MOLECULAR ECOLOGY
RESOURCES

Molecular Ecology Resources (2012) **12**, 990–998 doi: 10.1111/1755-0998.12001

DNA barcoding a regional fauna: Irish solitary bees

KARL N. MAGNACCA*† and MARK J. F. BROWN*‡

**Department of Zoology, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland, †State of Hawaii Division of Forestry and Wildlife, Hilo, HI 96720, USA, ‡School of Biological Sciences, Royal Holloway, University of London, Egham, TW20 0EX, UK*

Abstract

As the globally dominant group of pollinators, bees provide a key ecosystem service for natural and agricultural landscapes. Their corresponding global decline thus poses an important threat to plant populations and the ecosystems they support. Bee conservation requires rapid and effective tools to identify and delineate species. Here, we apply DNA barcoding to Irish solitary bees as the first step towards a DNA barcode library for European solitary bees. Using the standard barcoding sequence, we were able to identify 51 of 55 species. Potential problems included a suite of species in the genus *Andrena*, which were recalcitrant to sequencing, mitochondrial heteroplasmy and parasitic flies, which led to the production of erroneous sequences from DNA extracts. DNA barcoding enabled the assignment of morphologically unidentifiable females of the parasitic genus *Sphecodes* to their nominal taxa. It also enabled correction of the Irish bee list for morphologically inaccurately identified specimens. However, the standard COI barcode was unable to differentiate the recently diverged taxa *Sphecodes ferruginatus* and *S. hyalinatus*. Overall, our results show that DNA barcoding provides an excellent identification tool for Irish solitary bees and should be rolled out to provide a database for solitary bees globally.

Keywords: apidae, conservation, DNA barcodes, hymenoptera

Received 1 May 2012; revision received 17 July 2012; accepted 27 July 2012

Introduction

Pollination is a crucial ecosystem service for human nutrition and health as well as maintenance and functioning of natural ecosystems (Buchmann & Nabhan 1996; Costanza *et al.* 1997; Klein *et al.* 2007). While pollinator taxa are diverse, bees provide the majority of pollination services worldwide (Buchmann & Nabhan 1996; Klein *et al.* 2007). However, many bee populations are in decline across the world (Brown & Paxton 2009), posing a significant threat to natural and agricultural ecosystems (Steffan-Dewenter *et al.* 2005).

Most studies on the decline of bees have focused on highly eusocial bees (mostly honey bees, *Apis* spp. and bumblebees, *Bombus* spp. while the stingless bees currently lack similar studies; for research on non-social bees see Biesmeijer *et al.* 2006; Magnacca 2007; Patiny *et al.* 2009), but most (>95% of approximately 20 000) species of bee are solitary or only primitively eusocial (Michener 2007; for the purposes of this paper we consider the latter together and refer to them as 'solitary' as they share many of the same challenges outlined below).

Correspondence: Dr Mark J. F. Brown, Fax: +44-(0)1784-414224; E-mail: mark.brown@rhul.ac.uk

One reason for this bias is the taxon identification difficulties posed by solitary bees. Solitary bees are usually small and often morphologically cryptic. In addition, for many species males and females have not been associated (Sheffield *et al.* 2009), species show considerable morphological diversity, and keys often rely on quantitative characters or do not exist for many species (e.g., Batley & Hogendoorn 2009). All of these features mean that identification of solitary bees requires high levels of expertise (Packer *et al.* 2009b). However, accurate bioinventories, which are crucial to understanding patterns of solitary bee distribution, abundance and decline, require rapid processing and accurate identification (Packer *et al.* 2009a). One way to achieve this is to associate DNA barcodes with morphologically identified species and then use DNA barcoding for the identification in future surveys.

DNA barcoding has been used to identify cryptic bee species (Murray *et al.* 2008; Williams *et al.* 2012b), enhance taxonomic investigations (Droege *et al.* 2010), discover new species (Gibbs 2009; but see Kuhlmann *et al.* 2007), investigate the validity of morphological keys (Carolan *et al.* 2012), associate males and females within species (Sheffield *et al.* 2009) and create taxon-specific and region-based species banks (Sheffield *et al.* 2009;

© 2012 Blackwell Publishing Ltd

Magnacca & Brown 2010a; Williams *et al.* 2011, 2012a,b). Here, we conduct the first study of European solitary bees to develop a DNA barcoding resource for the island of Ireland. The Irish species list includes taxa for which only one sex can be accurately keyed out (G. Else, unpublished), owing to the general quantitative variation seen in cleptoparasitic taxa (Michener 2007), as well as a potential cryptic species complex—*Lasioglossum albipes/calceatum* (which is recognized as two distinct species in continental Europe, but appears morphologically as a single species in the British Isles; Fitzpatrick *et al.* 2006). We ask specifically whether DNA barcoding can resolve or, at the least, address these problems. Our results demonstrate the validity of barcoding for European solitary bees, while providing a starting point for the broader DNA barcoding of European bees.

Materials and methods

Taxon sampling

The objective was to include all species of Irish solitary bees ($N = 77$). We attempted to collect fresh ethanol-preserved specimens of as many species as possible in surveys across Ireland in 2008 to obtain high-quality DNA. These were supplemented with pinned specimens collected during the Irish bee survey (2004–2005) where necessary for taxonomic and geographic scope (see Table S1 for specimen information). All specimens were identified morphologically using the best available keys for the bee fauna of the British Isles (G. Else, unpublished). DNA was obtained for 57 of the 77 species of solitary bees recorded from Ireland (49 from at least one fresh specimen), from 248 individual specimens (mean = 4.3 specimens per species and median = 3 specimens per species). Most of the unsampled species (e.g. *Andrena rosae*, *Colletes daviesianus* and *Nomada argentata*) are extremely rare and/or highly localized in Ireland. Specimens earlier identified as *Andrena dorsata*, *A. helvola*, *A. ovatula* and *Megachile circumcinta* were sequenced and found to be the members of other species (see Results); consequently, these are not considered to be members of the Irish bee fauna. Definitive specimens of *A. helvola* and *A. ovatula* from England are included for comparison in case future specimens are collected in Ireland. Three species of *Andrena* proved to be recalcitrant in sequencing, leaving 55 Irish solitary bee species. Four species of *Bombus* are also included as a reference group for Apinae; bumblebee barcoding is currently been conducted at a global level by a consortium in BOLD (http://www.barcodinglife.org). Seven sequences of British and continental *Colletes succinctus* from the study of Kuhlmann *et al.* (2007) (GenBank nos. DQ085519–25) were included to confirm that the Irish specimens belong to *C. succinctus s.s.*

Molecular techniques

Total DNA was extracted using the DNeasy Blood & Tissue extraction kit (Qiagen Inc.), following the manufacturer's protocol. To preserve specimens as intact as possible, DNA was usually extracted from the mid- and hind right legs (see Appendix for details). In species that produced consistently polymorphic sequences, extractions were taken from abdominal tissue of the same individuals to check for the possibility of nonrandom haplotype segregation (Magnacca & Brown 2010b). For some species (e.g. *Andrena tarsata*), this was not possible owing to the need to preserve pinned museum specimens intact.

Primer sequences are listed in Table 1. The 'standard' barcoding fragment of cytochrome oxidase I was targeted, using primarily a version of the commonly used primer LCO (Folmer *et al.* 1994), modified for use in Hymenoptera, paired with either 'Nancy' (Simon *et al.* 1994) or NancyShort (C1-N-2171, Magnacca & Brown 2010a), missing the three bases. Two new primers, designated LCOLong and C1-N-2263, were used for a few difficult taxa. PCR was run using standard Taq (Invitrogen Corp.) with the following program: a starting denaturation at 94° for 180 s, followed by 35 cycles of 94° for 30 s, 46–50° for 45 s and 72° for 60 s, concluding with a final extension at 72° for 240 s.

PCR products were sent for sequencing by Macrogen (Seoul, South Korea) or sequenced in the School of

Table 1 Primer sequences used. Most sequences were amplified using LCOHym and either Nancy or Nancy Short (see text for further details)

Name	Direction	3′ base	Sequence	Reference
LCOHym	Forward	1514	TATCAACCAATCATAAAGATATTGG	(Magnacca & Brown 2010a)
LCOLong	Forward	1517	TCAACAAATCATAAAGATATTGGWAT	New
HCO	Reverse	2173	TAAACTTCAGGGTGACCAAAAAATCA	(Folmer *et al.* 1994)
Nancy	Reverse	2191	CCCGGTAAAATTAAAATATAAACTTC	(Simon *et al.* 1994)
NancyShort	Reverse	2194	CCCGGTAAAATTAAAATATAAAC	(Magnacca & Brown 2010a)
C1-N-2663	Reverse	2263	ACTATACCAATATTTCCAAATGTTTC	New

© 2012 Blackwell Publishing Ltd

Natural Sciences, Trinity College; both utilized ABI 3130xl capillary automated sequencers (Applied Biosystems Inc.). The PCR primers were used for sequencing. Most were sequenced from both directions, although sequences were of high enough quality that only one direction was necessary. Sequencing of *Andrena carantonica* with the C1-N-2263 primer consistently failed, despite the success of this primer in other taxa; as a result, the LCOHym/C1-N-2263 sequences of this species were available in the forward direction only. All polymorphic species were sequenced from both directions to ensure correct basecalling. Chromatograms were edited using FinchTV (Geospiza Inc.). Sequences were submitted to GenBank (accession numbers JQ909638–JQ909880), and sequences, trace files, and specimen data submitted to BOLD (http://www.barcodinglife.org, project Ireland Bee Barcoding Project [IBBP]).

Sequence analysis

Alignment of sequences was trivial as no gaps were present. After trimming the ends, a sequence of 654 base pairs was used for analysis. Only 18 specimens do not have the full sequence, and except for *A. carantonica* none is missing more than 28 bases. For tree construction, a Bayesian analysis was performed with MrBayes 3.1.2 (Huelsenbeck & Ronquist 2001), with the data partitioned into three sets by codon position. Models for each partition were selected using MrModeltest 2.2 (Nylander 2004); these were determined to be GTR + I + G for codon positions 1 and 2, and HKY + G for codon position 3. The analysis was run for 4 million generations with a burn-in of 1 million, with all parameters unlinked across partitions. For comparison with previous barcoding studies, trees were also generated in PAUP* 4.0b10 (Swofford 2003) using parsimony and neighbour-joining pairwise distance algorithm, using uncorrected (‘*P*’) distances. Bootstrap analyses were performed on these for 1000 and 10 000 replicates, respectively. Trees are arbitrarily displayed as rooted with *Hylaeus* at the top, as higher-level relationships are unreliable.

Results

Sequencing

Although most species were relatively easy to sequence, a surprising number (16/55) presented difficulties, particularly in the genus *Andrena* (Table 2). Several of the problematic species, including *Andrena barbilabris*, *A. carantonica*, *A. cineraria*, *A. clarkella*, *A. denticulata*, and *A. fucata*, co-amplified COI sequences of the intracellular parasite *Wolbachia* with the primer pair LCOHym/HCO. *Andrena barbilabris* failed to amplify at all with other primer combinations. *Andrena carantonica* and *A. clarkella* amplified well, but gave extremely poor sequences; the latter also produced a secondary band, but sequencing success was not improved by gel-purifying the proper-sized band. Only one clean sequence was obtained for *A. clarkella* of nine specimens available. All of these were extracted from fresh specimens, indicating that this failure to amplify or provide good sequences is not down to the quality of material (as might be the case if they were museum specimens). Sequences for *A. carantonica*, one of the most common Irish bees, were obtained using the primer pairs LCOHym/C1-N-2263 and LCOLong/Nancy, but the two combinations produce different sequences that are approximately 13% divergent. One of these is almost certainly a nuclear pseudogene (numt), but neither contains indels or stop codons that definitively mark it as such, and an unusually high proportion of the differences (70/82) is synonymous. On the basis of the comparison with other species for the amino acid changes that do occur, the LCOLong/Nancy sequence is probably the numt. Both are included in the analysis.

All leg extractions from three *A. cineraria* produced clean, high-quality sequences that appeared to be from a fly. The source is unknown, but is presumed to be a specific internal parasite because (i) no Diptera material is handled in the laboratory where the bees were extracted; (ii) no other bee species were contaminated in this way; and (iii) the three *A. cineraria* were extracted at different times and with no external signs of parasitism (e.g. eggs). The correct sequence for *A. cineraria* was obtained from an abdominal extraction. The fly sequence was identified through BOLD as *Myopa* sp. (Conopidae), with 98.5% identity to a sequence of *M. testacea* and 89–96% similarity to two other *Myopa* species. Conopids are internal parasites of insects including aculeate Hymenoptera, and at least five *Myopa* species, including *M. testacea*, are known to occur in Ireland (Speight 1978; Alexander 2011) Given the level of identity, it is uncertain whether this sample represents *M. testacea* or one of the other *Myopa* species. DNA barcoding of this group would provide a resolution to this question.

Six species, in *Andrena*, *Colletes* and *Lasioglossum*, exhibited moderately high levels of polymorphism in leg extractions, while producing clean sequences from abdominal extractions of the same individuals (Table 2). It is unclear whether the polymorphisms were because of heteroplasmy or numts. However, all of the polymorphic sequences had one or no amino acid changes. In contrast, *Andrena lapponica*, which was also polymorphic but had a higher rate of amino acid changes (about five per sequence), did not show any difference between leg and abdominal extractions. These result suggests that haplotype tissue segregation, which occurs in Hawaiian members of *Hylaeus* (Magnacca & Brown 2010b), is also

© 2012 Blackwell Publishing Ltd

Table 2 Summary of taxon sampling and sequencing problems encountered (n = numt; h = possible heteroplasmy; a = polymorphic leg extraction, abdominal clean; w = *Wolbachia* infection). X = specimens available but could not included as a result of sequencing failure. The list includes all solitary bee species that were believed to be present in Ireland prior to this study

Irish solitary bees	*N*	Problem	Notes
Andrena angustior	3	a	1 of 3 leg extractions polymorphic, abdominal extraction clean
Andrena apicata	2	a	Leg extractions polymorphic, abdominal extractions clean
Andrena barbilabris	X	w	Recalcitrant
Andrena bicolor	6		
Andrena carantonica	5	w, n	Produces poor quality polymorphic sequence with LCOHym-Nancy, 2 different clean sequences with other primer pairs
Andrena cineraria	1	w	Leg extractions produce Diptera sequence
Andrena clarkella	1	w, n	Visible double bands in PCR; even cut bands usually produce double sequence
Andrena coitana	8		
Andrena denticulata	3	w	
Andrena fucata	3	w	
Andrena fulva			
Andrena fuscipes	3		
Andrena haemorrhoa	4	a	All leg extractions polymorphic, abdominal extractions clean
Andrena humilis			
Andrena lapponica	X	n	Highly polymorphic in both leg and abdominal extractions
Andrena marginata			
Andrena minutula	2		
Andrena nigroaenea	8		
Andrena pilipes			
Andrena praecox	3		
Andrena rosae			
Andrena semilaevis			
Andrena stragulata			
Andrena subopaca	2		
Andrena tarsata	2	h	2 of 2 specimens polymorphic
Andrena wilkella	5		
Coelioxys elongata	1		
Coelioxys inermis			
Colletes daviesanus			
Colletes floralis	4	a	All leg extractions polymorphic, abdominal extractions clean
Colletes fodiens			
Colletes similis	5	a	All leg extractions polymorphic, abdominal extractions clean
Colletes succinctus	11	h	All sequences polymorphic; variable positions in the four Irish specimens different from those from Kuhlmann *et al.* (2007)
Halictus rubicundus	6	h	1 of 6 specimens polymorphic
Halictus tumulorum	5	h	1 of 5 specimens polymorphic
Hylaeus brevicornis			
Hylaeus communis	4		
Hylaeus confusus	2		
Hylaeus hyalinatus			
Lasioglossum albipes	15		
Lasioglossum calceatum	13		
Lasioglossum cupromicans	2		
Lasioglossum fratellum	3		
Lasioglossum lativentre	2		
Lasioglossum leucopus	2		
Lasioglossum nitidiusculum	1		
Lasioglossum punctatissimum	2		
Lasioglossum rufitarse			
Lasioglossum smeathmanellum			Identification questionable, may not be present in Ireland
Lasioglossum villosulum	4		
Megachile centuncularis	1		

© 2012 Blackwell Publishing Ltd

Table 2 (Continued)

Irish solitary bees	*N*	Problem	Notes
Megachile ligniseca			
Megachile maritima	2		
Megachile versicolor	3		
Megachile willoughbiella	2		
Nomada argentata			
Nomada fabriciana	2		
Nomada flavoguttata	6		
Nomada goodeniana	3		
Nomada leucophthalma	7		
Nomada marshamella	8		
Nomada obtusifrons	3		
Nomada panzeri	9		
Nomada ruficornis	12		
Nomada rufipes	2		
Nomada sheppardana			
Nomada striata			
Osmia aurulenta	2	h	1 of 2 specimens polymorphic
Osmia rufa	3		
Sphecodes crassus			Identification questionable, may not be present in Ireland
Sphecodes ephippius	8		
Sphecodes ferruginatus	3		
Sphecodes geoffrellus	5	h	2 of 5 specimens polymorphic
Sphecodes gibbus	6		
Sphecodes hyalinatus	5		
Sphecodes monilicornis	2		
Sphecodes pellucidus			

present in the six sequence-variable species. For these, only the clean sequence was included in the analysis as it was consistent across each species; because *A. lapponica* appears to co-amplify a numt, it was excluded. *Colletes succinctus* was polymorphic in both leg and abdominal extractions (of different individuals), but the number of polymorphisms was relatively low (12), all were synonymous, and sequencing chromatograms appeared significantly different between the two tissue types. This implies tissue segregation similar to that of the other *Colletes* species, but less strict.

Finally, an additional five species—*Andrena tarsata*, *Halictus rubicundus*, *H. tumulorum*, *Osmia aurulenta*, and *Sphecodes geoffrellus*—exhibited moderate polymorphism in some or all individuals. Both specimens of *A. tarsata* were polymorphic, but it is difficult to determine whether this is the result of heteroplasmy or a numt—there are a large number of polymorphisms, but very few result in amino acid changes, and none appear to be significant. For the remainder, only one or two individuals were polymorphic while the remainder sequenced cleanly, suggesting low rates of heteroplasmy in the populations.

Species identification

Virtually all (51/55) species were recovered as well-supported monophyletic branches using DNA barcodes, with only minimal intraspecific variation using all forms of analysis. The Bayesian and parsimony trees are nearly identical, differing primarily in the lower resolution of the parsimony analysis (see Fig. 1), and in that the parsimony tree resolves *Nomada leucophthalma* as a distinct clade, and consequently we show them as one tree (Fig. 1); the neighbour-joining tree (Fig. 2) differs in the placement of basal branches on the tree, but separates species equally well. There was a distinct break between interspecific and intraspecific genetic distance at 1% (Fig. 3), although variation at this point was continuous. All of the pairwise comparisons between 1.0% and 2.5% are between two pairs of cleptoparasitic nominal species, *Sphecodes ferruginatus* and *S. hyalinatus*, and *Nomada leucophthalma* and *N. panzeri*, which were also the least distinct on the tree. There was no correlation between the number of individuals sequenced within a species and the intraspecific genetic distance (Spearman's rank correlation, $r = 0.2$, $P = 0.159$).

Problematic taxa

Prior to barcoding, while males of *Sphecodes* spp. could be accurately keyed out to species, females could not be due to the quantitative variation in the characters used in the keys. DNA barcodes enabled the assignment of 11 female *Sphecodes* to their nominal taxa (*ephippius*, *ferruginatus*, *geoffrellus*, *hyalinatus* and *monilicornis*) by matching them to morphologically identified males. Both trees showed *S. ferruginatus* to be genetically a derived subgroup of *S. hyalinatus*. Given the clear interspecific distances across the remainder of the tree, this suggests that these taxa are in the process of divergence.

© 2012 Blackwell Publishing Ltd

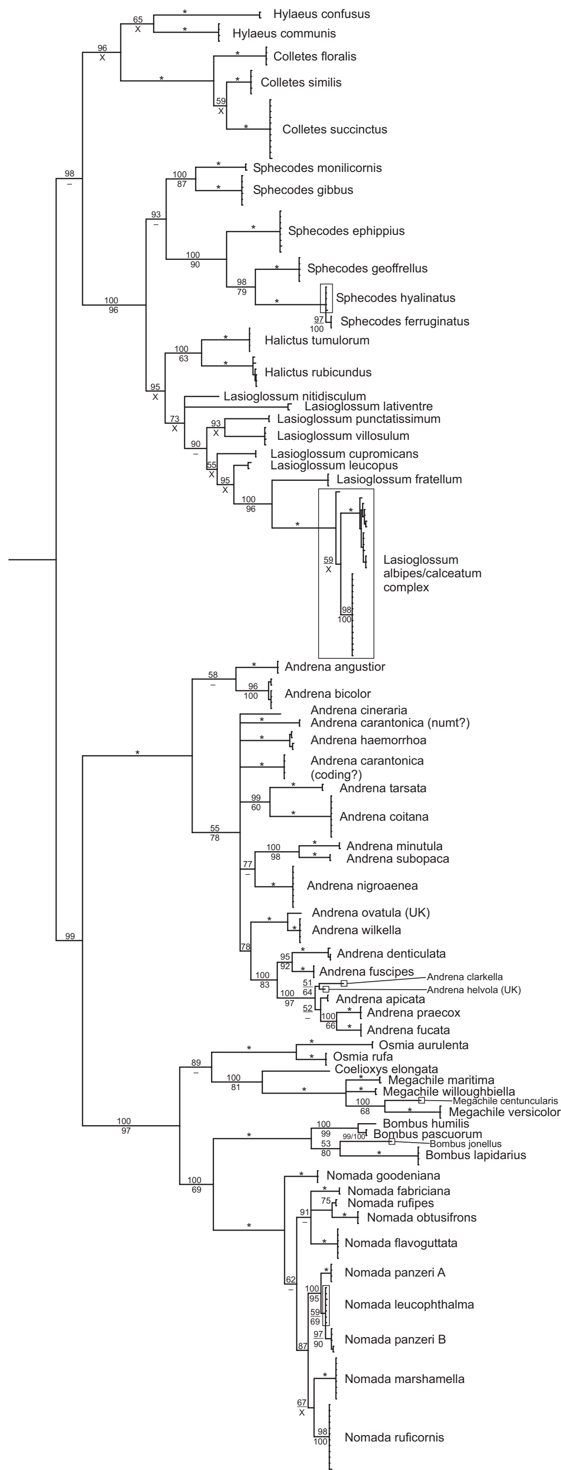

Fig. 1 Bayesian consensus tree. Numbers above branches are posterior probabilities, below are parsimony bootstrap percentages. 'X' indicates node that is unresolved in the parsimony consensus [except the *Hylaeus* + *Colletes* node, which the parsimony tree resolves as *Colletes* + (*Hylaeus* + Halictidae)]; '–'denotes nodes that are present but received <50% bootstrap support; and nodes that are 100% for both values are marked with an asterisk (*). Intraspecific support values not shown.

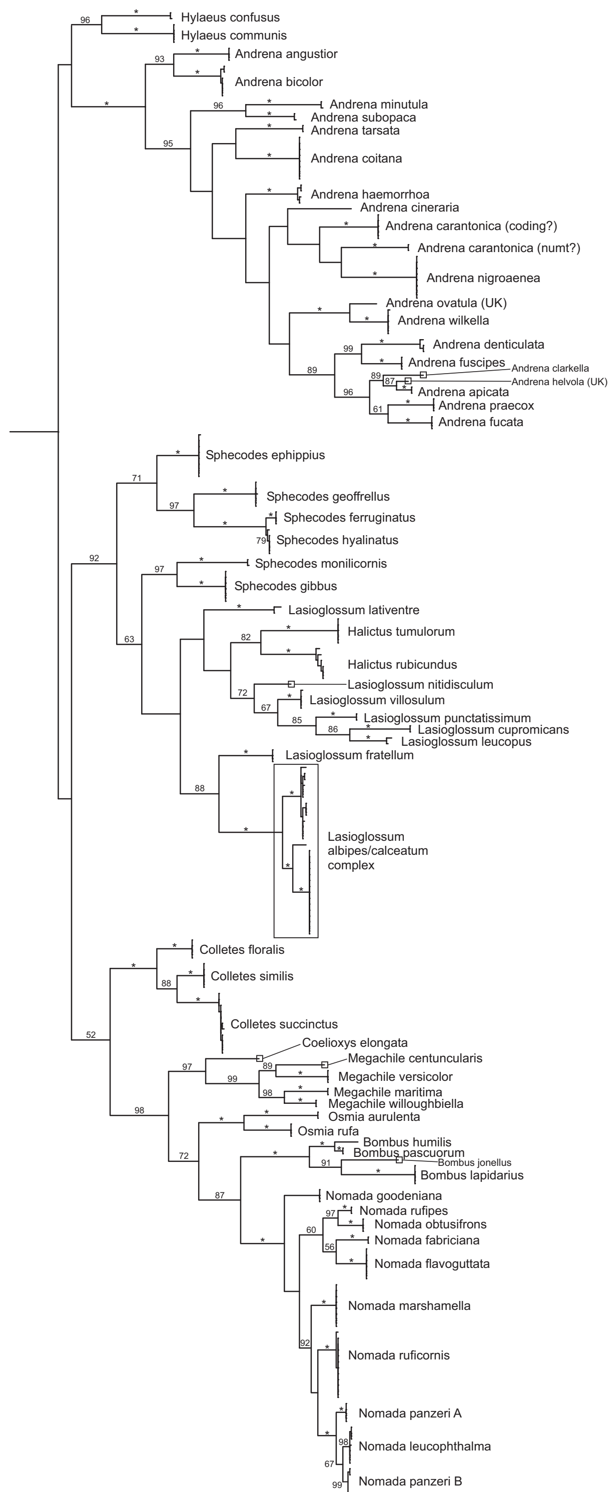

Fig. 2 Neighbour-joining tree with bootstrap values. Nodes with 100% support marked with an asterisk.

The putative cryptic species complex of *L. albipes/calceatum* separated into two distinct clusters plus one individual equally divergent from both clusters, in both the Bayesian and NJ trees. The major groups match the

© 2012 Blackwell Publishing Ltd

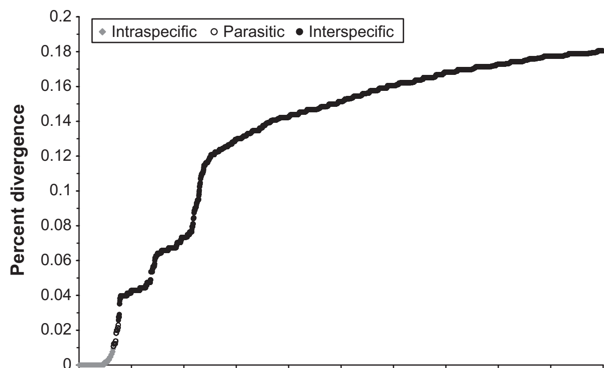

Fig. 3 Distribution of pairwise sequence differences, plotted as an accumulation curve (only the first 10 000 shown). Gray circles indicate intraspecific genetic distances. Open circles ('parasitic') are interspecific comparisons for the cleptoparasite species pairs *Sphecodes ferruginatus/hyalinatus* and *Nomada leucophthalama/panzeri*, which appear to be more similar to intraspecific comparisons (see text). Filled circles are all other interspecific pairwise distances.

L. albipes and *L. calceatum* of other sequences in BOLD and GenBank (the isolated individual matches the social *L. albipes* found by Danforth *et al.* 2003; see Discussion). However, none of the characters used to separate these species in available keys were qualitatively or quantitatively different between individuals in these clusters.

Discussion

DNA barcoding can provide a quick and reliable method for species identification in the European solitary bees. In addition, it may enable the development of more accurate morphological keys through the association between males and females within problematic species, as well as insight into the evolutionary status of species-pairs.

Our results largely support those of previous studies by Sheffield *et al.* (2009) of the Nova Scotia bee fauna and Williams *et al.* (2011, 2012a) in various *Bombus* subgenera, suggesting that DNA barcoding should be broadly applicable across the ecologically important bee fauna. The majority of species produced a clear DNA barcode with little intraspecific variation, and both trees resulted in the placement of species within congeneric groups. A barcoding approach could be particularly important in regions of the world where bee diversity remains to be enumerated (Brown & Paxton 2009; Eardley *et al.* 2009) and where keys are lacking (e.g., Batley & Hogendoorn 2009). Consequently, the current BEE-BOL project (http://www.bee-bol.org), to barcode the world's bees, and to which our work contributes, has both a high likelihood of success (Sheffield *et al.* 2009) and great importance from a conservation perspective. Still, the significant number of species in our data set that could not be either sequenced or produced polymorphic sequences suggests that this task may be more difficult than for other taxa, particularly in *Andrena*. One way to improve amplification success, and possibly sequence quality, may be to use polymerases with higher fidelity.

While our analyses recovered distinctive barcodes for the majority of species, there was no 'barcoding gap' (Hebert *et al.* 2003b) between intra- and interspecific divergence; most recent analyses suggest that such a barcoding gap is biologically unlikely to occur (Gibbs 2009). While intraspecific genetic distances fell below 1%, and interspecific distances above 1%, there was no break and divergence was continuous. This was surprising, particularly when comparing across a diverse multi-family grouping such as this, as many studies have found at least some overlap between the two categories (Ball *et al.* 2005; Hajibabaei *et al.* 2006; Papadopoulou *et al.* 2008). Additional sampling across the species' range in continental Europe would likely show such an overlap.

The sequencing resolved several questions regarding the presence of bee species in Ireland. Two specimens earlier identified morphologically as *Andrena dorsata* and *A. ovatula* were found to be fully identical to *A. wilkella* and distant from a British *A. ovatula*. Another, identified as *A. helvola*, was identical to *A. praecox*, quite removed from true *A. helvola*. Finally, a specimen previously identified as *Megachile circumcinta* was identical to *M. willughbiella*. All of the DNA identifications were confirmed by re-examination of the specimens, and therefore, these four species should be removed from the list of Irish bees (Fitzpatrick *et al.* 2006).

In practical terms, the most useful application of this data set is with the morphologically difficult Halictidae and *Nomada*. In the former, numerous morphologically ambiguous female specimens of *Sphecodes* were successfully associated with identified male specimens. Cleptoparasitic bees in general are known for high morphological variability and are notoriously difficult to identify (Sheffield *et al.* 2009). However, *S. ferruginatus* and *S. hyalinatus* did not resolve into distinct genetic clusters; rather, the former appears as a derived subgroup of the latter. Males conforming to the genitalic descriptions of each were included to ensure that each grouping was properly identified. These two species are likely to be recently diverged or in the process of speciation. At least two of the species we included, *S. ephippius* and *S. monilicornis*, are generalists on a variety of hosts but are believed to form host races based on female lineages (Bogusch *et al.* 2006) similar to cuckoo birds (Gibbs *et al.* 2000). However, we did not find any evidence of genetic differentiation within those species, and the distinct genitalia of *S. ferruginatus* and *S. hyalinatus* implies that they do not interbreed.

© 2012 Blackwell Publishing Ltd

We also found that the cryptic species *Lasioglossum albipes* and *L. calceatum* do indeed sort into two abundant and distinct genotypes, but also included a single individual that was equally different from both. Re-examination of the specimens in the light of their genetic determination still did not reveal any consistent morphological characters associated with them, including the sculpturing patterns of the head and propodeum that have been used to distinguish the two species (G. Else, unpublished). Furthermore, the specimens sequenced did not clearly segregate based on locality, collection date, host flower or other ecological characters that might lead to sympatric speciation (Kuhlmann *et al.* 2007); many were taken as mixed series. As a result, we currently cannot say which is *L. albipes* and which is *L. calceatum*. More detailed morphological investigations, including head measurements, may reveal consistent differences between them. It is important to note that the sample size was relatively small ($N = 29$ total) and limited to Ireland, and it is possible that *L. albipes* and *L. calceatum* are morphologically distinguishable species, but only one of the two is actually present in Ireland. However, the two major clades identified in our study correspond to sequences consistently identified as *L. albipes* and *L. calceatum* in BOLD and GenBank, including the *L. calceatum* and solitary *L. albipes* from Danforth *et al.* (2003). The single specimen that did not group with the others (which was taken in company with two putative '*L. calceatum*') is close only to a social bee identified as *L. albipes* in Danforth *et al.* (2003). Notably, all the Danforth *et al.* (2003) samples of these taxa came from France. More widespread genetic sampling throughout the full range may reveal that genetic differences apparent in the Irish population are consistent across the range and possibly correlated with social behaviour.

A similar situation exists with *Nomada leucophthalama* and *N. panzeri*. Specimens identified as *N. panzeri* segregate as two distinct clusters, while those identified as *N. leucophthalama* are intermediate and do not appear as a distinct clade on the Bayesian tree (Fig. 1; under parsimony they resolve similar to the NJ tree, Fig. 2). Given the propensity of cleptoparasitic bees, and *Nomada* in particular, for a high degree of morphological variation, it is even more likely in this situation that only a single species is involved. Again, however, a much larger and more diverse sample is needed before any taxonomic conclusions can be drawn. Moreover, both species are known to parasitize multiple *Andrena* hosts (Perkins 1919; Richards 1946) and could potentially form mitochondrially distinct host races (Gibbs *et al.* 2000).

For all three of the above-mentioned species pairs, the COI barcoding region, and mtDNA in general, may well be insufficient genetic markers for determining their status even with greater sampling. A previous study of the *Colletes succinctus* group (Kuhlmann *et al.* 2007) also failed to find clear species delineation with the DNA barcode region, whereas the ITS-2 and EF-1alpha regions clearly demonstrated the presence of multiple species. Given that the DNA barcode region was chosen, in part, owing to its stability within species (Hebert *et al.* 2003a), such errors are likely to occur in rapidly speciating groups. This reinforces the importance of using multiple diagnostic tools in taxonomic studies (Rubinoff *et al.* 2006).

To conclude, our results show that DNA barcoding can be used as a tool for bioinventories of solitary bees in Europe. While this may be relatively straightforward to achieve for most species, integrated studies of ecology, taxonomy and barcoding will be required to resolve problematic species and produce a complete barcode list for this fauna.

Acknowledgements

This study utilized the Irish Bee Survey collection, which was funded by the Higher Education Authority (Ireland) under their North-South Programme for Collaborative Research. We thank Úna Fitzpatrick, Tomás Murray and Robert Paxton, who contributed to this collection and gave their blessing for this study. Our study was funded by Science Foundation Ireland grant EEEOBF131 to MJFB.

References

Alexander KNA (2011) *An Invertebrate Survey of Coill Eoin, St. John's Wood, Co. Roscommon*. National Parks and Wildlife Service, Department of the Arts, Heritage and Gaeltacht, Dublin, Ireland.

Ball SL, Hebert PDN, Burian SK, Webb JM (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society*, **24**, 508–524.

Batley M, Hogendoorn K (2009) Diversity and conservation status of native Australian bees. *Apidologie*, **40**, 347–354.

Biesmeijer JC, Roberts SPM, Reemer M *et al.* (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*, **313**, 351–354.

Bogusch P, Kratochvil L, Straka J (2006) Generalist cuckoo bees (Hymenoptera: Apoidea: *Sphecodes*) are species-specialist at the individual level. *Behavioral Ecology and Sociobiology*, **60**, 422–429.

Brown MJF, Paxton RJ (2009) The conservation of bees: a global perspective. *Apidologie*, **40**, 410–416.

Buchmann SL, Nabhan GP (1996) *The Forgotten Pollinators*. Island Press, Washington, DC, USA.

Carolan JC, Murray TE, Fitzpatrick U *et al.* (2012) Colour patterns do not diagnose species: quantitative evaluation of a DNA barcoded cryptic bumblebee complex. *PLoS ONE*, **7**, e29251. doi:10.1371/journal.pone.0029251.

Costanza R, dArge R, De Groot R *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Danforth BN, Conway L, Ji S (2003) Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). *Systematic Biology*, **52**, 23–36.

Droege S, Rightmyer MG, Sheffield CS, Brady SG (2010) New synonymies in the bee genus *Nomada* from North America (Hymenoptera: Apidae). *Zootaxa*, **2661**, 1–32.

© 2012 Blackwell Publishing Ltd

Eardley CD, Gikungu M, Schwarz MP (2009) Bee conservation in Sub-Saharan Africa and Madagascar: diversity, status and threats. *Apidologie*, **40**, 355–366.

Fitzpatrick U, Murray TE, Byrne A, Paxton RJ, Brown MJF (2006) Regional red list of Irish bees. Available from http://www.tcd.ie/Zoology/research/Bees/Irish_bees_-_regional_red_list.pdf

Folmer O, 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**, 294–299.

Gibbs JJ (2009) New species in the *Lasioglossum petrellum* species group identified through an integrative taxonomic approach. *Canadian Entomologist*, **141**, 371–396.

Gibbs HL, Sorenson MD, Marchetti K *et al.* (2000) Genetic evidence for female host-specific races of the common cuckoo. *Nature*, **407**, 183–186.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 968–971.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 313–321.

Hebert PDN, Ratnasingham S, deWaard JR (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.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Klein AM, Vaissiere BE, Cane JH *et al.* (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B-Biological Sciences*, **274**, 303–313.

Kuhlmann M, Else GR, Dawson A, Quicke DLJ (2007) Molecular, biogeographical and phenological evidence for the existence of three western European sibling species in the *Colletes succinctus* group (Hymenoptera: Apidae). *Organisms Diversity & Evolution*, **7**, 155–165.

Magnacca KN (2007) Conservation status of the endemic bees of Hawai'i, *Hylaeus* (*Nesoprosopis*) (Hymenoptera: Colletidae). *Pacific Science*, **61**, 173–190.

Magnacca KN, Brown MJF (2010a) Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus* (*Nesoprosopis*) bees (Hymenoptera: Colletidae). *Bmc Evolutionary Biology*, **10**, 174.

Magnacca KN, Brown MJF (2010b) Tissue segregation of mitochondrial haplotypes in heteroplasmic Hawaiian bees: implications for DNA barcoding. *Molecular Ecology Resources*, **10**, 60–68.

Michener CD (2007) *The Bees of the World*. The Johns Hopkins University Press, Baltimore.

Murray TE, Fitzpatrick U, Brown MJF, Paxton RJ (2008) Cryptic species diversity in a widespread bumble bee complex revealed using mitochondrial DNA RFLPs. *Conservation Genetics*, **9**, 653–666.

Nylander JAA (2004) Mr Modeltest v2.2.

Packer L, Gibbs J, Sheffield C, Hanner R (2009a) DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources*, **9**(Suppl. 1), 42–50.

Packer L, Grixti JC, Roughley RE, Hanner R (2009b) The status of taxonomy in Canada and the impact of DNA barcoding. *Canadian Journal of Zoology*, **87**, 1097–1110.

Papadopoulou A, Bergsten J, Fujisawa T *et al.* (2008) Speciation and DNA barcodes: testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **363**, 2987–2996.

Patiny S, Rasmont P, Michez D (2009) A survey and review of the status of wild bees in the West-Palaearctic region. *Apidologie*, **40**, 313–331.

Perkins RCL (1919) The British species of *Andrena* and *Nomada*. *Transactions of the Entomological Society of London*, **67**, 218–320.

Richards OW (1946) *Nomada flava* Panzer and the strains of *N. panzeri* Lepeletier (Hymenoptera, Apidae). *Proceedings of the Royal Entomological Society of London Series B, Taxonomy*, **15**, 17–26.

Rubinoff D, Cameron S, Will K (2006) A genomic perspective on the shortcomings of mitochondrial DNA for 'barcoding' identification. *Journal of Heredity*, **97**, 581–594.

Sheffield CS, Hebert PDN, Kevan PG, Packer L (2009) DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. *Molecular Ecology Resources*, **9**, 196–207.

Simon C, Frati F, Beckenbach A *et al.* (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Speight MCD (1978) *Myopa curtirostris* new to Ireland and some other Irish conopid (Diptera) records. *The Irish Naturalists' Journal*, **19**, 276–278.

Steffan-Dewenter I, Potts SG, Packer L (2005) Pollinator diversity and crop pollination services are at risk. *Trends in Ecology & Evolution*, **20**, 651–652.

Swofford DL (2003) *PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods), version 4.0b10*. Sinauer Associates, Sunderland, Massachusetts.

Williams PH, An JD, Huang JX (2011) The bumblebees of the subgenus *Subterraneobombus*: integrating evidence from morphology and DNA barcodes (Hymenoptera, Apidae, *Bombus*). *Zoological Journal of the Linnean Society*, **163**, 813–862.

Williams PH, Brown MJF, Carolan JC *et al.* (2012a) Unveiling cryptic species of the bumblebee subgenus *Bombus* s. str. world-wide with COI barcodes (Hymenoptera: Apidae). *Systematics and Biodiversity*, **10**, 21–56.

Williams PH, Brown MJF, Carolan JC *et al.* (2012b) Cryptic bumblebee species: consequences for conservation and the trade in greenhouse pollinators. *PLoS ONE*, **7**, e32992.

M.J.F.B. conceived the study. K.N.M. conducted the molecular and taxonomic analyses. M.J.F.B. and K.N.M. wrote the manuscript.

Data Accessibility

DNA sequences: GenBank accessions JQ909638-JQ909880) Sequences, trace files, and specimen data can be found at BOLD (http://www.barcodinglife.org, project Ireland Bee Barcoding Project [IBBP]).

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1 Sample locations and specimen details.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2012 Blackwell Publishing Ltd