VIII ENCONTRO SOBRE ABELHAS
Biodiversidade e Uso Sustentado de Abelhas

23 a 26 de julho de 2008
Ribeirão Preto - SP - Brasil
Local: Hotel JP - Ribeirão Preto - SP
Barcoding the Bees of the World

Laurence Packer¹, Jason Gibbs¹, Cory Sheffield¹ and Peter Kevan²
¹Department of Biology, York University, 4700 Keele St., Toronto, ON, M3J 1P3, Canada
²Department of Environmental Biology, University of Guelph, Guelph, ON N1G 2W1, Canada
Corresponding Author: Laurence Packer, Department of Biology, York University, 4700 Keele St.,
Toronto, ON., M3J 1P3, Canada, 416 736 2100 ext 66524, laurencepacker@yahoo.com

Abstract

We demonstrate the utility of DNA barcoding as a taxonomic tool for the study of bee diversity
using two approaches. First, we demonstrate that the technique works well for the identification of all
bee species present in a geographically restricted region for which our taxonomic understanding of
bees was considered to be good. Secondly, we use this approach, combined with traditional morpho-
logical analysis, in a study of a taxonomically extremely difficult group of bees, the subgenus Dialictus
(family Halictidae; genus Lasio glossum). Here, DNA barcoding proved essential for the delimita-
tion of numerous species that were morphologically almost indistinguishable. We conclude that DNA
barcoding is efficient at the detection of cryptic species, associating the sexes of dimorphic species,
associating the castes of species with strong queen-worker dimorphism and as a generally useful tool for
basic identification. A global campaign to barcode the bees of the world has been initiated; its inaugural
meeting was held at York University in May, 2008 with Fernando Silveira representing Brazil.

Introduction

Identifying and describing biodiversity has long been the purview of taxonomists able to rec-
ognize subtle morphological differences among species. DNA barcoding is a new method that prom-
ises to speed taxonomic progress and allow accurate identification of species, even by those without
taxonomic expertise (Hebert et al., 2003).

DNA barcoding employs a short strand of a standard gene, the mitochondrial gene COI, to
identify species (Figure 1).

Figure 1. Schematic diagram of animal mitochondrial DNA showing the barcode region.
The principle behind DNA barcoding is that the nucleotide sequence diversity of the DNA barcode gene fragment is greater between species than within species. This indeed has been found to be the case in the vast majority of taxa upon which the method has been tested. Another useful feature is the relatively cheap costs associated with semi-automated identifications. Currently, it costs approximately $7 to obtain a DNA barcode from a pinned specimen, a small amount in comparison to the cost of a taxonomist’s time for a group requiring careful examination, and perhaps genitalic dissection. Moreover, the costs are expected to decrease as methods improve.

Accurate delimitation of species is crucial for understanding biodiversity. Especially problematic can be “easily recognized” species that may not be examined in as great detail as other, more difficult, species but may be a repository for cryptic species (Packer and Taylor, 1997). Integrative taxonomic approaches that combine morphological, molecular and other types of data are the best methods for describing biodiversity (Dayrat, 2005; Page et al., 2005), but DNA barcoding is increasing the rate at which new species are being discovered (Hebert et al., 2004) providing useful material for morphological comparison and description.

The large number of highly diverse genera and subgenera among the bees makes species recognition by traditional difficult even though the task is a crucially important one because of the role of bees in pollination and their potential for environmental monitoring (Kevan et al., 1997; Zayed and Packer, 2005; NAS 2007). In addition to the sexual dimorphism that often makes association of the sexes in rare insects difficult, many bees have the complexity of discrete or continuous caste-based variation. For example, in the halictine bees, differences between queens and workers of the same species are often much greater than the differences between sister species. The extreme situation is exhibited by the Brazilian species Halictus lanai (Moure): the workers are difficult to distinguish from those of H. hesperus Smith, but queens and workers of each species have non-overlapping size distributions and allometric variation that results in entirely different morphologies between the castes (Janjic and Packer, 2001).

We provide information on the current status of DNA barcoding of bees by making special reference to two detailed studies. One is an analysis of an entire bee fauna (the bees of Nova Scotia, Canada), a fauna that is considered to be well known taxonomically (Sheffield et al., 2003). The second is an analysis of a subgenus (Dialictus of the genus Lasioglossum), one that is exceptionally complex taxonomically because of morphological monotony, large numbers of single sex species descriptions and caste differences within many of the known species. The results indicate that DNA barcoding works well at both scales.

**Materials and Methods**

Bees were collected and pinned, or retained in ethanol or propylene glycol until required for DNA barcoding. One midleg was removed from each specimen and forwarded to the Biodiversity Institute of Ontario for sequencing. DNA extraction, PCR and sequencing were done at the Canadian Centre for DNA Barcoding at Guelph University (Guelph, Ontario) using standard protocols described elsewhere (Hebert et al., 2003) and available online at http://www.dnabarcoding.ca/pa/ge/research/protocols. Universal primers for amplifying the DNA barcode sequence for insects (LCO1490 and HCO2198; Folmer et al., 2004 or the variants LEPF and LEPR; Hebert et al., 2004) were used. All sequences were uploaded to the Barcode of Life Database (BOLD; Ratnasingham and Hebert, 2007) and are available for species identification purposes.
We present the results of both studies using a neighbour-joining (NJ) algorithm implemented using the analytical tools within BOLD. The NJ trees show the level of divergence among taxa. The use of such phenetic methods has been criticised, with some justification. Consequently, for the taxonomically intensive study of *Dialictus* we also show the data as a phylogenetic tree obtained from an equal weights parsimony analysis using Winclada (Nixon, 2002) and NONA (Goloboff, 1999). For clarity, only a small subset of the *Dialictus* species available were included.

**Results**

Data for the bees of Nova Scotia indicated that all 150 species could be readily distinguished using DNA barcoding. Sequence divergence within species was typically less than 1%, instances of larger differences proving to be misidentifications. Interspecific sequence divergences were mostly an order of magnitude higher, with only a few instances of divergence of less than 4%. An example of the results obtained for bees in one particularly difficult genus, *Andrena*, from Nova Scotia is presented in figure 2.

![Phenogram](image)

**Figure 2.** A neighbour joining phenogram for members of the genus *Andrena* sampled in Nova Scotia, Canada. Large sequence divergences between species are clear in this example.

Among other results obtained from this first geographic survey of bees using DNA barcoding are the discovery of a cryptic species of the genus *Ceratina* and a synonymy within the genus *Sphexodes*. In the latter, males of one species (known only from males) and the females of another (known only from females) were shown to have identical barcodes.

For the taxonomically more intensive study, we chose the subgenus *Dialictus* (Halictidae: *Lasioglossum*); one of the most taxonomically difficult groups amongst the bees. In North America, *Dialictus* are both speciose (over 270 currently recognized names) and often the most commonly collected subgenus of bee (MacKay and Knerer, 1979; Gritti and Packer, 2006; Campbell et al., 2007). *Dialictus* are also notoriously difficult to identify to species because of their highly monotonous morphology (Michener, 2007; Wheeler, 1928). In most cases, only extremely subtle differences can be used to dif-
ferentiate closely related species. Very little taxonomic progress has been made on this group in the last forty years (Mitchell, 1960; Knerer and Atwood, 1966), and there is much work that needs to be done.

An ongoing taxonomic revision of the North American species of *Dialictus* has revealed new species, new synonymies and the incorrectness of past synonymies. The eastern North American species of *Dialictus* were revised by Mitchell (1960). Four new species from Ontario were described by Knerer and Atwood (1966) but one of these must be synonymised. An integrative approach that uses DNA barcodes in combination with morphological comparison has identified an additional eleven new species in the eastern part of North America (Gibbs, in preparation). Twelve new species have also been found in western Canada but this area has not been as intensively studied, so this is a minimum estimate. Many species of *Dialictus* have been described from a single sex. Over twenty new synonymies have been identified among North American *Dialictus* and additional sampling (particularly of males) may allow additional associations. Less frequently, previous synonymies have been determined to be false. Five cases of these erroneous synonymies have been identified and in all cases the species can be differentiated by both morphology and DNA barcodes. Figure 3 shows an example of the complexity that can arise when dealing with this difficult group.
Figure 3. A small subset of the over 1500 sequences currently available for bees of the subgenus Dialictus. Figure 3A shows the results analysed using Neighbour-joining. Figure 3B shows the results of a parsimony analysis of the same taxa using L. rufitarse (Zetterstedt) as outgroup. In the latter example, only one individual was included for a species when there were multiple identical sequences. The existence of a cryptic species allied to L. pictum (Crawford) is shown here (as L. “pictum”). Subsequent study has revealed subtle diagnostic morphological differences between the two DNA sequence types. This figure also shows examples of synonymy detected using this method, within what will be known simply as L. semicaeruleum (Cockerell), two new synonyms for this species have been discovered so far (L. actuatum (Sandhouse) and L. pruinosiforme (Crawford)). All other results shown here are typical of the dataset as a whole: small intraspecific divergences and clearly larger interspecific differentiation.

Overall, the mean intraspecific divergence within Dialictus is less than 1% while the mean intrasubgeneric divergence is over 7%. Comparatively few species belong to clusters in which the DNA data and morphological data either separately or in combination, do not result in clear species boundaries. These amount to perhaps 3% of the total.

Discussion

The ability of DNA barcodes to distinguish between species depends on low levels of intraspecific variation and relatively high levels of interspecific variation. Within a geographically delimited survey of a whole community of bees, the minimal species level differentiation was close to 4% whereas within species variation was rarely more than 1%. The results were similar for a taxonomically intensive and geographically more extensive survey of Dialictus, a morphologically highly monotonous group of bees that have been the bane of North American melittologists for decades.

Dialictus presents a particular challenge for DNA barcoding because the group seems to have undergone rapid and frequent speciation. Dialictus comprises hundreds of species that likely evolved in the last 20 million years (Danforth et al., 2004). Such a rapid expansion, in combination with AT-biases in the third codon position (a common feature of many Hymenoptera), would be expected to cause low interspecific genetic variation in Dialictus. Nevertheless, in almost all cases there appears to be a distinct gap between the sequence divergences for within species variation and between species variation.

Even though often critisised for using phenetic methods, DNA barcode data is amenable to phylogenetic analysis, although this is not advisable for large datasets especially when there are many identical conspecific sequences. Using both maximum parsimony and Bayesian methods, we have shown that DNA barcode data are as useful as morphology in studies of Dialictus phylogeny (Gibbs, in preparation). Moreover, the results were of more than taxonomic interest in that the data show conclusively that parasitic lifestyles within this genus (these species were previously known under the generic name Paralictus as a result of their obvious loss of morphological adaptations for nest construction and pollen gathering) arose twice rather than once as previously supposed.

A global campaign to barcode the bees of the world has been initiated (see the website at: www.bee-bol.org). Preliminary surveys of numerous museum collections from around the world indicate that most specimens that were collected less than 15 years ago and kept dry in the interim yield good quality DNA barcode data. Even material older than that, including samples that had been subjected to relaxing for genitalic preparation, can provide useful information for species assignment even if the full length barcode is not attainable (Smith et al., in preparation).
The standard for sampling is to have from 5-10 specimens of each species sequenced. The locations of the samples should at least straddle the type locality (where known) and involve multiple sample sites. When cryptic species are discovered, additional samples will be required, including some from as close to the type locality as possible. For data to be entered into the database (see http://www.barcodinglife.org/views/login.php) full identification, identifier, label and repository information are required. At least one digital image per species is also expected.

Fernando Silveira is Brazil’s representative on the steering committee for barcoding the bees of the world.

References


Tucumán, Argentina.


Nixon KC (2002). *WinClada version 1.00.08.* Published by the author. Ithaca, NY, USA.


