NRC Research Press

ARTICLE

DNA barcoding a nightmare taxon: assessing barcode index numbers and barcode gaps for sweat bees

Jason Gibbs

Abstract: There is an ongoing campaign to DNA barcode the world's >20 000 bee species. Recent revisions of *Lasioglossum* (*Dialictus*) (Hymenoptera: Halictidae) for Canada and the eastern United States were completed using integrative taxonomy. DNA barcode data from 110 species of *L.* (*Dialictus*) are examined for their value in identification and discovering additional taxonomic diversity. Specimen identification success was estimated using the best close match method. Error rates were 20% relative to current taxonomic understanding. Barcode Index Numbers (BINs) assigned using Refined Single Linkage Analysis (RESL) and barcode gaps using the Automatic Barcode Gap Discovery (ABGD) method were also assessed. RESL was incongruent for 44.5% of species, although some cryptic diversity may exist. Forty-three of 110 species were part of merged BINs with multiple species. The barcode gap is non-existent for the data set as a whole and ABGD showed levels of discordance similar to the RESL. The *viridatum* species-group is particularly problematic, so that DNA barcodes alone would be misleading for species delimitation and specimen identification. Character-based methods using fixed nucleotide substitutions could improve specimen identification success in some cases. The use of DNA barcoding for species discovery for standard taxonomic practice in the absence of a well-defined barcode gap is discussed.

Key words: Apoidea, Hymenoptera, identification, species delimitation, taxonomy.

Résumé : Un effort est en cours pour réaliser le codage à barres de l'ADN pour les >20 000 espèces d'abeilles. De récentes révisions au sein du genre *Lasioglossum* (*Dialectus*) (Hymenoptera : Halictidae) au Canada et dans le nord-est des États-Unis ont été menées par une approche de taxonomie intégrative. Les données de codes à barres de l'ADN pour 110 espèces du genre *L.* (*Dialectus*) ont été examinées pour leur utilité en matière d'identification et de découverte d'une diversité taxonomique additionnelle. Le succès en matière d'identification des spécimens a été estimée à l'aide de la méthode de la plus grande correspondance (« best close match »). Le taux d'erreurs avoisinait les 20 % par rapport aux connaissances taxonomiques actuelles. La justesse de l'assignation des numéros d'index (BIN; « Barcode Index Numbers ») assignés en utilisant l'approche RESL (« Refined Single Linkage Analysis ») et de la séparation des codes à barres à l'aide de la méthode ABGD (« Automatic Barcode Gap Discovery ») a également été évaluée. Le RESL était congruent pour 44,5 % des espèces, bien qu'une certaine diversité cryptique pourrait exister. Quarante-trois des 110 espèces ont été assignées à des BIN composés de plusieurs espèces. Il n'y avait pas de séparation entre les codes à barres au sein des données dans leur ensemble et la méthode ABGD a montré des niveaux de discordance semblables à la méthode RESL. Le « groupe d'espèces *viridatum* » était particulièrement problématique, de telle manière que les codes à barres seuls seraient sujets à erreur pour la délimitation des espèces et l'identification des spécimens. Les méthodes basées sur les caractères utilisant des substitutions nucléotidiques fixées pourraient améliorer l'identification des spécimens dans certains cas. Les auteurs discutent de l'emploi du codage à barres de l'ADN comme méthode taxonomique standard pour la découverte d'espèces en l'absence d'une séparation bien définie des codes à barres. [Traduit par la Rédaction]

Mots-clés : Apoidea, Hymenoptera, identification, délimitation des espèces, taxonomie.

Introduction

DNA barcoding is a method for identifying biological samples to species using the sequence of a standardized gene fragment (Hebert et al. 2003*a*; Hebert and Gregory 2005). DNA barcoding standards for animals require sequencing a fragment of mitochondrial cytochrome *c* oxidase subunit 1 not less than 500 bp in length (Hebert et al. 2003*a*; Hubert et al. 2008), although shorter fragments also can be used for identification (Hajibabaei et al. 2006*b*) and individual nucleotide substitutions can be diagnostic (Burns et al. 2007; Gibbs 2009*b*). DNA barcodes are also commonly used to aid in the delimitation of species boundaries (Gibbs 2009*a*; Rehan and Sheffield 2011; González-Vaquero et al. 2016; Packer and Ruz 2017). I prefer to use the terms specimen identification and species discovery (Collins and Cruickshank 2012) to prevent confusion over the more ambiguous term species identification.

DNA barcoding for specimen identification works by comparing unknown samples to a database of sequences generated from identified material available on GenBank and the Barcode of Life Data System (BOLD; Ratnasingham and Hebert 2007). BOLD is the product of collaboration between computer programmers, molecular biologists, and taxonomists (Smith et al. 2005; Hajibabaei et al. 2005; Ratnasingham and Hebert 2007). Taxonomists provide the scientific context for sequence data used for specimen identification (Goldstein and DeSalle 2011). The advantage of this enterprise from the taxonomists' perspective is a potential wealth of molecular data that can be used to test hypotheses of species

Received 3 May 2017. Accepted 8 September 2017.

Corresponding Editor: S.J. Adamowicz.

J. Gibbs. University of Manitoba, Department of Entomology, 12 Dafoe Rd., Winnipeg, MB R3T 2N2, Canada.
Email for correspondence: jason.gibbs@umanitoba.ca.

limits (DeSalle et al. 2005). The combination of traditional taxonomic approaches with molecular methods can create a taxonomic feedback loop, which can lead to species discovery and more well-resolved taxonomies (Page et al. 2005). Inclusion of molecular data in taxonomic studies is one part of a broader integrative approach to the science sometimes referred to as integrative taxonomy (Dayrat 2005).

DNA barcoding success is often related to the presence of a barcoding gap (Meyer and Paulay 2005). If genetic divergence within species does not overlap with divergence between congeners then DNA barcodes can identify specimens effectively. DNA barcoding loses efficacy when the barcoding gap becomes small or absent (Meyer and Paulay 2005). Sequence divergence thresholds such as 2% or 3% have been suggested for grouping specimens into provisional species (Hebert et al. 2004*b*; Smith et al. 2005). High success rates using similar thresholds have been reported for some taxa (Hebert et al. 2003*b*, 2004*b*; Barrett and Hebert 2005), but would not be sufficient for some species complexes, including groups of butterflies (Hebert et al. 2004*a*), flies (Meier et al. 2006; Virgilio et al. 2012), and bees (Gibbs 2009*a*, 2009*b*; Almeida et al. 2009). DNA barcoding thresholds work best if matched to the specific taxon library (Meier et al. 2006). A regression method for determining ad hoc thresholds for DNA barcoding libraries has been proposed (Virgilio et al. 2012). An alternative to sequence divergence thresholds is to use diagnostic nucleotide substitutions or unique patterns of nucleotide polymorphism for identifying closely related species (DeSalle et al. 2005; Goldstein and DeSalle 2011; Gibbs et al. 2013).

Ratnasingham and Hebert (2013) recently developed the Barcode Index Number (BIN) system for categorizing DNA barcodes into operational taxonomic units (OTUs) in the absence of taxonomic information. BINs are assigned using Refined Single Linkage Analysis (RESL), an algorithm that does not use prior taxonomic knowledge (Ratnasingham and Hebert 2013). RESL uses a 2.2% threshold of sequence divergence to delimit preliminary OTUs and then refines them using a graphical Markov clustering analysis (Ratnasingham and Hebert 2013). A BIN can have four possible relationships for any species pair, which Ratnasingham and Hebert (2013) refer to as match, split, merge, and mixture (their fig. 1). When traditional taxonomy and BINs are concordant, they are said to match. Splits occur when single species are assigned multiple BINs. Merging occurs when a single BIN number is assigned to two or more species, what has long been referred to informally as lumping. In the mixture scenario, two BINs are assigned to two species, but sequences of at least one species fall into both BINs. Merging and mixture of BINs may occur in situations of introgression, incomplete lineage sorting, or if species are inappropriately assigned too many names (Rheindt et al. 2009). Other methods have been developed for assigning OTUs using sequence data (Pons et al. 2006), including the Automatic Barcode Gap Discovery (ABGD) method (Puillandre et al. 2012). ABGD is similar to RESL in that it defines preliminary OTUs by inferring a barcode gap and then refines partitions recursively allowing for different barcode gaps across the dataset. ABGD has been used in several barcoding studies, typically using default or slightly modified settings (Hendrixson et al. 2013; Kekkonen and Hebert 2014).

Bees (Hymenoptera: Apoidea: Anthophila) are the most important pollinators of flowering plants (Klein et al. 2007; Ollerton et al. 2011). Pollination services provided by bees are crucial for most terrestrial ecosystem functioning and much of agricultural production (Klein et al. 2007). Bees are also potentially valuable as indicator taxa of ecosystem health (Zayed et al. 2004), which stems from their increased extinction risk due to a haplodiploid sex determination mechanism (Zayed and Packer 2005; Zayed 2009). Haplodiploidy also has potential impacts on the rates of mitochondrial introgression (Patten et al. 2015), which may reduce the utility of DNA barcodes in species that hybridize (Nicholls et al. 2012; Patten et al. 2015). Nevertheless, DNA barcodes have been used successfully in bees for faunal studies (Sheffield et al. 2009; Magnacca and Brown 2012; Schmidt et al. 2015; Packer and Ruz 2017), species and subspecies discovery and delimitation (Gibbs 2009*a*; Rehan and Sheffield 2011; Pauly et al. 2015; González-Vaquero et al. 2016; Sheffield et al. 2016), and specimen identification (Sheffield et al. 2011).

I examined DNA barcode data generated for a group of taxonomically challenging bees in North America using BINs and barcode gaps. Sweat bees (Hymenoptera: Halictidae) have been called morphologically monotonous (Michener 1974) and the despair of taxonomists (Wheeler 1928), and the large subgenus *Lasioglossum* (*Dialictus*) is notoriously the most difficult to identify to species. *Lasioglossum* (*Dialictus*) are extremely abundant in surveys of North American bees (MacKay and Knerer 1979; Grixti and Packer 2006; Campbell et al. 2007; Droege et al. 2010; Ngo et al. 2013), making identification tools crucial for studies of bee diversity. Thousands of *L.* (*Dialictus*) are collected each year in North America, but taxonomic keys have only become available recently for a subset of species based on geographic regions (Gibbs 2009*b*, 2010*a*, 2011). Consequently, many studies have been published for which *L.* (*Dialictus*) specimens are unidentified or misidentified (Kalhorn et al. 2003; Giles and Ascher 2006; Grixti and Packer 2006; Kearns and Oliveras 2009; B.A. Smith et al. 2012; Wheelock and O'Neal 2016). A very small pool of taxonomists is available to provide reliable identifications to this group, limited primarily to species included in recent revisions (Gibbs 2010*a*, 2011; Gibbs et al. 2013). DNA barcodes have proven useful for facilitating species discovery within small species groups of *L.* (*Dialictus*) in the past (Gibbs 2009*a*, 2009*b*), and they have the potential to aid in the identification of difficult taxa for which taxonomic expertise is limited (Hebert et al. 2003*a*; Packer et al. 2009).

Lasioglossum (*Dialictus*) has a nearly cosmopolitan distribution including nearly the entire Nearctic, Neotropical, Palaearctic, and Afrotropical regions. The genus as a whole has a relatively recent origin (31 Mya, 95CI 24–48 Mya) (Gibbs et al. 2012), which has diversified rapidly into more than 1800 described species (Gibbs et al. 2012; Ascher and Pickering 2016). The genus has not been revised for many parts of the world (Michener 2007), so the total species richness may be much higher. The recent and rapid diversification of *Lasioglossum* is likely to blame for taxonomic challenges associated with the genus (Gibbs et al. 2012; Groom et al. 2013). In addition, sexual dimorphism and caste variation within *Lasioglossum* has led to taxonomic errors in the past (Knerer and Atwood 1964). It is unclear why *Lasioglossum* has diversified so rapidly, but their generalist nature seems to allow them to thrive in varied conditions. The small body size of *L.* (*Dialictus*) may also allow for geographic isolation of populations and subsequent speciation.

I restricted the study of *L.* (*Dialictus*) DNA barcoding to the species occurring in Canada and the United States for two reasons. There have been recent revisions of *L.* (*Dialictus*) in these two areas (Gibbs 2010*a*, 2011) and DNA barcode data for these species are most complete. The species in these regions as currently defined were based on an integrative approach using an evolutionary species concept. Data used in delimiting species included morphology, DNA barcodes, ecology, and geography (Gibbs 2009*a*, 2010*a*, 2011). These data were used in combination to corroborate taxonomic hypotheses (DeSalle et al. 2005). In cases where corroboration failed, individuals or populations were considered conspecific. Species were not defined by a single character where polymorphism could not be distinguished from species level differences. Some highly variable species were considered tentative species complexes in need of additional study (Gibbs 2010*a*).

Given that taxonomic information is lacking globally, including western North America, I explore the effectiveness of molecular methods for identifying specimens and inferring species boundaries in this group. Both RESL and ABGD have been reported to be relatively successful for delimiting species boundar-

ies using public data sets (Puillandre et al. 2012; Ratnasingham and Hebert 2013; Kekkonen and Hebert 2014), but these have often been across broader taxon groups where the taxonomy overall is less problematic. I focus attention on whether current DNA barcode data are sufficient for identification and delimitation of *L.* (*Dialictus*) specimens based upon the best close match method (Virgilio et al. 2012) and the BIN system. I also address remaining taxonomic issues in the group and the use of DNA barcodes for species delimitation in a difficult taxon.

Materials and methods

Data set

Compliant DNA barcodes for 110 species were generated concurrently with traditional taxonomic study that culminated in taxonomic monographs on *L.* (*Dialictus*) for Canada (Gibbs 2010*a*) and the eastern United States (Gibbs 2011). Species limits used herein were based on morphological study of many specimens, including name-bearing types, and through integrative approaches that incorporated barcode data with morphology and geography (Gibbs 2009*a*). Specimens were sequenced from the greatest possible geographic range of each species so that some sequences originate from specimens collected outside the geographic boundaries of Canada and the eastern United States. Whenever possible, samples were included from the type locality or as close as was possible. In delimiting species, sequence divergence from DNA barcodes in the absence of other supporting evidence was not considered sufficient to justify recognition of species (Gibbs 2010*a*, 2011). Instead, such situations were considered candidate species complexes in need of additional study, such as examination of nuclear loci (Carman and Packer 1996; Danforth et al. 1998).

The DNA barcoding methods were described in previous studies of *Lasioglossum* (Gibbs 2009*a*, 2009*b*; Gibbs et al. 2013). Briefly, sequencing was performed at the Canadian Centre for DNA Barcoding at the University of Guelph (Guelph, Ontario). DNA extractions were performed using standard methods for 96-well plates (Ivanova et al. 2006). Either a single hind leg or both the middle and hind leg (for small specimens) from the right side were used for extractions. Universal primer pairs were used to amplify the DNA barcode region; either LCO1490 and HCO2198 (Folmer et al. 1994) or the variants LepF1 and LepR1 (Hebert et al. 2004*a*). LepF1 and LepR1 are variations of the Folmer primers designed for Lepidoptera and often used for DNA barcoding Insecta. The LepR1 primer has since been shown to be a poor fit for halictid bees (M.A. Smith et al. 2012). Samples that failed to amplify were then reattempted using internal primer pairs LepF1 and C_AntMr1D-RonIIdeg_R and LepR1 with MLepF1. (Hajibabaei et al. 2006*b*; Smith et al. 2007). PCR and sequencing reactions followed standard protocols (Hajibabaei et al. 2005). Sequences were uploaded to the Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert 2007) and GenBank (accession numbers in prep.).

A full-length barcode sequence for these bees is 658 bp. Since internal primers were often necessary, many sequences did not approach this length. To limit spurious results from small sequence overlap, I limited most data analyses to barcode-compliant sequences greater than 500 bp, with two or more trace files, meeting minimum recommended standards for DNA barcoding (Hubert et al. 2008). Samples were also excluded from analyses if they had more than 1% ambiguous base calls. Any sequences with stop codons or showing evidence of contamination (e.g., best match sequence identity to non-bee taxa or other bee genera than to *Lasioglossum*) were also removed from the dataset. A larger number of non-compliant sequences and extra-limital species are also available on BOLD. These are not analyzed fully, but are referenced with respect to their inclusion in BINs and their potential effects on successful identification using DNA barcodes.

Sequences were downloaded from BOLD and realigned using MUSCLE (Edgar 2004) with default settings (gap penalty of −400 and extension penalty of −0.01) as implemented by MEGA7 (Kumar et al. 2016). Re-alignment was necessary to remove seven inappropriate insertion/deletions in the sequences generated in the BOLD alignment. The final alignment had no indels or stop codons. Aligned sequences were renamed in MESQUITE (Maddison and Maddison 2017) for compatibility with downstream analyses.

Specimen identification

Statistical analyses were conducted using R version 2.3.3 (R Core Team 2015), unless otherwise specified. The minimum interspecific and maximum intraspecific genetic distances for each species were calculated using the nonConDist and maxInDist functions in the spider package (Brown et al. 2012) using a Kimura two-parameter model of substitution (Kimura 1980). To assess the potential barcode gap, we examined the proportion of cases where minimum interspecific distance was greater than intraspecific distance for each sequence. The identification success rate was estimated using the best close match method (Meier et al. 2006) implemented by the threshOpt function. The ad hoc package (Sonet et al. 2013) was also used to assess the identification success rate and optimal threshold values for the DNA barcoding library based on the methods described by Virgilio et al. (2012). The checkDNAbcd function was used with a Kimura two-parameter model to calculate pairwise distances. Identification error rates using the best close match method were determined using the adhocTHR function. Diagnostic nucleotide substitutions were determined for species complexes using the function nucDiag in the spider package for species with merged BINs.

Species discovery

BOLD automatically implements RESL to group sequences into BINs (Ratnasingham and Hebert 2013). BIN contents were compared to currently accepted species limits to determine the congruence of RESL to integrative taxonomic assignment of species limits. In addition, the ABGD method (Puillandre et al. 2012) was used to estimate species limits using the online web application (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html, accessed 18 July 2017). The results of our optimal threshold analyses were used to improve a priori input values for intraspecific diversity limits (*P*) and minimum gap width (*X*). Default values of *P* between 0.001 and 0.1 were used. The number of steps was set to 20. Several values of *X* from 0.1 to 1.5 were evaluated. Simple distances are reported as these seemed to perform better than the Jukes–Cantor or Kimura 2-parameter method (Srivathsan and Meier 2012). ABGD output includes several options for partitioning the data over values of *P*. The preferred partition was selected using an intermediate value of *P* after an initial steep decline in number of species estimated (Puillandre et al. 2012).

Results

Data set

In total, 1279 DNA barcode-compliant sequences were generated for species of *L.* (*Dialictus*) revised in Canada and the eastern United States (supplementary data, Table S1[1]). Sequences had a strong AT bias in the 3rd codon position (GC content: 7.1 ± 0.07 SE). One hundred and ten species identified using integrative approaches were included (11.6 ± 1.3 sequences per species, range = 1–97). The mean number of haplotypes per species was 5.2 ± 0.46 (range 1–24). Mean maximum intraspecific distance was 1.69% ±

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2017-0096.

0.05% (range 0–7.5%), excluding singletons, while the mean minimum interspecific distance was 2.60% ± 0.21% (range 0–11.5%).

Specimen identification

The DNA barcoding threshold with the lowest relative error (20.3%) in comparison to current taxonomy was 1.4% based on the methods of Virgilio et al. (2012). The best match method implemented in SPIDER resulted in an optimal threshold of 0.2%–0.3%, but an overall error rate of 33.3%. Eleven species were represented by a single DNA barcode sequence, necessitating a false positive identification using the best match method. However, these species only made up a small proportion of the false positives and removing them from analyses did not improve results. For nine of these species the distance to their best match was greater than the ad hoc threshold of 1.4% (mean 3.2% ± 0.5%), suggesting that with additional sampling these species would be correctly identified. The singleton *L. reticulatum* (Robertson) differed from its closest match *L. bruneri* (Crawford) by 1.3%. Of the singletons, only *L. alachuense* (Mitchell) had no local gap with its closest match *L. oblongum* (Lovell). An analysis of species monophyly using neighbour joining revealed that 30 of the 110 species were non-monophyletic. Diagnostic nucleotide substitutions were found for a subset of problematic species that could allow identification using these characters (Table 3).

Species discovery

Ninety-four BINs corresponded to these 110 species in the data set, comprising 61 matches (Table 1), 6 splits (Table 2), 43 merged (Table 3), and 6 mixed BINs (Tables 2 and 3). Species with minimum interspecific distance greater than 1.4% usually (85%) showed concordance between traditional taxonomy and BINs (Fig. 1). Four species above this cut-off had multiple BINs suggestive of cryptic species and two species were morphologically distinct singletons grouped with the *viridatum* species complex.

The ABGD performed poorly with the Jukes–Cantor and Kimura 2-parameter when the relative gap width value (*X*) was greater than one. The results from simple p-distances were less impacted by values of *X* and more often estimated species richness similar to integrative taxonomy. The simple distance ABGD results with relative gap width set to 1.0 are summarized in Tables 1–3. The performance of ABGD was comparable to RESL. Examining pairwise distances across DNA barcode-compliant sequences reveals no definitive DNA barcode gap (Fig. 2).

Species with merged BINs were sometimes identifiable using fixed nucleotide substitutions, although they did not show clear barcode gaps (Table 3). For example, *L. knereri* Gibbs and *L. subversans* (Mitchell) were grouped together in both the ABGD and RESL analyses, but they are distinguished by four fixed nucleotide substitutions. In the case of the *L. viridatum* (Lovell) species complex, 19 morphospecies with barcode-compliant sequences were combined in a single BIN. Three of these could be distinguished from all others by fixed substitutions and four were treated as distinct groups in the ABGD analysis (Table 3). The remaining species lacked fixed differences with one or more species in the complex. One putative species in this complex, *L. subviridatum* (Cockerell), shows two distinct DNA barcode clusters, with 15 fixed nucleotide substitutions between them. Treated as a single species, as is currently the situation, the consensus sequence shows substantial overlap with other species in the *L. viridatum* group, with 0 fixed substitutions with 8 other species in the group (mean 2.1). If *L. subviridatum* were ultimately treated as two species, which seems plausible, only two other species in the *viridatum* complex would show zero fixed substitutions with either of the candidate species (mean 8.1). These results imply that even if *L. subviridatum* is maintained as a single species other taxa in the species group could be distinguished from it based on unique patterns of polymorphism.

In other species with merged BINs, many are distinguishable by one or more fixed nucleotide substitutions. The exceptions are *L. callidum* (Sandhouse) and *L. versatum* (Robertson), which show identical barcode haplotypes; *L. leucocomum* (Lovell) and *L. succinipenne* (Ellis); *L. puteulanum* Gibbs and *L. tegulare* (Robertson); and the group of *L. hitchensi* Gibbs, *L. leviense* (Mitchell), and *L. weemsi* (Mitchell). In combination with the *L. viridatum* group above, 21 of 110 species are not readily identifiable using DNA barcodes even when considering single base pair substitutions. In the case of *L. puteulanum* and *L. tegulare*, the majority of individuals are readily distinguishable, but four individuals that seem to match *L. tegulare* morphologically group with *L. puteulanum* based on barcodes. The range of these individuals might indicate that there may be unresolved aspects of the taxonomy of this complex. Similarly, *L. weemsi* is distinguishable from all but one individual of each of *L. leviense* and *L. hitchensi*. The morphological identifications have been carefully checked multiple times, so the discrepancy is either due to unreliability of the distinguishing anatomical traits or the barcode sequences. Cross-contamination of samples within a plate is a possibility, although attempts were made to keep closely related species and putative conspecifics well separated within a plate.

Some species have BINs and ABGD groups congruent with taxonomic species delimitation, but may have incongruence with extra-limital species. The eastern North American species *L. trigeminum* Gibbs has distinct barcodes from the morphologically similar species *L. versatum* and *L. callidum*. However, *L. connexum* (Cresson), a species found in Texas and therefore not treated in previous revisions, has only a single nucleotide difference from *L. trigeminum*, but due to insufficient amplification success in *L. connexum*, it remains unclear if that difference is fixed.

Discussion

The value of DNA barcodes in an integrative taxonomic approach is difficult to quantify retroactively. Species delimitation is a process that occurs over time as taxonomists familiarize themselves with a taxon and gather evidence from multiple sources. In the context of *L.* (*Dialictus*), the availability of an independent data set was invaluable even in cases were RESL and ABGD were deemed unsuccessful. As an illustration, I describe the situation of *L. ephialtum* Gibbs, which has a specific epithet that literally means nightmare in Greek. *Lasioglossum ephialtum* was not described until 2010, but it is a relatively common species, particularly in urban areas (Gibbs 2010*a*). The species was often misidentified as *L. versatum* (or the junior synonym *rohweri* Ellis) by earlier workers leading to confusion over the diagnostic characters for *L. versatum*. Although DNA barcodes of *L. ephialtum* are not easily distinguished from those of members of the *L. viridatum* species-group, they are clearly distinct from *L. versatum*, which helped to clarify the latter species' limits. Other advantages of incorporating DNA barcodes into these revisions was the ability to associate dimorphic sexes and female castes (Gibbs 2010*a*). Many *Lasioglossum* are social with queens and workers, which can be so different that they have been described as distinct species (Knerer and Atwood 1964), and nearly two thirds of North American *L.* (*Dialictus*) prior to recent revisionary studies were described from a single sex (Moure and Hurd 1987).

DNA barcodes may also benefit taxonomy by showing clear variation in morphologically distinct species that were often not recognized due to insufficient study. Species with distinct characters in otherwise problematic taxa may not be scrutinized as closely and incorrectly lumped (Packer and Taylor 1997). DNA barcodes sequence variation can highlight the variation in a species complex leading to more careful taxonomic study. Examples in *L.* (*Dialictus*) include the *L. tegulare* and *L. petrellum* (Cockerell) species-groups (Gibbs 2009*a*, 2009*b*). When DNA barcodes suggest splitting species it could be evidence of cryptic species, but it might also represent highly divergent barcodes within a single species. Heteroplasmy (i.e., the presence of multiple haplotypes within an individual) has been shown to occur in the bee genus *Hylaeus* (Magnacca and Brown 2010). Heteroplasmic individuals

Table 1. Taxonomically concordant Barcode Index Numbers (BINs) based on analysis of 1279 barcode-compliant sequences and morphology of voucher specimens. BIN members include non-compliant DNA barcodes.

Species	BIN	ABGD group	Region	Total BIN members
Lasioglossum abundipunctum Gibbs	BOLD:AAE1318	1	CAN	5
L. achilleae (Mitchell)	BOLD:AAW9967	2	CAN/EUS	1
L. albipenne (Robertson)	BOLD:AAB6738	5	CAN/EUS	11
L. albohirtum (Crawford)*	BOLD:ABZ7019	6	CAN	8
L. anomalum (Robertson)	BOLD:AAA7867	7	CAN/EUS	6
L. apopkense (Robertson)	BOLD:AAE1289	8	EUS	5
L. batya Gibbs	BOLD:AAE1190	10	EUS	2
L. brunneiventre (Crawford)	BOLD:AAB7689	12	CAN	14
L. coreopsis (Robertson)	BOLD:AAC1424	17	EUS	16
L. creberrimum (Robertson)	BOLD:AAC2514	18	EUS	12
L. dashwoodi Gibbs	BOLD:AAJ1718	25	CAN	2
L. disparile (Cresson)	BOLD:AAC8559	26	EUS	9
L. ebmerellum Gibbs	BOLD:AAD7578	27	CAN	6
L. gotham Gibbs	BOLD:AAK8338 (NC)	NA	EUS	1
L. halophitum (Graenicher)	BOLD:AAC5496	33, 34	EUS	9, 2
L. hartii (Robertson)	BOLD:AAL9893	35	EUS	2
L. hemimelas (Cockerell)	BOLD:ABZ6179	36	EUS	1
L. heterognathum (Mitchell)	BOLD:ACE5057	37	CAN/EUS	3
L. hyalinum (Crawford)	BOLD:AAE1411	44	CAN	4
L. illinoense (Robertson)	BOLD:AAD2971	45	CAN/EUS	5
L. imitatum (Smith)	BOLD:AAA7868	47	CAN/EUS	26
L. incompletum (Crawford)	BOLD:AAB8298	48	CAN	13
L. izawsum Gibbs	BOLD:ABZ7461	49	CAN/EUS	2
L. laevissimum (Smith)	BOLD:AAA7513	52	CAN/EUS	42
L. lepidii (Graenicher)	BOLD:ABY5206	53	EUS	3
L. lineatulum (Crawford)	BOLD:AAA2141	57, 58	CAN/EUS	13, 1
L. lionotum (Sandhouse)	BOLD:AAE5826	59	CAN/EUS	3
L. longifrons (Baker)	BOLD:AAC8294	60	EUS	9
L. macroprosopum Gibbs	BOLD:AAE1153	61	CAN	2
L. marinum (Crawford)	BOLD:AAD0611	62	EUS	8
L. michiganense (Mitchell)	BOLD:ACE8520	63	CAN/EUS	3
L. nevadense (Crawford)	BOLD:AAB0372	64, 65	CAN	21, 1
L. nigroviride (Graenicher)	BOLD:AAD3968	66	CAN/EUS	1
L. nymphale (Smith)	BOLD:AAB6637	68	EUS	17
L. occidentale (Crawford)	BOLD:AAC7155	69, 70	CAN	10, 1
L. oceanicum (Cockerell)	BOLD:AAE1244	71, 72	CAN/EUS	1, 3
L. packeri Gibbs	BOLD:AAE5057	73	CAN	4
L. pavoninum (Ellis)	BOLD:AAF4051	74	CAN	3
L. perpunctatum (Ellis)	BOLD:AAC1122	75	CAN/EUS	13
L. platyparium (Robertson)	BOLD:ACE6636	77	CAN/EUS	11
L. prasinogaster Gibbs	BOLD:AAB3612	78	CAN	17
L. punctatoventre (Crawford)	BOLD:AAD9279	80	CAN	5
L. raleighense (Mitchell)	BOLD:AAE1191	82	EUS	3
L. rozeni Gibbs	BOLD:AAD7869	84	CAN/EUS	6
L. rufulipes (Cockerell)	BOLD:AAE5185	85	CAN/EUS	4
L. sedi (Sandhouse)	BOLD:ABZ3764	94	CAN	16
L. semicaeruleum (Cockerell)	BOLD:AAB1546	95	CAN	28
L. simplex (Robertson)	BOLD:ACE5372	97	EUS	2
L. smilacinae (Robertson)	BOLD:AAE1871	98	CAN/EUS	1
L. cf. *stuartense* (Mitchell)	BOLD:AAW9918	16	EUS	1
L. tamiamense (Mitchell)	BOLD:AAX0169	100	EUS	1
L. tarponense (Mitchell)	BOLD:AAE5196	101	EUS	4
L. tenax (Sandhouse)	BOLD:AAB6998	102	CAN/EUS	19
L. timothyi Gibbs	BOLD:AAE1693	103	CAN/EUS	3
L. versans (Lovell)	BOLD:ABZ6180	105, 106, 107, 108	CAN/EUS	6, 1, 1, 1
L. vierecki (Crawford)	BOLD:AAB4651	109	CAN/EUS	13
L. zephyrum (Smith)	BOLD:AAD0272	111	CAN/EUS	6
L. n. sp. (ex. Alberta)	BOLD:AAX015	99	CAN	1

Note: BIN numbers only represented by non-compliant barcodes are indicated with NC. ABGD groups assigned to compliant sequences are included. Species are categorized by regions based on the taxonomic revisions used to delimit species boundaries: Canada (CAN) or eastern United States (EUS) (see Gibbs 2010*a*, 2011).

*One BIN but geographic and morphological variation consistent with two species.

may still be identified using DNA barcodes, but the reference database must include multiple divergent haplotypes for each species. The best example of cryptic species among bees is in the *Halictus ligatus* Say/*Halictus poeyi* Lepeletier species pair (Packer et al. 2016). Originally separated based upon extensive allozyme data (Carman and Packer 1996), these two species were subsequently distinguished using both nuclear and mitochondrial DNA sequences including DNA barcodes (Danforth et al. 1998, 1999). The two are sympatric in a narrow area with almost no heterozygosity at the allozyme loci (never more than one of seven loci out of hundreds of individuals from sympatric populations sampled) (Packer 1999). No morphological differences have been discovered

Table 2. Morphologically identifiable species represented by multiple BINs based on analysis of 1279 barcode-compliant sequences.

Species	Included BINs	ABGD group	Region	Total BIN members
Lasioglossum cressonii (Robertson)	BOLD:AAA5974, BOLD:AAP0913, BOLD:ABY7290, BOLD:ACE7675, BOLD:ABY8200 (NC), BOLD:AAA5973	19, 20, 21, 22, 23, 24	CAN/EUS	5, 1, 18, 1, 1, 1
L. hitchensi Gibbs	BOLD:AAE5827, BOLD:ABY8629, BOLD:AAA3782*	38, 39, 40	CAN/EUS	73, 1, 1
L. hudsoniellum (Cockerell)	BOLD:AAB5530, BOLD:ACE7697, BOLD:ABZ0512 [singleton]	41, 43, 43	CAN	16, 2, 1
L. leucocomum (Lovell)	BOLD:ACE3786, BOLD:ACE9642, BOLD:ACF2785*	54, 55, 56	CAN/EUS	12, 3, 2
L. pruinosum (Robertson)	BOLD:AAC1101, BOLD:ABY7039	79	CAN/EUS	8
L. ruidosense (Cockerell)	BOLD:AAA5893, BOLD:ACE3475, BOLD:ACE3476, BOLD:AAA5895, BOLD:AAA5894 [singleton]	86, 87, 88, 89, 90, 91	CAN	3, 9, 3, 9, 2, 1

Note: Total BIN members includes non-compliant DNA barcodes with less than 500 bp or greater than 1% ambiguous nucleotides. Additional BINs based on non-compliant sequences are indicated by NC. ABGD group numbers are provided.
*Discordant with extra-limital species in BOLD.

despite intensive investigation, although there are slight phenological differences in sympatry (Dunn et al. 1998). DNA barcodes have recently recovered a third Neotropical species in this complex, presumably *H. townsendi* Cockerell based on geography (Packer et al. 2016).

Lasioglossum (*Dialictus*) is taxonomically challenging and it is likely that species-level diversity has not been completely described even within areas that have received detailed study (Gibbs 2010*b*, 2011). *Lasioglossum* is the most species rich of all bee genera, but few revisionary studies exist. This group of small bees has diversified rapidly to occupy virtually every terrestrial habitat on the planet (Michener 1979, 2007). Although broadly generalist and highly adaptable, *L.* (*Dialictus*) are small-bodied species that are likely limited in their dispersal ability. As a result, this subgenus seems able to successfully establish and persist in novel and often marginal habitats, but is also susceptible to physical barriers to gene flow. Species-rich taxa with recent diversification are not expected to perform well with single locus automated species delimitation methods (Puillandre et al. 2012).

The highest level of intraspecific divergence in this study, 6.9%, is seen in *L. ruidosense*, putatively a single species with an enormous north–south distribution, ranging from southern New Mexico to Alaska (Gibbs 2010*a*). Southern populations of *L. ruidosense* are primarily limited to high elevations, which have the potential to isolate populations leading to allopatric speciation (Coyne and Orr 2004) as seems to be the case with another species, *L. boreale* Svensson, Ebmer and Sakagami, with a similar distribution in North America (Packer and Taylor 2002). Allozyme data from *L. boreale* showed little genetic variation from individuals spanning large geographical distances, but high elevation populations in the southwestern United States were found to have fixed differences. Additional study is required to determine and describe the extent of diversity in this apparent complex. In the *L. ruidosense* case, the incongruence between current taxonomy and DNA barcoding is likely caused by insufficient taxonomic study. High intraspecific genetic divergence is also present in *L. cressonii* (Robertson) and *L. sagax* (Sandhouse). Neither of these two species has a distinct geographical or morphological pattern correlated with sequence divergence. In fact, highly divergent sequences have been found from a single locality during the same collecting event for both species (Gibbs 2010*a*). Substitutions between haplotypes are strongly biased in the 3rd codon position, which does not suggest amplification of a nuclear pseudogene, although a recently derived pseudogene is possible. A more detailed examination of these species that includes morphology and nuclear DNA is required. In the meantime, they can at most be considered unconfirmed candidate species (Padial et al. 2010). *Lasioglossum cressonii* is a common, distinctive species and similar cases in the past have turned out to be cryptic species (Danforth et al. 1998; Gibbs 2009*a*). *Lasioglossum sagax* is a less well-known species that until recently was only known from the holotype (Wolf and Ascher 2009; Gibbs 2010*a*). Future work that includes multiple loci will be productive for delimiting cryptic species where morphological characters are lacking. It should be noted that DNA barcoding may overestimate the number of species (Dasmahapatra et al. 2010), so the deep divergences within species like *L. cressonii* may not be true evidence of speciation.

Mitochondrial DNA has proven capable of identifying specimens and clarifying species boundaries in other taxonomically challenging bees (Danforth et al. 1998; Murray et al. 2007; Bertsch 2009; Sheffield et al. 2009; Gibbs 2009*b*; Magnacca and Brown 2010; Rehan and Sheffield 2011). However, DNA barcoding is only moderately successful for specimen identification in *L.* (*Dialictus*) using either DNA barcode gaps or BINs. The mean maximum intraspecific value of 1.69% reported here is higher than mean intraspecific values reported for some taxa, including bats (Clare et al. 2007), birds (Hebert et al. 2004*b*), marine fish (Ward et al. 2005), Lepidoptera (Hajibabaei et al. 2006*a*), bumble bees (Bertsch 2009), and a general bee fauna (Sheffield et al. 2009) but does correspond to some previous studies of insect taxa, such as mayflies (Alexander et al. 2009) and black flies (Rivera and Currie 2009). Direct comparisons are not always reliable due to differences in

Table 3. Taxonomically discordant Barcode Index Numbers (BINs) that merge multiple morphospecies into a single BIN.

BIN	Included taxa	ABGD group	Diagnostic substitutions	Sample size
BOLD:AAA3781	*Lasioglossum callidum* (Sandhouse), *L. versatum* (Robertson)	13	0	33+97
BOLD:AAA3782	*L. hitchensi* Gibbs, *L. leviense* (Mitchell), *L. weemsi* (Mitchell)	102	0	73+4+25
BOLD:AAA8348*	*L. knereri* Gibbs, *L. subversans* (Mitchell)	51	4 (270, 384, 504, 631)	36+3
BOLD:AAB0760*	*L. ellisae* (Sandhouse), *L. imbrex* Gibbs	28, 46	7 (81, 282, 312, 315, 573, 603, 640)	21, 9
BOLD:AAB2245*	*L. trigeminum* Gibbs	104	NA	29
BOLD:AAB6361	*L. floridanum* (Robertson), *L. pilosum*† (Smith)	31 (76)	0	4+18, (1)
BOLD:AAC1104*	*L. ceanothi* (Mitchell)†, *L. foveolatum* (Robertson), *L. pictum* (Crawford)	15, 112, 32	6, 1, 0 (219, 315, 327, 360, 402, 462, 594, 618, 625, 634)	2, 1, 9+9
BOLD:AAB6998	*L. cattellae* (Ellis) (NC), *L. tenax* (Sandhouse)	102	0 (NC)	19
BOLD:AAC7048	*L. bruneri* (Crawford), *L. reticulatum* (Robertson)	11, 83	8 (159, 189, 201, 234, 360, 363, 432, 634)	11, 1
BOLD:AAD7708*	*L. puteulanum* Gibbs, *L. tegulare* (Robertson)	46	0	20+26
BOLD:AAD9325	*L. arantium* Gibbs, *L. miniatulum* (Mitchell)	9	3 (66, 177, 585)	5, 6
BOLD:AAW9735*	*L. flaveriae* (Robertson)	30	NA	1
BOLD:ABY5205*	*L. carlinvillense* Gibbs	14	NA	3
BOLD:ABZ0652*	*L. abanci* (Crawford)†, *L. admirandum* (Sandhouse), *L. apocyni* (Mitchell), *L. atwoodi* Gibbs, *L. dreisbachi* (Mitchell), *L. ephialtum* (Gibbs)‡, *L. fattigi* (Mitchell), *L. georgeickworti* Gibbs, *L. katherineae* Gibbs†, *L. marinense* (Michener), *L. novascotiae* (Mitchell)†, *L. obscurum* (Robertson), *L. pacatum* (Sandhouse), *L. paradmirandum* (Knerer & Atwood), *L. planatum* (Lovell), *L. sagax* (Sandhouse)‡, *L. subviridatum* (Cockerell), *L. taylorae* Gibbs, *L. viridatum* (Lovell)†	(3), 3 (29), (50), (67), (110)	0 except *abanci* 3 (200, 201, 212), *admirandum* 2 (310, 561), and *novascotiae* 6 (25, 141, 201, 384, 387, 390)	1, 8+16+6+4+(44, 1)+4+8+(6)+9+(1)+10+20+8+8+(18, 4)+38+3, (6)
BOLD:ABZ0968	*L. alachuense* (Mitchell), *L. oblongum* (Lovell)	4	1 (180)	1+10
BOLD:ACF2785	*L. leucocomum* (Lovell)‡, *L. succinipenne* (Ellis)	54, 55, 56	0	7+5, (3, 2)
BOLD:ACF5350*	*L. sandhousiellum* Gibbs	93	NA	7
BOLD:ACF5692	*L. sheffieldi* Gibbs, *L. yukonae* Gibbs	96	2 (267, 561)	7+3

Note: The status of Automatic Barcode Gap Discovery (ABGD) groups for these taxa are provided. The pattern of discordance between taxonomy and either BINs or ABGD groups is not identical. The number of diagnostic substitutions calculated with the nucDiag function in the R package spider are provided and where available the base pair positions relative to the standard 658 bp DNA barcode fragment for bees are provided. In BINs with more than two species, there may be additional pairwise diagnostic characters.

*Discordant with extra-limital species in BOLD.

†ABGD groups discordant with BINs.

‡ABGD groups partially discordant with BINs.

Fig. 1. Summary of minimum interspecific distance and maximum intraspecific distance for 110 species of *Lasioglossum* with DNA barcode-compliant sequences. Horizontal dashed line indicates 1.4% interspecific distance. Diagonal solid line indicates the expected barcode gap where interspecific divergence is greater than intraspecific divergence. Shape and shading indicate relationship of BIN to traditional taxonomic identification.

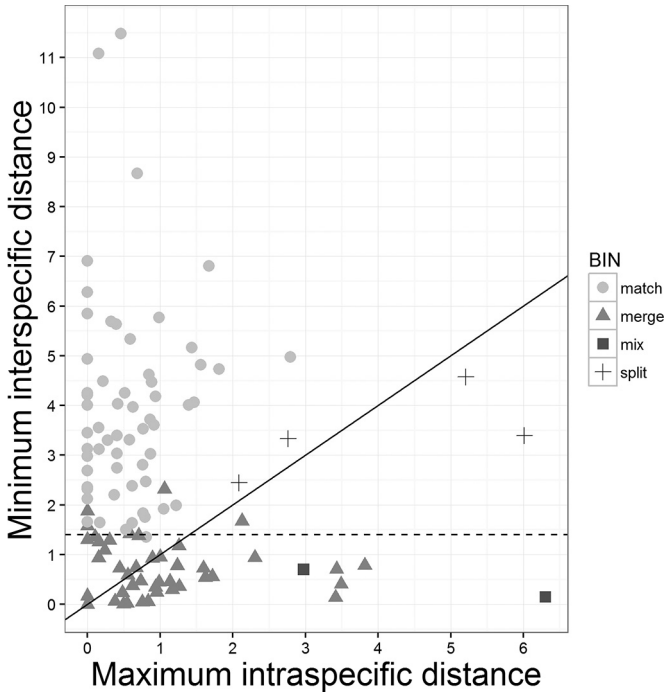
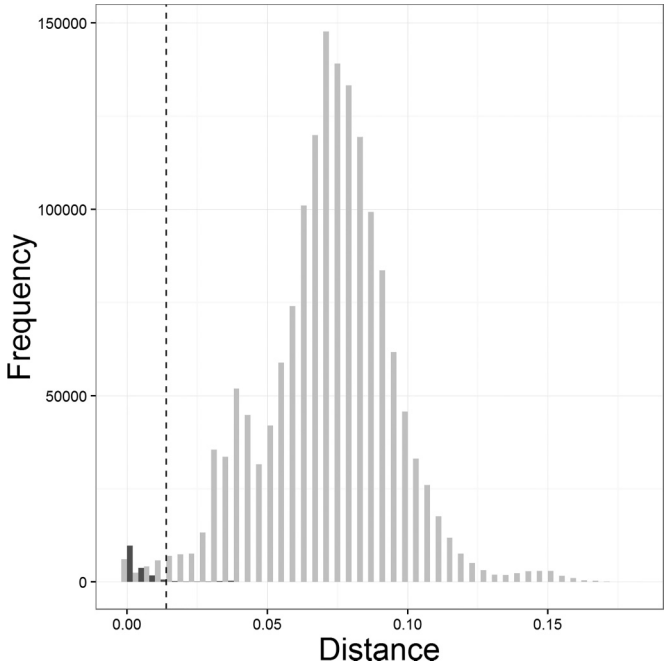


Fig. 2. Summary of DNA barcode gap from all pairwise comparison of 110 species of *Lasioglossum* with DNA barcode-compliant sequences. Most pairwise comparisons are between different species with large genetic distances (light gray bars). Intraspecific distances (dark gray bars) overlap extensively with interspecific distances, preventing the use of DNA barcode gap for correctly, reliably diagnosing some species. Vertical dashed line indicates 1.4% genetic distance.



sampling effort and the use of mean intraspecific divergences in some studies (e.g., Barrett and Hebert 2005) and the more relevant value, the mean maximum intraspecific divergence used here (Collins and Cruickshank 2012). Minimum distances between species pairs was on average four times the mean intraspecific genetic divergence but less than double the mean maximum intraspecific divergence. This gap between inter- and intraspecific genetic divergences is less than those of most barcoding studies and falls well below the 10-fold difference suggested in an earlier study of birds (Hebert et al. 2004*b*). Importantly, intra- and interspecific divergence varies among taxa and in recently diverged groups any simple criterion for delimiting species will have flaws.

Cases where DNA barcodes do not differ among species help to emphasize the importance of traditional taxonomic practices. In another group of bees, DNA barcodes do not distinguish morphologically similar but ecologically distinct species in the *Colletes succinctus* (L.) group, although differences in nuclear sequences were present (Kuhlmann et al. 2007). The *L. viridatum* species-group defined by Gibbs (2010*a*) is comprised of many of the most taxonomically challenging *L.* (*Dialictus*) species in Canada and the eastern United States. Members of this complex are often distinguishable based on multiple lines of evidence including ecological differences (e.g., *L. georgeickworti* Gibbs is the only member restricted to coastal areas of the northeastern United States and *L. subviridatum* nests in logs unlike the ground nests of most other species; Gibbs 2011). Unfortunately, few members of the *L. viridatum* group have had their nests discovered, and none has had their nesting biology described in detail. The nesting biology of *L. viridatum* described by Atwood (1933) is unreliable (Zarrillo et al. 2016), and may pertain to *L. laevissimum*. DNA barcodes of species in the *L. viridatum* group are among the most difficult to interpret, including a surprising 19 species lumped into a single BIN. It is tempting to consider this as a case of traditional taxonomy oversplitting species. However, this BIN includes morphologically distinct, geographically separated populations that occupy distinct ecological niches (Gibbs 2010*a*, 2011). It is possible that the species complex in its entirety has not been accurately delimited, but multiple lines of evidence suggest that it is not a single species (Gibbs 2010*a*, 2011). The absence of clear DNA barcode clusters is presumably due to incomplete lineage sorting or introgression in a species complex that has undergone recent and rapid diversification. A neighbour-joining analysis suggested that 27% of species in this study had non-monophyletic DNA barcode clusters. Neighbour-joining is not a preferred method of phylogenetic analysis, being generally outperformed by other methods of phylogenetic analysis that are better able to account for evolutionary rate variation (Huelsenbeck 1995; Felsenstein 2004). However, neighbour-joining was assessed here because it is commonly used to examine DNA barcode sequences and is the standard tree building algorithm used in BOLD, due to its speed and relatively good performance when sequences have recently diverged (Holder and Lewis 2003). The high level of paraphyly found here for *Lasioglossum* should be taken into account when interpreting results from neighbour-joining trees as part of future DNA barcoding efforts.

Although DNA barcodes may be insufficient for species identification using RESL or ABGD in some cases, fixed substitutions can be informative in an integrative taxonomic approach. *Lasioglossum* (*Dialictus*) is a recently derived subgenus that has diversified into many hundreds of species (Gibbs et al. 2012; Ascher and Pickering 2016). Closely related species that have recently separated will have evolved fewer neutral mutations than other species. Morphological characters, if under selection, can evolve rapidly (Owen and Harder 1995; Thompson 1998) and could result in clearly defined species with little genetic divergence in mitochondrial haplotypes. In addition, the possibility of introgression is presumably higher in closely related species and could confound specimen identification and species discovery using DNA barcodes by reducing intraspecific variation among species (Pentinsaari et al.

2014). *Lasioglossum hitchensi* and *L. weemsi* are closely related, but distinguishable based on hair patterns of the metasoma; however, a single individual of *L. hitchensi* was found to have a DNA barcode consistent with *L. weemsi*. This example is a candidate for introgression between closely related species.

In some cases, character-based methods of identification may still allow correct determination of species where distance-based methods are misleading (Burns et al. 2007). Previous results from species of *Lasioglossum* suggest that unique fixed substitutions or unique polymorphism patterns may be sufficient to distinguish species in the absence of a clear DNA barcode gap (Gibbs 2009*a*, 2009*b*; Gibbs et al. 2013). Similar methods have also been used in the *Bombus lucorum* (L.) species complex of bumble bees (Murray et al. 2007; Bertsch 2009) and Hawaiian *Hylaeus* (Magnacca and Brown 2010). All of the species recognized in the taxonomic revisions are diagnosable using morphological characters (Gibbs 2010*b*, 2011). If voucher specimens are retained, then additional morphological or geographic characters could be used in cases where DNA barcodes fail entirely. For example, a DNA barcode sequence matching that of *L. versatum* and *L. callidum*, which share an identical DNA barcode haplotype, can be differentiated by examining morphological characters, including the shape of the female mandible and protrochanter and male facial pubescence (Gibbs 2010*a*, 2011). Although DNA barcodes do not distinguish *L. tenax* (Sandhouse) and *L. cattellae* (Ellis), these two species can be distinguished by geographical distribution alone, in addition to aspects of their microsculpture, colour, and pubescence (Gibbs 2010*a*). *Lasioglossum tenax* is an alpine/boreal species restricted to Canada and the Rocky Mountains, while *L. cattellae* is largely an open grassland species in the mid-Atlantic and Midwestern states of the United States. The BIN system on BOLD allows for expert opinion to annotate BINs believed to include multiple species (Ratnasingham and Hebert 2013). Until BINs are fully annotated and taxonomic revisions are complete, users would be wise to treat BINs with care and use alternative data, including morphology and geographical distribution, to make species identifications using DNA barcodes.

In conclusion, DNA barcodes in *L.* (*Dialictus*) lack a well-defined barcode gap that can be used for delimiting species, and RESL and ABGD worked unambiguously in only about half the cases. Nevertheless, DNA barcodes are sufficient for identifying specimens of many species in this taxonomically challenging group even with simple distance-based methods. Character-based identifications have the potential for a greater level of success in identifying specimens with low interspecific genetic divergence. Unique combinations of nucleotide substitutions may allow specimen identification even when interspecific divergences are low and fixed nucleotide substitutions are absent (Gibbs et al. 2013). Although some species pairs cannot be identified using DNA barcodes, the sequences are nevertheless informative for reducing the list of possible species names. Simple annotations to the BOLD system (e.g., *viridatum* complex) could flag such species for additional study or sequencing of additional loci. For this reason, I reiterate the need of morphology-based taxonomy and integrative approaches using multiple loci as further tests of species identity. DNA barcodes, even with a weak barcode gap, are nevertheless a useful taxonomic tool for bees when used in conjunction with morphology, behaviour, and other data.

Acknowledgements

Laurence Packer (York University) was instrumental to the success of this project, providing full support for the research at all stages and providing valuable comments and additions to the manuscript. The editor and two anonymous reviewers also provided useful comments that helped to improve the manuscript greatly. This research was supported through funding to the Canadian Barcode of Life Network from Genome Canada, NSERC, and other sponsors listed at www.BOLNET.ca. Ontario Graduate Scholarships in Science and Technology awarded during the time of data collection were greatly appreciated.

References

Alexander, L.C., Delion, M., Hawthorne, D.J., Lamp, W.O., and Funk, D.H. 2009. Mitochondrial lineages and DNA barcoding of closely related species in the mayfly genus *Ephemerella* (Ephemeroptera: Ephemerellidae). J. N. Am. Benthol. Soc. **28**(3): 584–595. doi:10.1899/08-150.1.

Almeida, E., Laurenne, N., Kuhlmann, M., and Quicke, D. 2009. Molecular phylogeny and historical biogeography of the bee genus *Colletes* Latreille, 1802 (Hymenoptera: Apiformes: Colletidae), based on mitochondrial COI and nuclear 28S sequence data. Insect Syst. Evol. **40**(3): 291–318. doi:10.1163/139956009X12475840653733.

Ascher, J.S., and Pickering, J. 2016. Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). Draft-45. Available from http://www.discoverlife.org/mp/20q?guide=Apoidea_species.

Atwood, C.E. 1933. Studies on the Apoidea of western Nova Scotia with special reference to visitors to apple bloom. Can. J. Res. **9**(5): 443–457. doi:10.1139/cjr33-101.

Barrett, R.D.H., and Hebert, P.D.N. 2005. Identifying spiders through DNA barcodes. Can. J. Zool. **83**(3): 481–491. doi:10.1139/z05-024.

Bertsch, A. 2009. Barcoding cryptic bumblebee taxa: *B. lucorum*, *B. crytarum* and *B. magnus*, a case study. Beiträge zur Entomol. **59**(2): 287–310.

Brown, S.D.J., Collins, R.A., Boyer, S., Lefort, M.-C., Malumbres-Olarte, J., Vink, C.J., and Cruickshank, R.H. 2012. SPIDER: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. **12**(3): 562–565. doi:10.1111/j.1755-0998.2011.03108.x. PMID:22243808.

Burns, J.M., Janzen, D., Hajibabaei, M., Hallwachs, W., and Hebert, P.D.N. 2007. DNA barcodes of closely related (but morphologically and ecologically distinct) species of skipper butterflies (Hesperiidae) can differ by only one to three nucleotides. J. Lepid. Soc. **61**: 138–153.

Campbell, J.W., Hanula, J.L., and Waldrop, T.A. 2007. Effects of prescribed fire and fire surrogates on floral visiting insects of the blue ridge province in North Carolina. Biol. Conserv. **134**: 393–404. doi:10.1016/j.biocon.2006.08.029.

Carman, G.M., and Packer, L. 1996. A cryptic species allied to *Halictus ligatus* Say (Hymenoptera: Halictidae) detected by allozyme electrophoresis. J. Kans. Entomol. Soc. **69**: 168–176.

Clare, E.L., Lim, B.K., Engstrom, M.D., Eger, J.L., and Hebert, P.D.N. 2007. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Mol. Ecol. Notes, **7**(2): 184–190. doi:10.1111/j.1471-8286.2006.01657.x.

Collins, R.A., and Cruickshank, R.H. 2012. The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. **13**(6): 969–975. doi:10.1111/1755-0998.12046. PMID:23280099.

Coyne, J.A., and Orr, H.A. 2004. Speciation. Sinauer Associates, Inc., Sunderland, Mass.

Danforth, B.N., Mitchell, P.L., and Packer, L. 1998. Mitochondrial DNA differentiation between two cryptic *Halictus* (Hymenoptera: Halictidae) species. Ann. Entomol. Soc. Am. **91**(4): 387–391. doi:10.1093/aesa/91.4.387.

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**(3): 605–618. doi:10.1006/mpev.1999.0670. PMID:10620417.

Dasmahapatra, K.K., Elias, M., Hill, R.I., Hoffman, J.I., and Mallet, J. 2010. Mitochondrial DNA barcoding detects some species that are real, and some that are not. Mol. Ecol. Resour. **10**(2): 264–273. doi:10.1111/j.1755-0998.2009.02763.x. PMID:21565021.

Dayrat, B. 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. **85**(3): 407–415. doi:10.1111/j.1095-8312.2005.00503.x.

DeSalle, R., Egan, M.G., and Siddall, M. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos. Trans. R. Soc. B Biol. Sci. **360**(1462): 1905–1916. doi:10.1098/rstb.2005.1722.

Droege, S., Tepedino, V.J., LeBuhn, G., Link, W., Minckley, R.L., Chen, Q., and Conrad, C. 2010. Spatial patterns of bee captures in North American bowl trapping surveys. Insect Conserv. Divers. **3**: 15–23. doi:10.1111/j.1752-4598.2009.00074.x.

Dunn, M., Mitchell, P.L., and Packer, L. 1998. Phenology and social biology of two sibling species of *Halictus* in an area of sympatry. Can. J. Zool. **76**(12): 2207–2213. doi:10.1139/z98-163.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5): 1792–1797. doi:10.1093/nar/gkh340. PMID:15034147.

Felsenstein, J. 2004. Inferring phylogenies. Sinauer Associates, Inc., Sunderland, Mass.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. **3**(5): 294–299. PMID:7881515.

Gibbs, J. 2009*a*. Integrative taxonomy identifies new (and old) species in the *Lasioglossum* (*Dialictus*) *tegulare* (Robertson) species group (Hymenoptera, Halictidae). Zootaxa, **2032**: 1–38.

Gibbs, J.J. 2009*b*. New species in the *Lasioglossum petrellum* species group identified through an integrative taxonomic approach. Can. Entomol. **141**(1): 371–396. doi:10.4039/n09-020.

Gibbs, J. 2010*a*. Revision of the metallic species of *Lasioglossum* (*Dialictus*) in Canada (Hymenoptera, Halictidae, Halictini). Zootaxa, **3073**: 1–382.

Gibbs, J. 2010*b*. Atypical wing venation in *Dialictus* and *Hemihalictus* and its implications for subgeneric classification of *Lasioglossum*. Psyche, **2010**: 605390. doi:10.1155/2010/605390.

Gibbs, J. 2011. Revision of the metallic *Lasioglossum* (*Dialictus*) of eastern North America (Hymenoptera: Halictidae: Halictini). Zootaxa, **3073**: 1–216.

Gibbs, J., Brady, S.G., Kanda, K., and Danforth, B.N. 2012. Phylogeny of halictine bees supports a shared origin of eusociality for *Halictus* and *Lasioglossum* (Apoidea: Anthophila: Halictidae). Mol. Phylogenet. Evol. **65**(3): 926–939. doi:10.1016/j.ympev.2012.08.013. PMID:22982437.

Gibbs, J., Packer, L., Dumesh, S., and Danforth, B.N. 2013. Revision and reclassification of *Lasioglossum* (*Evylaeus*), *L.* (*Hemihalictus*) and *L.* (*Sphecodogastra*) in eastern North America (Hymenoptera: Apoidea: Halictidae). Zootaxa, **3672**: 1–117. doi:10.11646/zootaxa.3672.1.1. PMID:26146702.

Giles, V., and Ascher, J.S. 2006. A survey of the bees of the Black Rock Forest Preserve, New York (Hymenoptera: Apoidea). J. Hymenopt. Res. **15**(2): 208–231.

Goldstein, P.Z., and DeSalle, R. 2011. Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. Bioessays, **33**(2): 135–147. doi:10.1002/bies.201000036. PMID:21184470.

González-Vaquero, R.A., Roig-Alsina, A., and Packer, L. 2016. DNA barcoding as a useful tool in the systematic study of wild bees of the tribe Augochlorini (Hymenoptera: Halictidae). Genome, **59**(10): 889–898. doi:10.1139/gen-2016-0006. PMID:27673501.

Grixti, J.C., and Packer, L. 2006. Changes in the bee fauna (Hymenoptera: Apoidea) of an old field site in southern Ontario, revisited after 34 years. Can. Entomol. **138**(2): 147–164. doi:10.4039/n05-034.

Groom, S.V.C., Stevens, M.I., and Schwarz, M.P. 2013. Diversification of Fijian halictine bees: Insights into a recent island radiation. Mol. Phylogenet. Evol. **68**(3): 582–594. doi:10.1016/j.ympev.2013.04.015. PMID:23629054.

Hajibabaei, M., deWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., et al. 2005. Critical factors for assembling a high volume of DNA barcodes. Philos. Trans. R. Soc. B Biol. Sci. **360**(1462): 1959–67. doi:10.1098/rstb.2005.1727.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W., and Hebert, P.D.N. 2006*a*. DNA barcodes distinguish species of tropical Lepidoptera. Proc. Natl. Acad. Sci. **103**(4): 968–971. doi:10.1073/pnas.0510466103. PMID:16418261.

Hajibabaei, M., Smith, M.A., Janzen, D.H., Rodriguez, J.J., Whitfield, J.B., and Hebert, P.D.N. 2006*b*. A minimalist barcode can identify a specimen whose DNA is degraded. Mol. Ecol. Notes, **6**(4): 959–964. doi:10.1111/j.1471-8286.2006.01470.x.

Hebert, P.D.N., and Gregory, T.R. 2005. The promise of DNA barcoding for taxonomy. Syst. Biol. **54**(5): 852–859. doi:10.1080/10635150500354886. PMID:16243770.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., and Jeremy, R. 2003*a*. Biological identifications through DNA barcodes. Proc. Biol. Sci. **270**(1512): 313–321. doi:10.1098/rspb.2002.2218. PMID:12614582.

Hebert, P.D.N., Ratnasingham, S., and DeWaard, J.R. 2003*b*. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. **270**(S1): S96–S99. doi:10.1098/rsbl.2003.0025.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., and Hallwachs, W. 2004*a*. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. **101**(41): 14812–14817. doi:10.1073/pnas.0406166101. PMID:15465915.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. 2004*b*. Identification of birds through DNA barcodes. PLoS Biol. **2**(10): e312. doi:10.1371/journal.pbio.0020312. PMID:15455034.

Hendrixson, B.E., DeRussy, B.M., Hamilton, C.A., and Bond, J.E. 2013. An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, Aphonopelma). Mol. Phylogenet. Evol. **66**(1): 327–340. doi:10.1016/j.ympev.2012.10.004. PMID:23092751.

Holder, M., and Lewis, P.O. 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet. **4**: 275–284. doi:10.1038/nrg1044. PMID:12671658.

Hubert, N., Hanner, R., Holm, E., Mandrak, N.E., Taylor, E., Burridge, M., et al. 2008. Identifying Canadian freshwater fishes through DNA barcodes. PLoS ONE, **3**(6): e2490. doi:10.1371/journal.pone.0002490. PMID:22423312.

Huelsenbeck, J.P. 1995. The robustness of two phylogenetic methods: four-taxon simulations reveal a slight superiority of maximum likelihood over neighbor joining. Mol. Biol. Evol. **12**(5): 843–849. doi:10.1093/oxfordjournals.molbev.a040261. PMID:7476130.

Ivanova, N.V., DeWaard, J.R., and Hebert, P.D.N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Mol. Ecol. Notes, **6**(4): 998–1002. doi:10.1111/j.1471-8286.2006.01428.x.

Kalhorn, K.D., Barrows, E.M., and LaBerge, W.E. 2003. Bee (Hymenoptera: Apoidea: Apiformes) diversity in an Appalachian shale barrens. J. Kans. Entomol. Soc. **76**(3): 455–468.

Kearns, C.A., and Oliveras, D.M. 2009. Boulder County Bees revisited: a resampling of Boulder Colorado bees a century later. J. Insect Conserv. **13**(6): 603–613. doi:10.1007/s10841-009-9211-8.

Kekkonen, M., and Hebert, P.D.N. 2014. DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. Mol. Ecol. Resour. **14**(4): 706–715. doi:10.1111/1755-0998.12233. PMID:24479435.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**(2): 111–120. doi:10.1007/BF01731581. PMID:7463489.

Klein, A.-M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C., and Tscharntke, T. 2007. Importance of pollinators in changing landscapes for world crops. Proc. R. Soc. B Biol. Sci. **274**(1608): 303–313. doi:10.1098/rspb.2006.3721.

Knerer, G., and Atwood, C.E. 1964. Further notes on the genus *Evylaeus* Robertson (Hymenoptera: Halictidae). Can. Entomol. **96**(7): 957–962. doi:10.4039/Ent96957-7.

Kuhlmann, M., Else, G., Dawson, A., and Quicke, D. 2007. Molecular, biogeographical and phenological evidence for the existence of three western European sibling species in the *Colletes succinctus* group (Hymenoptera: Apidae). Org. Divers. Evol. **7**(2): 155–165. doi:10.1016/j.ode.2006.04.001.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. **33**(7): 1870–1874. doi:10.1093/molbev/msw054. PMID:27004904.

MacKay, P.A., and Knerer, G. 1979. Seasonal occurrence and abundance in a community of wild bees from an old field habitat in southern Ontario. Can. Entomol. **111**(3): 367–376. doi:10.4039/Ent111367-3.

Maddison, W.P., and Maddison, D.R. 2017. Mesquite: a modular system for evolutionary analysis. Version 3.2. Available from http://mesquiteproject.org.

Magnacca, K.N., and Brown, M.J.F. 2010. Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus* (*Nesoprosopis*) bees (Hymenoptera: Colletidae). BMC Evol. Biol. **10**: 174. doi:10.1186/1471-2148-10-174. PMID:20540728.

Magnacca, K.N., and Brown, M.J.F. 2012. DNA barcoding a regional fauna: Irish solitary bees. Mol. Ecol. Resour. **12**(6): 990–998. doi:10.1111/1755-0998.12001. PMID:22931682.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L., and Hedin, M. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. **55**(5): 715–728. doi:10.1080/10635150600969864. PMID:17060194.

Meyer, C.P., and Paulay, G. 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. **3**(12): e422. doi:10.1371/journal.pbio.0030422. PMID:16336051.

Michener, C.D. 1974. The social behavior of the bees. Belknap Press, Cambridge, Mass.

Michener, C.D. 1979. Biogeography of the bees. Ann. Mo. Bot. Gard. **66**(3): 277. doi:10.2307/2398833.

Michener, C.D. 2007. The bees of the world. 2nd Edition. The Johns Hopkins University Press, Baltimore, Md.

Moure, J.S., and Hurd, P.D. 1987. An annotated catalog of the halictid bees of the Western Hemisphere (Hymenoptera: Halictidae). Smithsonian Institution Press, Washington, D.C.

Murray, T.E., Fitzpatrick, Ú., Brown, M.J.F., and Paxton, R.J. 2007. Cryptic species diversity in a widespread bumble bee complex revealed using mitochondrial DNA RFLPs. Conserv. Genet. **9**(3): 653–666. doi:10.1007/s10592-007-9394-z.

Ngo, H.T., Gibbs, J., Griswold, T., and Packer, L. 2013. Evaluating bee (Hymenoptera: Apoidea) diversity using Malaise traps in coffee landscapes of Costa Rica. Can. Entomol. **145**(4): 435–453. doi:10.4039/tce.2013.16.

Nicholls, J.A., Challis, R.J., Mutun, S., and Stone, G.N. 2012. Mitochondrial barcodes are diagnostic of shared refugia but not species in hybridizing oak gallwasps. Mol. Ecol. **21**(16): 4051–4062. doi:10.1111/j.1365-294X.2012.05683.x. PMID:22724511.

Ollerton, J., Winfree, R., and Tarrant, S. 2011. How many flowering plants are pollinated by animals? Oikos, **120**(3): 321–326. doi:10.1111/j.1600-0706.2010.18644.x.

Owen, R.E., and Harder, L.D. 1995. Heritable allometric variation in bumble bees: opportunities for colony-level selection of foraging ability. J. Evol. Biol. **8**(6): 725–738. doi:10.1046/j.1420-9101.1995.8060725.x.

Packer, L. 1999. The distribution of *Halictus ligatus* Say and *H. poeyi* Lep. (Hymenoptera: Halictidae) in North America. Univ. Kansas Nat. Hist. Museum Spec. Publ. **24**: 81–84.

Packer, L., and Ruz, L. 2017. DNA barcoding the bees (Hymenoptera: Apoidea) of Chile: species discovery in a reasonably well known bee fauna with the description of a new species of *Lonchopria* (Colletidae). Genome, **60**(5): 414–430. doi:10.1139/gen-2016-0071. PMID:28177842.

Packer, L., and Taylor, J.S. 1997. How many cryptic species are there? An application of the phylogenetic species concept to genetic data for some comparatively well known bee "species." Can. Entomol. **129**(4): 587–594. doi:10.4039/Ent129587-4.

Packer, L., and Taylor, J.S. 2002. Genetic variation within and among populations of an arctic/alpine sweat bee (Hymenoptera: Halictidae). Can. Entomol. **134**(5): 619–631. doi:10.4039/Ent134619-5.

Packer, L., Gibbs, J., Sheffield, C., and Hanner, R. 2009. DNA barcoding and the mediocrity of morphology. Mol. Ecol. Resour. **9**(s1): 42–50. doi:10.1111/j.1755-0998.2009.02631.x. PMID:21564963.

Packer, L., Ali, E., Dumesh, S., and Walker, K. 2016. The identification of pollinators: Where are we and where should we go? *In* Pollination services to agriculture: Sustaining and enhancing a key ecosystem service. *Edited by* B. Gemmill-Herren. Routledge, New York. pp. 57–73.

Padial, J.M., Miralles, A., De la Riva, I., and Vences, M. 2010. The integrative future of taxonomy. Front. Zool. **7**: 16. doi:10.1186/1742-9994-7-16. PMID: 20500846.

Page, T.J., Choy, S.C., and Hughes, J.M. 2005. The taxonomic feedback loop: symbiosis of morphology and molecules. Biol. Letters **1**: 139–142. doi:10.1098/rsbl.2005.0298.

Patten, M.M., Carioscia, S.A., and Linnen, C.R. 2015. Biased introgression of mitochondrial and nuclear genes: a comparison of diploid and haplodiploid systems. Mol. Ecol. **24**(20): 5200–5210. doi:10.1111/mec.13318. PMID:26173469.

Pauly, A., Devalez, J., Sonet, G., Nagy, Z.T., and Boevé, J.-L. 2015. DNA barcoding and male genital morphology reveal five new cryptic species in the West Palearctic bee *Seladonia smaragdula* (Vachal, 1895) (Hymenoptera: Apoidea: Halictidae). Zootaxa, **4034**(2): 257. doi:10.11646/zootaxa.4034.2.2. PMID: 26624441.

Pentinsaari, M., Mutanen, M., and Kaila, L. 2014. Cryptic diversity and signs of mitochondrial introgression in the *Agrilus viridis* species complex (Coleoptera: Buprestidae). Eur. J. Entomol. **111**(4): 475–486. doi:10.14411/eje.2014.072.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., et al. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. **55**(4): 595–609. doi:10.1080/10635150600852011. PMID: 16967577.

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol. Ecol. **21**(8): 1864–1877. doi:10.1111/j.1365-294X.2011.05239.x. PMID:21883587.

R Core Team. 2015. R: a language and environment for statistical computing. R version 3.2.2. R Foundation for Statistical Computing.

Ratnasingham, S., and Hebert, P.D.N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Mol. Ecol. Notes, **7**(3): 355–364. doi:10.1111/j.1471-8286.2007.01678.x. PMID:18784790.

Ratnasingham, S., and Hebert, P.D.N. 2013. A DNA-based registry for all animal species: the barcode index number (BIN) system. PLoS ONE, **8**(7): e66213. doi:10.1371/journal.pone.0066213. PMID:23861743.

Rehan, S.M., and Sheffield, C.S. 2011. Morphological and molecular delineation of a new species in the *Ceratina dupla* species-group (Hymenoptera: Apidae: Xylocopinae) of eastern North America. Zootaxa, **2873**: 35–50.

Rheindt, F.E., Christidis, L., and Norman, J.A. 2009. Genetic introgression, incomplete lineage sorting and faulty taxonomy create multiple cases of polyphyly in a montane clade of tyrant-flycatchers (*Elaenia*, Tyrannidae). Zool. Scr. **38**(2): 143–153. doi:10.1111/j.1463-6409.2008.00369.x.

Rivera, J., and Currie, D.C. 2009. Identification of Nearctic black flies using DNA barcodes (Diptera: Simuliidae). Mol. Ecol. Resour. **9**(s1): 224–236. doi:10.1111/j.1755-0998.2009.02648.x. PMID:21564982.

Schmidt, S., Schmid-Egger, C., Morinière, J., Haszprunar, G., and Hebert, P.D.N. 2015. DNA barcoding largely supports 250 years of classical taxonomy: identifications for Central European bees (Hymenoptera, Apoidea partim). Mol. Ecol. Resour. **15**(4): 985–1000. doi:10.1111/1755-0998.12363. PMID:25588628.

Sheffield, C.S., Hebert, P.D.N., Kevan, P.G., and Packer, L. 2009. DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Mol. Ecol. Resour. **9**(s1): 196–207. doi:10.1111/j.1755-0998.2009.02645.x. PMID:21564979.

Sheffield, C.S., Dumesh, S., and Cheryomina, M. 2011. *Hylaeus punctatus* (Hymenoptera: Colletidae), a bee species new to Canada, with notes on other non-native species. J. Entomol. Soc. Ontario, **142**: 29–43.

Sheffield, C.S., Richardson, L., Cannings, S., Ngo, H., Heron, J., and Williams, P.H. 2016. Biogeography and designatable units of *Bombus occidentalis* Greene and *B. terricola* Kirby (Hymenoptera: Apidae) with implications for conservation status assessments. J. Insect Conserv. **20**(2): 189–199. doi:10.1007/s10841-016-9853-2.

Smith, B.A., Brown, R.L., LaBerge, W., and Griswold, T.L. 2012. A faunistic survey of bees (Hymenoptera: Apoidea) in the Black Belt Prairie of Mississippi. J. Kans. Entomol. Soc. **85**(1): 32–47. doi:10.2317/JKES111025.1.

Smith, M.A., Fisher, B.L., and Hebert, P.D.N. 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos. Trans. R. Soc. B Biol. Sci. **360**(1462): 1825–1834. doi:10.1098/rstb.2005.1714.

Smith, M.A., Wood, D.M., Janzen, D.H., Hallwachs, W., and Hebert, P.D.N. 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. Proc. Natl. Acad. Sci. **104**(12): 4967–4972. doi:10.1073/pnas.0700050104. PMID:17360352.

Smith, M.A., Bertrand, C., Crosby, K., Eveleigh, E.S., Fernandez-Triana, J., Fisher, B.L., et al. 2012. Wolbachia and DNA barcoding insects: patterns, potential, and problems. PLoS ONE, **7**(5): e36514. doi:10.1371/journal.pone.0036514. PMID:22567162.

Sonet, G., Jordaens, K., Nagy, Z.T., Breman, F., de Meyer, M., Backeljau, T., and Virgilio, M. 2013. Adhoc: an R package to calculate ad hoc distance thresholds for DNA barcoding identification. Zookeys, **365**: 329–336. doi:10.3897/zookeys.365.6034.

Srivathsan, A., and Meier, R. 2012. On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics, **28**: 190–194. doi:10.1111/j.1096-0031.2011.00370.x.

Thompson, J.N. 1998. Rapid evolution as an ecological process. Trends Ecol. Evol. **13**(8): 329–332. doi:10.1016/S0169-5347(98)01378-0. PMID:21238328.

Virgilio, M., Jordaens, K., Breman, F.C., Backeljau, T., and De Meyer, M. 2012. Identifying Insects with Incomplete DNA Barcode Libraries, African Fruit Flies (Diptera: Tephritidae) as a Test Case. PLoS ONE, **7**(2): e31581. doi:10.1371/journal.pone.0031581. PMID:22359600.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., and Hebert, P.D.N. 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. B Biol. Sci. **360**(1462): 1847–1857. doi:10.1098/rstb.2005.1716.

Wheeler, W.M. 1928. The social insects: their origin and evolution. Kegan Paul, Trench, Trubner and Co., Ltd. London, UK.

Wheelock, M.J., and O'Neal, M.E. 2016. Insect pollinators in Iowa cornfields: Community identification and trapping method analysis. PLoS ONE, **11**(7): e0143479. doi:10.1371/journal.pone.0143479. PMID:27459648.

Wolf, A.T., and Ascher, J.S. 2009. Bees of Wisconsin (Hymenoptera: Apoidea: Anthophila). Gt. Lakes Entomol. **41**: 129–168.

Zarrillo, T.A., Ascher, J.S., Gibbs, J., and Stoner, K.A. 2016. New and noteworthy records of bees (Hymenoptera: Apoidea: Anthophila) for Connecticut. J. Kans. Entomol. Soc. **89**(2): 138–157. doi:10.2317/0022-8567-89.2.138.

Zayed, A. 2009. Bee genetics and conservation. Apidologie, **40**(3): 237–262. doi: 10.1051/apido/2009026.

Zayed, A., and Packer, L. 2005. Complementary sex determination substantially increases extinction proneness of haplodiploid populations. Proc. Natl. Acad. Sci. **102**(30): 10742–10746. doi:10.1073/pnas.0502271102. PMID:16020532.

Zayed, A., Roubik, D.W., and Packer, L. 2004. Use of diploid male frequency data as an indicator of pollinator decline. Proc. R. Soc. B Biol. Sci. **271**(Suppl.): S9–S12. doi:10.1098/rsbl.2003.0109.