 0.67), and 16S (1 MP tree, tree length = 448, CI = 0.56, RI = 0.60). Numbers above branches are bootstrap values; numbers below are decay indices.

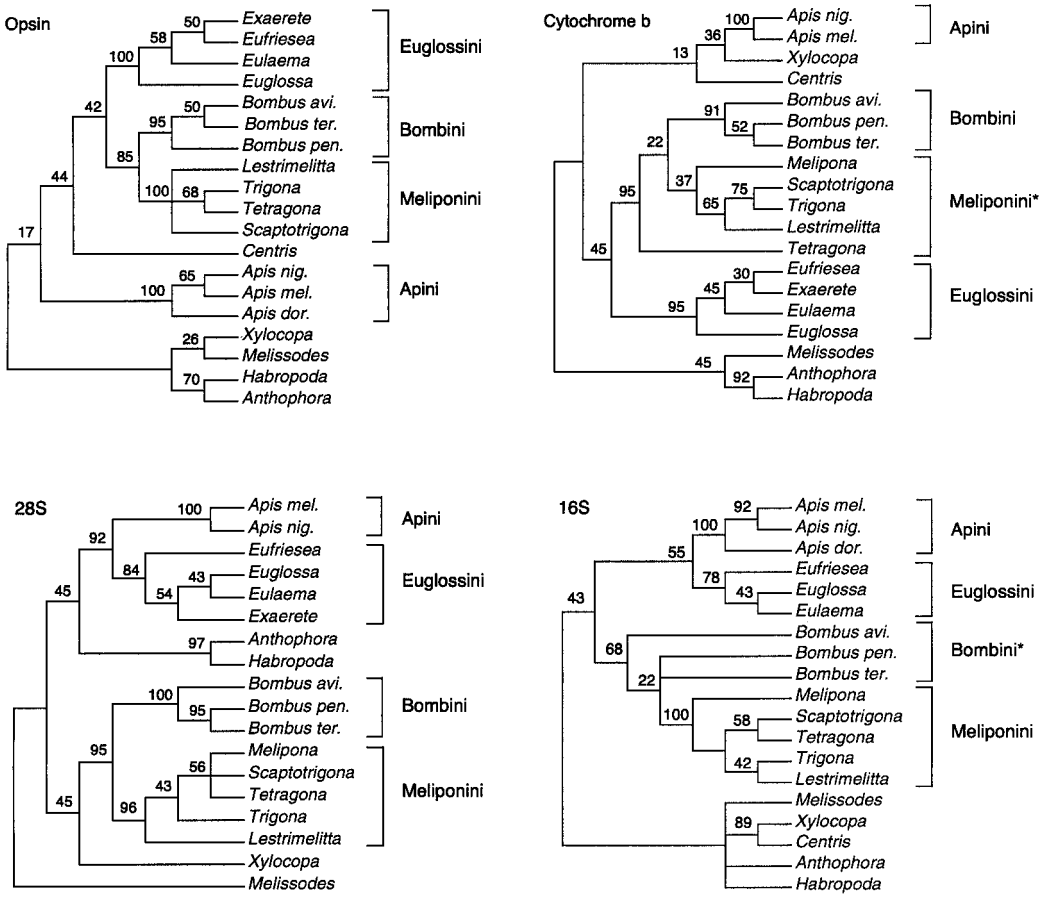

FIGURE 3. Most-parsimonious rooted tree topologies estimated from each of the four gene fragments using all outgroup taxa: opsin (weighted parsimony analysis, strict consensus of 2 trees, tree length = 1067, CI = 0.63, RI = 0.72), cyt b (1 tree, tree length = 984, CI = 0.44, RI = 0.45), 28S (2 trees, tree length = 754, CI = 0.73, RI = 0.59), and 16S (3 trees, tree length = 758, CI = 0.49, RI = 0.51). An asterisk located next to a tribe name means that this tribe does not emerge as monophyletic in the resulting tree (although non-monopoly is poorly supported).

not possible to infer the Apini + Meliponini clade, which was strongly supported by the morphological data.

Rooting the Molecular Trees

Parsimony analyses of all molecular data sets, including all outgroup sequences (but 28S sequence not available for *Centris*), resulted in MP trees (Fig. 3) that are fully compatible with the tribal relationships estimated from the unrooted trees (Fig. 2), with one exception: the tribe Meliponini is paraphyletic in the cyt b tree of Figure 3. In the opsin and cyt b trees, the root is placed along the branch leading to Apini. The 28S and 16S trees are rooted along the branch separating the (Bombini + Meliponini) and the (Apini + Euglossini) clades. However, the 28S, opsin, and cyt b topologies show a paraphyletic ingroup with respect to one or more outgroup taxa (Fig. 3). Because we have no a priori reason to believe that the corbiculate bees are paraphyletic, and because the pattern of paraphyly observed varies among the different genes, we investigated the possibility that ingroup paraphyly might be an artifact of our selection of outgroups.

Outgroup rooting can often be problematic, when little information is available on the phylogenetic position of the taxa outside the ingroup. Even if this information is available, the taxon identified as most closely related to the ingroup can still be distant enough from the ingroup taxa to make it difficult to identify the position of the root. Table 3 shows the mean p-distance (uncorrected distance) between each outgroup taxon and the

TABLE 3. Mean p-distances from each outgroup taxon to all ingroup taxa.

Outgroup	Opsin	Cyt b	28S	16S
Centris	0.127	0.207	—	0.362
Melissodes	0.172	0.199	0.120	0.223
Xylocopa	0.158	0.202	0.123	0.205
Anthophora	0.161	0.240	0.134	0.193
Habropoda	0.160	0.240	0.168	0.208

ingroup taxa. These distance values provide an a priori indication of which outgroup taxa are more distant from the ingroup, independent of their phylogenetic position (the outgroup sequence most closely related to the ingroup sequences could, nevertheless, be more distant from the ingroup than an outgroup sequence that shares an older common ancestor with the ingroup). However, as Table 3 shows, an outgroup taxon that appears to be close to the ingroup in one data set may be the most distant outgroup taxon in another. For instance, *Melissodes* is the closest outgroup in the cyt b data set but the most distant outgroup in the opsin data set. Therefore, we performed additional parsimony analyses on each data set, using different combinations of the original five outgroup taxa, excluding the outgroup sequences displaying the highest p-distances in Table 3. We excluded *Melissodes* from the opsin matrix, *Anthophora* and *Habropoda* from cyt b, *Habropoda* from 28S, and *Centris* and *Melissodes* from 16S. The resulting trees (not shown) were all fully compatible with the tribal relationships estimated from both rooted (Fig. 3) and unrooted (Fig. 2) trees. A portion of these trees still shows a paraphyletic ingroup with respect to one or more outgroup taxa. The opsin and cyt b trees continue to be rooted along the branch leading to Apini, and the 28S and 16S trees are rooted along the branch separating (Bombini + Meliponini) and (Apini + Euglossini).

As already discussed, the monophyly of the corbiculate apines is well supported by morphological evidence (Roig-Alsina and Michener, 1993). Furthermore, the pattern of paraphyly is inconsistent within and between genes and depends on the outgroup combination. Paraphyly of the ingroup disappears when the DNA data sets are combined in a global parsimony analysis (see below). Moreover, when the GTR ML model is applied to the AT-rich mtDNA sequences, ingroup paraphyly disappears in the case of cyt b (Table 4). Applying the HKY85 ML model to all gene sequences had no effect on ingroup paraphyly (Table 4). Thus, either the corbiculate bees are indeed monophyletic, or we cannot consistently determine how they are not.

TABLE 4. Results of analyses assessing the monophyly of the ingroup.

Gene	Analysis	Model[a]	Ingroup monophyletic or paraphyletic
mtDNA			
16S	Parsimony	UW	Ingroup monophyletic
Cyt b	Parsimony	UW	Ingroup paraphyletic
Cyt b	ML	HKY85	Does not remove paraphyly
Cyt b	ML	GTR	Removes paraphyly
Nuclear			
Opsin	Parsimony	UW	Ingroup paraphyletic
Opsin	ML	HKY85	Does not remove paraphyly
28S	Parsimony	UW	Ingroup paraphyletic
28S	ML	HKY85	Does not remove paraphyly

[a]UW, unweighted parsimony; HKY, Hasegawa, Kishino, and Yano (1985) model; GTR, General Time Reversible model (Yang, 1994).

Parsimony analysis of the reduced 28S data set (variable-length region excluded, 38 characters) produced essentially the same tree topologies as those derived from the intact data set. However, analyses of the reduced 16S data set (196 characters excluded) resulted in poorly resolved topologies, suggesting that most of the phylogenetic signal occurs in the variable-length regions of those data. This finding is consistent with prior analyses of 16S sequences of the corbiculate bees (Cameron, 1993).

Parsimony analyses for each data set, executed while constraining the ingroup to be monophyletic, produced trees rooted in the same position as those of the previous unconstrained analyses: on the Apini branch (opsin and cyt b) or on the internal branch dividing (Bombini + Meliponini) and (Apini + Euglossini) (16S and 28S). Similarly, constrained ML analyses (HKY85 for opsin and 28S, GTR for 16S and cyt b; all parameters estimated with ML, invariant sites estimated with outgroups excluded) recovered the same relationships for the respective gene

TABLE 5. Percentage of variable sites estimated with SplitsTree and ML.

	Outgroups included		Outgroups excluded	
	SplitsTree	ML	SplitsTree	ML
Opsin	52	—	57	50
Cyt b	64	—	61	60
28S	75	—	80	100
16S	58	—	52	50

as ML (also the same as MP) without constraining the ingroup to be monophyletic. These two different rootings were also obtained when the outgroups were constrained to the pattern of relationships reported by Roig-Alsina and Michener (1993).

The proportions of variable sites estimated by using SplitsTree and ML, with and without outgroups, are shown in Table 5. Spectral analyses on the LogDet distance matrices, performed several times for each gene (once with each outgroup taxon and a different proportion of invariant sites removed), all gave the greatest support for the same root position, that is, the root lying along the branch leading to *Apis* for the cyt b and opsin data sets, and along the branch separating the tree into (Bombini + Meliponini) and (Apini + Euglossini) for the 16S and 28S data sets. Moreover, as more sites were removed (from 0% to 40%, 50%, and 60%), the support for the rooting split increased. This was the outcome for all four genes. Therefore, correcting for differences in base composition and positional rate heterogeneity among taxa did not resolve the rooting conflict observed between the protein-coding genes and the ribosomal genes.

We used parametric bootstrapping according to Huelsenbeck (1997) to investigate whether the outgroup rooting inferred from opsin and cytb data resulted from attraction between the long branch leading to *Apis* and another long branch leading to one or more of the outgroup taxa. Starting with the most-parsimonious tree (for each data set), a new tree was obtained by reassigning the position of the root on the branch separating (Bombini + Meliponini) and (Apini + Euglossini) (i.e., the rooting supported by the 28S and 16S data sets) and by making the ingroup monophyletic. For each of these new trees, 100 data sets were simulated with Sequence-Generator (v1.1; Rambaut and Grassly, 1997), using an HKY85 model of evolution (Hasegawa et al., 1985) incorporating site-specific rate heterogeneity. Parameters of this model (branch lengths, TI/TV ratio, proportion of invariable sites, shape parameter of the gamma distribution) were estimated from the original data set for each new tree by using the likelihood criterion. A parsimony analysis was then performed on each of the 100 simulated data sets. In both cases, a majority of the resulting trees recovered the rooting of the model tree that was used to generate the simulated data sets, suggesting that the opsin and cyt b rooting is unlikely to be the result of a long-branch attraction phenomenon. This reconfirms the results of the spectral analysis method (see previous paragraph) of Steel et al. (2000), which also tested for biases attributable to rate heterogeneity among lineages.

To investigate whether the outgroup rooting inferred with 16S and 28S data might have been an artifact of our alignment procedures, we realigned the data sets by using CLUSTAL X. For this alignment we defined a priori the guide tree as ((((Bombini + Meliponini) + Euglossini) + Apini) outgroups), which corresponds to the rooted tree supported by the two protein-coding genes. Parsimony analyses conducted on these realigned data sets resulted in the same tribal relationships and position of the root as the original trees for those data. The inferred topologies are therefore stable with respect to the alignment procedure used.

Analysis of Conflict Between Data Sets

The *P*-values from the ILD tests are shown in Table 6. None of the paired comparisons of the DNA data sets are significantly incongruent at the significance level ($P < 0.01$) proposed by Cunningham (1997). When

TABLE 6. *P*-values resulting from the ILD tests. Upper matrix values were obtained analyzing ingroup taxa alone, lower matrix values were obtained by including outgroup taxa in the analyses. Two additional tests, performed to compare all DNA data against the morphology data and protein coding data (opsin + cyt b) against ribosomal DNA (16S + 28S) returned *P*-values of 0.001 and 0.044, respectively.

	Opsin	28S	16S	Cyt b
Opsin	—	0.716	0.402	0.860
28S	0.064	—	0.415	0.597
16S	0.200	0.176	—	0.050
Cytb	0.992	0.011	0.088	—

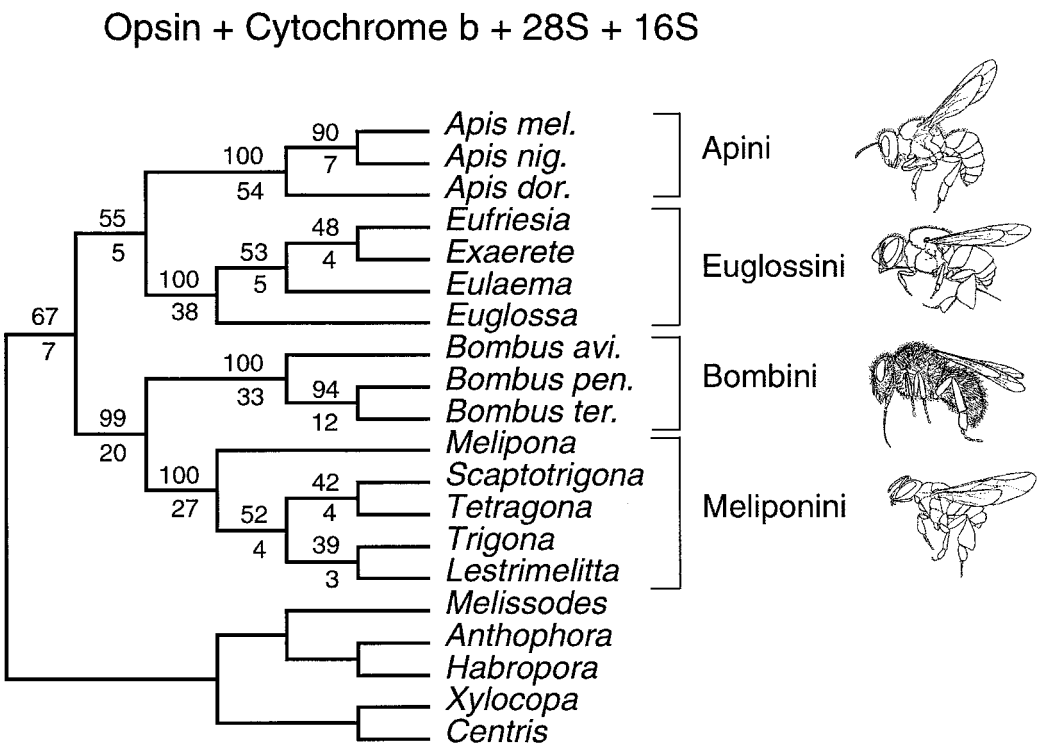

FIGURE 4. Phylogeny of the four tribes of corbiculate bees estimated from parsimony analysis of sequences of all four genes (opsin, cyt b, 28S, and 16S) combined: 1 MP-tree, tree length = 3036, CI = 0.54, RI = 0.52). Numbers above branches are bootstrap values; those below are decay indices.

outgroup taxa are excluded from the DNA data sets, the congruence among them increases, probably the result of eliminating the rooting conflict between the opsin/cyt b and the 16S/28S data sets. Equally weighted parsimony analysis of the total DNA data set (all five outgroup taxa included) produced the tree shown in Figure 4. The (Meliponini + Bombini) clade is strongly supported (bootstrap value 99%), just as when each gene was analyzed separately. However, the (Apini + Euglossini) clade, initially supported by only the 16S and 28S data sets, is supported by a low bootstrap value (55%). Note that the ingroup appears to be monophyletic for the total DNA tree.

The morphology data set is strongly incongruent with the molecular data set ($P = 0.001$). As a result, combining the DNA and morphological data into a single matrix for a global parsimony analysis could be misleading or uninformative (Cunningham, 1997). Nonetheless, we have performed such an analysis, and the tree obtained is shown in Figure 5. The (Bombini + Meliponini) clade is retained, but with less bootstrap support (73%). The large decrease in support of the (Bombini + Meliponini) clade is probably the result of the incompatibility of this clade with the topology supported by the

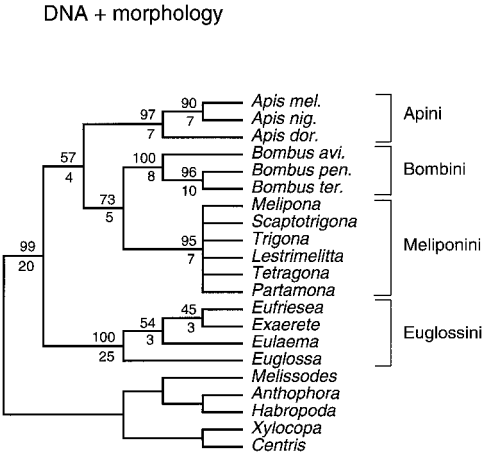

FIGURE 5. Phylogeny of the corbiculate tribes estimated from parsimony analysis of all DNA data sets plus morphology (strict consensus of 9 MP trees, tree length = 3253, CI = 0.54, RI = 0.52).

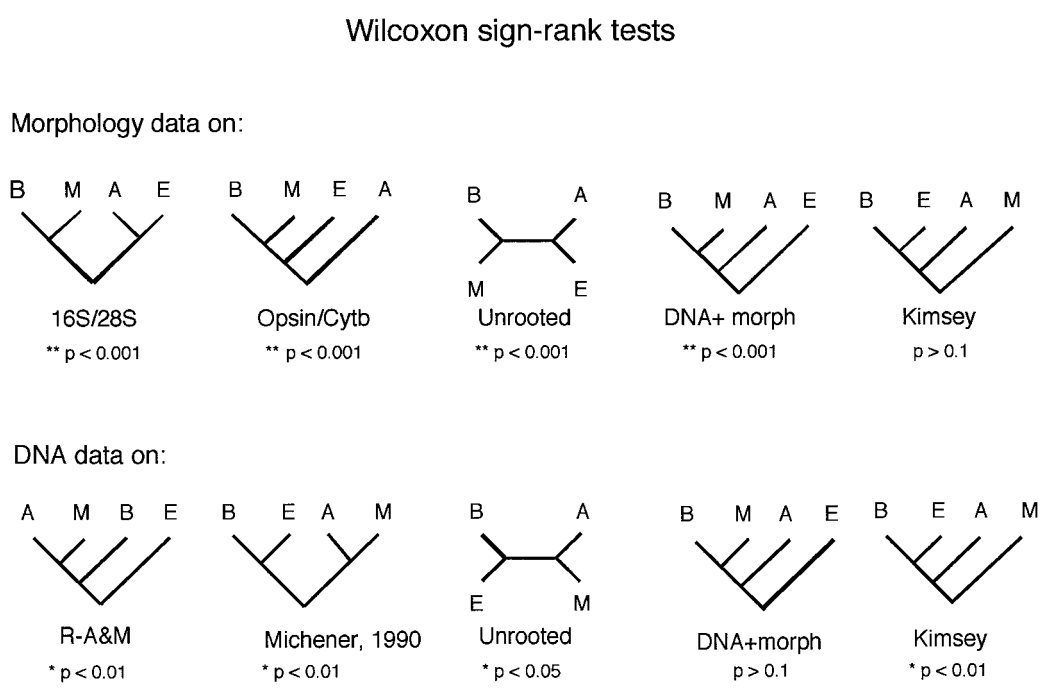

FIGURE 6. Wilcoxon sign-rank tests applied to alternative topologies relative to the optimal (MPtree) topology for the morphological and genetic data partitions. P values are listed below each alternative topology. A, Apini; M, Meliponini; B, Bombini; E, Euglossini. R-A&M, Roig-Alsina and Michener (1993).

morphological data. A clade appears (Apini + Bombini + Meliponini; bootstrap value 57%) that is not supported by any of the DNA or morphology data partitions alone and therefore may be misleading.

Results of two-tailed Wilcoxon tests are shown in Figure 6. All statistical examinations of morphology data sets optimized onto alternative molecular topologies (and vice versa) resulted in rejection of the null hypothesis that the character changes are equally parsimonious on both the optimal and the alternative trees. When the morphological character changes were optimized onto the two differently rooted trees estimated from DNA data (one tree for 16S/28S, a second for cyt b/opsin), the null hypothesis was rejected at $P = 0.001$. Reciprocal analyses mapping the DNA data onto two optimal morphological topologies (Roig-Alsina and Michener's, and another obtained by Michener in 1990) resulted in rejection of the null hypothesis at $P = 0.01$. The same results were obtained for unrooted analyses (outgroups excluded), except that when morphological data are mapped onto the unrooted DNA topology the result was much less significant ($P < 0.05$) than when the DNA data were mapped onto the unrooted morphology tree ($P < 0.001$). The morphology data do, however, fit trees from prior morphology-based analyses that imply a nonsister group relationship between Apini and Meliponini, namely, the tree topologies from Winston and Michener (1977) and from Kimsey (1984). The differences in topology between the morphology and DNA data sets cannot be attributed to random error; the data sets strongly conflict. Hence the topology derived from morphology is not close to optimal for the molecular data, and vice versa.

DISCUSSION

These analyses clearly show that the molecular data conflict strongly with the morphological data. These two major partitions result in significantly different tree topologies for the corbiculate bees, and both the ILD tests and the Wilcoxon tests indicate that the two partitions are indeed conflicting in character support. Under these

circumstances we confront three options: (1) Accept the morphological topology, (2) accept one or both of the rooted molecular topologies, or (3) accept the overall combined-partitions topology.

We are not inclined to accept the results from the morphological data set only, which was originally assembled for a higher-level investigation of all long-tongued bees (Apidae and Megachilidae) and therefore suffers from insufficient ingroup taxon sampling for detailed analysis of the corbiculate tribes. Furthermore, the inconsistency among tree topologies produced from the many earlier morphological studies is unsettling; it could be attributed to large amounts of homoplasy among the characters or, in some cases, to different outgroups influencing the rooting. To explore the behavior of the morphological characters further, we used the method of Ronquist (1994), which examines the effects of potentially convergent characters on phylogenetic relationships by excluding them from analysis. We hypothesized that the Apini + Meliponini relationship might have resulted from convergent similarities associated with the highly eusocial mode of life, thus obscuring a Bombini + Meliponini relationship. To test this hypothesis, we eliminated from the data matrix all putative apomorphies shared by Apini + Meliponini and not occurring in either Bombini or Euglossini. This resulted in a rerooting of the tree along the Meliponini branch (Fig. 7), such that Apini + Meliponini no longer formed a clade; however, a Bombini + Meliponini clade did not appear instead. From this we conclude that the morphological synapomorphies supporting Apini + Meliponini were not, in fact, obscuring an underlying Bombini + Meliponini relationship. Nonetheless, the problems of insufficient taxonomic sampling within the corbiculate tribes and discordance among many of the former trees cast doubt on the validity of the morphology-based phylogenetic conclusions.

The result of the overall combined-partitions analysis of the molecular and morphological data is of concern because of the strong character conflict between these two partitions. In our view, the global tree is a compromise between the two respective topologies from the separate analyses. Although this rejects the Apini + Meliponini hypothesis, supported by morphology alone, and favors the Bombini + Meliponini hypothesis from molecular data, it also suggests a weakly supported clade (Apini + (Bombini + Meliponini)) that is not recovered in any of the analyses of individual data sets and thus is likely to be misleading.

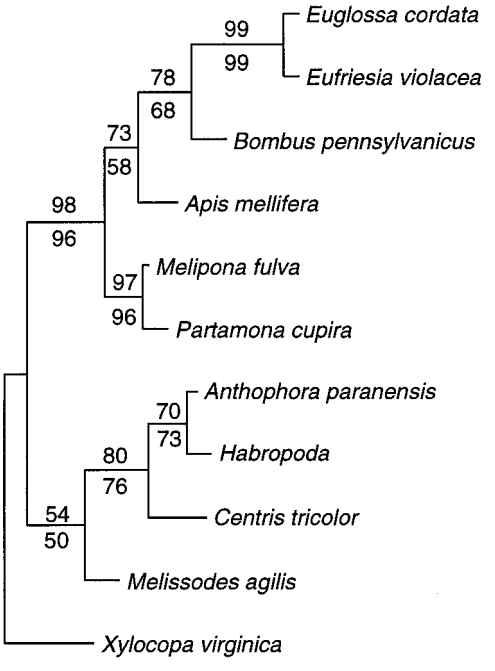

FIGURE 7. MP tree of the corbiculate bees estimated from morphological characters, but excluding all synapomorphies supporting (Apini + Meliponini). Exhaustive search, all characters unordered. Numbers above branches are bootstrap values based on both adult and larval characters; numbers below are based on adult characters only. Tree length = 170, CI = 0.57, RI = 0.64.

By default, we are left to consider option (2): Accept one or both of the rooted molecular topologies. If we take into account only the results of the ILD tests, which indicated no significant incongruence among the molecular data sets but highly significant incongruence between molecular and morphological data, one could argue that the most rigorously supported estimate of corbiculate bee relationships should use all of the molecular data (Fig. 4). This amounts to acceptance of the conditional data combination approach of Bull et al. (1993), in which each partition is assessed against the others for evidence of character congruence and hence combinability. Because incongruent data sets fail the test of combinability, it follows from Bull et al. that the morphology data should

be excluded from analysis of the combined molecular data. We support this conclusion in part. The remainder of the discussion will focus on the rationale underlying our conclusions.

When confronted with strong conflict among data sets, statistical tests such as the ILD test do not resolve the conflict. They do not provide criteria for selecting between one or another of the conflicting partitions. Wilcoxon tests provide some insights regarding the behavior of data on alternative (suboptimal) topologies, potentially offering insights into the actual location (branch or node) of character conflict. A clear pattern emerges from optimization of the morphology data onto the various topologies shown in Figure 6; namely, the morphology data do not fit a tree that includes a Bombini + Meliponini clade, but will fit a tree that does *not* include an Apini + Meliponini clade (i.e., the Kimsey [1984] and Winston and Michener [1977] hypotheses of dual origin of Apini and Meliponini). On the other hand, the molecular data, when optimized onto the various rooted and unrooted suboptimal topologies in Figure 6; will simply not fit a tree that does not contain a Bombini + Meliponini clade. This is not particularly surprising, given the clear and consistent results from the unrooted analyses of each of the four genes, indicating that regardless of where the tree is rooted, there is absolutely no way to obtain a tree that contains an Apini + Meliponini clade. The rooted analyses are fully consistent with this result.

Although unrooted analyses of genes resulted in identical tribal topologies, rooting with outgroups resulted in two topologies. One is supported by the rRNA genes (16S and 28S; the ribosomal tree), the other by the protein-coding genes (cyt b and opsin; the protein tree). The ribosomal tree is rooted on the branch that divides (Apini + Euglossini) and (Bombini + Meliponini), whereas the protein tree is rooted on the branch leading to the Apini. Combining the data does not help resolve this conflict. Rather, it provides a compromise between the two conflicting topologies. Possible sources of bias not yet considered in the data include (1) secondary structure, known to impose selective constraints on nucleotide substitutions (Buckley et al., 2000), possibly leading to constraints on substitutions in the 16S and 28s rRNA molecules, resulting in a lack of independence among different informative sites (but see Dixon and Hillis, 1993); and (2) sorting of ancestral polymorphism among lineages, leading to a lack of congruence between gene trees and species trees (Pamilo and Nei, 1988; Takahata, 1989). The first potential source of error was examined in the 16S gene by fitting a Bombini sequence to the 16S secondary-structure model of Gutell (1993) and comparing the fit with that of other Hymenoptera (see Whitfield and Cameron, 1998). Analyses using sequences aligned to secondary structure were not different from those already presented. The second source of error, differential lineage sorting, cannot explain the incongruence between the 16S tree and the cyt b tree because both genes are linked in the nonrecombining haploid mitochondrial genome and should therefore record the same organismal histories. Given the lack of strong evidence at this time that would allow us to choose between the ribosomal and protein trees, in the remainder of the discussion we consider both topologies as alternative, equally supported phylogenetic hypotheses.

Implications for Social Evolution

Many of the earlier investigations of social evolution in the corbiculate bees operated on the more or less unstated (certainly untested) assumption that the evolution of eusociality followed a ladderlike progression, from solitary behavior to the "pinnacle" of the most complex societies. This assumption has probably influenced the strongly held view that the highly eusocial tribes share a common ancestor that was itself highly eusocial. Indeed, honey bees and stingless bees share many similarities in their complex social organization, including large, perennial colonies that accommodate a morphologically and behaviorally distinct queen modified for egg-laying and female offspring (workers) that exhibit a high degree of task specialization and complex communication (Michener, 1974; Roubik, 1989). Nonetheless, the mechanics of their respective social systems, including colony founding, nest architecture, and recruitment to resources, are strikingly different (Sakagami, 1971; Winston and Michener, 1977). For instance, honey bees recruit nestmates to food sources and nest sites by way of a symbolic dance-language and food odors, whereas stingless bees use a system of trail pheromones to

guide recruits directly to the resource. When honey bees initiate a new colony, the old queen leaves her nest accompanied by a swarm of workers who search for an appropriate new nest site. Stingless bees instead send forth a young queen from the old nest. She will take up residence in a new nest that has been diligently constructed over several weeks by workers from the old colony. Whether these differences in highly eusocial behavior are the result of numerous modifications from a common highly eusocial ancestor or instead reflect two independent origins of elaborate social organization can best be assessed with a robust phylogeny.

We have argued that the most robust hypothesis of relationships for the corbiculate tribes is based on the summation of molecular data (Figs. 2–4). These data suggest that either highly eusocial behavior evolved twice independently in Apini and Meliponini or their hypothetical common ancestor was highly eusocial, with subsequent reversals in Bombini and Euglossini. Unfortunately, the phylogeny alone cannot help to distinguish which of the two evolutionary scenarios is most-parsimonious. Either alternative requires three evolutionary changes on the molecular tree. Support for one or the other of these two contrasting interpretations will ultimately come from a greater knowledge of behavior, ecology, and physiology of many more species of these bees. A greater understanding of potential behavioral homology through detailed studies of specific components of social behavior (e.g., design of nest architecture, colony size, perenniality, division of labor and polyethism, caste determination) and their ecological context will be required to advance beyond current (untested) assumptions. For example, the assumption that advanced eusociality is (or is not) an irreversible stage in social evolution requires rigorous investigation of the plasticity of sociality in the corbiculate bees, particularly in the Bombini and Meliponini.

Interestingly, there are no cases of a loss or reversal of the highly eusocial state to a solitary or even primitively eusocial condition in either of the two highly eusocial apine tribes. Even within the large and diverse tribe of stingless bees, all known species are classified as highly eusocial. Likewise, there are no known examples of loss of sociality in the primitively eusocial Bombini. In contrast, the more plastically eusocial sweat bees have had instances of reversal to the solitary state (Michener, 1974; Wcislo and Danforth, 1997). Social parasitism, with its linked loss of the worker caste, is also entirely absent in the two highly eusocial apine tribes, although this phenomenon occurs in the primitively eusocial Bombini and other groups of social bees, such as the sweat bees. Once the obligate highly eusocial condition has evolved, therefore, it appears unlikely to undergo a reversal. In fact, this argument has been made in the past by Winston and Michener (1977). One can easily conceive that the evolution of a strict, obligate division of labor between queen and worker castes (through loss of individual totipotency), such as that singly characterizing Apini and Meliponini, proceeds in one direction without the possibility of reversal, and it is hard to imagine that the morphological and behavioral interdependency existing at their level of social organization could undergo reversal. For instance, the queen would need to develop the anatomical and behavioral mechanisms for an independent existence, including sufficient wing size and musculature for flight, corbiculae for collecting pollen, nest initiation capability, and brood care. It would no longer be adaptive for workers to retain strong age polyethism (successive development of behavioral classes based on age) or the hormonal mediation of that modus operandi. This is not to say that the conditions for such evolutionary changes could never occur, but that they are improbable and there is no evidence that they have in the past. To the contrary, as already mentioned, there are no extant or fossil representatives of Apini or Meliponini that are primitively eusocial or solitary. Definitive answers to questions of reversal are ultimately obtainable only through new conceptual advances and investigations into the evolutionary genetics and ecology of highly eusocial traits.

If indeed the highly eusocial behavior of Apini and Meliponini evolved twice independently—the hypothesis we support as being more likely—then the comparative behavior of these bees takes on added interest from an evolutionary perspective. The apparent convergent similarities in behavior between these two tribes suggest that highly eusocial organization has limited permutations. Each tribe has been channeled along a similar behavioral track, responding in similar fashion to similar contingencies.

For instance, with the evolution of large nests organized around a highly modified queen unable to survive alone, colony reproduction by swarming is most efficient, and large nests require closely (efficiently) packed comb and perenniality, to allow for the creation of large colonies. Our recognition of patterns of this sort is only partially complete, with serious shortcomings in our factual information for many species. There is also still much to be gained from further study of morphology of the corbiculate bees. As we suggested above, the evolution of social behavior can be fully understood only through knowledge of the evolutionary history of species. In turn, knowledge of the environmental context of that history, together with a more complete picture of the elements of comparative social behavior, would promise new evolutionary insights.

ACKNOWLEDGMENTS

We thank Stella Koulianos, Gard Otis, David Roubik, and Robbin Thorp for assistance with taxon collection; Rob Brooks for identification of *Eufriesea caerulescens;* Roy Snelling for identification of *Centris inermis;* and Jim Whitfield for line drawings of bees. We especially thank Jim Whitfield for many helpful discussions and careful reading of the manuscript and Peter Lockhart for suggesting some of the analyses performed in this study. This research was supported by NSF (GER-9450117) and USDA (NRICGP-9501893) grants to S.A.C. and J. Whitfield, respectively.

REFERENCES

BELSHAW, R., AND D. L. J. QUICKE. 1997. A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae). Mol. Phylogenet. Evol. 7:281–293.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

BUCKLEY, T. R., C. M. SIMON, P. K. FLOOK, AND B. MISOF. 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. Insect Mol. Biol. 9:565–580.

BULL, J. J., J. P. HUELSENBECK, C. W. CUNNINGHAM, D. L. SWOFFORD, AND P. J. WADDELL. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42:384–397.

CAMERON, S. A. 1993. Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. Proc. Natl. Acad. Sci. USA 90:8687–8691.

CAMERON, S. A., J. N. DERR, A. D. AUSTIN, J. B. WOOLEY, AND R. A. WHARTON. 1992. The application of nucleotide sequence data to phylogeny of the Hymenoptera: A review. J. Hymenop. Res. 1:63–79.

CHANG, B. S. W., D. AYERS, W. C. SMITH, AND N. E. PIERCE. 1996. Cloning of the gene encoding honeybee long-wavelength rhodopsin: A new class of insect visual pigments. Gene 173:215–219.

CHARLESTON, M. A., AND R. D. M. PAGE. 1997. Spectrum 2.0. Distributed by the authors (http://taxonomy.zoology.gla.ac.uk/~mac/spectrum/spectrum.html), Univ. of Glasgow, U.K.

CHAVARRÍA, G., AND J. M. CARPENTER. 1994. "Total evidence" and the evolution of highly social bees. Cladistics 10:229–258.

CUNNINGHAM, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14:733–740.

DESALLE, R., AND A. V. Z. BROWER. 1997. Process partitions, congruence, and the independence of characters: Inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. Syst. Biol. 46:751–764.

DIXON, M. T., AND HILLIS, D. M. 1993. Ribosomal RNA secondary structure: Compensatory mutations and implications for phylogenetic analysis. Mol. Biol. Evol. 10:256–267.

DOWTON, M., AND A. D. AUSTIN. 1994. Molecular phylogeny of the insect order Hymenoptera: Apocritan relationships. Proc. Natl. Acad. Sci. USA 91:9911–9915.

EICKWORT, G. C., J. M. EICKWORT, J. GORDON, AND M. A. EICKWORT. 1996. Solitary behavior in a high-altitude population of the social sweat beei *Halictus rubicundu* (Hymenoptera: Halictidae). Behav. Ecol. Sociobiol. 38:227–233.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1994. Testing significance of incongruence. Cladistics 10:315–319.

FELSENSTEIN, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

GARÓFALO, C. A., E. CAMILLO, S. C. AUGUSTO, B. M. V. DE JESUS, AND J. C. SERRANO. 1998. Nest structure and communal nesting in *Euglossa* (*Glossura*) *annectans* Dressler (Hymenoptera, Apidae, Euglossini). Rev. Bras. Zool. 15:589–596.

GILBERT, D. G. 1996. SeqPup version 0.6. Published electronically on the Internet, available via anonymous ftp to ftp.bio.indiana.edu.

GRAYBEAL, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47:9–17.

GUTELL, R. R., 1993. Comparative studies of RNA: Inferring higher-order structure from patterns of sequence variation. Curr. Opin. Struct. Biol. 3:313–322.

HANCOCK, J. M., D. TAUTZ, AND G. A. DOVER. 1988. Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster.* Mol. Biol. Evol. 5:393–414.

HASEGAWA, M., AND T. HASHIMOTO. 1993. Ribosomal RNA trees misleading? Nature 361:23.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

HENDY, M. D., AND PENNY, D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38:297–309.

HILLIS, D. M. 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. 44:3–16.

HILLIS, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3–8.

HUELSENBECK, J. P. 1997. Is the Felsenstein zone a fly trap? Syst. Biol. 46:69–74.

HUELSENBECK, J. P., J. J. BULL, AND C. W. CUNNINGHAM. 1996. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11:152–158.

HUSON, D. H. 1998. SplitsTree: A program for analyzing and visualizing evolutionary data. Bioinformatics 14:68–73.

KIMSEY, L. 1984. A re-evaluation of the phylogenetic relationships in the Apidae (Hymenoptera). Syst. Entomol. 9:435–441.

KLUGE, A. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38:7–25.

KOULIANOS, S., R. SCHMID-HEMPEL, D. W. ROUBIK, AND P. SCHMID-HEMPEL. 1999. Phylogenetic relationships within the corbiculate Apinae (Hymenoptera) and the evolution of eusociality. J. Evol. Biol. 12:380–384.

LARSON, A. 1994. The comparison of morphological and molecular data in phylogenetic systematics. Pages 371–390 *in* Molecular approaches to ecology and evolution (B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds.). Birkhauser, Basel.

LOCKHART, P. J., AND S. A. CAMERON. 2001. Trees for bees. Trends Ecol. Evol. 16:84–88.

LOCKHART, P. J., A. W. D. LARKUM, M. A. STEEL, P. J. WADDELL, AND D. PENNY. 1996. Evolution of chlorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis. Proc. Natl. Acad. Sci. USA 93:1930–1934.

LOCKHART, P. J., M. A. STEEL, A. C. BARBROOK, D. H. HUSON, M. A. CHARLESTON, AND C. J. HOWE. 1998. A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. Mol. Biol. Evol. 15:1183–1188.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, AND D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 12:503–513.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: Analysis of phylogeny and character evolution. Sinauer, Sunderland, Massachusetts.

MARDULYN, P., AND S. A. CAMERON. 1999. The major opsin in bees (Insecta: Hymenoptera): A promising nuclear gene for higher level phylogenetics. Mol. Phylogenet. Evol. 12:168–176.

MARDULYN, P., AND J. B. WHITFIELD. 1999. Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of microgastrinae (Hymenoptera: Braconidae): Evidence of a high diversification rate in this group of parasitoids. Mol. Phylogenet. Evol. 12:282–294.

MCGINLEY, R. J. 1981. Systematics of the Colletidae based on mature larvae with phenetic analysis of apoid larvae. Univ. Calif. Publ. Entomol. 91:1–307.

MICHENER, C. D. 1944. Comparative external morphology, phylogeny, and a classification of the bees. Bull. Am. Mus. Nat. Hist. 82:151–326.

MICHENER, C. D. 1974. The social behavior of the bees: a comparative study. Harvard Univ. Press, Cambridge, Massachusetts.

MICHENER, C. D. 1990. Classification of the Apidae (Hymenoptera). Univ. Kansas Sci. Bull. 54:75–164.

MICHENER, C. D., AND R. W. BROOKS. 1984. Comparative study of the glossae of bees. Contrib. Am. Entomol. Inst. 22:1–73.

NIXON, K. C., AND J. M. CARPENTER. 1996. On simultaneous analysis. Cladistics 12:221–241.

PACKER, L. 1990. Solitary and eusocial nests in a population of *Augochlorella striata* (Provancher) (Hymenoptera: Halictidae) at the northern edge of its range. Behav. Ecol. Sociobiol. 27:339–344.

PAMILO, P., AND M. NEI. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5:568–583.

PENNY, D. M., D. HENDY, P. J. LOCKHART, AND M. A. STEEL. 1996. Corrected parsimony, minimum evolution, and Hadamard conjugations. Syst. Biol. 45:596–606.

PLANT, J. D., AND H. F. PAULUS. 1987. Comparative morphology of the postmentum of bees (Hymenoptera: Apoidea) with special remarks on the evolution of the lorum. Z. Zool. Syst. Evol.-Forsch. 25:81–103.

POE, S. 1998. Sensitivity of phylogeny estimation to taxonomic sampling. Syst. Biol. 47:18–31.

PRENTICE, M. 1991. Morphological analysis of the tribes of Apidae. Pages 51–69 *in* Diversity in the genus *Apis* (D. R. Smith, ed.). Westview Press, Boulder, Colorado.

RAMBAUT, A., AND N. C. GRASSLY. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13:235–238.

ROIG-ALSINA, A., AND C. D. MICHENER. 1993. Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). Univ. Kansas Sci. Bull. 55:124–162.

RONQUIST, F. 1994. Evolution of parasitism among closely related species: Phylogenetic relationships and the origin of inquilinism in gall wasps (Hymenoptera, Cynipidae). Evolution 48:241–266.

ROUBIK, D. W. 1989. Ecology and natural history of tropical bees. Cambridge Univ. Press, Cambridge, UK.

SAKAGAMI, S. F. 1971. Ethosoziologischer Vergleich zwischen Honigbienen und stachellosen Bienen. Z. Tierpsychol. 28:337–350.

SANDERSON, M. J., M. J. DONOGHUE, W. PIEL, AND T. ERIKSSON. 1994. TreeBASE: A prototype database of phylogenetic analyses and an interactive tool for browsing the phylogeny of life. Am. J. Bot. 81: 183.

SCHÖNITZER, K. 1986. Comparative morphology of the antenna cleaner in bees. Z. Zool. Syst. Evol.-Forsch. 24:35–51.

SCHULTZ, T. R., M. S. ENGEL, AND M. PRENTICE. 1999. Resolving conflict between morphological and molecular evidence for the origin of eusociality in the corbiculate bees: A hypothesis-testing approach. Pp. 125-138 *in* Entomological contributions in memory of Byron A. Alexander (G. W. Byers, R. H. Hagen, and R. W. Brooks, eds.). Univ. Kans. Nat. Hist. Mus. Spec. Publ. 24, 251 pp.

SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU, AND P. FLOOK. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87:651–701.

SORENSON, M. D. 1996. TreeRot computer software. Museum of Zoology, Univ. of Michigan, Ann Arbor.

STEEL, M., D. H. HUSON, AND P. J. LOCKHART. 2000. Invariable sites models and their use in phylogeny reconstruction. Syst. Biol. 49:225–232.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogeny reconstruction. Pages 407–514 *in* Molecular systematics, 2nd ed. (D. M. Hillis, C. Moritz and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

SWOFFORD, D. L. 1998. *PAUP** 4.0. Phylogenetic analysis using parsimony. Beta version 4.0b1 for Power PC, Sinauer, Sunderland, Massachusetts.

TAKAHATA, N. 1989. Gene genealogy in three related populations: Consistency probability between gene and population trees. Genetics 122:957–966.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876–4882.

WARDLAW, A. C. 1985. Practical statistics for experimental biologists. Wiley, Chichester, U.K.

WCISLO, W. T., AND B. DANFORTH. 1997. Secondarily solitary: The evolutionary loss of social behavior. Trends Ecol. Evol. 12:468–474.

WHEELER, W. C., AND D. GLADSTEIN. 1995. Malign 2.5. American Museum of Natural History, New York.

WHITFIELD, J. B., AND S. A. CAMERON. 1998. Hierarchical analysis of variation in the mitochondrial 16S rRNA gene among Hymenoptera. Mol. Biol. Evol. 15:1728–1743.

WINSTON, M. L., AND C. D. MICHENER. 1977. Dual origin of highly social behavior among bees. Proc. Natl. Acad. Sci USA 74:1135–1137.

YANG, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol. Biol. Evol. 10:1396–1401.

YANG, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:105–111.

Received 5 November 1999; accepted 13 April 1999
Associate Editor: A. Brower

APPENDIX 1. MORPHOLOGICAL CHARACTERS AND CHARACTER STATES

All characters are treated as unordered. Abbreviations in parentheses refer to the character number in Roig-Alsina and Michener (RAM), 1993.

Head

1. (RAM2). Anterior tentorial pit; 0: high on epistomal suture; 1: at or below middle of lateral part of epistomal suture.

2. (RAM3) Integument of paraocular area; 0: not differentiated; 1: punctures sparser and smaller.

3. (RAM4) Paraocular carina; 0: absent; 1: present.

4. (RAM5) Condyle of anterior mandibular articulation; 0: contiguous with lateral clypeal margin; 1: partly covered by lateral clypeal margin.

5. (RAM6) Lateral part of lower portion of clypeus; 0: not bent backward; 1: strongly and abruptly bent.

6. (RAM10) Tuft on apical margin of labrum; 0: absent; 1: present.

7. (RAM13) Dorsal sheet of anterior tentorial arm; 0: without spur reaching eye; 1: with spur reaching eye.

8. (RAM14) Union of anterior tentorial arm to head wall below antennal socket; 0: reaching lower margin; 1: forming triangular space.

9. (RAM16) Lateral expansion of internal thickening above epistomal ridge; 0: at least as wide as half-width of socket diameter; 1: reduced, less than half-width of socket diameter.

10. (RAM17) Clypeus with apical inflection; 0: present; 1: reduced to narrow band.

11. (RAM21) Postoccipital pouch below foramen magnum; 0: absent; 1: shallow; 2: distinct and deep.

12. (RAM22) Fan-shaped posterior sheets of tentorium; 0: well developed; 1: small to absent.

13. (RAM23) Attachment of secondary tentorial bridge to posterior wall of head; 0: above and separate from hypostoma; 1: as in 0, but vertical line wider, representing wider septum; 2: secondary bridge fused directly to hypostoma.

14. (RAM24) Epistomal suture below anterior tentorial pits; 0: nearly straight, or gently curved; 1: extending straight down, then angulate laterad.

Mouthparts

15. (RAM29) Maxillary stipes with comb in concavity on distal posterior margin; 0: absent; 1: present.

16. (RAM30) Maxillary stipes with ridge on outer surface; 0: absent; 1: present.

17. (RAM36) Maxillary galeal blade; 0: uniformly sclerotized except sometimes extreme apex; 1: posterior margin broadly desclerotized almost to base.

18. (RAM39) Maxillary lacinia; 0: rounded; 1: elongate.

19. (RAM40) Stipital sclerite of maxilla; 0: distinct; 1: fused to stipes.

20. (RAM41) Galeal blade with internal sclerotized surface; 0: as wide as external surface; 1: at most two-thirds as wide as external surface; 2: at least three-fourths as wide as external surface but still narrower.

21. (RAM43) Lorum and mentum; 0: united; 1: separated from one another.

22. (RAM44) Base of lorum; 0: simple; 1: with longitudinal fissure on each side.

23. (RAM46) Subligular process of prementum; 0: fully sclerotized and united to rest of prementum; 1: separated from prementum; 2: weakly sclerotized.

24. (RAM50) First segment of labial palpus; 0: without membranous margin; 1: with membranous inner margin.

25. (RAM52) Glossal rod; 0: absent; 1: present, but not enclosing bacular canal; 2: present, surrounding bacular canal.

26. (RAM54) Posterior surface of flabellum; 0: smooth or nearly so; 1: with a cobblestone pattern.

27. (RAM55) Annular hairs of glossa; 0: extending to base of flabellum; 1: separated from flabellum by nonannulate shank.

28. (RAM60) Mandible of female; 0: slender; 1: pollex expanded to form two to several teeth or an edentate margin above rutellum.

29. (RAM61) Mandibular grooves and ridges on outer surface; 0: distinct; 1: largely absent.

Mesosoma

30. (RAM62) Pronotum with ventrolateral extensions; 0: fused midventrally; 1: separated midventrally.

31. (RAM63) Lateral carina separating exposed part of propleuron; 0: present; 1: absent.

32. (RAM64) Apophyseal arms of prosternum; 0: fused along median crest; 1: separate from one another.

33. (RAM65) Apophyseal pit of prosternum; 0: present, near middle of prosternum; 1: expanded to posterior extremity of prosternum as broad groove; 2: absent.

34. (RAM66) Prosternal shape; 0: not or only moderately constricted medially; 1: strongly constricted.

35. (RAM69) Internal scrobal ridge from mesepisternal scrobe posteriorly to intersegmental suture; 0: absent; 1: present.

36. (RAM71) Distance between metapleural pits: height of metapleuron; 0: >0.20; 1: 0.10–0.19; 2: 0.09 or less.

37. (RAM72) Membrane closing space behind metasternum and hind coxae; 0: arises above free apex of metasternum; 1: arises from apical margin of metasternum.

38. (RAM74) Profile of metanotum; 0: subhorizontal or slanting; 1: vertical, not overhung by scutellum; 2: vertical, strongly overhung by scutellum.

39. (RAM75) Lower extremity of metapostnotum internally; 0: with vertical longitudinal ridge; 1: with longitudinal ridge extending downward to propodeal margin; 2: with ridge extends beyond marginal area of propodeum; 3: absent.

40. (RAM78) Conjuctiva between metasternum, hind coxae, and S1; 0: entirely membranous; 1: with sclerotized bars.

Legs

41. (RAM80) Hind trochanter with inner basal surface; 0: angulate; 1: rounded.

42. (RAM81) Isolation of hind tibial spur; 0: none to partial; 1: almost complete; 2: complete with sclerotized bridge.

43. (RAM83) Outer hind tibial spur of female; 0: finely serrate or ciliate; 1: coarsely serrate; 2: absent

44. (RAM84) Basitibial plate; 0: present, at least in female; 1: absent.

45. (RAM85) Hind tibial scopa with corbicula; 0: no; 1: yes.

46. (RAM86) Apex of inner surface of hind tibia (female); 0: without comb of bristles; 1: with comb of bristles (rastellum).

47. (RAM87) Apex of hind tibia (female); 0: not expanded dorsally; 1: expanded dorsally.

48. (RAM88) Base of hind basitarsus (female); 0: not broadened; 1: widened to form auricle.

49. (RAM89) Hind basitarsus (female); 0: producing second tarsomere at apex; 1: projecting distad as process without apical brush; 2: projecting distad as in (1) but with apical brush (penicillus).

50. (RAM90) Shape of hind basitarsus (female); 0: > 3× as long as wide; 1: 1.5× as long as wide; 2: 1.6–2.9× as long as wide.

51. (RAM91) Under surface of middle tibia (female); 0: with oblique longitudinal ridge bearing a longitudinal brush or hairs; 1: flat, with scattered hairs.

52. (RAM92) Middle tibial spur; 0: finely serrate or ciliate; 1: coarsely serrate; 2: serrate but ending in two to several large teeth or spines.

53. (RAM95) Trunk of anterior tibial spur; 0: simple; 1: with low expansion at right angle to velum; 2: with strong expansion at right angle to velum.

54. (RAM96) Velum of anterior tibial spur; 0: narrow, 1.5× long as wide; 1: broad, 1.1–1.45× long as wide; 2: about as long as broad.

55. (RAM98) Arolia; 0: present; 1: absent.

Wings

56. (RAM100) Number of submarginal cells in forewing; 0: three; 1: two; 2: none clearly defined.

57. (RAM101) Wing vestiture; 0: hairy throughout; 1: partly bare.

58. (RAM102) Length of marginal cell of forewing; 0: equal to or longer than distance from its apex to wing tip; 1: shorter than distance from its apex to wing tip.

59. (RAM103) Apex of marginal cell of forewing; 0: pointed, on wing margin; 1: separated from wing margin, pointed; 2: separated from wing margin, rounded; 3: open, or closed by weak vein.

60. (RAM104) Stigma of forewing; 0: longer than broad, margin within marginal cell convex to straight; 1: longer than broad, margin within marginal cell concave; 2: small, as long as broad; 3: narrow, almost parallel-sided.

61. (RAM105) Jugal lobe of hindwing; 0: long, 0.5× as long as vannal lobe; 1: short, 0.26–0.49× as long as vannal lobe; 2: shorter, 0.25× as long as vannal lobe; 3: absent.

Metasoma

62. (RAM113) Surface of T5 of female; 0: with prepygidial fimbria; 1: without prepygidial fimbria.

63. (RAM116) Pygidial plate of T6 of female; 0: present; 1: absent.

64. (RAM119) Apex of T7 of male; 0: entire; 1: with two conical points.

65. (RAM120) S7 of male; 0: with 2 or 4 apical lobes; 1: without apical lobes; 2: short or transverse, without apical lobes; 3: disc to whole sternum membranous.

66. (RAM121) S8 of male; 0: with single apical projection; 1: without apical projection; 2: bilobed apically; 3: almost completely absent.

Male Genitalia

67. (RAM122) Gonobase; 0: forming a complete ring; 1: not evident ventrally; 2: almost absent.

68. (RAM125) Gonostylus; 0: articulated to gonocoxite; 1: indistinctly fused to gonostylus (appearing absent); 2: double.

69. (RAM126) Volsella; 0: distinct, chelate; 1: a free sclerite but not chelate; 2: absent, or fused to gonocoxite.

70. (RAM127) Dorsal bridge of penis valves; 0: short or absent; 1: expanded posteriorly as spatha.

Larvae

71. (RAM larval 2) Spiculation on dorsal surface of labrum; 0: absent; 1: present.

72. (RAM larval 3) Epipharyngeal spiculation; 0: present; 1: absent.

73. (RAM larval 9) Anterior tentorial pit; 0: high; 1: low.

74. (RAM larval 11) Posterior thickening of head wall; 0: well developed; 1: weakly developed; 2: absent medially.

75. (RAM larval 14) Median longitudinal thickening of head wall; 0: absent; 1: developed only dorsally; 2: extending forward to level of epistomal suture.

76. (RAM larval 15) Hypostomal ridge; 0: well developed; 1: weak.

77. (RAM larval 17) Angle of hypostomal ridge to posterior thickening of head wall; 0: obtuse; 1: perpendicular.

78. (RAM larval 18) Pleurostomal ridge; 0: well developed; 1: weak.

79. (RAM larval 19) Epistomal ridge or depression.; 0: well below level of antennae; 1: arched upward to or above antennal level

80. (RAM larval 30) Mandibular apex; 0: simple; 1: bidentate with dorsal tooth or teeth subequal; 2: bidentate with ventral tooth longer.

81. (RAM larval 35) Teeth on dorsal apical edge of mandible; 0: present; 1: absent.

82. (RAM larval 37) Mandibular apical concavity; 0: weakly to moderately developed; 1: strongly developed.

83. (RAM larval 38) Mandibular concavity; 0: oblique, not scooplike; 1: scooplike.

84. (RAM larval 43) Maxillary palpus; 0: elongate, usually twice as long as basal diameter; 1: apparently absent; 2: shorter than basal diameter.

85. (RAM larval 46) Galea; 0: absent; 1: present.

86. (RAM larval 48) Labial palpus; 0: shorter than maxillary palpus; 1: subequal to or longer than maxillary palpus.

87. (RAM larval 55) Hypopharyngeal groove; 0: distinct; 1: absent or indistinct.

88. (RAM larval 56) Body integument; 0: with patches or transverse rows of conspicuous spicules or setae; 1: without conspicuous spicules or setae.

89. (RAM larval 57) Body integument; 0: apparently nonsetose; 1: seemingly conspicuously setose.

90. (RAM larval 60) Body form; 0: robust to moderately robust; 1: slender.

91. (RAM larval 61) Body, as seen in side view; 0: widest medially; 1: widest posteriorly.

92. (RAM larval 63) Dorsal conical tubercles, two per segment, on thorax and at least first abdominal segment; 0: absent; 1: present.

93. (RAM larval 66) Venter of abdominal segment X; 0: rounded, not produced; 1: produced.

94. (RAM larval 67) Dorsum of abdominal segment X; 0: without transverse line or ridge; 1: with transverse ridge; 2: with transverse line.

95. (RAM larval 74) Spiracular atrial rim; 0: present; 1: absent.

APPENDIX 2. MATRIX OF MORPHOLOGICAL CHARACTER STATES. POLYMORPHISMS ARE UNDERLINED AND SUBSCRIPTED

Xylocopa vir	00010	10100	01011	00110	10001	11000	00000	00020	10000	00010
Nomada sp.	0011$\underline{0_1}$	00010	01000	00011	00000	00000	00000	00100	10010	00000
Melissodes ag	01110	10100	20201	10011	00011	10001	10000	00020	02000	00022
Anthophora pa	01011	00000	20201	10002	01011	00000	01000	00130	00000	00021
Habropoda	01011	00001	20201	10002	00011	?0000	01000	00130	00000	00021
Centris tri	00010	00000	00101	10001	01011	10001	01000	10130	10000	00011
Bombus penn	00001	00101	00201	00101	10111	00100	10000	11201	12011	11101
Apis mell	00000	00111	01201	00?01	10202	00110	10210	00201	1?211	11101
Melipona ful	00000	00111	01200	00001	10212	00110	11210	11200	1?211	11011
Partamona cup	10000	00?11	01200	10001	10202	00110	11210	01200	1?211	11011
Euglossa cor	10111	01?00	20211	11001	01011	11100	10001	10221	12011	11111
Eufriesia viol	10111	11?01	20111	11001	01011	11100	10001	10221	12011	11101

Xylocopa vir	10021	01012	11002	11121	01101	00101	01120	01001	00000
Nomada sp.	10000	00000	21001	0101$\underline{_21}$	00021	10100	00000	01000	00000
Melissodes ag	10110	00022	00000	21021	?????	?????	?????	?????	?????
Anthophora pa	00020	01122	10011	01020	01001	00000	01120	00100	00000
Habropoda	00020	01021	10001	01220	01001	00001	01120	00100	00000
Centris tri	01011	01122	00011	01220	00000	?0001	01101	00110	00001
Bombus penn	10211	00021	31101	01011	01102	00012	11100	10011	11121
Apis mell	10210	00023	01102	12120	11011	11110	10021	11101	10011
Melipona ful	10010	20033	01101	32021	11110	11110	00121	11000	01011
Partamona cup	10010	20030	11101	32021	?????	?????	?????	?????	?????
Euglossa cor	10211	00022	21101	01011	?????	?????	?????	?????	?????
Eufriesia viol	11211	00022	31101	01011	00001	00102	11101	01011	11120