Mitochondrial DNA Differentiation Between Two Cryptic *Halictus* (Hymenoptera: Halictidae) Species

BRYAN N. DANFORTH,[1] PAULA LEVIN MITCHELL,[2] AND LAURENCE PACKER[3]

Ann. Entomol. Soc. Am. 91(4): 387-391 (1998)

ABSTRACT *Halictus ligatus* Say and its cryptic sibling species *H. poeyi* Lepeletier were sampled from an area of sympatry and typed for mitochondrial DNA sequence differentiation, as was 1 individual of each species from allopatric populations. Two other species, *H. rubicundus* (Christ) and *H. farinosus* Smith, were included to root the tree. Phylogenetic analysis of the aligned sequences showed that *H. ligatus* and *H. poeyi* are sister species and differ from each other at 33 of the 797 aligned nucleotide positions. Sequence divergence within species was <0.6%, whereas that between species exceeded 4%. These data support an earlier report that *H. ligatus* and *H. poeyi* represent distinct species (based on nuclear allozyme markers) and are sufficient to cast doubt on a recent glaciation-induced speciation event as the cause of the differentiation of these 2 taxa.

KEY WORDS *Halictus*, sibling species, sympatry, mDNA divergence, speciation

WITH A RANGE extending from southern Canada through the United States, Mexico, and Central America to Colombia and Venezuela, *Halictus ligatus* Say has been considered to be the most widespread halictine species in the New World (Michener and Bennett 1977). However, recent electrophoretic analysis (Carman and Packer 1997) has demonstrated that 2 species are involved: one tentatively referred to as *H. poeyi* Lepeletier, is found in the southeastern United States, whereas true *H. ligatus* is found to the north and west, the 2 being sympatric throughout the Piedmont region southeast of the Appalachian Mountains. To which species the Mexican, Caribbean, and Central and South American populations belong is not yet known.

Carman and Packer (1997) documented that the 2 species are separated by 7 fixed electrophoretic differences out of 34 allozyme loci surveyed. They suggested that *H. poeyi* may have become separated from *H. ligatus* in the Ocala highlands of Florida during a past glaciation, as has been suggested for other taxa (Ellsworth et al. 1994).

The purpose of this article was to investigate the level of mitochondrial DNA (mDNA) differentiation between the 2 species. Consequently, samples were obtained from an area of sympatry. Additionally, 1 individual of each species from as far away from this area as possible was sequenced.

Materials and Methods

Most bees were obtained from Rock Hill, SC, in May 1996, collected as part of a survey of the phenological and behavioral differences between the species in sympatry (L.P. and P.L.M., unpublished data). All individuals used were female and all Rock Hill bees were overwintered foundresses. The thoracic musculature of all the bees collected at Rock Hill was first typed by allozyme electrophoresis for 3 loci that unambiguously discriminate between the species in sympatry: fumarate hydratase, glucose phosphate isomerase, and esterase (as described in Carman and Packer [1997]). Additional sequences were obtained from 1 *H. poeyi* sampled from Marathon, FL, and 1 *H. ligatus* from Ithaca, NY. Although these last 2 individuals were not surveyed electrophoretically, both were obtained from areas where large samples of bees have indicated that only 1 species is present (Carman and Packer 1997; unpublished data). These 2 individuals were chosen from sites as far away from the area of sympatry as possible to obtain a preliminary indication of the level of variation that might exist within each species.

Heads were separated and stored in 95% EtOH. Each head from Rock Hill was assigned a letter (A–I) identifying its species affiliation and then sent to Danforth for DNA sequencing. Danforth was unaware of the species affiliation of all bees from Rock Hill until long after the sequencing had been completed.

DNA was extracted by standard protocols. Specimens were ground in individual 1.5-ml Eppendorf tubes in the presence of 2× CTAB extraction buffer and 100 μg of proteinase K. Tubes were incubated for 2 h at 55°C, extracted with chloroform-isoamylalcohol, digested for 30 min in the presence of 10 μg RNase, and then extracted again with phenol-chloroform-

[1] Department of Entomology, Cornell University, Ithaca, NY 14853.
[2] Department of Biology, Winthrop University, Rock Hill, SC 29733.
[3] To whom correspondence should be addressed: Department of Biology and Faculty of Environmental Studies, York University, North York, ON, M3J 1P3 Canada.

0013-8746/98/0387-0391$02.00/0

Table 1. Nucleotide differences between and within *H. ligatus* and *H. poeyi*

	Variable Positions
	* ********* ******** ********** ****
	00000000122222233333333344444444556667 7
	00111278711579901233778901123899679585 5
	37235787136864535839257624735368767478 9
H. poeyi (FL)	TCCCTTAACATTCCCCCCCCACTACACTTCCTAATCCGT
H. ligatus (NY)	CTTTAATTTCCTTTTTTTTTATCTTTTAATTAATCTAAT
Rock Hill A	TCCCTTAACATTCCCCCCCCACTACACTTCCTAATCCAT
Rock Hill B	TCCCTTAACATTCCCCCCCCGCTACACTTCCTGATCCAT
Rock Hill C	CCTTAATTTCCTTTTTTTTTATCTTTTAATTAATCTAAC
Rock Hill D	TCCCTTAACATTCCCCCCCCGCTACACTTCCTGATCCAT
Rock Hill E	CTTTAATTTCCTTTTTTTTTATCTTTTAATTAATCTAAT
Rock Hill F	CCTTAATTTCCTTTTTTTTTATCTTTTAATTAATCTAAT
Rock Hill G	CCTTAACTTCCCTTTTTTTTATCTTTTAATTAATCTAAC
Rock Hill H	TCCCTTAACATTCCCCCCCCGCTACACTTCCTGATCCAT
Rock Hill I	TCCCTTAACATTCCCCCCCCCACTACACTTCCTAATCCAT
Rock Hill J	CCTTAATTTCCTTTTTTTTTATCTTTTAATTAATCTAAC

Differences between *H. ligatus* and *H. poeyi* are indicated with an asterisk. Other changes represent intraspecific polymorphism.

isoamylalcohol and chloroform-isoamylalcohol, in that order. The DNA was precipitated with 2.5 volumes of ice cold EtOH and 1/10 volume of 3 M sodium acetate, washed once in 80% EtOH, and resuspended in 50 μl Tris-EDTA pH 7.6 buffer.

We polymerase chain reaction (PCR)-amplified a specific region of the mitochondrial cytochrome oxidase I gene (COI) by using 2 primers: C1-J-2183 (alias Jerry; Simon et al. 1994) and TL2-N-3014 (alias Pat; Simon et al. 1994). This primer pair produces a single bright band in all halictid bees tested corresponding to positions 2,481–3,380 in the published *Apis mellifera* L. mitochondrial genome (Crozier and Crozier 1993). After amplification we gel-purified the PCR products by using low-melting-point agarose and the Promega Wizard PCR preps DNA purification system (Promega, Madison, WI).

Gel-purified PCR products were sequenced in both directions by using the original PCR primers (C1-J-2183 and TL2-N-3014) either by automated sequencing on an Applied Biosystems 373A sequencer or by manual sequencing using the Amersham Thermo Sequenase Cycle Sequencing Kit (Amersham, Cleveland, OH) and standard protocols. Both methods gave identical results for broadly overlapping regions.

Sequences were aligned using MegAlign in the Lasergene software package (DNASTAR, Madison, WI). The total aligned sequence included 797 nucleotides (from position 2,517 to position 3,314 in the published *Apis* sequence; Crozier and Crozier [1993]) and spanned 265 codons plus 2 bases at the 3′ end.

Analyses of sequence composition, sequence divergence, and phylogenetic relationships were performed using a test version of PAUP*4 (PAUP version 4.0.0d55; D. Swofford [1996], see Swofford [1993] for details on earlier versions of the program) for the Power MacIntosh. All analyses were performed using the exhaustive search option and trees were rooted either by midpoint rooting or with 2 other North American species of *Halictus* selected as outgroups: *H. rubicundus* (Christ), *H. farinosus* Smith (Michener 1978, Pesenko 1985).

Voucher specimens are retained at York University; additional vouchers will be housed at Cornell and the Royal Ontario Museum. We submitted 6 sequences from this study to GenBank (accession numbers AF045372–AF045377). These included sequences for the 2 outgroups (*H. farinosus* and *H. rubicundus*), 2 sequences for *Halictus ligatus* (1 from Ithaca, NY, and 1 from Rock Hill, SC), and 2 sequences for *H. poeyi* (1 from Marathon, FL, and 1 from Rock Hill, SC). Because the Rock Hill *H. ligatus* and *H. poeyi* each showed some intraspecific variation, we submitted consensus sequences with polymorphic sites indicated as variable. Interested readers will be able to reconstruct the original sequence data from these GenBank sequences and from Table 1.

Results

Nucleotide Composition. Overall, the sequences show strong A/T bias with 74.5% of the sites either A or T (Fig. 1). This is characteristic of insect mitochondrial genomes (Crozier and Crozier 1993, Simon et al. 1994). First and 2nd positions are not obviously dif-

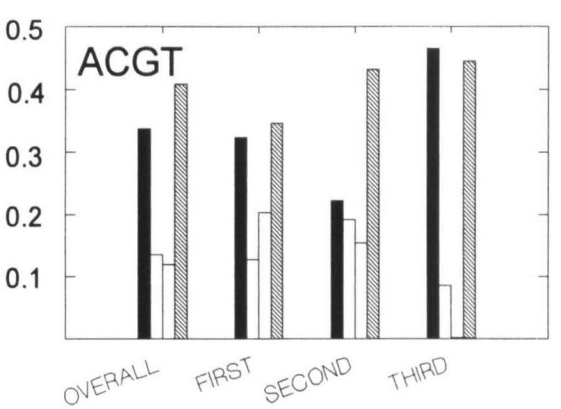

Fig. 1. Base composition of the aligned sequences for all 797 nucleotide positions (overall) and for each site within codons.

Table 2. Sequence divergence among species and populations of *H. farinosus*, *H. rubicundus*, *H. ligatus*, and *H. poeyi*

	far	*rub*	*lig* NY	*lig* RH	*poe* FL	*poe* RH
H. farinosus	—	2.8	5.6	5.6	6.0	5.6
H. rubicundus	10.7	—	6.0	6.0	6.5	6.0
H. ligatus (NY)	12.7	12.9	—	0.0	0.4	0.0
H. ligatus (RH)	12.5	12.6	0.0	—	0.4	0.0
H. poeyi (FL)	12.9	13.5	4.5	4.3	—	0.4
H. poeyi (RH)	12.5	13.2	4.4	4.2	0.1	—

Values above the diagonal represent uncorrected amino acid divergences; values below the diagonal represent uncorrected nucleotide divergences. RH, Rock Hill locality of *H. ligatus* and *H. poeyi*.

ferent from the overall nucleotide composition but 3rd positions show very strong A/T bias, 91.2% of the nucleotides being either an A or a T.

Sequence Divergence. We aligned the 14 *Halictus* sequences by using MegAlign. Alignments were unambiguous and no insertion or deletion mutations were detected. The uncorrected, pairwise nucleotide divergences among all specimens ranged from 0.0 to 13.5% (*H. poeyi* versus *H. rubicundus;* Table 2).

Among the *ligatus/poeyi* samples, the uncorrected pairwise divergences fell into 2 categories: values <0.01 (1% divergence) and values >0.04 (4% divergence) (Fig. 2). Low values represent intraspecific comparisons and high values correspond to interspecific comparisons (as determined by allozyme electrophoresis). Comparing the *ligatus/poeyi* sequences we detected a total of 39 variable sites (Table 1).

Phylogenetic Analysis. We analyzed the aligned sequences phylogenetically to distinguish between nucleotide sites that differentiate the 2 species and nucleotide sites that vary within species. By using the exhaustive search option in PAUP*4, we obtained 1 tree based on parsimony analysis with a length of 40 steps and a consistency index *(ci)* of 1.0 (Fig. 3A). The analysis yielded 33 nucleotide differences between species and 3–4 nucleotide differences within species (Table 2). All the within-species polymorphisms were transitions (3 G/A and 4 T/C). One position (758 A–G) was autapomorphic for the specimen of *H. poeyi* from Florida. Most interspecific differences (21 of 33 = 64%) were transitions and all of these changes were T/C polymorphisms (Table 3). The remaining 12

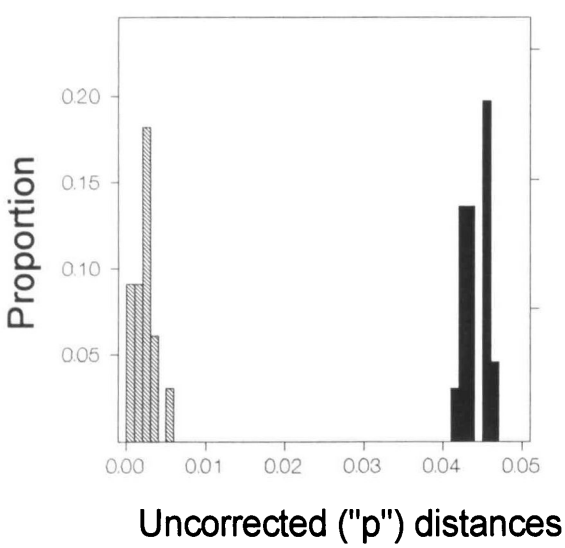

Fig. 2. Uncorrected pairwise sequence divergences among the 12 specimens of *H. ligatus/poeyi*. Shaded bars represent intraspecific comparisons and black bars represent interspecific comparisons.

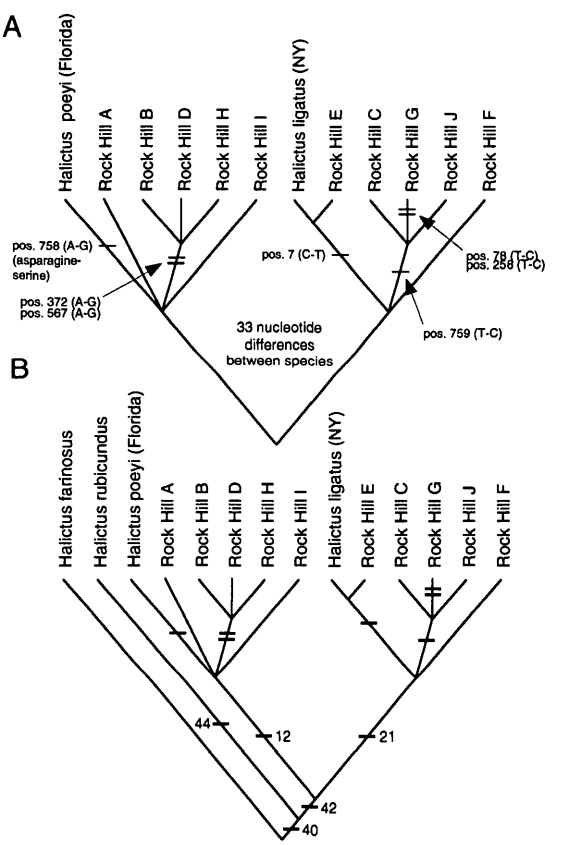

Fig. 3. (A) Phylogenetic tree of the 12 COI sequences showing characters mapped on. This tree was rooted by midpoint rooting so should be viewed as an unrooted network. (B) Phylogenetic tree of the 14 COI sequences with *H. rubicundus* and *H. farinosus* selected as outgroups for the analysis. Numbers along nodes indicate the number of character state changes.

Table 3. Nucleotide differences between *H. ligatus* and *H. poeyi* (33 total differences)

Position	Change	Codon position	Change type (tn/tv)
3	T/C	3	tn
12	T/C	3	tn
13	T/C	1	tn
15	T/A	3	tv
27	T/A	3	tv
78*	T/A	3	tv
87	T/A	3	tv
171	T/C	3	tn
213	A/C	3	tv
216	T/C	3	tn
276	T/C	3	tn
294	T/C	3	tn
295	T/C	1	tn
303	T/C	3	tn
315	T/C	3	tn
328	T/C	1	tn
333	T/C	3	tn
339	T/C	3	tn
375	T/C	3	tn
387	T/C	3	tn
396	A/T	3	tv
402	T/C	3	tn
414	T/A	3	tv
417	T/C	3	tn
423	T/A	3	tv
435	T/A	3	tv
483	T/C	3	tn
496	T/C	1	tn
498	T/A	3	tv
576	T/A	3	tv
597	T/C	3	tn
654	T/C	3	tn
687	A/C	3	tv

Also shows T/C transition within *H. ligatus* (Fig. 3A).

interspecific differences were transversions involving either T/A (10) or A/C (2) changes. Twenty-nine of the 33 interspecific changes were at 3rd positions (Table 3). Most of the nucleotide changes were synonymous changes, but 2 gave rise to altered amino acid sequence when translated: position 758 (A–G) in *Halictus poeyi* from Florida (Fig. 3A) and position 687 (A–C) between *H. ligatus* and *H. poeyi*.

In a 2nd analysis with 2 *Halictus* species included as outgroups we obtained a single tree (Fig. 3B) with *ci* = 0.95, retention index *(ri)* = 0.96, and a length of 166 steps. In this tree the 33 differences between *H. ligatus* and *H. poeyi* were resolved into 12 autapomorphies of *H. poeyi* and 21 autapomorphies of *H. ligatus*.

Discussion

The nucleotide composition of the COI region analyzed in our study is not noticeably different from that of other bee sequences. Insect mitochondrial genomes are universally A/T rich (Simon et al. 1994; Crozier and Crozier 1992, 1993) and the most extreme A/T biases in protein coding genes occur in 3rd positions. Arias and Sheppard (1996), for example, reported that 91.1% of 3rd position sites contain an A or T in *A. mellifera*. As with most studies of insect mitochondrial DNA, we observed a biased transition:transversion ratio (2:1 in our study). Overall, our results are in close agreement with previous studies of mDNA (DeSalle 1992).

Our mDNA differences mirror those found in allozymes (Carman and Packer 1997) and support species status for these 2 *Halictus* species. Sequence divergence in *ligatus/poeyi* is bimodal with pairwise divergences of either <1% (intraspecific variation) or >4% (interspecific variation). Vogler et al. (1993) reviewed data on mitochondrial sequence divergence among insects within the same population, among populations of the same species, and among distinct species. The average interpopulation sequence divergence did not exceed 4%, and the interspecific divergence (distantly related species as well as sibling species) generally exceeded 5% (in figure 1 Vogler et al. [1995]). Arias and Sheppard (1996) reported a maximal sequence divergence of 2.13% among geographically separated subspecies of honeybees based on mitochondrial NADH dehydrogenase subunit 2. This is consistent with the view that *H. ligatus* and *H. poeyi* are indeed distinct species, more divergent from each other than are widely separated populations of *A. mellifera* in Europe and Africa.

Mitochondrial protein coding genes are estimated to diverge at roughly 2%/million years, at least over divergences of <8%, where saturation at synonymous sites was detected (DeSalle et al. 1987). Based on this imperfect rate estimate, these 2 *Halictus* species may have been separated for >2 million years. This would seem to be inconsistent with 1 of Carman and Packer's (1997) suggestions, that the 2 species may have separated during a recent glaciation event that occurred no more than 1 million years ago. However, given the inaccuracies inherent in any clock-based model of sequence divergence, our data are not inconsistent with differentiation occurring during an earlier ice age in the Quaternary. An alternative explanation of the distribution of these 2 species in the United States has *H. poeyi* becoming differentiated in Mexico or Central America and dispersing to the southeastern United States either along the Gulf Coast or via the Caribbean from the Yucatan peninsula. Testing this requires analysis of bees from west of Alabama, the Caribbean, and the neotropics.

Acknowledgments

We are very grateful to Shuqing Ji for providing technical assistance in this project and Luke Ballard for reading sequencing gels. Two anonymous reviewers provided helpful comments on earlier versions of the manuscript. This project was supported by a National Science Foundation Research Grant in Systematic Biology (DEB-9508647) to B.N.D. (Cornell) and an NSERC Research Grant to L.P. (York).

References Cited

Arias, M. C., and W. S. Sheppard. 1996. Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) in-

ferred from mitochondrial DNA sequence. Mol. Phylogenet. Evol. 5: 557–566.

Carman, G. M., and L. Packer. 1997. A cryptic species allied to *Halictus ligatus* Say (Hymenoptera: Halictidae) detected by allozyme electrophoresis. J. Kans. Entomol. Soc. 69: 168–176.

Cornuet, J. M., L. Garnery, and M. Solignac. 1991. Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. Genetics 1128: 393–403.

Crozier, R. H., and Y. C. Crozier. 1992. The cytochrome b and ATPase genes of honeybee mitochondrial DNA. Mol. Biol. Evol. 9: 474–482.

1993. The mitochondrial genome of the honeybee *Apis mellifera:* complete sequence and genome organization. Genetics 133: 97–117.

DeSalle, R. 1992. The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. Mol. Phylogenet. Evol. 1: 31–40.

DeSalle, R., T. Freedman, E. M. Prager, and A. C. Wilson. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawiian *Drosophila.* J. Mol. Evol. 26: 157–164.

Ellsworth, D. L., R. L. Honeycutt, N. J. Silvy, J. W. Bickham, and W. D. Klimstra. 1994. Historical biogeography and contemporary patterns of mitochondrial DNA variation in White-Tailed Deer from the Southeastern United States. Evolution 48: 122–136.

Michener, C. D. 1978. The classification of halictine bees: tribes and old world nonparasitic genera with strong venation. Univ. Kans. Sci. Bull. 51: 501–538.

Michener, C. D., and F. D. Bennett. 1977. Geographical variation in nesting biology and social organization of *Halictus ligatus.* Univ. Kans. Sci. Bull. 51: 233–260.

Pesenko, Y. A. 1985. A subgeneric classification of bees of the genus *Halictus* Latreille, sensu stricto (Hymenoptera: Halictidae). Entomol. Rev. 63: 1–20.

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.

Swofford, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign, IL.

1996. PAUP: phylogenetic analysis using parsimony, version 4.0.0d55, beta-test version. Smithsonian Institution, Washington, DC.

Vogler, A. P., and R. Desalle. 1993. Phylogeographic patterns in coastal North American tiger beetles (*Cicindela dorsalis* Say) inferred from mitochondrial DNA sequences. Evolution 47: 1192–1202.

1994. Diagnosing units of conservation management. Conserv. Biol. 8: 354–363.

Vogler, A. P., R. DeSalle, T. Assman, C. B. Knisley, and T. D. Schultz. 1993. Molecular population genetics of the endangered tiger beetle, *Cicindela dorsalis* (Coleoptera: Cicindellidae). Ann. Entomol. Soc. Am. 86: 142–152.

Received for publication 27 August 1997; accepted 18 February 1998.