Molecular Phylogenetics and Evolution
Vol. 19, No. 1, April, pp. 76–93, 2001
doi:10.1006/mpev.2001.0911, available online at http://www.idealibrary.com on IDEAL®

Phylogenetic Utility of the Major Opsin in Bees (Hymenoptera: Apoidea): A Reassessment

John S. Ascher,[1] Bryan N. Danforth, and Shuqing Ji

Department of Entomology, Comstock Hall, Cornell University, Ithaca, New York 14853-0901

Received April 5, 2000; revised November 13, 2000

Major opsin (LW *Rh*) DNA sequence has been reported to provide useful data for resolving phylogenetic relationships among tribes of corbiculate bees based on analyses of 502 bp of coding sequence. However, the corbiculate tribes are believed to be of Cretaceous age, and strong support for insect clades of this age from small data sets of nucleotide sequence data has rarely been demonstrated. To more critically assess opsin's phylogenetic utility we generated an expanded LW *Rh* data set by sequencing the same gene fragment from 52 additional bee species from 24 tribes and all six extant bee families. Analyses of this data set failed to provide substantial support for monophyly of corbiculate bees, for relationships among corbiculate tribes, or for most other well-established higher-level relationships among long-tongued bees. However, monophyly of nearly all genera and tribes is strongly supported, indicating that LW *Rh* provides useful phylogenetic signal at lower taxonomic levels. When our expanded LW *Rh* data set is combined with a morphological and behavioral data set for corbiculate bees, the results unambiguously support the traditional phylogeny of the corbiculate bee tribes: (Euglossini + (Bombini + (Meliponini + Apini))). This implies a single origin of advanced eusocial behavior among bees rather than dual origins, as proposed by several recent studies.

Key Words: **phylogeny; molecular systematics; parsimony; maximum-likelihood; Apidae; corbiculate bees; *Apis; Bombus*.**

INTRODUCTION

DNA sequences obtained from protein-coding regions of nuclear genes have provided useful data for phylogenetic studies in a variety of insect taxa (Baker and DeSalle, 1997; Brower and DeSalle, 1994; Soto-Adames *et al.*, 1994; Shimada *et al.*, 1995; Tatarenkov *et al.*, 1999). Several nuclear genes have been developed for studying family-level divergences in Lepidoptera (Friedlander *et al.*, 1994, 1996, 1998; Mitchell *et al.*, 1997; Fang *et al.*, 1997; Brower and DeSalle, 1998; Regier *et al.*, 1998; Campbell *et al.*, 2000). Of these, EF-1α and PEPCK have been used in studies of bee relationships (Danforth *et al.*, 1999; Leijs, 2000; Danforth and Ji, 2001; Sipes and Wolf, 2001). Mardulyn and Cameron (1999) report on the utility of another nuclear gene, long-wavelength rhodopsin (also known as green or major opsin; hereafter referred to as LW *Rh*). Their conclusion that LW *Rh* recovers relationships among corbiculate bee tribes seems promising, since sources of phylogenetic data for recovering Mesozoic divergences in insects are few. LW *Rh* is a member of a multigene family encoding visual pigments (Chang *et al.*, 1996; Townson *et al.*, 1996) and occurs in three forms in bees, although the paralogous sequences are very different and orthology can usually be determined unambiguously.

Corbiculate bees are a monophyletic group of long-tongued bees (hereafter, L-T bees) within Apidae: Apinae (Roig-Alsina and Michener, 1993) and comprise four extant tribes, each of which is monophyletic (Michener, 2000 and references therein). These are the Euglossini (orchid bees, 175+ known species, five genera), Bombini (bumble bees, 239 species, all in the genus *Bombus* [Williams, 1998]), Apini (honey bees, at least 7 extant species, all in the genus *Apis* [Engel, 1999c]), and Meliponini (stingless bees, several hundred species, numerous genera). The corbiculate bee tribes are thought to be of Mesozoic age because a fossil meliponine has been described from the late Cretaceous (~80 mya; Michener, 2000 and references therein; Grimaldi, 1999; Engel, 2000a). Corbiculate bees are of interest because they include the only advanced eusocial bees, the only bees to store harvestable honey, and the most important pollinators of field and greenhouse crops.

The phylogenetic relationships among corbiculate bee tribes remain extraordinarily controversial. Nine of the 15 theoretically possible rooted trees for these four taxa have been published as potential phylogenies since 1984 (8 are diagrammed by Schultz *et al.*, 1999, Fig. 1; the 9th is Fig. 2c of Koulianos *et al.*, 1999; see

[1] To whom correspondence should be addressed. Fax: (607) 255-0939. E-mail: ja41@cornell.edu.

1055-7903/01 $35.00

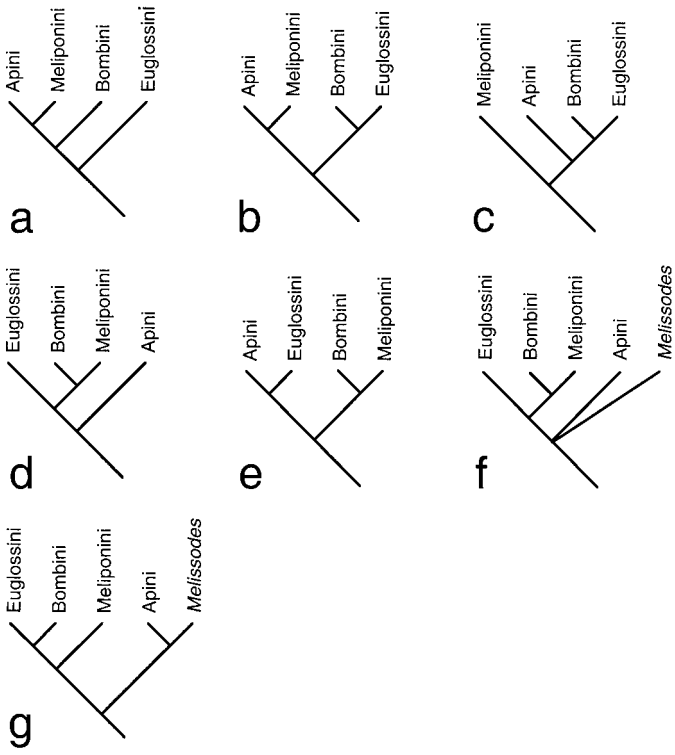

FIG. 1. Previously published phylogenies of corbiculate bee tribes. (a) Preferred phylogeny (referred to in the text as M44) (Michener, 1944, 1990 (in part, when characters of Prentice, 1991 were considered); Maa, 1953; Prentice, 1991; Roig-Alsina and Michener, 1993; Chavarría and Carpenter, 1994; Engel, 2000b, 2001; Noll, 2001; Schultz *et al.*, 1999 (morphology and overall conclusion); present study). (b) Second-best morphological phylogeny (Michener, 1974, 1990 (in part, not preferred over M44 when recoded by Chavarría and Carpenter, 1994); Schultz *et al.*, 1999 (combined analyses). (c) Morphological phylogeny implying dual origin of (or reversal in) advanced eusocial behavior in bees (Winston and Michener, 1977; Kimsey, 1984; Sakagami and Maeta, 1984). (d) Preferred molecular phylogeny of Mardulyn and Cameron (1999) (Sheppard and McPheron, 1991; Cameron, 1993 (published tree excluding *Exoneura*, one of two mP trees including *Exoneura*); Mardulyn and Cameron, 1999 (equal weights parsimony and overall conclusion); Koulianos *et al.*, 1999 (amino acid sequence). (e) Alternative molecular phylogeny (Cameron, 1991, 1993 (one of two MP trees including *Exoneura*); Schultz *et al.*, 1999 (reanalyses of Cameron, 1993); Koulianos *et al.*, 1999 (nucleotide sequence). (f) Mardulyn and Cameron (1999) (weighted parsimony analysis). (g) Mardulyn and Cameron (1999) (ML analysis).

also Prentice, 1991 and our Fig. 1). However, all recent morphological analyses of extant corbiculates (Prentice, 1991; Roig-Alsina and Michener, 1993) and extant plus fossil corbiculates (Engel, 2000b, 2001), a cladistic analysis of behavior (Noll, 2001), and simultaneous analyses of multiple data sets (Chavarría and Carpenter, 1994; Schultz *et al.*, 1999) support the phylogeny proposed by Michener (1944): (Euglossini + (Bombini + (Apini + Meliponini))) (Fig. 1a). This phylogeny, hereafter referred to as M44, is the only phylogeny consistent with a single origin of obligate eusociality among corbiculates and of advanced eusociality among bees (Michener, 1974; Prentice, 1991, Fig. 3.6a).

Mardulyn and Cameron's (1999) confidence in LW *Rh*'s phylogenetic utility was primarily due to topological congruence between their LW *Rh* tribal phylogeny of corbiculate bees (Fig. 1d) and the phylogenies (Figs. 1d and 1e) resulting from independent analyses of short (≤550-bp) nucleotide sequences obtained from three other genes (mitochondrial 16S rDNA [495 bp, 171 parsimony-informative sites], Cameron, 1991, 1993 [reanalyzed by Chavarría and Carpenter, 1994; Schultz *et al.*, 1999]; nuclear 28S rDNA [237 bp, 7 parsimony-informative sites], Sheppard and McPheron, 1991; and mitochondrial cytochrome *b* [520 bp, 167 parsimony-informative sites], Koulianos *et al.*, 1999). Mardulyn and Cameron (1999) consider their preferred corbiculate tribal phylogeny (Fig. 1d) to be "a well corroborated molecular phylogeny" despite its partial incongruence with their weighted parsimony (Fig. 1f) and maximum-likelihood (ML) (Fig. 1g) analyses of LW *Rh*, partial incongruence with tribal phylogenies based on analyses of other genes (Fig. 1e), and complete incongruence with all tribal phylogenies based on morphological characters or simultaneous analyses of morphological, behavioral, and molecular characters (Figs. 1a–1c). In fact, not a single morphological character has ever been reported that would support Bombini + Meliponini, the only clade recovered consistently by the molecular studies.

Our study sought to evaluate more fully the phylogenetic utility of LW *Rh* at two different taxonomic levels. First, we sought to examine how well LW *Rh* performed at the tribal, subfamilial, and familial levels within bees (focussing on the L-T clade). Our criteria for "performance" included (1) using jackknife, bootstrap, and Bremer support values to assess nodal support and (2) evaluating topological congruence between our results based on LW *Rh* and recent morphological studies of bees (e.g., Roig-Alsina and Michener, 1993). Assessing the level at which LW *Rh* provides useful phylogenetic information is very important because there are no previously published molecular studies of higher-level bee phylogeny (e.g., among families, subfamilies, and noncorbiculate tribes). Furthermore, precise information about the taxonomic level (and age of divergence) for which LW *Rh* provides phylogenetic signal is useful for those choosing genes for studies of other insect groups. Second, we sought to evaluate the phylogenetic results obtained by previous molecular studies of corbiculate bees, in particular the LW *Rh* study of Mardulyn and Cameron (1999). We investigated whether relationships obtained with only two outgroups (neither of which is considered to be the sister group to the corbiculates; Roig-Alsina and Michener, 1993) would hold when we added 51 noncorbiculate taxa from 23 tribes and all extant bee families. We further assessed corbiculate tribal relationships

based on LW *Rh* in simultaneous analysis with a morphological data set. The two goals of our study are complementary in that LW *Rh*'s utility at higher levels (among L-T bees in general) reflects its utility for recovering relationships among corbiculate tribes.

MATERIALS AND METHODS

We amplified and sequenced a ±700-bp fragment (including three exons and two introns) of the LW *Rh* gene from 52 bee species. Combining our 52 sequences with the 16 sequences obtained by Mardulyn and Cameron (1999) plus the complete *Apis* sequence generated by Chang *et al.* (1996), we produced an expanded data set of 68 species in total. We also included our expanded LW *Rh* data set in simultaneous analysis with a data set of morphological and behavioral characters derived from Schultz *et al.* (1999) with minor additions and modifications.

DNA Sequencing

Bees for this study were collected by the authors in New York, Arizona, New Mexico, California, and Italy (except for *Melipona* sp., which was collected by K. B. Miller in Bolivia). Most specimens used for sequencing were preserved in 95% EtOH, but recently collected pinned specimens were also used. Pinned specimens older than 3–5 years were not suitable for DNA extractions, but those collected more recently provided high-molecular-weight DNA for PCR. Outgroup and ingroup taxa included in this study, locality data, specimen voucher numbers, and GenBank accession numbers are listed in Table 1. Voucher specimens are housed in the Cornell University Insect Collection.

DNA extractions followed standard protocols detailed in Danforth *et al.* (1999). Two sets of PCR products were used to generate the data set. Primers for PCR and sequencing (LWRhF and LWRhR) were developed by Donat Angosti of the American Museum of Natural History and are listed in Mardulyn and Cameron (1999). These primers produced multiple bands in many taxa. However, there was usually a ±700-bp PCR product that appeared considerably stronger than all others. This was presumed to be the long-wavelength paralog. An additional nested forward primer LWRhF3 (5′-AGA TAC AAC GTR ATC GTS AAR GGT-3′) was used to amplify *Ceratina, Exomalopsis,* and *Tetraloniella* sequences, because LWRhF did not work for these taxa.

PCR amplifications were carried out following standard protocols (Palumbi, 1996), with the following cycle conditions: 94°C, 1 min denaturation; 50–56°C, 1 min annealing; 72°C, 1 min to 1 min 30 s extension. Prior to sequencing, PCR products were gel-purified in low-melting-point agarose (FMC, Rockland, ME) overnight at 4°C and then purified using the Promega (Madison, WI) Wizard PCR Preps DNA Purification kit. Automated sequencing of PCR products was performed on an ABI 377 automated sequencer using the LWRhF, LWRhR, and LWRhF3 primers. All sequences were confirmed in both directions.

Taxon Sampling

We sequenced LW *Rh* in 52 species, 1 of which was represented from two localities (giving a total of 53 sequences). These sequences were combined with the 15 sequences published by Mardulyn and Cameron (1999) plus the complete *Apis mellifera* coding sequence published by Chang *et al.* (1996) (GenBank Accession No. U26026). The total data set consisted of 68 species (Table 1) (69 sequences) belonging to 44 genera, 27 tribes, 10 subfamilies, and all six bee families. Our data set includes many taxa within the L-T bee clade, which is derived from a paraphyletic group of primitive bees referred to as the short-tongued bees (S-T bees; Alexander and Michener, 1995).

This is the first molecular study of the L-T bees to include representatives of any of the nine tribes (Centridini, Rhathymini, Ericrocidini, Melectini, Anthophorini, Tetrapediini, Osirini, Protepeolini, and Isepeolini) that appear closest to the corbiculates in the best available cladograms of L-T bees (Roig-Alsina and Michener, 1993). Inclusion of the Centridini, Protepeolini, Melectini, and Anthophorini in our study is particularly important because these tribes are believed to be more closely related to corbiculates (Roig-Alsina and Michener, 1993) than are the tribes/subfamilies used as outgroups in earlier molecular studies. Centridini are believed to be (or belong to) the sister group to the corbiculates. Use of exemplars representing 12 apine tribes allows us to test ingroup monophyly, whereas the inclusion of multiple genera of selected tribes (e.g., three genera of Eucerini and of Anthidiini) and multiple species of selected genera (e.g., four species of *Diadasia* and three species of *Nomada, Anthophora,* and *Hoplitis*) allows us to assess performance of LW *Rh* at lower taxonomic levels.

We rooted our trees with *Colletes skinneri* because we hypothesize that the Colletidae (*sensu* Michener, 1944; including the stenotritine genera *Stenotritus* and *Ctenocolletes*) is the monophyletic sister group to the rest of the bees.

Parsimony Analyses

Parsimony analyses were performed using PAUP*4 (PAUP v. 4.0b2a; Swofford, 1999), Winclada (Beta) ver. 0.99 (Nixon, 1999), and NONA (Goloboff, 1994). For PAUP analyses we used heuristic searches with TBR branch swapping, random addition sequence for taxa, and 50 replicates per search. Most analyses were performed based on equal weights, but we also examined the consequences of downweighting third position transversions by 1/4 (as recommended by Mardulyn and Cameron [1999]). Bootstrap support (Felsenstein,

TABLE 1

Taxa Included in the Analysis Arranged by Family, Subfamily, and Tribe

Voucher No.		Species	Author	Family	Subfamily	Tribe	GenBank Access. No.	Locality data
Ansp	643	*Andrena* (*Callandrena*) sp.		Andrenidae	Andreninae	Andrenini	AF344618	AZ: Cochise Co. 20 mi S. Animas 17 Sept. 1999
Cafr	515	*Calliopsis fracta*	(Rozen)	Andrenidae	Panurginae	Calliopsini	AF344587	CA: Santa Clara Co. San Antonio Junction 28 May 1999
Capu	509	*Calliopsis pugionis*	Cockerell	Andrenidae	Panurginae	Calliopsini	AF344588	CA: Riverside Co. San Jacinto WA 18 May 1993
Pnca	514	*Panurgus calcaratus*	Scopoli	Andrenidae	Panurginae	Panurgini	AF344612	ITALY: Rome 7 June 1998
Anfu	658	*Anthophora furcata*	(Panzer)	Apidae	Apinae	Anthophorini	AF344615	NY: Tompkins Co. Ithaca 22 August 1999
Anmo	642	*Anthophora montana*	Cresson	Apidae	Apinae	Anthophorini	AF344616	AZ: Cochise Co. 20 mi S. Animas 17 Sept. 1999
Anur	504	*Anthophora urbana*	Cresson	Apidae	Apinae	Anthophorini	AF344585	CA: Santa Clara Co., Del Puerto Cyn. 27 May 1999
		Apis dorsata	Fabricius	Apidae	Apinae	Apini	AF091733	
		Apis mellifera	Linnaeus	Apidae	Apinae	Apini	AF091732	
		Apis nigrocincta	Smith	Apidae	Apinae	Apini	AF091728	
		Bombus avinoviellus	Skorikov	Apidae	Apinae	Bombini	AF091719	
		Bombus pensylvanicus	(DeGeer)	Apidae	Apinae	Bombini	AF091727	
		Bombus terrestris	(Linneaus)	Apidae	Apinae	Bombini	AF091722	
Cnho	503	*Centris hoffmanseggiae*	Cockerell	Apidae	Apinae	Centridini	AF344590	CA: Kern Co. 5 mi S. Mojave 13 June 1999
Dibi	490	*Diadasia bituberculata*	(Cresson)	Apidae	Apinae	Emphorini	AF344594	CA: Contra Costa Co. Mitchell Cyn. 5 June 1999
Didi		*Diadasia diminuta*	(Cresson)	Apidae	Apinae	Emphorini	AF344592	UT: Cache Co., West of Hyrum Dam 3 Sept. 1996 (S. Sipes)
Dima	486	*Diadasia martialis*	Timberlake	Apidae	Apinae	Emphorini	AF344595	CA: San Diego Co. Borrego Springs 18 March 1999
Dini		*Diadasia nigrifrons*	(Cresson)	Apidae	Apinae	Emphorini	AF344593	UT: Cache Co., Logan Canyon 5 July 1996 (S. Sipes)
Ptsp	629	*Ptilothrix* sp.		Apidae	Apinae	Emphorini	AF344629	AZ: Cochise Co. 12 mi NE Portal 12 Sept. 1999
Ptsp	648	*Ptilothrix* sp.		Apidae	Apinae	Emphorini	AF344630	NM: Hidalgo Co. Rodeo 22 Sept. 1999
Mede	485	*Melissodes desponsa*	Smith	Apidae	Apinae	Eucerini	AF344603	NY: Tompkins Co. Ithaca, 29 July 1997
		Melissodes rustica	(Say)	Apidae	Apinae	Eucerini	AF091731	
Mesp	229	*Melissodes* sp.		Apidae	Apinae	Eucerini	AF344606	NM: Cochise Co. Rodeo 20 Aug. 1997
Svma	617	*Svastra machaerantherae*	(Cockerell)	Apidae	Apinae	Eucerini	AF344631	AZ: Cochise Co. 12 mi NE Portal 19 Sept. 1999
Svob	631	*Svastra obliqua*	(Say)	Apidae	Apinae	Eucerini	AF344632	NM: Luna Co. 34 mi NE Deming 13 Sept. 1999
Ttsp	646	*Tetraloniella* sp.		Apidae	Apinae	Eucerini	AF344636	NM: Hidalgo Co. 20 mi S. Animas 17 Sept. 1999
		Eufriesea caerulescens	(Lepeletier)	Apidae	Apinae	Euglossini	AF091725	
		Euglossa imperialis	Cockerell	Apidae	Apinae	Euglossini	AF091720	
		Eulaema meriana	(Olivier)	Apidae	Apinae	Euglossini	AF091721	
		Exaerete frontalis	Guerin	Apidae	Apinae	Euglossini	AF091718	
Excm	627	*Exomalopsis completa*	Cockerell	Apidae	Apinae	Exomalopsini	AF344622	AZ: Cochise Co. 12 mi NE Portal 8 Sept. 1999
Exru	625	*Exomalopsis rufiventris*	Timberlake	Apidae	Apinae	Exomalopsini	AF344623	AZ: Cochise Co. 14 mi W Douglas 11 Sept. 1999

TABLE 1—*Continued*

Voucher No.		Species	Author	Family	Subfamily	Tribe	GenBank Access. No.	Locality data
Xmca	499	*Xeromelecta californica*	Timberlake	Apidae	Apinae	Melectini	AF344613	CA: Santa Clara Co. Del Puerto Cyn. 27 May 1999
Zoma	650	*Zacosmia maculata*	(Cresson)	Apidae	Apinae	Melectini	AF344637	NM: Grant Co. Hachita 24 Sept. 1999
Mpsp	522	*Melipona* sp.		Apidae	Apinae	Meliponini	AF344607	BOLIVIA: Dept. Santa Cruz, San Jose' 26 June 1999
		Lestrimelitta limao	(Smith)	Apidae	Apinae	Meliponini	AF091723	
		Scaptotrigona depilis	(Moure)	Apidae	Apinae	Meliponini	AF091729	
		Trigona dorsalis	Smith	Apidae	Apinae	Meliponini	AF091726	
		Trigona necrophaga	Camargo & Roubik	Apidae	Apinae	Meliponini	AF091724	
Lpsn	651	*Leiopodus singularis*	(Lin. & Mich.)	Apidae	Apinae	Protepeolini	AF344624	NM: Grant Co. Hachita 24 Sept. 1999
Orba	637	*Oreopasites barbarae*	Rozen	Apidae	Nomadinae	Ammobatini	AF344626	AZ: Cochise Co. 14 mi SW Apache 10 Sept. 1999
Epsc	489	*Epeolus scutellaris*	Say	Apidae	Nomadinae	Epeolini	AF344596	NY: Tompkins Co. Ithaca 22 August 1997
Trrz	635	*Triepeolus "rozeni"*	Hurd, ms. name	Apidae	Nomadinae	Epeolini	AF344634	AZ: Cochise Co. Chiricahua Monument 14 Sept. 1999
Trvb	641	*Triepeolus verbesinae*	(Cockerell)	Apidae	Nomadinae	Epeolini	AF344635	AZ: Cochise Co. 14 mi SW Apache 10 Sept. 1999
Hoca	519	*Holcopasites calliopsidis*	(Linsley)	Apidae	Nomadinae	Holcopasitini	AF344600	NY: Schuyler Co. Valois 18 June 1999
Horu	511	*Holcopasites ruthae*	Cooper	Apidae	Nomadinae	Holcopasitini	AF344602	CA: Riverside Co. San Jacinto WA 18 May 1993
Noim	502	*Nomada imbricata*	Smith	Apidae	Nomadinae	Nomadini	AF344608	NY: Tompkins Co. Ithaca 24 April, 1999
Nomc	501	*Nomada maculata*	Cresson	Apidae	Nomadinae	Nomadini	AF344609	NY: Tompkins Co. Ithaca 3 May, 1999
Noob	492	*Nomada obliterata*	Cresson	Apidae	Nomadinae	Nomadini	AF344610	NY: Tompkins Co. Ithaca 3 May, 1999
Pnve	652	*Paranomada velutina*	Linsley	Apidae	Nomadinae	Nomadini	AF344627	AZ: Cochise Co. 2 mi E. Apache 10 Sept. 1999
Trpn	653	*Triopasites penniger*	(Cockerell)	Apidae	Nomadinae	Nomadini	AF344633	NM: Grant Co. Hachita 24 Sept. 1999
Cedu	656	*Ceratina calcarata*	Robertson	Apidae	Xylocopinae	Ceratinini	AF344620	NY: Tompkins Co. Ithaca 4 August 1999
Xytb	500	*Xylocopa tabaniformis*	Smith	Apidae	Xylocopinae	Xylocopini	AF344614	CA: Santa Clara Co. Mt. Hamilton 27 May 1999
		Xylocopa virginica	(Linnaeus)	Apidae	Xylocopinae	Xylocopini	AF091730	
Cosk	632	*Colletes skinneri*	Viereck	Colletidae	Colletinae	Colletini		AZ: Cochise Co. Chiricahua Mts. 14 Sept. 1999
Noht	634	*Dieunomia heteropoda*	Say	Halictidae	Nomiinae	Nomiini	AF344625	AZ: Cochise Co. Chiricahua Monument 14 Sept. 1999
Prve	649	*Dieunomia nevadensis*	Cresson	Halictidae	Nomiinae	Nomiini	AF344628	AZ: Cochise Co. Apache 22 Sept. 1999
Ansp	630	*Anthidiellum notatum*	(Latreille)	Megachilidae	Megachilinae	Anthidiini	AF344617	AZ: Cochise Co. 12 mi NE Portal 12 Sept. 1999
Atob	505	*Anthidium oblongatum*	(Illiger)	Megachilidae	Megachilinae	Anthidiini	AF344586	NY: Tompkins Co. Ithaca 1 July 1999
Ansp	645	*Anthidium porterae*	Cockerell	Megachilidae	Megachilinae	Anthidiini	AF344619	NM: Hidalgo Co. 20 mi S. Animas 17 Sept 1999
Paju	495	*Paranthidium jugatorium*	(Say)	Megachilidae	Megachilinae	Anthidiini	AF344611	NY: Tompkins Co. Ithaca 31 July 1997
Cxal	487	*Coelioxys alternata*	Say	Megachilidae	Megachilinae	Megachilini	AF344591	NY: Tompkins Co. Ithaca 29 July 1997
Mepg	595	*Megachile pugnata*	Say	Megachilidae	Megachilinae	Megachilini	AF344605	NY: Schuyler Co. Valois 14 July 1999
Chfu	496	*Chelostoma fuliginosum*	(Panzer)	Megachilidae	Megachilinae	Osmiini	AF344589	NY: Tompkins Co., Ithaca 24 June 1997

TABLE 1—*Continued*

Voucher No.		Species	Author	Family	Subfamily	Tribe	GenBank Access. No.	Locality data
Hoal	507	*Hoplitis albifrons*	(Kirby)	Megachilidae	Megachilinae	Osmiini	AF344598	CA: Contra Costa Co. Donner Cyn. 30 May 1999
Hobi	493	*Hoplitis biscutellae*	(Cockerell)	Megachilidae	Megachilinae	Osmiini	AF344599	CA: Riverside Co. Corn Springs 16 March 1997
Hopi	506	*Hoplitis pilosifrons*	(Cresson)	Megachilidae	Megachilinae	Osmiini	AF344601	NY: Tompkins Co. Ithaca 22 May 1999
Hela	488	*Hesperapis larreae*	(Cresson)	Melittidae	Dasypodainae	Dasypodaini	AF344597	CA: Los Angeles Co. Palmdale 13 June 1999
Meew	508	*Melitta eickworti*	Snelling & Stage	Melittidae	Melittinae	Melittini	AF344604	NY: Tompkins Co. Ithaca 25 June 1997

Note. GenBank accession numbers are indicated for all sequences, and locality data are given for all specimens sequenced by us. We have revised the spelling of one of Mardulyn and Cameron's (1999) exemplars and suggest a revised identification for another. *B. pensylvanicus* (De Geer) is spelled with one "n" (Williams, 1998). "*Trigona hypogea*" is *T. necrophaga* Camargo and Roubik (Engel and Schultz, 1997; Schultz *et al.*, 1999) if "*T. hypogea*" of Mardulyn and Cameron (1999) refers to the same DNA extraction and voucher as that of Cameron (1991, 1993).

1985) was calculated based on 100 replicates with 10 random addition sequences per replicate. Bremer support (BS; also known as branch support or the decay index; Bremer, 1988, 1994) was calculated by saving trees one to four steps longer than the most parsimonious trees. Parsimony jackknifing (Farris *et al.*, 1996) was performed using Winclada (Beta) ver. 0.99 and NONA. Jackknife frequencies were calculated based on 1000 replicates with 10 random sequence additions per replicate, 10 trees held, and swapping to completion on the shortest trees obtained using the command max*.

Simultaneous Analysis of Morphological and Molecular Data

Simultaneous analyses (equivalent to the total evidence approach of Chavarría and Carpenter, 1994; see Nixon and Carpenter, 1996; Kluge, 1998; Gatesy *et al.*, 1999) were performed on a data set consisting of our expanded LW *Rh* data in combination with a data set of morphological and behavioral characters (Appendix) derived from that of Schultz *et al.* (1999), with characters that were uninformative for tribal relationships deleted and with modifications that reflect our revised interpretation of the homology of certain characters and different views about cladistic character coding included. These simultaneous analyses were limited to testing relationships among the corbiculate tribes because monophyly of the corbiculates and of each corbiculate tribe is uncontroversial and because consideration of morphological characters that were informative within these tribes and within the outgroups is beyond the scope of this study. We did not include all the molecular studies published previously because their sampling of noncorbiculates is very limited, because the 16S and 28S rDNA data sets have already been analyzed in simultaneous analyses (Chavarría and Carpenter, 1994; Schultz *et al.*, 1999), and because work underway specifically addresses this problem (S. A. Cameron and P. Mardulyn, unpublished).

Maximum-Likelihood Analyses

Maximum-likelihood analyses were performed using PAUP*4 (PAUP v. 4.0b2a; Swofford, 1999) to replicate the analyses of Mardulyn and Cameron (1999) and to explore the degree to which the results obtained under equal weights parsimony are due to rate heterogeneity among sites, biased base composition, or unequal transition/transversion (ti/tv) rates, parameters that are incorporated in the models discussed below. For the maximum-likelihood analyses we initially used the equal weights parsimony trees to estimate the log likelihood of each tree under 20 possible models of sequence evolution (Sullivan and Swofford, 1997; Frati *et al.*, 1997; Huelsenbeck and Crandall, 1997), from the simplest (Jukes–Cantor) to the most parameter rich (general time-reversible). We accounted for rate heterogeneity among sites in five ways: (1) no rate heterogeneity, (2) proportion of invariant sites (I), (3) gamma-distributed rates (G) (with four rate categories), (4) proportion of invariant sites + gamma distributed rates (G + I), and (5) site-specific rates (SSR), whereby rates were estimated for each codon position.

When performing branch swapping in ML we used the ML parameters obtained from the equal weights parsimony tree as starting parameters for tree searching. We then performed a series of branch swapping iterations, in the following order: NNI, SPR(1), SPR(2), TBR(1), and TBR(2). At each iteration the ML parameters were reestimated based on the trees currently in memory and applied to the next round of branch swapping. The parameter estimates resulting from this search algorithm are discussed below. Four models were selected for tree search: K2P + SSR, K2P + G + I, HKY + SSR, and GTR + SSR. Bootstrap values were calculated using 100 replicates.

TABLE 2

Nucleotide Frequencies Based on Overall LW *Rh* Data Set

	A	C	G	T	*P* value
Overall	24.75	23.73	22.27	28.83	0.999
First	32.58	14.60	25.40	27.39	0.999
Second	22.41	22.59	18.52	36.48	0.999
Third	19.27	34.00	24.09	22.62	0.0029

RESULTS

Alignment

We aligned the 15 sequences published by Mardulyn and Cameron (1999) with the published coding sequence for *A. mellifera* (Chang *et al.*, 1996) plus an additional 53 bee sequences using MegAlign in the Lasergene software package (DNASTAR Inc., Maddison, WI). Because we sequenced genomic DNA our sequences included two introns of variable length (as did those of Mardulyn and Cameron, 1999, p. 171). Intron/exon boundaries were identified by comparison with the published coding sequence for *A. mellifera*. Because alignments within the introns were unclear, we excluded the introns from our analyses (as did Mardulyn and Cameron, 1999). However, sequences submitted to GenBank include the entire genomic sequence obtained, including introns. All of the eucerine genera sequenced (*Svastra, Melissodes, Tetraloniella*) and amplified (*Xenoglossa* and *Peponapis*) have exceptionally large introns in comparison to those of other L-T bees.

Some of our aligned sequences were incomplete. In four taxa (*Ceratina, Tetraloniella,* and both species of *Exomalopsis*) 110 bp were missing from the 5′ end of our sequences because the nested primer LWRhF3 was used. Smaller numbers of base pairs were also missing both in the sequences generated by us (up to 9 bp missing in *Melissodes desponsa*) and in the sequences generated by Mardulyn and Cameron (1999) (up to 50 bp missing in *Apis nigrocinctus*).

The aligned data set included a total of 495 aligned sites rather than the 502 bp reported by Mardulyn and Cameron for their data set. This discrepancy is due to our exclusion of 7 bp at the 5′ end of our data set due to the lack of data for the majority of sequences, including 7 of the 15 generated by Mardulyn and Cameron.

Nucleotide alignments are available from the corresponding author.

Base Composition

The overall base composition and the base composition broken down by codon position are shown in Table 2. Overall, the base composition was only slightly A/T biased (53.5%). Our results are comparable to those of Mardulyn and Cameron (1999) in that first position sites tend to show the greatest A/T bias (59.9% A/T in our data set and 60.6% A/T in the Mardulyn and Cameron [1999] data set).

There was significant heterogeneity in base composition among taxa only in third position sites ($P = 0.0029$), which tended overall to show a slight G/C bias (58.1% G/C). We investigated third position base composition among tribes of corbiculate bees by first calculating overall base composition for noncorbiculate Apinae and then comparing each of the four corbiculate tribes to this null hypothesis. While base composition for Bombini, Meliponini, and Euglossini did not differ significantly from that for the noncorbiculate apines ($P > 0.05$), Apini showed highly significant ($P < 0.001$) deviation from the noncorbiculate Apinae in general (Fig. 2). These differences are due to the substantial A/T bias in the three *Apis* species (41.9% G/C) as compared to the overall G/C content in the noncorbiculate apines (56.5% G/C).

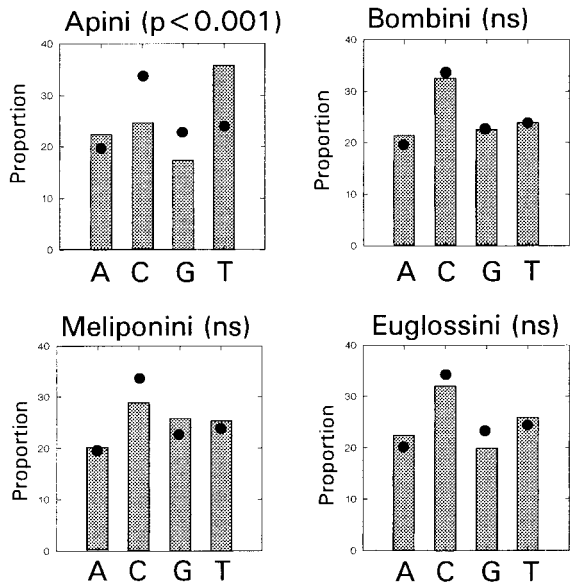

FIG. 2. Base composition at third position sites for the four tribes of corbiculate bees compared to the composition of third position sites in the noncorbiculate Apinae. Base composition among the noncorbiculate Apinae (indicated by the closed circles) was as follows: 19.82% A, 33.83% C, 22.63% G, and 23.72% T.

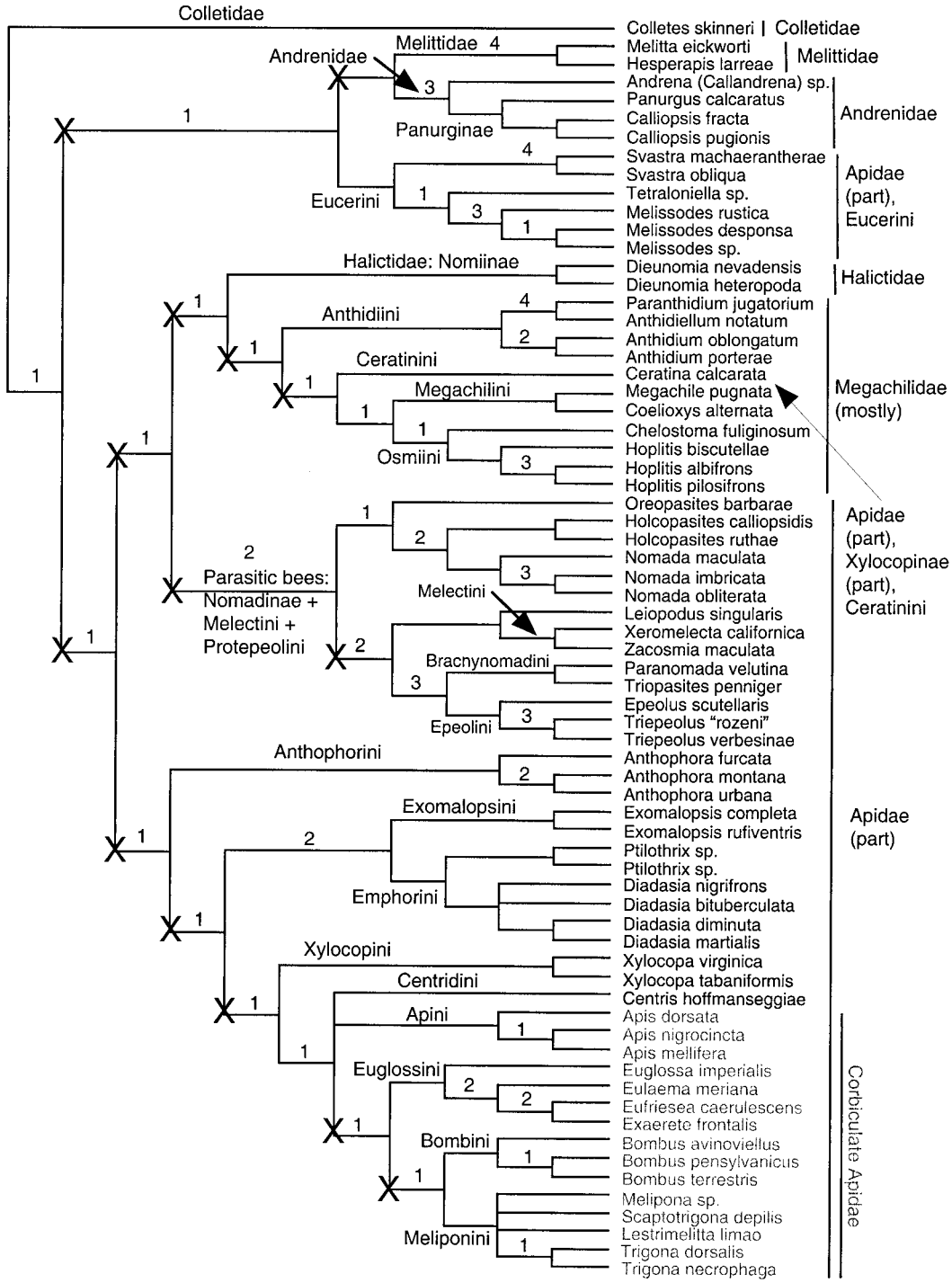

FIG. 3. Consensus tree of eight equally parsimonious trees obtained from unweighted analysis of nucleotide data (228 parsimony-informative sites, CI = 0.2473, RI = 0.6373, length = 1686). *Colletes skinneri* was used to root the tree. Bremer support (BS) values are shown above the branches (BS ≥ 5 for unlabeled nodes). Branches subtending nodes strongly incongruent with previous morphological studies are marked with an X. Corbiculate bees are indicated in gray.

Parsimony Analyses

Equal weights parsimony analyses of the entire 495-bp alignment (228 sites were parsimony informative) using either PAUP or NONA resulted in the same eight trees of length (L) 1686 steps (Fig. 3). Parsimony-informative sites were mostly in third positions (Fig. 4a) but first and second positions accounted for approximately one third of all parsimony-informative sites.

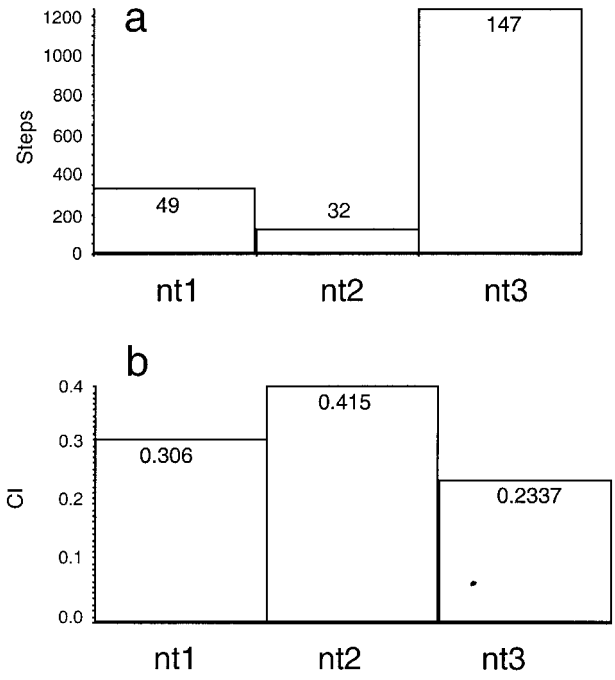

FIG. 4. Analysis of sequence variation based on codon position. (a) Histogram showing numbers of steps on the most parsimonious trees for nt1, nt2, and nt3. Values on bars indicate numbers of parsimony-informative nucleotide sites by position. (b) Histogram showing the CI for each nucleotide position. Values are indicated on bars.

Homoplasy was highest in third position sites and lowest in second position sites, as expected for protein-coding genes (Fig. 4b). No significant incongruence among sites or between nt1 + nt2 and nt3 was detected using the incongruence length difference test (Farris *et al.*, 1994).

Well-supported nodes in equal weights parsimony analysis were identified as those with >50% bootstrap and jackknife support (Fig. 5) and Bremer support greater than 3. Bootstrap and jackknife frequencies were similar, as expected. All 16 genera and 19 tribes that are represented by multiple exemplars are resolved as monophyletic groups and are supported by jackknife frequencies of ≥57%. Tribes are supported by jackknife frequencies of 98% or more, except for Epeolini (93%) and Osmiini (72%). Osmiini is the tribe with the least morphological support (Roig-Alsina and Michener, 1993) and may prove to be paraphyletic with respect to Megachilini (Engel, 1999b) when additional taxa are sampled. All genera are supported by jackknife values of ≥84%, except for *Anthidium* (60%) and *Trigona* (57%). Most potential relationships within genera and tribes are also resolved and congruent with our expectations based on adult morphology (e.g., relationships between our exemplars of *Bombus* (Williams, 1994), *Nomada,* and *Hoplitis*). Support for relationships at or below the tribal level is substantial, as demonstrated by jackknife frequencies of 98% or greater at 23 nodes (excluding the outgroup node). In contrast, relationships among tribes, subfamilies, and families are unresolved in our bootstrap and jackknife consensus trees, excepting the subfamily Megachilinae (to which all 10 of our megachilid exemplars belong), the subfamily Panurginae (represented in our analyses by members of the derived panurgine tribes Panurgini and Calliopsini (Ruz, 1987)), and Melittidae. All recovered clades are fully congruent with morphological and behavioral data except relationships within the corbiculate tribe Euglossini, which are inconsistent with morphology-based phylogenies (Kimsey, 1982, 1987; Michener, 1990; Engel, 1999a).

Bremer support values generally correspond with bootstrap and jackknife frequencies (Fig. 3). All clades with jackknife frequencies >72% have Bremer support of 2 or more. All tribes (except Osmiini, BS = 1) and all genera (except *Trigona,* BS = 1, *Anthidium,* BS = 2, and *Triepeolus,* BS = 3) have Bremer support of 5 or more, as does the subfamily Panurginae. Three clades were not present in the bootstrap or jackknife 50% consensus tree but have Bremer support of 3 or more. One of these, the Andrenidae (BS = 3), is congruent with morphological data, whereas the other two (Brachynomadini + Epeolini, BS = 3; Protepeolini + Melectini, BS ≥ 5) are not (Rozen, 1996; Roig-Alsina and Michener, 1993). The latter two results are suspect due to the extremely long branch subtending *Leiopodus* (see below). Relationships among the corbiculate bee tribes are unresolved in the consensus of trees one step longer than the most parsimonious trees (i.e., BS = 1 for these clades).

Some clades present in the strict consensus tree (Fig. 3), but receiving no or minimal nodal support (not appearing in the bootstrap or jackknife 50% consensus trees nor with BS > 2), are those that one would expect based on morphology (e.g., Megachilini + Osmiini). However, most such clades are spurious, as they cannot be reconciled with morphological data (indicated with X on Fig. 3). For example, the Andrenidae appear as sister to the Melittidae, the Megachilidae (a family of L-T bees) is sister to Halictidae (a family of S-T bees), the Eucerini (a tribe of L-T bees in the family Apidae) is sister to the Andrenidae plus Melittidae (two families of S-T bees), and *Ceratina* (an apid) arises from within the Megachilidae and is far removed from *Xylocopa* (another member of the Xylocopinae). Placement of *Leiopodus* and the Melectini (both cleptoparasitic apines) within Nomadinae is also incongruent with morphological studies (Roig-Alsina and Michener, 1993), as are relationships among tribes of Nomadinae (Rozen, 1996) and the clade *Exaerete* + *Eufriesea* (Kimsey, 1982, 1987; Michener, 1990; Engel, 1999a). Furthermore, several clades strongly supported by morphological data, such as L-T bees, Melittidae + L-T

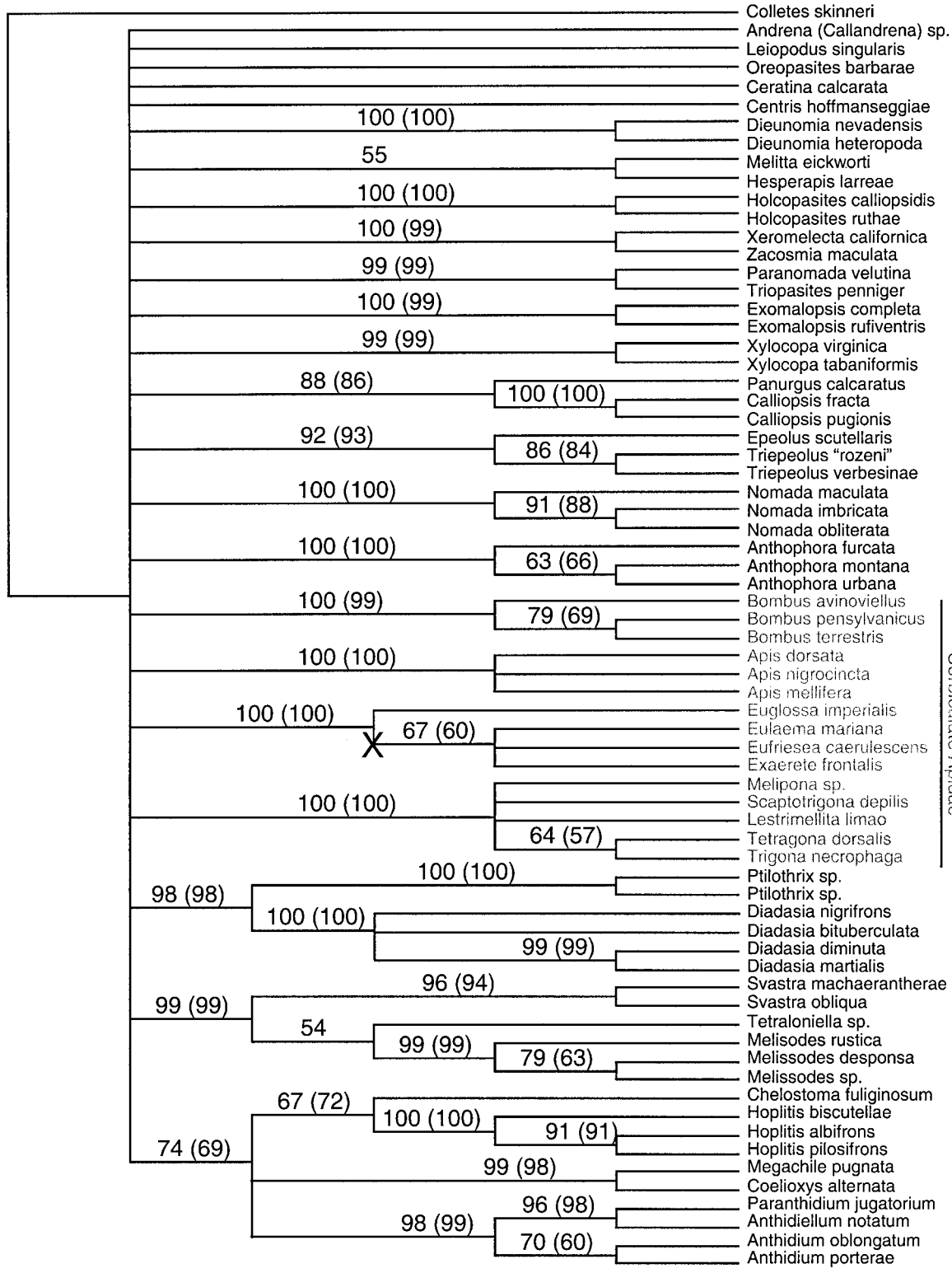

FIG. 5. The 50% bootstrap consensus tree based on 100 bootstrap replicates with 10 random addition sequences. Jackknife frequencies (in parentheses) were calculated using 1000 replicates with 10 random sequence additions per replicate, hold/10, and max*. The single clade inconsistent with previous morphological studies is marked with an X. Corbiculate bees are indicated in gray.

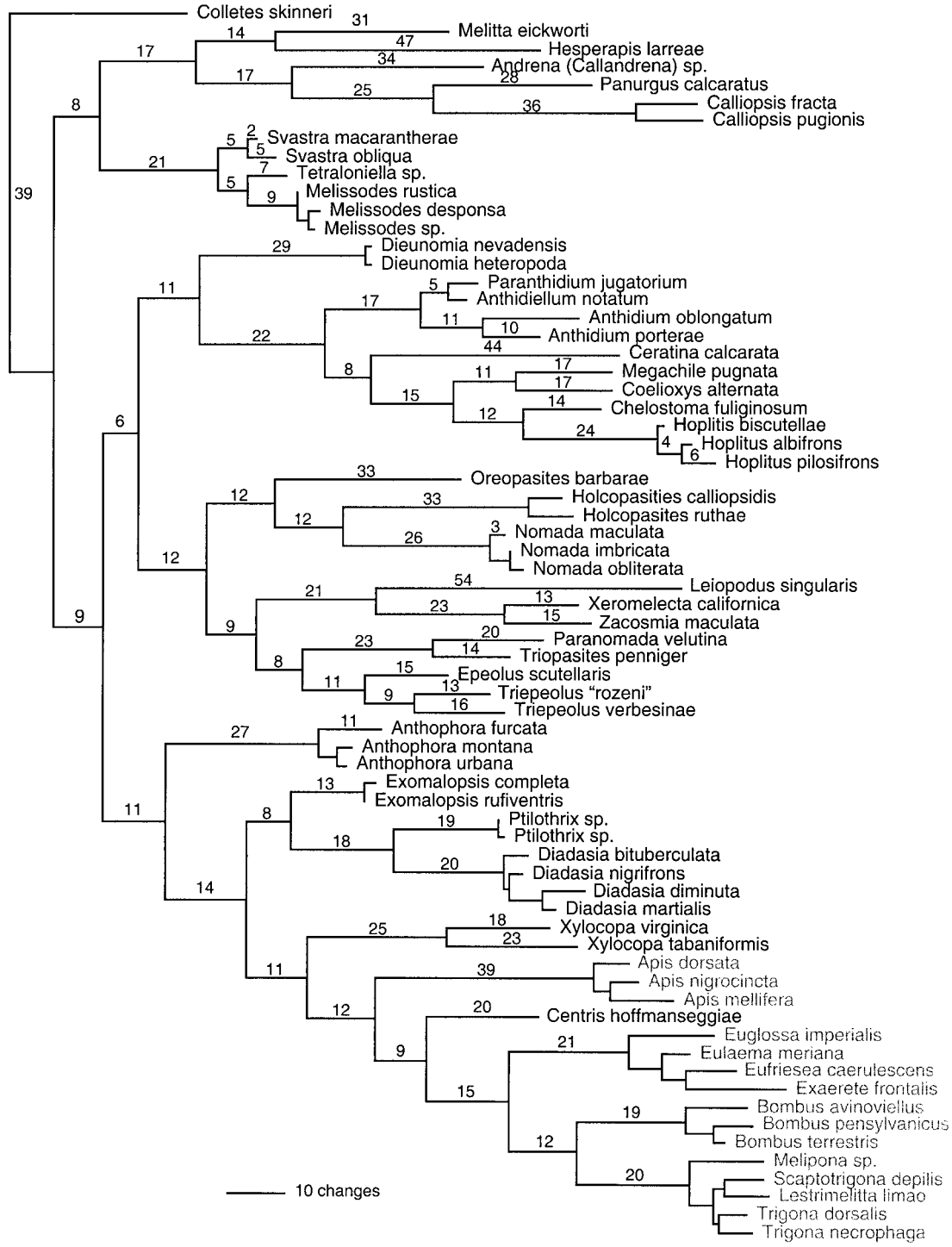

FIG. 6. One of the eight equally parsimonious trees. Branch lengths are proportional to character state changes under ACCTRAN optimization. Numbers above branches indicate number of characters supporting each node of the tree. Corbiculate bees are indicated in gray.

bees, Apidae, Nomadinae, and Xylocopinae, are not recovered in our consensus of most parsimonious trees.

Figure 6, one of the eight equally parsimonious trees, is included to show relative branch lengths under ACCTRAN optimization and should not be viewed as a preferred phylogenetic hypothesis. One of two equally parsimonious resolutions of relationships among corbiculate tribes (Fig. 6) is anomalous in that the corbiculates are paraphyletic with respect to *Centris*, which we expect to be the sister taxon to the corbicu-

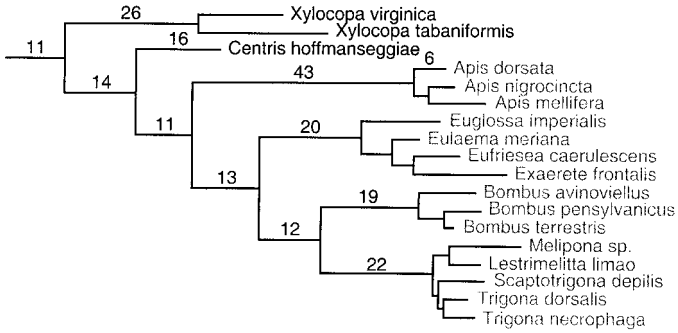

FIG. 7. The other equally parsimonious resolution of relationships among the corbiculate bee tribes. (See Fig. 6 for the alternative resolution). Corbiculate bees are indicated in gray.

lates. The other (Fig. 7) corresponds to the preferred molecular phylogeny of Mardulyn and Cameron (1999; Fig. 1d). The exceptionally long branches subtending *Apis*, *Ceratina*, and especially *Leiopodus* indicate that the anomalous placement of these taxa on the cladogram may be unreliable. For example, *Apis* shows remarkably high sequence divergence (Fig. 8) and significantly biased third position base composition (see above) compared to all other corbiculate genera.

Downweighting of third position transitions by 1/4 (as recommended by Mardulyn and Cameron, 1999) yielded trees similar to the equal weights tree (for example, corbiculates remained paraphyletic) but with elevated levels of bootstrap support for certain nodes (support for Bombini + Meliponini increased from <50% to 76%).

Maximum-Likelihood Analysis

Figure 9 shows the −Ln likelihood scores of the equal weights parsimony trees based on the 20 models

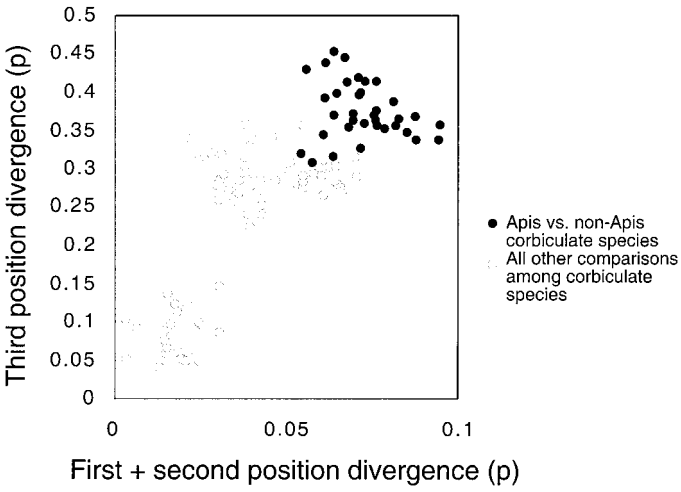

FIG. 8. Third position uncorrected divergences plotted against 1st + 2nd position uncorrected divergences for pairwise comparisons among taxa of corbiculate bees. Closed circles indicate comparisons between *Apis* and non-*Apis* corbiculate species.

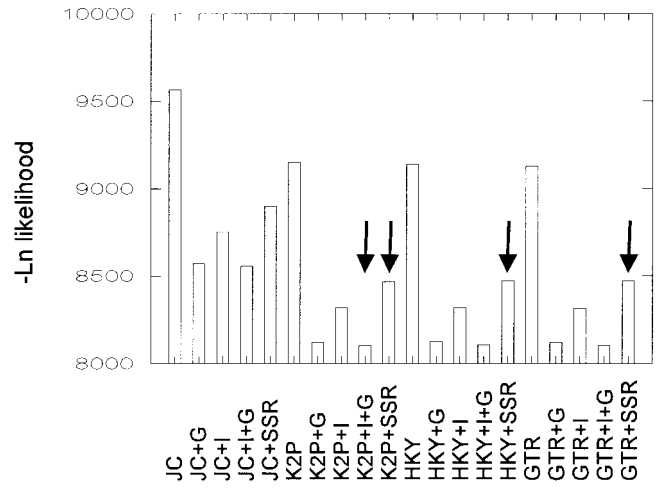

FIG. 9. −Ln likelihood scores based on 20 possible models of sequence evolution. The K2P + SSR, K2P + I + G, HKY + SSR, and GTR + SSR models were selected for branch swapping (indicated by arrows).

of sequence evolution used. Because neither the incorporation of empirical base frequencies (using the HKY model) nor the incorporation of a six-parameter rate matrix (using the GTR model) substantially improved the −Ln likelihood score when −Ln likelihoods were calculated on the equal weights parsimony trees, we initially selected the Kimura two-parameter (K2P) model as an appropriate model for branch swapping.

Using the K2P model with a transition/transversion ratio estimated from the data (ti/tv ratio = 2.43) and site-specific rates estimated from the original equal weights parsimony trees and refined through several rounds of tree search (relative rates: first, 0.517; second, 0.188; third, 2.30), we obtained one tree (Fig. 10). This same topology was obtained irrespective of which of the original eight equal weights parsimony trees were used to start. Additional ML analyses incorporating the HKY + SSR and the GTR + SSR models also resulted in this topology. Our ML analyses failed to support corbiculate monophyly. Indeed, *Centris*, *Ceratina*, and *Xylocopa* fall out *within* the corbiculate bees, and *Apis* is apparently sister to *Xylocopa*. As with the parsimony analysis, *Leiopodus*, *Ceratina*, and *Apis* are subtended by long branches. *Ceratina* is in a different but equally anomalous position. (Dramatic instability in the placement of *Ceratina* presumably results from the absence of 110 missing base pairs at the 5′ end of its sequence.) If one disregards placement of *Ceratina* and poorly supported nodes, the ML tree closely resembles the consensus of the most parsimonious trees. In addition, relationships among the four species of *Diadasia* were resolved and congruent with results based on other genes and more complete taxon sampling (Sipes and Wolf, 2001). As in the parsimony analysis, little bootstrap support exists below the level of tribe

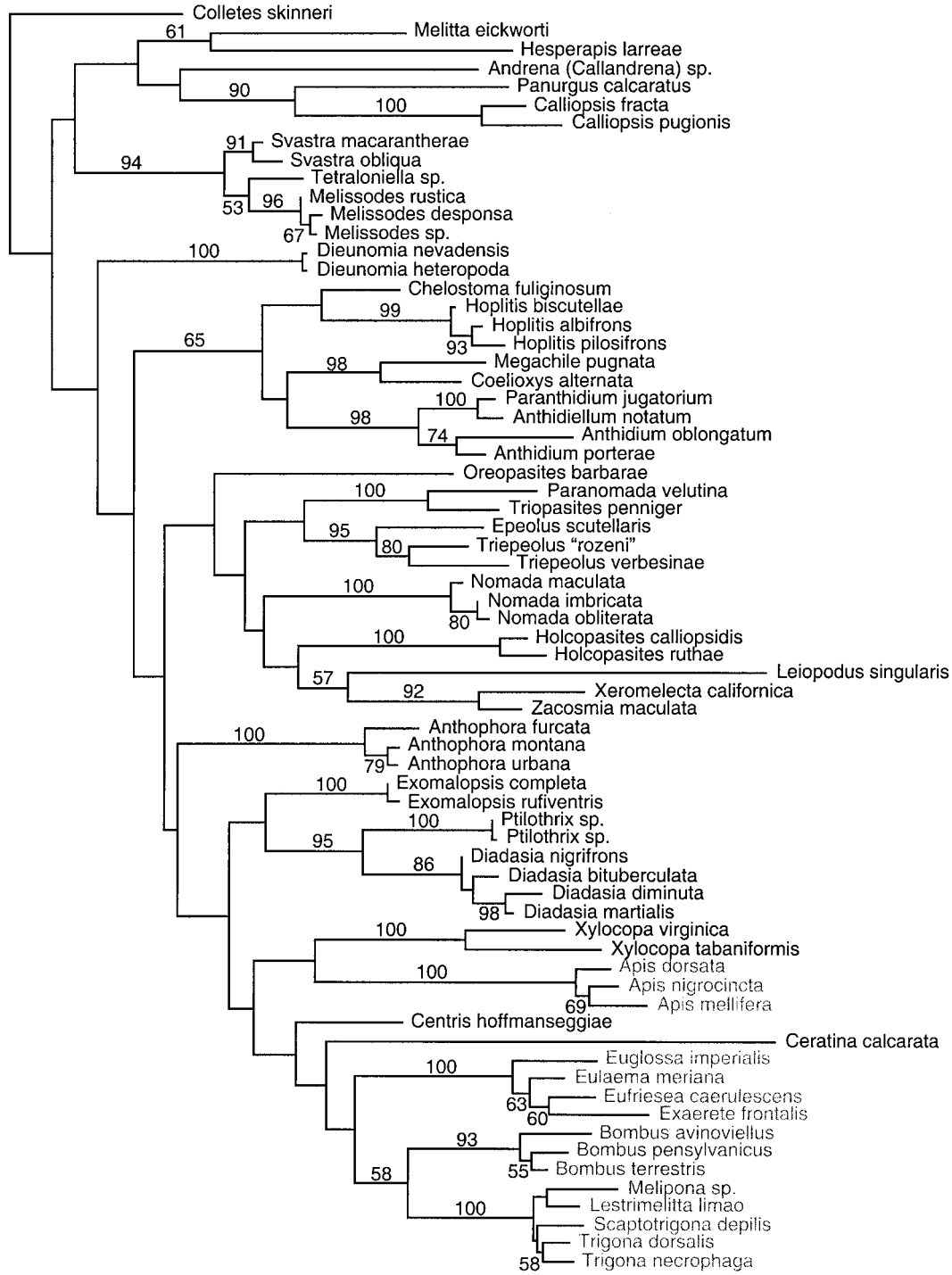

FIG. 10. ML tree based on the K2P + SSR model. Bootstrap values (based on 100 bootstrap replicates) are shown above nodes. Corbiculate bees are indicated in gray.

(Fig. 10), although support for Bombini + Meliponini is greater than 50%. Analyses incorporating more complex models of rate heterogeneity (K2P + G + I) resulted in slightly different tree topologies, but were identical to Fig. 10 for nodes supported by >50% bootstrap values.

Simultaneous Analysis of Morphological and Molecular Data

A simultaneous analysis of two data partitions, our LW *Rh* data set and a morphological and behavioral data set (Appendix), resulted in a consensus of most

parsimonious trees (of L = 1717 steps) congruent with M44 (Fig. 1a). The most parsimonious trees obtained from simultaneous analysis are only 4 steps (0.0024%) longer than the sum of the most parsimonious tree lengths obtained from separate analyses of each partition.

An analysis of the morphological and behavioral data partition alone also yielded M44 (Fig. 1a; tree length = 27 steps, CI = 0.815, RI = 0.969). Complete congruence between this topology and that obtained from simultaneous analysis suggests that incongruence introduced by combining our data partitions is due entirely to a slight increase in homoplasy in the LW *Rh* partition. Tree length for the morphological and behavioral data partition increases by 67% (18 steps) relative to that for the MP tree (L = 27 steps) when constrained to Mardulyn and Cameron's (1999) preferred molecular phylogeny (Fig. 1d). This dramatic increase in tree length is not surprising, since all morphological or behavioral characters are incongruent with this topology.

In contrast, our LW *Rh* molecular data partition is only very weakly incongruent with M44. Tree length for this molecular partition increases by only 0.0024% (4 steps) relative to the MP tree (L = 1686 steps) in simultaneous analysis (which supports M44) and by a mere 0.0012% (2 steps) when the advanced eusocial clade (Apini + Meliponini) is constrained to be monophyletic. This result suggests that nearly as many molecular characters are congruent with morphology-based tribal phylogenies of corbiculate bees (Figs. 1a and 1b) as with the preferred molecular phylogeny of Mardulyn and Cameron (1999; our Fig. 1d). A high degree of congruence among morphological and behavioral characters (RI = 0.953 for the data set of Schultz *et al.* (1999); RI = 0.969 for our data set) and the ambiguity of molecular data sets for corbiculates explain why morphological characters largely determined the outcome of simultaneous analysis in this and previous (Chavarría and Carpenter, 1994; Schultz *et al.*, 1999) studies.

DISCUSSION

Assessment of LW Rh for Resolving Relationships among Bee Tribes, Subfamilies, and Families

LW *Rh* sequence data are efficient at resolving lower-level (i.e., relatively recent) divergences among bees as demonstrated by substantial clade support (based on jackknife, bootstrap, and Bremer support values) for nearly all genera and tribes and for many relationships within genera and tribes that are consistent with morphological and behavioral data. On the other hand, our data set failed to unambiguously resolve most higher-level relationships, including those among tribes, subfamilies, and families (exceptions: Panurginae and Melittidae). Since higher-level relationships among L-T bees are also difficult to resolve using morphological data (as demonstrated by bootstrap values of less than 50% for most basal clades when we reanalyzed Roig-Alsina and Michener's (1993) data set of adult morphology and by topological instability when this data set was modified by Silveira (1993)), lack of basal resolution in our analyses of LW *Rh* sequences may reflect the inherent difficulty of recovering clades of Cretaceous age separated by short internodes. Alternatively, LW *Rh* may be subject to patterns of homoplasy that make it a poor data set for phylogenetic analysis at this level (see below).

Assessment of LW Rh for Resolving Relationships within the Corbiculate Bees

Relationships among the corbiculate bee tribes were unresolved based on the absence of nodal support in equal weights parsimony analysis. Weighted parsimony and ML had the effect of elevating levels of support for Bombini + Meliponini. However, corbiculate monophyly was only ambiguously supported in our equal weights analysis, and neither weighted parsimony nor maximum-likelihood (under any of the four models that we tested) recovered corbiculate monophyly (see also Mardulyn and Cameron, 1999; our Figs. 1f and 1g). The failure of LW *Rh* to recover corbiculate monophyly is a significant problem, given the fact that corbiculates are among the most strongly supported clades of bees (Roig-Alsina and Michener, 1993). In all cases it is *Apis* that falls outside the remaining corbiculates. What factors could explain the anomalous placement of *Apis* in our study? There are two (related) attributes of the *Apis* sequences included here that could produce highly anomalous results: (1) highly skewed third position base composition (Fig. 2) and (2) high sequence divergence when compared to other corbiculate bees (Fig. 8), resulting in an extremely long branch subtending the three *Apis* species (Figs. 6, 7, and 10). We suspect that together these attributes of the *Apis* sequences result in incorrect placement of *Apis* and possibly in incorrect rooting of the corbiculate clade.

Although our LW *Rh* data set alone failed to resolve corbiculate relationships, our combined analysis with morphological data was consistent with the traditional morphological phylogeny (referred to as M44 above; Fig. 1a). Previous studies based on combinations of morphological and molecular data (Chavarría and Carpenter, 1994; Schultz *et al.*, 1999) reached the same conclusion. Furthermore, detailed cladistic analyses of behavior (Noll, 2001) and of Baltic amber corbiculate fossils (Engel, 2000b, 2001) provide additional support for M44. In light of the substantial corroboration for M44, we view the hypothesis of dual origins of advanced eusociality (Winston and Michener, 1977; Kimsey, 1984 [Fig. 1c]; Cameron, 1991, 1993; Mardulyn

and Cameron, 1999; Koulianos *et al.*, 1999 [Figs. 1d and 1e]) in corbiculate bees as unsupported by the totality of the phylogenetic data.

For Bombini + Meliponini to be accepted as a well-corroborated hypothesis, the following criteria should be met by any future molecular studies: (1) substantial support for corbiculate monophyly (as measured by bootstrap and Bremer support values) should be demonstrated when an adequate sampling of tribes potentially most closely related to the corbiculates is included (see discussion of taxon sampling above), (2) the data should be shown to fit the hypothesis of Bombini + Meliponini significantly better than the alternative hypothesis (M44), and (3) some attempt must be made to reconcile the remarkable incongruity of the morphological and behavioral data with Bombini + Meliponini.

General Conclusions about the Phylogenetic Utility of LW Rh in Insects

Homoplasy in our LW *Rh* (rhodopsin) data set results in lack of nodal support at deeper levels (Figs. 3 and 5) and spurious results in our most parsimonious and maximum-likelihood trees (Figs. 3 and 10). This problem was refractory to correction by any of the models that we employed (including weighted parsimony and widely employed ML models). Interestingly, Chang and Campbell (2000) found strong support for incorrect nodes when analyzing a data set of 25 vertebrate rhodopsin sequences using parsimony, distance, and likelihood methods. They suggest that convergent evolution at hydrophobic amino acid sites, an evolutionary process not included in existing model-based phylogenetic approaches, may account in part for this problem. Future molecular evolutionary studies may allow for more specific and realistic modeling of rhodopsin sequence evolution. However, the utility of complex models (containing numerous parameters) for phylogenetic inference may be limited (Farris, 1999), and taxon-specific effects (such as the significantly biased third position base composition in *Apis*) cannot be accounted for in current ML models. In our view, progress in higher-level insect molecular systematics will be better achieved by generating data sets from genes demonstrably useful for recovering ancient divergences rather than by using complex models to extract limited phylogenetic signal from suboptimal data sets. At present, genes such as elongation factor 1α (EF-1α), wingless, phosphoenolpyruvate carboxykinase (PEPCK), and dopa-decarboxylase (DDC) (references listed above) appear more promising than does LW *Rh* for resolving Cretaceous-age divergences in insects, including those among corbiculate bees.

APPENDIX

Morphological and Behavioral Data Matrix (Modified from That of Schultz *et al.*, 1999)

	Characters																				
											1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**2**
Taxa	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**0**
Panurginae	0	?	?	0	0	0	0	0	0	?	0	0	?	0	0	1	0	0	0	0
other outgroups	0	?	?	0	0	0	0	0	0	?	0	0	?	0	*	0	0	0	0	0
Xylocopini: *Xylocopa*	0	?	?	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Centridini: *Centris*	0	?	?	0	0	0	0	0	0	?	0	0	?	0	1	0	0	0	0	0
Euglossini	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0
Bombini: *Bombus*	0	0	0	0	1	0	0	1	0	1	0	1	1	1	1	0	0	0	0	0
Meliponini	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1
Apini: *Apis*	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1

	2	**2**
	1	**2**
Panurginae	0	–
other outgroups	0	–
Xylocopini: *Xylocopa*	0	–
Centridini: *Centris*	0	–
Euglossini	0	–
Bombini: *Bombus*	1	0
Meliponini	1	1
Apini: *Apis*	1	1

Morphological and behavioral character descriptions.

1: Mandibular grooves. (0) Present. (1) Absent (*Apis*) or limited to the oblique groove of Michener and Fraser (1978; Meliponini). The mandibles of Apini differ from the mandibles of Meliponini, but are more similar to each other than either is to the mandibles of any other bees (see Michener and Fraser, 1978; note that both taxa have a large cap of the rutellum). The secondary mandibular ridges and associated grooves of Bombini are not primarily homologous with those of Euglossini (Michener and Fraser, 1978) and should not be considered to belong to the same character state "very well developed" unless a more precise homology can be described.

2: Hypopharyngeal plate. (0) Sensory lobes elongate. (1) Sensory lobes short, transverse.

3: Stipites. (0) Posteriorly produced into a dorsal flange overlapping cardines laterally in response. (1) Not produced posteriorly as a flange.

4: "Basigaleal bar." (0) Not differentiated. (1) Differentiated as a distinct process.

5: Postmentum. (0) Continuous. (1) Divided into "mentum" and lorum. The "mentum" of bees is not homologous with the mentum of generalized insects (Plant and Paulus, 1987).

6: Prosternum. (0) Not constricted, apophyseal pit present. (1) Constricted, apophyseal pit absent. A strongly constricted prosternum seems incompatible with a well-developed apophyseal pit. Therefore, we regard Schultz *et al.*'s (1999) characters 7 and 8 as attributes of a single complex character.

7: Prosternal setae. (0) Present. (1) Absent.

8: Basisternum. (0) Normally developed. (1) Enlarged.

9: Antero-lateral mesoscutal process. (0) With parascutal carina present. (1) With parascutal carina absent.

10: Metapleural ridge. (0) Extending to postero-lateral corner of mesopleuron. (1) Curved before postero-lateral corner of mesopleuron.

11: Metatibial spurs. (0) Present. (1) Absent.

12: Auricle. (0) Absent. (1) Present. The auricle and the penicillum (unique to Meliponini) are functional equivalents and are not known to cooccur (Prentice, 1991). Therefore they could be regarded as nonindependent derived states of the character "pollen-packing structure," which would be inapplicable in taxa lacking a corbicula. If coded in this way, absence of an auricle no longer provides evidence for excluding Meliponini from the group (Apini + Euglossini + Bombini).

13: Strigilis. (0) Without anterior velum. (1) With anterior velum.

14: Jugal lobe. (0) Present. (1) Absent (Bombini) or limited to a jugal comb of long, stiff bristles (Euglossini). The derived state could be partitioned into two states, but is coded as a single state here so that any bias in character coding will favor the alternative morphological phylogeny over M44.

15: Alar papillae/stigma size/body size. (0) Absent/large/small. (1) Present/small/large. Each of the several unrelated bee taxa with alar papillae (e.g., Caupolicanini, stenotritine Colletidae, Oxaeinae, Xylocopini, and Centridini) also have small stigmas and are large in size. The cooccurrence of these and other (e.g., reduced arolia) characters in these large bees likely reflects size scaling due to biophysical constraints (Danforth, 1989). The characters involved cannot be considered to provide strong independent evidence of phylogenetic relationships unless more precise homologies can be defined. Characters correlated with body size are especially problematic when assessing relationships among the corbulicate tribes because both generalized Euglossini (e.g., *Eulaema*) and *Bombus* queens are exceptionally large bees and because both the currently accepted (Centridini and other noncorbiculate Apinae) and the formerly accepted (Xylocopinae) sister groups to the corbiculates also include very large bees, but vary with respect to body size.

16: Gonobase. (0) Normally developed. (1) Reduced or absent.

17: SVII and SVIII of male. (0) Normally developed. (1) Reduced or absent.

18: Cuticle, bands of uneven sclerotization. (0) Absent. (1) Present.

19: Larval food. (0) Not highly supplemented with pharyngeal gland secretions. (1) Highly supplemented with pharyngeal gland secretions.

20: First recurrent vein (1r-m). (0) Longer, oblique, not angulate or moderately so. (1) Short and angulate.

21: Eusociality. (0) Absent. (1) Present. Characters 21 and 22 are logically dependent and are coded so as to avoid redundancy of information. These can be divided into multiple independent characters, increasing quantitative support for M44 (Noll, 2001), but the independence of some of these characters is questionable. Schultz *et al.* (1999) excluded the character "social behavior" from their analysis on the grounds that including this character might be construed as circular. We believe that all characters should be included in cladistic analysis, including those of special interest (Luckow and Bruneau, 1997; Wenzel, 1997; Zrzavy, 1997).

22: Degree of eusociality. (−) Inapplicable (noneusocial taxa). (0) Primitive. (1) Advanced.

ACKNOWLEDGMENTS

We thank M. S. Engel for help with references, access to unpublished manuscripts, and discussions regarding extant and fossil corbiculate bees. S. Sipes contributed two *Diadasia* sequences and M. G. Rightmyer generously donated the specimen of *Leipodus singularis*. K. C. Nixon provided a beta test version of Winclada. The following people provided comments on earlier versions of this paper: Karl Magnacca, Kelly Tilmon, Chung-Ping Lin, Curtis Ewing, Sedonia Sipes, and two anonymous reviewers. This project was supported by a National Science Foundation Research Grant in Systematic Biology (DEB-9508647) to B.N.D.

REFERENCES

Alexander, B. A., and Michener, C. D. (1995). Phylogenetic studies of the families of short-tongued bees (Hymenoptera: Apoidea). *Univ. Kansas Sci. Bull.* **55:** 377–424.

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

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

Bremer, K. (1994). Branch support and tree stability. *Cladistics* **6:** 369–372.

Brower, A. V. Z., and DeSalle, R. (1994). Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Ann. Entomol. Soc. Am.* **87:** 702–716.

Brower, A. V. Z., and 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 Mol. Biol.* **7:** 73–82.

Campbell, D. L., Brower, A. V. Z., and Pierce, N. E. (2000). Molecular evolution of the *wingless* gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera: Papilionoidea). *Mol. Biol. Evol.* **17:** 684–696.

Cameron, S. A. (1991). A new tribal phylogeny of the Apidae inferred from mitochondrial DNA sequences. *In* "Diversity in the genus *Apis*" (D. R. Smith, Ed.), pp. 71–87. Westview Press, Boulder, CO.

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

Chang, B. S. W., and Campbell, D. L. (2000). Bias in phylogenetic reconstruction of vertebrate rhodopsin sequences. *Mol. Biol. Evol.* **17:** 1220–1231.

Chang, B. S. W., Ayers, D., Smith, W. C., and Pierce, N. E. (1996). Cloning of the gene encoding honeybee long-wavelength rhodopsin: A new class of insect visual pigments. *Gene* **173:** 215–219.

Chavarría, G., and Carpenter, J. M. (1994). Total evidence and the evolution of highly social bees. *Cladistics* **10:** 229–258.

Danforth, B. N., and Ji, S. (2001). Australian *Lasioglossum* + *Homalictus* form a monophyletic group: Resolving the "Australian enigma." *Syst. Biol.*, **50:** 1–16.

Danforth, B. N., Sauquet, H., and Packer, L. (1999). Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae) based on parsimony and likelihood analyses of nuclear EF-1α sequence data. *Mol. Phylogenet. Evol.* **13:** 605–618.

Engel, M. S. (1999a). The first fossil *Euglossa* and phylogeny of the orchid bees (Hymenoptera: Apidae; Euglossini). *Am. Mus. Novit.* **3272:** 1–14.

Engel, M. S. (1999b). *Megachile glaesaria,* the first megachilid bee fossil from amber (Hymenoptera: Megachilidae). *Am. Mus. Novit.* **3296:** 1–11.

Engel, M. S. (1999c). The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; *Apis*). *J. Hymenop. Res.* **8:** 165–196.

Engel, M. S. (2000a). A new interpretation of the oldest fossil bee (Hymenoptera: Apidae). *Am. Mus. Novit.* **3296:** 1–11.

Engel, M. S. (2000b). Fossils and phylogeny: A paleontological perspective on social bee evolution. *In* "Anais do IV Encontro Sobre Abelhas" (M. M. G. and K. Hartfelder, Eds.), pp. 217–224. Univ. de São Paulo, Ribeirão Preto.

Engel, M. S. (2001). Monograph of the Baltic amber bees and evolution of the Apoidea (Hymenoptera). *Bull. Am. Mus. Nat. Hist.*, **259:** 1–192.

Fang, Q. Q., Cho, S., Regier, J. C., Mitter, C., Matthews, M., Poole, R. W., Friedlander, T. P., and Zhao, S. (1997). A new nuclear gene for insect phylogenetics: Dopa decarboxylase is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). *Syst. Biol.* **46:** 269–283.

Farris, J. S. (1999). Likelihood and inconsistency. *Cladistics* **15:** 199–204.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., and Kluge, A. G. (1996). Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12:** 99–124.

Farris, J. S., Källersjö, M., Kluge, A. G., and Bult, C. (1994). Testing significance of incongruence. *Cladistics* **10:** 315–319.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Frati, F., Simon, C., Sullivan, J., and Swofford, D. L. (1997). Evolution of the mitochondrial cytochrome oxidase II gene in Collembola. *J. Mol. Evol.* **44:** 145–158.

Friedlander, T. P., Horst, K. R., Regier, J. C., Mitter, C., Peigler, R. S., and Fang, Q. Q. (1998). Two nuclear genes yield concordant relationships within Attacini (Lepidoptera: Saturniidae). *Mol. Phylogenet. Evol.* **9:** 131–140.

Friedlander, T. P., Regier, J. C., and Mitter, C. (1994). Phylogenetic information content of five nuclear gene sequences in animals: Initial assessment of character sets from concordance and divergence studies. *Syst. Biol.* **43:** 511–525.

Friedlander, T. P., Regier, J. C., Mitter, C., and Wagner, D. L. (1996). A nuclear gene for higher level phylogenetics: Phosphoenolpyruvate carboxykinase tracks Mesozoic-age divergences within Lepidoptera (Insecta). *Mol. Biol. Evol.* **13:** 594–604.

Gatesy, J., O'Grady, P., and Baker, R. H. (1999). Corroboration among data sets in simultaneous analysis: Hidden support for phylogenetic relationships among higher level artiodactyl taxa. **15:** 271–313.

Goloboff, P. (1994). NONA: A tree searching program. Program and documentation. Published by the author, Tücuman, Argentina.

Grimaldi, D. A. (1999). The co-radiations of insects and angiosperms in the Cretaceous. *Ann. Missouri Bot. Gard.* **86:** 373–406.

Huelsenbeck, J. P., and Crandall, K. A. (1997). Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* **28:** 437–466.

Kimsey, L. S. (1982). Systematics of bees of the genus *Eufriesea* (Hymenoptera: Apidae). *Univ. Calif. Publ. Entomol.* **95:** 1–125.

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

Kimsey, L. S. (1987). Generic relationships within the Euglossini (Hymenoptera: Apidae). *Syst. Entomol.* **12:** 63–72.

Kluge, A. G. (1998). Total evidence or taxonomic congruence: Cladistics or consensus classification. *Cladistics* **14:** 151–158.

Koulianos, S., Schmid-Hempel, R., Roubik, D. W., and Schmid-Hempel, P. (1999). Phylogenetic relationships within the corbiculate Apinae (Hymenoptera) and the evolution of eusociality. *J. Evol. Biol.* **12:** 380–384.

Leijs, R. (2000). "Evolution of the Large Carpenter Bees (*Xylocopa*): Molecular Phylogenies, Historical Biogeography, Mating Strategies and Sociality." Ph.D. thesis, Flinders University of South Australia, Bedford Park, Australia.

Luckow, M., and Bruneau, A. (1997). Circularity and independence in phylogenetic tests of ecological hypotheses. *Cladistics* **13:** 145–151.

Maa, T. (1953). An inquiry into the systematics of the tribus Apidini or honeybees. *Treubia* **21:** 525–640.

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

McGinley, R. J. (1980). Glossal morphology of the Colletidae and recognition of the Stenotritidae at the family level (Hymenoptera: Apoidea). *J. Kansas Entomol. Soc.* **53:** 539–552.

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 Bees: A Comparative Study," Harvard Univ. Press, Cambridge, MA.

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

Michener, C. D. (2000). "The Bees of the World," Johns Hopkins Press, Baltimore.

Michener, C. D., and Brooks, R. W. (1984). Comparative study of the glossae of bees (Apoidea). *Contr. Am. Entomol. Inst.* **22:** 1–73.

Michener, C. D., and Fraser, A. (1978). A comparative anatomical study of the mandibular structure in bees. *Univ. Kansas Sci. Bull.* **51:** 463–482.

Mitchell, A., Cho, S., Regier, J. C., Mitter, C., Poole, R. W., and Matthews, M. (1997). Phylogenetic utility of elongation factor-1α in Noctuoidea (Insecta: Lepidoptera): The limits of synonymous substitution. *Mol. Biol. Evol.* **14:** 381–390.

Nixon, K. C. (1999). Winclada (beta) ver. 0.99, Published by the author, Trumansburg, NY.

Nixon, K. C., and Carpenter, J. M. (1996). On simultaneous analysis. *Cladistics* **12:** 221–241.

Noll, F. B. (2001). Behavioral phylogeny of corbiculate Apidae (Hymenoptera: Apinae), with special reference to social behavior. *Cladistics*, in press.

Palumbi, S. R. (1996). The polymerase chain reaction. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), 2nd ed., pp. 205–247. Sinauer, Sunderland, MA.

Plant, J. D., and Paulus, H. F. (1987). Comparative morphology of the postmentum of bees (Hymenoptera: Apoidea) with special remarks of the evolution of the lorum. *Sonderdr. Zool. Systematik Evolutionsforsch.* **25:** 81–103.

Prentice, M. (1991). Morphological analysis of the tribes of Apidae. *In* "Diversity in the genus *Apis*" (D. R. Smith, Ed.), pp. 51–69. Westview Press, Boulder, CO.

Regier, J. C., Fang, Q. Q., Mitter, C., Peigler, R. S., Friedlander, T. P., and Solis, M. A. (1998). Evolution and phylogenetic utility of the *period* gene in Lepidoptera. *Mol. Biol. Evol.* **15:** 1172–1182.

Roig-Alsina, A., and Michener, C. D. (1993). Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). *Univ. Kansas Sci. Bull.* **55:** 123–162.

Rozen, J. G., Jr. (1996). Phylogenetic analysis of the cleptoparasitic bees belonging to the Nomadinae based on mature larvae (Apoidea: Apidae). *Am. Mus. Novit.* **3180:** 1–39.

Ruz, L. (1987). "Classification and Phylogenetic Relationships of the Panurgine Bees (Hymenoptera - Andrenidae)." Ph.D. thesis, Univ. of Kansas, Lawrence.

Sakagami, S. F., and Maeta, Y. (1984). Mutilfemale nests and rudimentary castes in the normally solitary bee *Certina japonica* (Hymenoptera: Xylocopinae). *J. Kansas Entomol. Soc.* **57:** 639–656.

Schultz, T. R., Engel, M. S., and Prentice, M. (1999). Resolving conflict between morphological and molecular evidence for the origin of eusociality in the corbiculate bees (Hymenoptera: Apidae): A hypothesis-testing approach. *Univ. Kansas Nat. Hist. Mus. Spec. Publ.* **24:** 110–123.

Sheppard, W. S., and McPheron, B. A. (1991). Ribosomal DNA diversity in Apidae. *In* "Diversity in the Genus *Apis*" (D. R. Smith, Ed.), pp. 89–102. Westview Press, Boulder, CO.

Shimada, T., and Kurimoto, Y. (1995). Phylogenetic relationship of slikmoths inferred from sequence data of the arylphorin gene. *Mol. Phylogenet. Evol.* **4:** 223–234.

Silveira, F. A. (1993). Phylogenetic relationships of the Exomalopsini and Ancylini (Hymenoptera: Apidae). *Univ. Kansas Sci. Bull.* **55:** 163–173.

Sipes, S. D., and Wolf, P. G. (2001). Phylogenetic relationships within *Diadasia*, a group of specialist bees. *Mol. Phylogenet. Evol.* **19:** 144–156.

Soto-Adames, F. N., Robertson, H. M., and Berlocher, S. H. (1994). Phylogenetic utility of partial DNA sequences of G_6pdh at different taxonomic levels in Hexapoda with emphasis on Diptera. *Ann. Entomol. Soc. Am.* **87:** 723–736.

Sullivan, J., and Swofford, D. L. (1997). Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. *J. Mammal. Evol.* **4:** 77–86.

Swofford, D. L. (1999). "PAUP Version 4.0b2a: Phylogenetic Analysis Using Parsimony." Sinauer, Sunderland, MA.

Tatarenkov, A., Kwiatovski, J., Skarecky, D., Barrio, E., and Ayala, F. J. (1999). On the evolution of *Dopa decarboxylase* (Ddc) and *Drosophila* systematics. *J. Mol. Evol.* **48:** 445–462.

Townson, S. M., Chang, B. S., Salcedo, E., Chadwell, L. V., Pierce, N. E., and Britt, S. G. (1998). Honeybee blue- and ultraviolet-sensitive opsins: Cloning, heterologous expression in *Drosophila*, and physiological characterization. *J. Neurosci.* **18:** 2412–2422.

Wenzel, J. W. (1997). When is a phylogenetic test good enough? *In* "The Origin of Biodiversity in Insects: Phylogenetic Tests of Evolutionary Scenarios" (P. Grandcolas, Ed.), pp. 31–45. Vol. 173, Mémoir Museum National d'Histoire Naturelle, Paris.

Williams, P. H. (1998). An annotated checklist of bumble bees with an analysis of patterns of description (Hymenoptera: Apidae, Bombini). *Bull. Nat. Hist. Mus. Lond. (Entomol.)* **67:** 79–152.

Winston, M. L., and Michener, C. D. (1977). Dual origin of highly social behavior among bees. *Proc. Natl. Acad. Sci. USA* **74:** 1135–1137.

Zrzavy, J. (1997). Phylogenetics and ecology: All characters should be included in the cladistic analysis. *Oikos* **80:** 186–192.