*Zootaxa* 1713: 27–38 (2008)
www.mapress.com/zootaxa/


ISSN 1175-5326 (print edition)
**ZOOTAXA**
ISSN 1175-5334 (online edition)

# DNA barcoding of new world cicada killers (Hymenoptera: Crabronidae)

JON M. HASTINGS[1], PATRICK J. SCHULTHEIS[1], MAGGIE WHITSON[1], CHARLES W. HOLLIDAY[2], JOSEPH R. COELHO[3] & ANGELA M. MENDELL[1]

[1] *Department of Biological Sciences, Northern Kentucky University, Highland Heights, KY 41099, USA. E-mail: hastings@nku.edu; schultheisp@nku.edu; whitsonma@nku.edu; Angela.Mendell@CCHMC.org*
[2] *Department of Biology, Lafayette College, Easton, PA 18042, USA. E-mail: hollidac@lafayette.edu*
[3] *Institute for Franciscan Environmental Studies, Biology Program, Quincy University, 62301. E-mail: coelhjo@quincy.edu*

## Abstract

We are engaged in a comprehensive study of cicada killers (*Sphecius* spp.), including their behavioral ecology. At one location, we observed interactions among three putative species of *Sphecius*, and used DNA barcoding to help clarify relationships among them. For this, we sequenced a fragment of the mitochondrial cytochrome *c* oxidase subunit I gene. During our study, a new taxonomic key of New World cicada killers, based on morphology, was published, and we expanded the barcoding project to test the congruence between barcodes and this key. In general, barcoding evidence supports morphological distinctions among species; sequence divergences between individuals of different species were within the range expected for congeneric Hymenoptera. However, two conflicts between barcoding and morphological evidence were noted. 1) Haplotypes of *Sphecius grandis* Say fall into two highly divergent clades, suggesting they are cryptic species. 2) Two clades of *S. convallis* Patton were found, and the sequences of one clade are virtually identical to those of *S. speciosus* Drury, suggesting that this clade of *S. convallis* is conspecific with *S. speciosus*. Alternative explanations are possible for this result, including hybridization and introgression between the two species. We conclude that our DNA barcoding evidence should be interpreted with caution, but that it has generated interesting questions we hope to resolve with field research coupled with analysis of suitable nuclear gene sequences.

**Key words:** cryptic species, introgression, Big Bend National Park, sympatric species, cytochrome *c* oxidase subunit I

## Introduction

Cicada killers (*Sphecius* spp., Hymenoptera: Crabronidae) are ground-nesting wasps; nests may be solitary or in aggregations of up to several hundred (Evans & O'Neill 2007). These large provisioning wasps are named after the parental behavior of females, which capture and paralyze cicadas (Homoptera: Cicadidae) in nearby trees and carry them to their nests. There are currently 21 species in the genus worldwide (Pulawski 2006), five of which occur in the Western Hemisphere (Bohart 2000). Since 2004, we have collaborated on a comprehensive study of cicada killer biology, including field research of *Sphecius* physiological and behavioral ecology in Big Bend National Park (BBNP) in Texas, USA. At this location we have observed individuals of three putative species (*S. convallis* Patton, *S. grandis* Say, and *S. speciosus* Drury) interacting in the same trees and nesting areas, and females of the latter two preying on the same species of cicadas. This led us to question whether or not populations of these wasps were reproductively isolated from one another. The initial intent of this study was to use molecular evidence to help clarify the species boundaries among these sympatric wasps. We also wanted to determine the degree to which molecular evidence supported the morphological distinctions used in a recently published key to New World cicada killers (Holliday & Coelho 2006).

DNA barcoding uses relatively short, standard DNA sequences to identify species. A fragment from the 5' end of the mitochondrial cytochrome *c* oxidase subunit I gene (COX I) has shown much promise as a barcod-

*Accepted by L. Packer: 16 Jan. 2008; published: 27 Feb. 2008*

ing standard (Dasmahapatra & Mallet 2006**)**. The utility of this gene stems from the observation that COX I sequences are consistent within species, but quite different between congeneric species, even those that have diverged apparently recently (Hebert *et al*. 2003a). Hebert *et al*. (2004a) noted that intraspecific differences in COX I in North American birds averaged only 0.43%, while differences between closely related species averaged 7.93%. In a study of COX I sequences of spiders, Barrett and Hebert (2005) found that the mean percent sequence divergence between congeneric species was 16.4%, which was more than an order of magnitude greater than the mean of 1.4% found between conspecific individuals. In fact, Hebert *et al.* (2004a) have proposed the arbitrary value of 10x the average within species difference in COX I as a possible standard threshold for species distinction.

The use of DNA barcoding as a means of species identification has generated much criticism (e.g., Moritz & Cicero 2004; Will & Rubinoff 2004). DeSalle *et al.* (2005) point out that some barcoding studies have revealed high variability of COX I sequences within species; such variability could result in an overestimation of species diversity when insufficient sampling has occurred. In a test of the ability of COX I barcoding to correctly identify Diptera, Meier *et al*. (2006) revealed a high rate of misidentification. This was attributed to extensive overlap between intra- and interspecific genetic variability. Another criticism of COX I barcoding is the contention that we poorly understand the biology of mitochondria and the processes that impact their evolution (Ballard & Rand 2005), and as a consequence, caution should be employed when making conclusions about uncorroborated mitochondrial barcoding evidence.

Despite these criticisms, genetic information, including COX I sequences, has often been successfully used for assessment of taxonomic diversity within, as well as between, specimen collection sites (Danforth *et al*. 1998; Hebert *et al*. 2004b; Smith *et al*. 2005). Diversity measures based on barcoding evidence often exceed measures based solely on morphology. The existence of highly divergent barcode groupings within morphospecies has been used to corroborate electrophoretic (Danforth *et al*. 1998) and ecological evidence (Hebert *et al*. 2004b; Smith *et al*. 2007) to help identify sympatric cryptic species. Additionally, the discovery of highly divergent barcodes within morphospecies has prompted closer examination of morphological variation, more detailed study of ecological specialization, or additional gene sequence analysis for possible identification of sympatric cryptic species (Smith *et al*. 2005).

We sequenced COX I fragments of a sample of four putative cicada killer species from one location in Argentina and 19 locations within North America, including individuals of the three morphospecies found in BBNP. We determined the degree of sequence divergence between all pairs of individuals to compare the percent divergence within and among species. To help clarify relationships among these wasps, we also used these sequences to construct a phylogeny. We predicted that the phylogeny would reveal one clade for each species sampled (S. *speciosus, S. grandis, S. convallis* and *S. spectabilis* Taschenberg). Furthermore, if the results of Hebert *et al.* (2003b) and Hebert *et al.* (2004a) hold true, we expected to find divergences of COX I sequences among these clades to be at least an order of magnitude greater than they are within clades. For the BBNP samples, we expected to find three distinct barcode groupings, each corresponding to one of the three morphospecies, *S. convallis, S. grandis,* and *S. speciosus,* found there.

## Materials and methods

**Sampling**. DNA was extracted from four morphospecies of *Sphecius* and a total of 56 individuals from 20 geographic locations (Table 1; Fig. 1). Most specimens were captured by the authors, preserved in 95% ethanol and stored in a -20º C freezer. Other specimens were mailed dry to CWH for identification by collectors and museums in response to a solicitation effort. Specimen vouchers and voucher DNA are retained in the insect collection at Northern Kentucky University. Voucher DNA and specimens can be obtained by contacting JMH; upon his retirement, specimen vouchers will be deposited in the Academy of Natural Sciences entomology collection, Philadelphia, PA.

**TABLE 1.** GenBank accession numbers, isolate identification, latitude & longitude of collection site, and species & haplotype of specimens used in analysis.

| species, haplotype | isolate | GenBank accession number | latitude, longitude |
|---|---|---|---|
| *S. convallis*, A | 308ScAZ | EF203763 | 31.26, -111.11 |
| | 309ScAZ | EF203764 | 31.26, -111.11 |
| *S. convallis*, B | 307ScCA | EF203752 | 33.82, -116.53 |
| | 304ScCA | EF203752 | 33.82, -116.53 |
| | 303ScCA | EF203753 | 33.82, -116.53 |
| | 302ScCA | EF203754 | 33.82, -116.53 |
| | 802ScTX | EF203755 | 29.18, -103.00 |
| | 803ScTX | EF203756 | 29.18, -103.00 |
| | 801ScTX | EF203757 | 29.18, -103.00 |
| | 403ScsAZ | EF203758 | 31.87, -109.10 |
| | 305ScCA | EF203759 | 33.82, -116.53 |
| | 502ScgcAZ | EF203760 | 36.10, -112.16 |
| | 507ScAZ | EF203761 | 36.77, -111.66 |
| | 505ScgcAZ | EF203762 | 36.86, -111.58 |
| S. *grandis*, A | 213SgCO | EF201983 | 38.05 -102.95 |
| | 212SgCO | EF203712 | 38.05, -102.95 |
| | 201SgCO | EF203713 | 38.05, -102.95 |
| | 57SgTX | EF203714 | 29.20, -102.92 |
| | 219SgCO | EF203715 | 38.05, -102.95 |
| | 9SgTX | EF203716 | 29.18, -103.00 |
| | 8SgTX | EF203717 | 29.18, -103.00 |
| | 16SgTX | EF203718 | 29.18, -103.00 |
| | 69SgTX | EF203719 | 29.20, -102.92 |
| | 50SgTX | EF203720 | 29.20, -102.92 |
| | 29SgTX | EF203721 | 29.20, -102.92 |
| | 15SgTX | EF203722 | 29.18, -103.00 |
| | 14SgTX | EF203723 | 29.18, -103.00 |
| | 13SgTX | EF203724 | 29.18, -103.00 |
| | 28SgTX | EF203725 | 29.20, -102.92 |
| | 604SgsAZ | EF203726 | 31.55, -109.66 |
| S. *grandis*, B | 51SgTX | EF203727 | 29.20, -102.92 |
| | 11SgTX | EF203728 | 29.18, -103.00 |
| | 601SgCA | EF203729 | 33.72, -115.40 |
| | 602SgCA | EF203730 | 34.84, -117.02 |
| | 503SggcAZ | EF203731 | 36.24, -112.52 |
| | 504SggcAZ | EF203732 | 36.86, -111.58 |
| | 506SggcAZ | EF203733 | 36.86, -111.58 |
| | 901SgID | EF203734 | 43.66, -116.70 |
| | 605SgWA | EF203735 | 46.58, -119.15 |
| S. *speciosus* | 30SsTX | EF203736 | 29.20, -102.92 |
| | 18SsTX | EF203737 | 29.20, -102.92 |

| | | | |
|---|---|---|---|
| | 26SsTX | EF203738 | 29.18, -103.00 |
| | 25SsTX | EF203739 | 29.18, -103.00 |
| | 230SsCO | EF203740 | 38.05, -102.95 |
| | 228SsCO | EF203741 | 38.05, -102.95 |
| | 215SsCO | EF203742 | 38.05, -102.95 |
| | 103SsKY | EF203743 | 39.03, -84.46 |
| | 105SsKY | EF203744 | 39.03 -84.46 |
| | 134SsKY | EF203745 | 39.03 -84.46 |
| | 115SsKY | EF203746 | 39.03 -84.46 |
| | 114SsKY | EF203747 | 39.03 -84.46 |
| | 137SsKY | EF203748 | 39.03 -84.46 |
| | 701SsPA | EF203749 | 40.70 -75.21 |
| | 702SsPA | EF203750 | 40.70 -75.21 |
| S. *spectabilis* | 1004SstARG | EF203765 | -26.96, -66.14 |
| | 1002SstARG | EF203766 | -26.96, -66.14 |

**Molecular methods.** DNA was extracted by two methods. For most alcohol-preserved specimens DNA was extracted from flight muscle tissue using DNAzol DIRECT (Molecular Research Center, Inc., Cincinnati, OH) following the procedure recommended by the manufacturer. For most dried specimens, DNA was isolated from ground flight muscle by digestion of the tissue in 100 mM Tris-HCl (pH 8.5), 10 mM EDTA, 200 mM NaCl, 0.2% SDS, and 0.75 µg/µl proteinase K followed by phenol/chloroform extraction and alcohol precipitation. A fragment, approximately 650 base pairs long, near the 5' end of the COX I gene was amplified by PCR using an AccuPrime Taq DNA Polymerase System (Invitrogen life technologies, Carlsbad, CA). PCR primers were designed by locating the target gene-fragment within a published mitochondrial genome of the Italian honey bee, *Apis mellifera ligustica* Spinola available in GenBank (accession L06178), and selecting complementary sequences that would be suitable for primers. The forward primer for all samples was 5'–TATCAACCAATCATAAAAATATTG–3', and the reverse primer was 5'–TAAACTTCTGGATGACCAAAAAATCA–3' which are similar to the Folmer primer set, LCO1490 and HCO2198 (Folmer *et al.* 1994).

PCR was carried out with a Model PTC–100 Programmable Thermal Controller (GMI, Inc., Ramsey, MN). The same PCR protocol was followed for each sample: 1) 94° C for 2:00 min, 2) 94° C for 0:30 min, 3) 45° C for 1:30 min, 4) 70° C for 1:00, 5) return to step 2 for 5 cycles, 6) 94° C for 0:30 min, 7) 51° C for 1:30 min, 8) 70° C for 1:00, 9) return to step 6 for 35 cycles, 10) 70° C for 5:00 min. PCR products were identified by gel electrophoresis and were then isolated and purified from the gels with Qiagen MinElute Gel Extraction Kits (Qiagen Sample & Assay Technologies, Valencia, CA). Unidirectional sequencing of the PCR products was done by Retrogen, Inc. (San Diego, CA)

COX I sequences have been deposited in GenBank along with specimen collection and identification information. Sequence accession numbers are EF201983 and EF203712 - EF203767. Simultaneous alignment of all sequences can be observed by locating the data sets within "TaxBrowser" on the National Center for Biotechnology Information (NCBI) website, and then selecting the "popset" function.

**Data analysis**. COX I sequences from 56 individuals of *Sphecius* were aligned by eye, resulting in a final alignment of 648 base pairs. There were no insertions or deletions. Sequences from *Bembix troglodytes* Handlirsch and *Sceliphron caementarium* Drury (Hymenoptera: Sphecidae) were used as outgroups; we generated the former sequence from a specimen collected at BBNP (GenBank accession EF203767), the latter sequence was obtained from GenBank (accession EF032291). These taxa represent the family to which *Sphecius* currently belongs (Crabronidae), and the family from which it was recently removed (Melo 1999), respectively. In the final data set, only two sequences were missing bases from the 3' ends: *S. convallis* CA5 was missing the last 26 bases, and *S. convallis* CA1 the last six bases. This brought the amount of missing or ambiguous data to about 4% (27 bases).

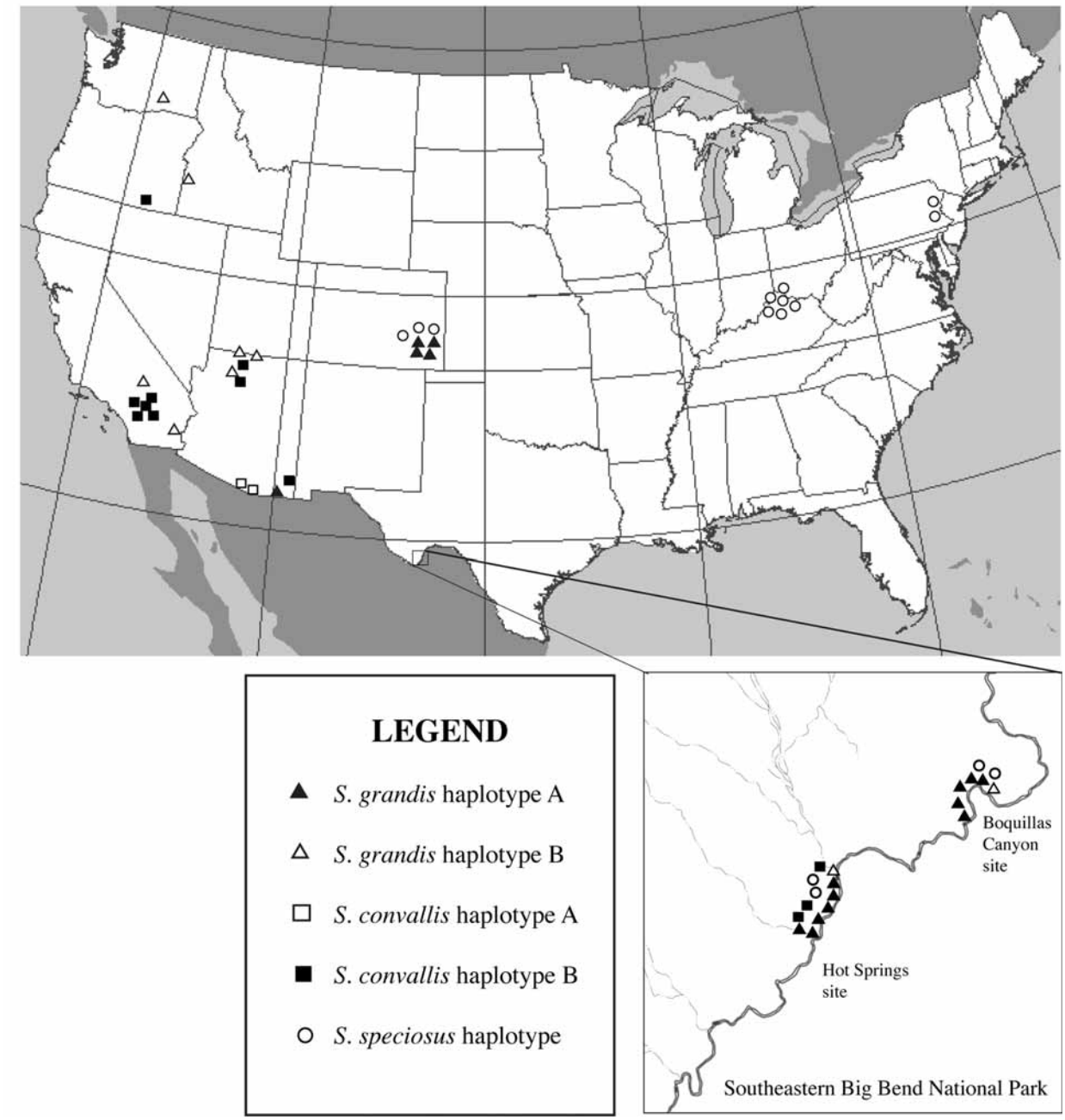


**FIGURE 1**. Geographic distribution of mitochondrial haplotypes for individuals of the three *Sphecius* species sampled from the USA. Approximate localities for all 54 individuals are shown. (Note: *S. spectabilis* does not occur in the USA; the two individuals sampled were Argentinean).

The software PAUP* 4.0b10 was used for analysis of DNA data (Swofford 2002). Neighbor-joining (NJ) and maximum likelihood analyses were run. For all analyses, the number of trees which could be saved was not limited.

During NJ, any ties were broken randomly. NJ was used for the set of samples from BBNP, as well as for all 56 COX I sequences sampled in this study. Bootstrapping for these data sets was done using 10,000 replicates and NJ as the optimality criterion.

Maximum likelihood analysis was used to further examine the data set of the 24 unique COX I sequences. The program Modeltest 3.7 (Posada & Crandall 1998) was used to determine the best maximum likelihood model for analysis. Of the models tested, GTR+G best fit the data and was used for the maximum likelihood analysis. As maximum likelihood bootstrap analyses are too time consuming for general use, parsimony was used as the optimality criterion for bootstrapping this data set. The analysis consisted of 10,000 bootstrap replicates with the starting tree generated by 100 rounds of random addition.

**Results**

**BBNP data set: NJ analysis and percent divergences.** Samples fell into three major clades (Fig. 2), each with 100% bootstrap support. Two of these clades (A, B) were formed by *S. grandis* sequences, and these did not cluster together, rather clade B grouped with the sequences for the other two sympatric species. These subgroups within species will hereafter be referred to as haplotypes. The third major clade was formed by individuals of *S. convallis* and *S. speciosus*. The percent divergences of COX I sequences among these three clades (Table 2) is much more than an order greater than the average within-clade divergence, which, for this data set is about 0.15%.

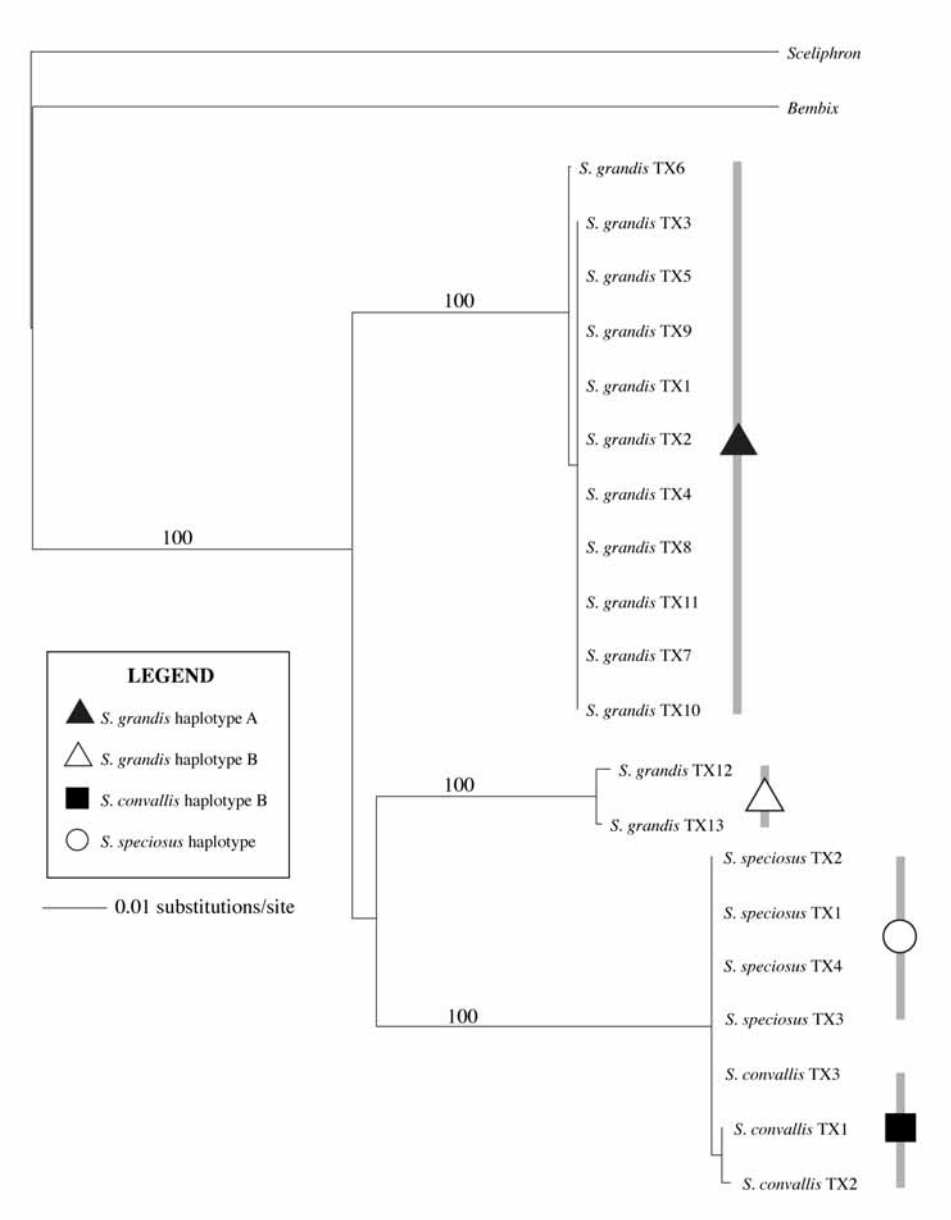


**FIGURE 2**. COX I gene tree resulting from NJ analysis of the BBNP data set. Bootstrap values greater than 70% are provided above the branches. Haplotype symbols match those used in Fig. 1. Note that *S. speciosus* and the majority of *S. convallis* individuals have the same haplotype. The different symbols here emphasize that these are different species.

**TABLE 2.** Pairwise differences (% range) in COX I sequences among species of cicada killers from Big Bend National Park. The two clades of *S. grandis* have been separated for this analysis. Numbers in bold are ranges within groups. This table is modeled after one in Scheffer *et al*. (2006).

| | *S. convallis* (n = 3) | *S. grandis* A (n = 11) | *S. grandis* B (n = 2) | *S. speciosus* (n = 4) |
|---|---|---|---|---|
| *S. convallis* | **0.15 – 0.30** | 9.40 – 9.60 | 8.80 – 9.30 | 0 – 0.30 |
| *S. grandis* A | | **0 – 0.30** | 7.60 – 7.70 | 9.30 – 9.40 |
| *S. grandis* B | | | **0.30** | 9.01 |
| *S. speciosus* | | | | **0** |

**NJ analysis: full data set.** Eight clades with 100% bootstrap support were found (Fig. 3). These included the three seen in the BBNP tree and the patterns for these in the two analyses are congruent. Hereafter, if not specified, the clades discussed have 100% bootstrap support.

As before, *S. grandis* had two distinct haplotypes (A, B). Within *S. grandis* haplotype B, there were two subgroups: one formed by individuals from Texas, the other by individuals from other states. The average COX I sequence divergence between these two subgroups was about 1.9% . Broader sampling of *S. convallis* resulted in two haplotype clades for that species as well (A, B), with only weak support (74%) for these clades being sister to each other. Clade A consisted of two samples from Arizona, while clade B contained individuals from Arizona and other states. All individuals of *S. speciosus* grouped with the *S. convallis* clade B, and one *S. convallis* B sequence (gc4) was actually identical to a *S. speciosus* (TX1) sequence.

**Maximum likelihood analysis: pruned data set.** The one maximum likelihood tree had a -ln likelihood of 2461.55323 (Fig. 4) and was congruent with both full and BBNP-only neighbor-joining trees. *Sphecius* was supported as monophyletic (100%), but within the genus there was little support along the backbone of the tree, so relationships among clades remain unclear. As before, there was strong bootstrap support (90-100%) for all of the haplotype clades, but no support for either a monophyletic *S. grandis* or *S. convallis*.

**Percent divergence: full data set.** Within the five major haplotype clades (*S. convallis* A, *S. grandis* A & B, *S. spectabilis*, and the combined *S. speciosus*/*S. convallis* B clade), sequences generally were less than 2% divergent; average within-clade divergence was about 0.40% (Table 3). Sequence divergence among these clades ranged from 5.8 – 9.2%, which exceeds the divergence threshold for species distinction suggested by Hebert *et al.* (2004a), which for this sample of *Sphecius* would be about 4.00%.

**TABLE 3.** Pairwise differences (% range) in COX I sequences among clades of New World cicada killers. Numbers in bold are ranges within clades. *S. speciosus* and *S. convallis* B are listed spearately

| | *S. convallis* A (n = 2) | *S. convallis* B (n = 12) | *S. grandis* A (n = 16) | *S. grandis* B (n = 9) | *S. speciosus* (n = 15) | *S. spectabilis* (n = 2) |
|---|---|---|---|---|---|---|
| *S. convallis* A | **0** | 6.13–8.35 | 7.99–8.36 | 6.00–7.04 | 8.35–8.37 | 6.44–6.60 |
| *S. convallis* B | | **0–1.12** | 9.34–10.51 | 8.30–10.85 | 0–1.12 | 7.36–8.87 |
| *S. grandis* A | | | **0–0.05** | 7.39–7.98 | 9.01–9.47 | 7.20–7.33 |
| *S. grandis* B | | | | **0–2.12** | 8.30–8.90 | 5.83–6.57 |
| *S. speciosus* | | | | | **0–0.02** | 7.21–7.33 |
| *S. spectabilis* | | | | | | **0.05** |

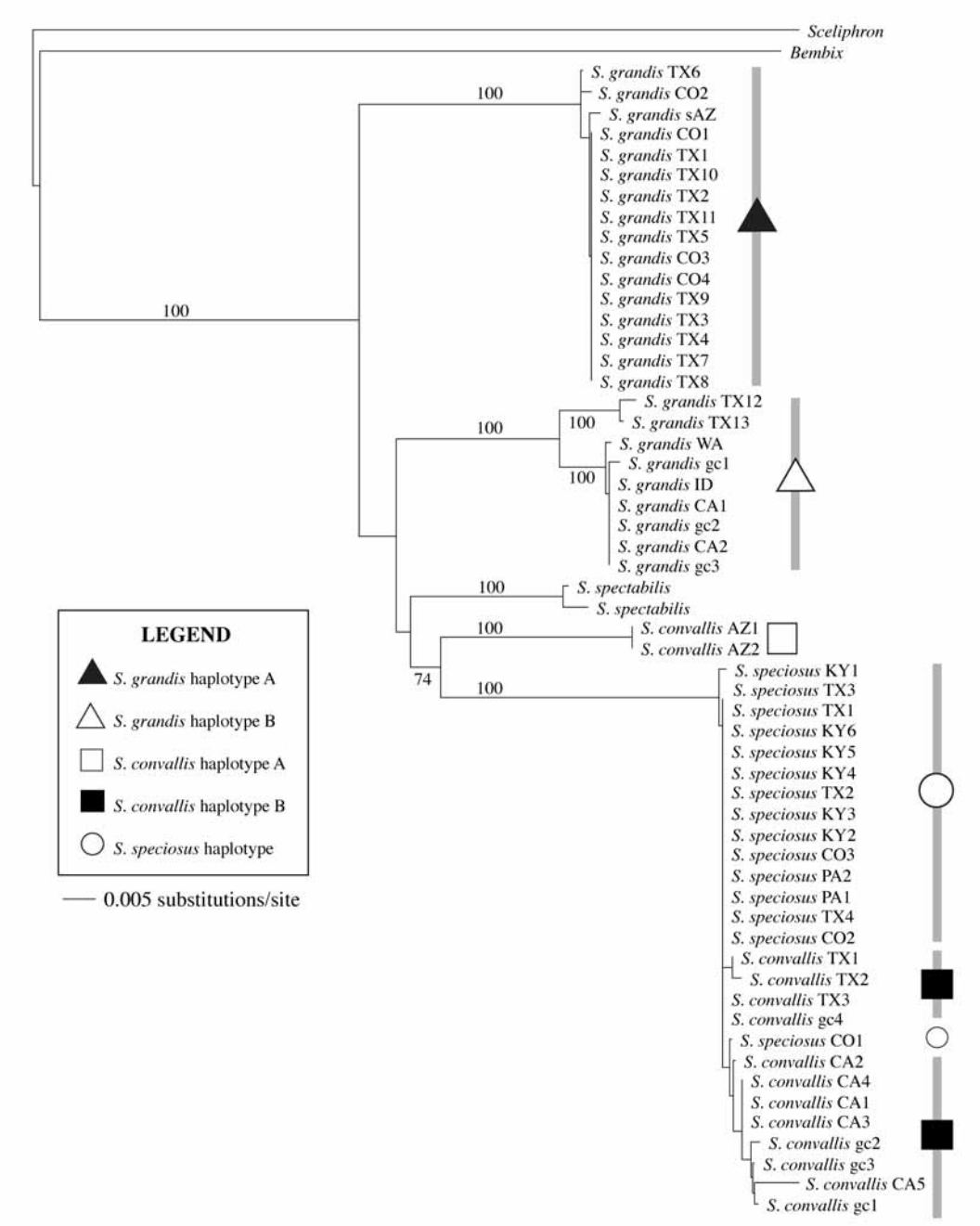


**FIGURE 3**. COX I gene tree resulting from NJ analysis of the full data set, including 56 *Sphecius* sequences. Bootstrap values greater than 70% are provided above the branches. Haplotype symbols match those used in Fig. 1. Abbreviations after each name indicate the state where the individual was collected (and gc indicates Grand Canyon National Park). When multiple individuals from a state were used in the analysis, they are numbered.

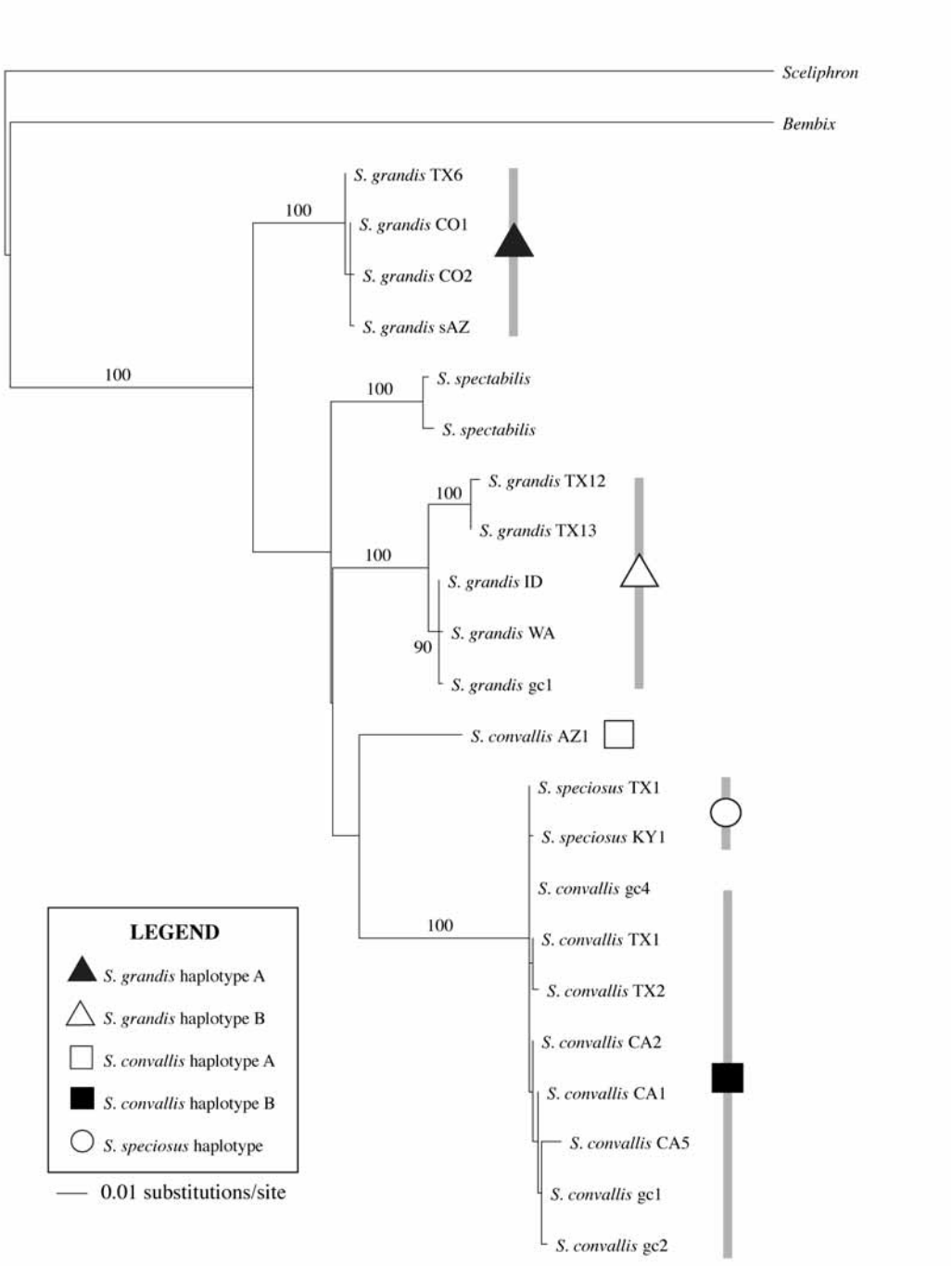


**FIGURE 4**. COX 1 gene tree resulting from maximum likelihood analysis of the pruned data set, with duplicate *Sphecius* sequences removed. Bootstrap values greater than 70% are provided above the branches. Haplotype symbols match those used in Fig. 1.

## Discussion

**Sympatric populations of *Sphecius* in BBNP**: Our prediction that barcodes of cicada killer wasps from BBNP would fall into three distinct groups was upheld, but the barcode groups do not correspond to the three named species collected there. The arbitrary threshold for species distinction of Hebert *et al.* (2004a) supports recognition of *S. grandis*, but the similarity of the *S. convallis* and *S. speciosus* barcodes favors collapsing these two species into one. Additionally, the existence of two divergent clades within *S. grandis* suggests that this morphospecies actually consists of two cryptic species, a common finding of barcoding studies of sympatric multispecies assemblages (Hebert, *et al*. 2004b; Smith, *et al*. 2005).

For evaluation of a putative species, classical taxonomy attempts to assess variation among many individuals from many locations within its geographic range. DeSalle *et al.* (2005) point out that broad sampling is as important to barcoding studies as it is to studies based on morphology. Though the original intent of our study was to clarify relationships among populations of *Sphecius* in BBNP, the inclusion of individuals from a variety of locations provided a better understanding of these relationships.

**The big picture.** Analysis of the full data set affects interpretations in two ways: 1) it confounds the status of *S. convallis* and its relationship to *S. speciosus*, and 2) it reinforces the suggestion that there are two cryptic species within *S. grandis*. As before, *S. grandis* wasps from the larger sample fell into two haplotype clades. Outside of BBNP, these clades are allopatric, with haplotype A found only south and east of the Rocky Mountains, while haplotype B occurs west of this mountain range (Fig. 1). The potential isolating influence of the mountains might account for the evolution of different lineages of *S. grandis* (Smith *et al*. 2005).

As barcoding analysis can result in overestimation of species diversity, the results presented here are, by no means, sufficient to conclude that cryptic species exist within *S. grandis*. For example, highly divergent COX I sequences within populations of the leafmining flies, *Liriomyza trifolii* Burgess (Agromyzidae), suggested cryptic species but, preliminary analysis of nuclear genes has not supported this conclusion (Scheffer *et al.* 2006). Though not common, divergent mitochondrial lineages can be retained within species. Conclusions based on barcoding data are more convincing when supported by ecological evidence (Dasmahapatra & Mallet 2006). Reports of ecological evidence corroborating barcoding identification of cryptic species include the Smith *et al*. (2006) study of parasitoid Tachinid flies and the Hebert *et al*. (2004b) study of tropical butterflies. Currently, we have found no evidence of ecological separation between groups of *S. grandis*.

Highly divergent COX I sequences within *S. convallis* raise a similar question of relationships, complicated by the similarity of one of the two *S. convallis* haplotypes to that of *S. speciosus*. One interpretation of this result is that *S. convallis* and *S. speciosus* are, indeed, separate species, but that introgression of the mitochondrial lineage of the latter has recently occurred within the population of the former. Introgression of mtDNA from one population into another is not uncommon, especially if frequent hybridization occurs and the introgressed mitochondrial genome has a selective advantage (Melo-Ferreira *et al*. 2005). As the geographic ranges of *S. convallis* and *S. speciosus* overlap, such opportunities exist.

Alternatively, the similarity of the COX I sequences of *S. speciosus* and *S. convallis* B could be explained by recent speciation and lack of time for evolutionary divergence. The *S. convallis* A sequences could simply be a different haplotype retained from the ancestral population. A third possible explanation is that *S. convallis* A and *S. speciosus* are separate species, and that *S. convallis* B is merely a morphological variant of the latter. Additional ecological research or analysis of suitable nuclear gene sequences (Gompert *et al*. 2006) could help resolve the issue of the status of *S. convallis* and its relationship to *S. speciosus*. Though nuclear genes generally evolve at a slower rate than mitochondrial genes, fragments of several nuclear genes, such as phosphoenolpyruvate carboxykinase (PEPCK) (Leys *et al.* 2002), elongation factor-1α (EF-1α) (Danforth *et al*. 1999; Leys *et al.* 2002), and the D1–D2 region of the large subunit ribosomal RNA gene (LSU rDNA D1–D2) (Sonnenberg *et al*. 2007) have shown promise for resolving species. In fact, analysis of COX I sequences combined with that of the LSU rDNA D1–D2 gene fragment has proved useful in detecting mitochondrial transfer between species (Sonnenberg *et al*. 2007).

**Barcodes and the new taxonomic key.** The new key to New World *Sphecius* (Holliday & Coelho 2006) uses morphological traits and geographic distributions to distinguish among five named species of *Sphecius*, including the four species, *S. convallis, S. grandis, S. spectabilis,* and *S. speciosus*, that were part of this study. The barcodes support identification of *S. spectabilis* and *S. grandis* as species distinctly different from *S. speciosus* and *S. convallis*. However, the barcoding evidence regarding the status of *S. convallis* and its relationship to *S. speciosus* is ambiguous. Additionally, the existence of highly divergent barcode groupings within *S. grandis* raises the possibility that this morphospecies actually represents two cryptic species. In their discussion of the shortcomings of the earlier keys of Fox (1895) and of Bohart (2000), Holliday and Coelho (2006) point to the failure of these keys to correctly identify a large percentage of *S. convallis* and *S. grandis*. It is noteworthy that the conflicts between the barcoding evidence presented here and the new key involve these same species. Furthermore, the new key describes a wide range of variation in color and markings of gastral terga within *S. grandis*. If, indeed, there are cryptic species of *S. grandis*, as suggested by the barcode evidence, then some of this variation could actually be due to inter-, rather than intraspecific differences among individuals. An advantage of the barcoding evidence is that, unlike the morphological distinctions made in the new and in the earlier keys to the genus, it is independent of sex. We do not consider the barcoding evidence presented here to be any more or any less useful to the assignment of New World *Sphecius* specimens to a particular species than is the evidence used in the new taxonomic key. However, if combined with morphological and distributional data, barcodes would likely enhance our understanding of the relationships among populations of these wasps.

The results of this study provide evidence of the utility of COX I barcoding. Despite the limitations of the method as a sole means of species identification and delineation, it has generated two new lines of inquiry: 1) the status of the two barcode groups of *S. grandis*, and 2) the relationship between *S. convallis* and *S. speciosus*. Future work to resolve these two questions will follow several lines of investigation, including analysis of nuclear genes, closer examination of morphological variation, and field studies of the behavioral ecology of sympatric populations of these wasps.

**Acknowledgments.**

We thank Allan Hook for identification of the *Bembix*, and the collectors who sent specimens to us for this project. We are grateful to the National Park Service for permission to conduct research in Big Bend National Park under NPS permit # BIBE-2005-SCI-003, and to M. Alex Smith and an anonymous reviewer for their useful comments on the manuscript This work was supported in part by grants from the Northern Kentucky University Center for Integrative Natural Sciences and Mathematics (JMH and PJS), by grant number P20RR-16481 from the National Center for Research Resources, a component of the National Institutes of Health (PJS), the Quincy University Faculty Development and Welfare Committee (JRC), and the Lafayette College Academic Research Committee (CWH). The maps used for Figure 1 were made available by the National Atlas of the United States (http://www.nationalatlas.gov/) and the National Park Service maps website from the Harpers Ferry Center of the U.S. Department of the Interior ( http://www.nps.gov/carto/).

## References

Ballard, J.W.O. & Rand, D.M. (2005) The population biology of mitochondrial DNA and its phylogenetic implications. *Annual Review of Ecology and Systematics*, 36, 621–642.

Barrett, R. & Hebert, P.D.N. (2005) Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*, 83, 481–491.

Bohart, R.M. (2000) A review of the Gorytini in the neotropical region (Hymenoptera: Sphecidae: Bembicinae). *Contributions on Entomology International*, 4, 111–259.

Danforth, B.N., Mitchell, P.L. & Packer, L. (1998) Mitochondrial DNA differentiation between two cryptic *Halictus* (Hymenoptera: Halictidae) species. *Annals of the Entomological Society of America*, 91, 387-391.

Danforth, B.N., Sauquet, H. & Packer, L. (1999) Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae) based

on parsimony and likelihood analyses of nuclear EF-1α sequence data. *Molecular Phylogenetics and Evolution*, 13(3), 605-618.
Dasmahapatra, K.K. & Mallet, J. (2006**)** DNA barcodes: recent successes and future prospects. *Heredity,* 97, 254–255.
DeSalle, R., Egan, M.G. & Siddall, M. (2005) The unholy trinity: taxonomy, species delimitation, and DNA barcoding. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 360, 1905–1916.
Evans, H.E. & O'Neill, K.M. (2007) *The Sand Wasps: Natural History and Behavior.* Harvard University Press. Cambridge, Massachusetts, 340 pp.
Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3, 294–297.
Fox, W.J. (1895) Synopsis of the Stizini of Boreal America. *Proceedings of the Academy of Natural Sciences of Philadelphia* 47, 264–268.
Gompert, Z., Nice, C.C., Fordyce, J.A., Forister, M.L. & Shapiro, A.M. (2006) Identifying units for conservation using molecular taxonomy: the cautionary tale of the Karner blue butterfly. *Molecular Ecology,* 15, 1759–1768.
Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological Sciences,* 270, 313–321.
Hebert, P.D.N., Ratnasingham, S. & de Waard, J.R. (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B: Biological Sciences,* 270, S96–S99.
Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S. & Francis, C.M. (2004a) Identification of birds through DNA barcodes. *PLoS Biology,* 2(10), e312.
Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H. & Hallwachs, W. (2004b) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proceedings of the National Academy of Sciences of the United States of America,* 101, 14812–14817.
Holliday, C.W. & Coelho, J.R. (2006) Improved key to New World species of *Sphecius* (Hymenoptera: Crabronidae). *Annals of the Entomological Society of America,* 99, 793–798.
Leys, R., Cooper, S.J.B., Schwarz, M.P. (2002) Molecular phylogeny and historical biogeography of the large carpenter bees, genus *Xylocopa* (Hymenoptera: Apidae). *Biological Journal of the Linnean Society,* 77, 249–266.
Meier, R., Shiyang, K., Vaidys, G. & Ng, P.K.L. (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology,* 55(5), 715–728.
Melo, G.A.R. (1999) Phylogenetic relationships and classification of the major lineages of Apoidea (Hymenoptera), with emphasis on the crabronid wasps. *Scientific Papers, Natural History Museum of the University of Kansas,* 14, 1–55.
Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N. & Alves, P.C. (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology,* 14, 2459–2464.
Moritz, C. & Cicero, C. (2004) DNA barcoding: promise and pitfalls. *PloS Biology,* 2(10), e354.
Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,* 14(9), 817–818.
Pulawski, W.J. (2006, 3 August version) Catalog of Sphecidae. http://www.calacademy.org/research/entomology/Entomology_Resources/Hymenoptera/sphecidae.htm. 14 November 2007.
Scheffer, S.J., Lewis, L.L. & Joshi, C.J. (2006) DNA barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. *Annals of the Entomological Society of America,* 99(2), 204–210.
Smith, M.A., Fisher, B.L. & Hebert, P.D.N. (2005) DNA Barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 360, 1825-1834.
Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W. & Hebert, P.D.N. (2006) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences of the United States of America,* 103, 3657–3662.
Smith, M.A., Wood, D.M., Janzen, D.H., Hallwachs, W., & Hebert, P.D.N. (2007) DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 4967–4972.
Sonnenberg, R., Nolte, A.W., & Tautz, D. (2007) An evaluation of LSU rDNA D1–D2 sequences for their use in species identification. *Frontiers in Zoology,* 4, 6.
Swofford, D.L. (2002) Phylogenetic Analysis Using Parsimony (And Other Methods). 4.0 Beta. Sunderland, MA: Sinauer.
Will, K.W. & Rubinoff, D. (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics,* 20, 47–55.