DNA BARCODES TO EXPLORE DIVERSITY IN APHIDS
(HEMIPTERA APHIDIDAE AND ADELGIDAE)

INTRODUCTION

Many aphid species are important pests in agriculture and forestry, both through direct damage and by acting as vectors of numerous plant diseases (Blackman and Eastop, 2000, Eastop, 1977; Van Emden and Harrington, 2007; Minks and Harrewijn, 1987). As a result of ease of transport and parthenogenetic reproduction, they are important as invasive pests throughout the world (for example, Foottit et al., 2006; Messing et al., 2007; Teulon and Stufkens, 2002). Reliable identification of species is essential for the integrated management of pest aphids and for the early detection and risk analysis of newly introduced species (Miller and Foottit, 2009). However, an evolutionary tendency towards the loss of taxonomically useful characters, polymorphism, and morphological plasticity due to host influences and other environmental factors, complicate the recognition of species and the analysis of relationships at all levels (Foottit, 1997).

DNA barcoding, using a short standardized DNA sequence, has been proposed as a approach to the characterization of life forms. The presence of different morphological forms of a single species on different hosts and at different times of the year makes it difficult to consistently associate routinely collected field samples with particular species definitions. DNA barcoding has been proposed as a standardized approach to the characterization of life forms. We have tested the effectiveness of the standard 538-bp barcode fragment from the 5′ end of the mitochondrial cytochrome c oxidase 1 gene (COI) to differentiate among species of aphids and adelgids. Results are presented for a preliminary study on the application of DNA barcoding in which approximately 3600 specimens representing 568 species and 169 genera of the major subfamilies of aphids and the adelgids have been sequenced. Examples are provided where DNA barcoding has been used as a tool in recognizing the existence of cryptic new taxa, linking life stages on different hosts of adelgids, and as aid in the delineation of species boundaries. The future use of these DNA barcodes for the routine detection of invasive species, the resolution of pest complexes, and the analysis of diverse faunas is discussed.

KEY WORDS: Aphididae, Adelgidae, species identification, DNA barcodes, COI

MATERIALS AND METHODS

Aphid samples were collected between 1991 and 2008 into liquid nitrogen or into 95% ethanol. Voucher specimens from each collection were mounted on microscope slides and deposited in the Canadian National Collection of Insects (Agriculture and Agri-Food, Ottawa, Ontario, Canada) or at the Irrigated Agriculture Research and Extension Center, Washington State University (Prosser, Washington, USA). Samples of aphid DNA (particularly Hormaphidinae: Cerataphidini) were provided by D. Stern (vouchers in natural History Museum, London). DNA extraction, amplification and sequencing for most specimens were done at the Biodiversity Institute of Ontario (Guelph, Ontario, Canada) using techniques described by deWaard et al. (2008). Additional samples were processed at Agriculture and Agri-Food Canada (Ottawa, Ontario, Canada) or by Nathan Havill at Yale University (New Haven, Connecticut, USA). The primer pairs LepF and LepR or M13-tailed alternates, LCO1490_t1 and HCO2198_t2 (primer sequences available in BOLD – the Barcode of Life Data System) were used to amplify an approximately 700 bp DNA fragment of mitochondrial COI, which was subsequently sequenced in both directions using either LepF and LepR or M13F and M13R (primer sequences given in BOLD).

Collection information and sequence data have been entered into BOLD – the Barcode of Life Data System. Subsets discussed in Foottit et al., (2008, 2009) are publicly available on BOLD. Additional sets will be made available as they are validated and published.
Pairwise distances were calculated using the Kimura 2-parameter model of base substitution (Kimura, 1980) and patterns of variation among samples were examined using neighbor-joining (Saitou & Nei, 1987) as implemented on BOLD.

Classification and nomenclature of aphid taxa follows Remaudière and Remaudière (1997) and Nieto Nafría et al. (1998).

RESULTS

GENERAL RESULTS

A total of 3630 specimens of 568 species in 169 genera have been sampled, representing the families Adelgidae, Phylloxeridae and 18 of the 24 subfamilies of Aphididae (Table 1). About one third of the species are represented by at least 3 samples and one fifth by 10 or more samples. The majority of the specimens are from North America, but samples of Aphididae from Micronesia and New Zealand, and Adelgidae from Japan, China and Europe are also included in the current data set.

Maximum within-species pairwise distances varied from 0 to 5.5% in Aphididae and 0 to 8.75% in Adelgidae, with a mean maximum of 0.58%. However, maximum within-species distance was less than 2% in 94% of the species represented by more than 3 samples, and mean maximum pairwise distance in these species is 0.35%. Within the more variable species, distribution of pairwise distances was distinctly bimodal or trimodal, suggesting possible cryptic species, although this may also be explained as a sampling artefact or evidence of introgression. In particular the maximum intraspecific divergence of 8.87% observed in species differing by less than 1% include Apis (Barapsis) varians Patch versus A. manitobensis Robinson and Rojanavongse, members of the Acerthosiphon kondoi group (kondoi Shinji, caraganae (Cholodkovsky) and churchillense Robinson), Ericapis fimbrita (Richards) versus E. scamelli (Mason), and certain groups of Uroleucon species, Illinois species, Macrosiphum species and Kakimia (genus Nasonoa) species. Although some of these species may prove to be synonyms, most are distinguishable by host association and by morphology (albeit with some difficulty), and thus likely represent recently separated species. Additional sampling of low divergence species pairs is required to determine whether the small observed differences are fixed sequence differences or artefacts of the sample set.

Table 1 – Summary of taxonomic distribution of material sampled. Classification follows Remaudière and Remaudière (1997) and Nieto Nafría et al. (1998).

<table>
<thead>
<tr>
<th>Genera</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adelgidae</td>
<td>2 17</td>
</tr>
<tr>
<td>Phylloxeridae</td>
<td>3 3</td>
</tr>
<tr>
<td>Aphididae</td>
<td>1 2</td>
</tr>
<tr>
<td>Anocicicinai</td>
<td>23 55</td>
</tr>
<tr>
<td>Chaitophorinae</td>
<td>3 14</td>
</tr>
<tr>
<td>Drepanosiphinae</td>
<td>2 6</td>
</tr>
<tr>
<td>Eriosomatinae</td>
<td>20 44</td>
</tr>
<tr>
<td>Greenseidinae</td>
<td>2 3</td>
</tr>
<tr>
<td>Hormaphidinae</td>
<td>9 13</td>
</tr>
<tr>
<td>Lachnia</td>
<td>12 59</td>
</tr>
<tr>
<td>Lizerinai</td>
<td>1 1</td>
</tr>
<tr>
<td>Mindarinae</td>
<td>1 2</td>
</tr>
<tr>
<td>Phloeomyzinae</td>
<td>1 1</td>
</tr>
<tr>
<td>Phylaphidinae</td>
<td>2 3</td>
</tr>
<tr>
<td>Pterocommatinae</td>
<td>2 6</td>
</tr>
<tr>
<td>Salusaphidinae</td>
<td>5 9</td>
</tr>
<tr>
<td>Tamalinae</td>
<td>1 4</td>
</tr>
<tr>
<td>Taiwanaaphidinae</td>
<td>1 1</td>
</tr>
<tr>
<td>Thelaxinae</td>
<td>1 1</td>
</tr>
<tr>
<td>total</td>
<td>169 568</td>
</tr>
</tbody>
</table>

STATUS OF FORMS OF PENTALONIA NIGRONERVOSA (APHIDIDAE: APHIDINAE)

Pentalonia nigronervosa Coquerel, the banana aphid, is known as a pest of banana and various crop plants in families Araceae and Zingiberaceae. Van der Goot (1917) described P. caladii, distinguished by differences in length of the ultimate rostral article and siphunculus. However, the status of this entity has not been adequately addressed. Later authors treated it as a synonym or ‘form’ of P. nigronervosa. Remaudière and Remaudière (1997) continue to list it as a form, and biological studies often make no explicit distinction (for example, Robson et al., 2007). Barcode data for sixty P. nigronervosa samples from Polynesia, Australia and Florida (USA) falls into two uniform groups differing by about 3% sequence divergence. All samples from one group (1 exception) occur on banana, and all samples from the other group on Araceae and Zingiberaceae (again 1 exception only). Subsequent investigation showed that the groups correspond exactly to fixed differences in sequences for nuclear elongation factor 1 alpha introns, and to non-overlapping morphometric differences. Given the economic importance of this aphid, more formal recognition and definition of the form caladii is desirable (manuscript in preparation).

GEOGRAPHIC AND HOST-ASSOCIATED VARIATION IN CHAITOPHORUS POPULICOLA (APHIDIDAE: CHAITOPHORINAE)

Chaitophorus populicola Thomas is widespread in North America on various species of Populus. There is considerable variation in coloration, body size and setal shape, but the extent to which the observed variation is of genetic or environmental origin is unclear. Hille Ris Lambers (1960) described subspecies patchae for specimens with relatively longer, acuminate setae. Chaitophorus brunneri Williams is currently considered a synonym.
Maximum pairwise distance among barcodes for 237 collections from across Canada and north-central United States is 3.16%, or if two samples from Populus deltoides in north-central USA are excluded, 2.01%. The minimum pairwise distance to the most similar Chaitophorus species (Ch. nudus Richards) is 7.46%. Forty-two haplotypes were observed, of which 16 were replicated. Each replicated haplotype (along with similar unique haplotypes) is restricted to a particular geographic area (eastern, central or western part of the continent) and is specific to a particular Populus section (sections Populus, Tacamahaca and Agerat) (4 exceptions). Thus, in each region, sympatric host-specific haplotypes occur. We are currently examining nuclear genes and morphometric characters to confirm and assess the taxonomic significance of this variation.

POSSIBLE CRYPTIC SPECIES WITHIN PEMPHIGUS BETAE (APHIDIDAE: ERIOSOMATINAE)

Pemphigus betae Doane, the sugar-beet root aphid is widespread in North America. Barcode sequences reveal three distinct sympatric groups among 190 samples identified as this species. Pairwise genetic distances among the groups are of the same magnitude as distances among other Pemphigus species. The barcode-defined groups correspond exactly to groups previously identified using several nuclear markers. Morphometric analysis of this group is now being pursued.

EXPLORATION OF PATTERNS OF VARIATION AMONG CINARA SPECIES (APHIDIDAE: LACHNINAE)

Conifer aphids in the genus Cinara are a diverse group in North America and appear to have undergone extensive postglacial radiation on a number of hosts. While there is a wide range in morphological characters, many of these characters form a continuum across putative species. It is thus unclear which components of the observed variation are genetically based, and thus of taxonomic value, and which are the result of environmental effects. We have begun to use DNA barcoding as a tool for the exploration of species diversity in this group.

A neighbour-joining tree summarizing preliminary results for Cinara species is shown in Figure I. Species are well differentiated: within-species pairwise distances range from 0 to 2.03% (mean 0.52%) (assumes C. obscura groups discussed below are separate species) in contrast to minimum between-species differences of 2.7 to 13.2% (mean 7.91%). Clusters based on barcode sequences correspond to major host groups and to morphologically defined species groups. Three potential morphologically cryptic species are suggested within Cinara obtusa (Gillette and Palmer) on Picea, C. obscura Bradley on Picea and C. curvipes (Patch) on Abies. The separation of the C. obtusa groups shows an especially deep sequence divergence (labelled C. obtusa and C. sp. nr obscura in Figure I). The taxonomic status of these potential new species is currently being investigated using molecular and morphometric approaches.

Figure I – Neighbour-joining tree summarizing mitochondrial cytochrome oxidase subunit 1 sequence (DNA barcode) distances for species of Cinara. Host-plant groups are indicated on the right (Cup = Cupressaceae). Identical haplotypes are collapsed to a single node with number (n) of identical samples indicated for n>1.

ASSOCIATION OF MORPHS ON ALTERNATE HOSTS IN ADELGES TSUGAE (ADELGIDAE)

Adelges tsugae Annand, the hemlock woolly adelgid, is found in east Asia and in western North America, and has been introduced to eastern North America where it is a serious pest. As a result of morphological differences
between stages of the complex life cycles of adelgids, combined with the reduced number of characters for delineating species, it is often difficult to link life cycle forms of a species on different hosts. Host alternation in *A. tsugae* has been reported previously in Japan based on circumstantial evidence (Inouye 1953, McClure 1996), but this species has otherwise been collected only on *Tsuga*. Barcodes sequences for samples of *Adelges* collected from galls on spruce (*Picea*) in Japan and China match barcodes for *A. tsugae* populations on *Tsuga* in the respective region, supporting the presence of host alternation in this species.

**DISCUSSION**

The current expanded data set confirms the trends reported earlier by Foottit et al. (2008, 2009), namely that in general species of aphids and adelgids correspond to clusters of similar barcode sequences distinct from neighbouring clusters. Thus barcode sequences are useful as an identification tool in these groups. Congeneric species of aphids and adelgids usually form cohesive barcode assemblages, indicating that DNA barcodes can also be used for genus level identification of species not included in the barcode library. There are, however, some limitations for species identification in some species currently radiated groups such as *Illinios*, *Uroleucon* and *Macrospirum*.

Beyond simple identification, DNA barcodes can also provide a means for detecting and resolving taxonomic problems. In the case of banana aphid (*Pentalonia nigronervosa*), the use of barcodes has helped to clarify the status of host specific forms. Even in well-studied aphid groups, cryptic species are present. For example, cryptic entities within the economically important sugar-beet root aphid (*Pemphigus betae*) are clearly evident through DNA barcode analysis. Barcodes provide a simple method of identifying members of this group and thus will be useful in the monitoring and management of this pest.

With our work on *Canara* and other speciose genera shown that the use of DNA barcodes is an efficient means for surveying taxonomic diversity in particular groups or in particular geographic regions. Barcodes also provide a reliable means of associating life cycle forms and immature stages, the identification of which has been a difficult area in aphid taxonomy and an impediment to effective pest management.

**ACKNOWLEDGEMENTS**

This work was funded by Agriculture and Agri-Food Canada, the USDA Forest Service and USDA CSREES, and through grants to the Canadian Barcode of Life Network from Genome Canada (through the Ontario Genomics Institute), NSERC, the Moore Foundation and other sponsors listed at www.BOLNET.ca.

We thank the staff of the Biodiversity Institute of Ontario for processing samples and data management. We wish to thank the numerous people who provided material as acknowledged by Foottit et al. (2008 and 2009).

**REFERENCES**


Kimura M., 1980 – *A simple method for estimating evolutionary rate of base substitutions through
comparative studies of nucleotide sequences. - Journal of Molecular Evolution, 16: 11-120.


Bianca