



## DNA Barcoding: Mixed results for big-headed flies (Diptera: Pipunculidae)

JEFFREY H. SKEVINGTON<sup>1</sup>, CHRISTIAN KEHLMAIER<sup>2</sup> and GUNILLA STÅHLS<sup>3</sup>

1. Invertebrate Biodiversity, Agriculture and Agri-Food Canada, 960 Carling Avenue, K.W. Neatby Building, Ottawa, ON, K1A 0C6, Canada; e-mail: skevingtonj@agr.gc.ca; url: www.canacoll.org. Telephone: 613-759-1647. Fax: 613-759-1927.

2. State Natural History Collections Dresden, Museum of Zoology, Königsbrücker Landstrasse 159, D-01109 Dresden, Germany; e-mail: kehlmaier@web.de.

3. Finnish Museum of Natural History, Entomology Department, PO Box 26 (Teollisuuskatu 23), FI-00014, University of Helsinki, Finland; e-mail: gunilla.stahls@helsinki.fi.

### Abstract

Sequence data from 658 base pairs of mitochondrial cytochrome c oxidase I (*cox1*) were analysed for 28 described species of Pipunculidae (Diptera) in an effort to test the concept of DNA Barcoding on this family. Two recently revised but distantly related pipunculid lineages with presumed different evolutionary histories were used for the test (*Clistoabdominalis* Skevington, 2001 and *Nephrocerus* Zetterstedt, 1838). An effort was made to test the concept using sister taxa and morphologically similar sibling species swarms in these two genera. Morphological species concepts for *Clistoabdominalis* taxa were either supported by *cox1* data or found to be too broad. Most of the discordance could be accounted for after reassessing morphological characters. In these cases, the molecular data were invaluable in assisting taxonomic decision-making. The radiation of Nearctic species of *Nephrocerus* could not be diagnosed using *cox1*. The ability of *cox1* to recover phylogenetic signal was also tested on *Clistoabdominalis*. Morphological data for *Clistoabdominalis* were combined with the molecular data set. The pipunculid phylogeny from molecular data closely resembles the published phylogeny based on morphology. Partitioned Bremer support is used to localize areas of conflict between the datasets.

**Key words:** Pipunculidae, Diptera, DNA Barcoding, phylogeny, mitochondrial rDNA, *cox1* gene, COI, molecular systematics, morphology, total evidence

### Introduction

It has been proposed that the mitochondrial gene cytochrome c oxidase I (*cox1*—also referred to as COI) can be used as the core of a global identification system for animals (Hebert et al. 2003a). We decided to test this concept of ‘DNA barcoding’ on big-headed flies (Pipunculidae) using two recently revised genera. Six hundred and fifty-eight base pairs of *cox1* were sequenced for 17 of 29 species of Australian *Clistoabdominalis* and all six species of Nearctic *Nephrocerus*. Three Palaearctic species of *Nephrocerus* were included in the analysis to help interpret the results for this lineage.

#### *Background and Philosophy of DNA Barcoding*

Molecular data has long been used to help with species delimitation (Sperling 2003). In bacterial systematics, 16S rRNA has been the standardized dataset used for new species description for years (Lipscomb et al. 2003; Theron and Cloete 2000; Vandamme et al. 1996). In bacteriology, it has often also been used as the only dataset for phylogeny reconstruction and classification and as a result has many detractors (e.g. Gest 1999).

Only recently has there been a push to use a specific standardized molecular dataset across the animal kingdom (Hebert et al. 2003a). A call to focus on only four genes, including COI, across all insects was made

earlier by Caterino et al. (2000). Hebert et al.'s (2003a) initiative was coined 'DNA barcoding' and a number of claims were made during its formative period. These included the use of DNA barcodes to drive species identification, species definition, and phylogeny estimation (Hebert et al. 2003a). These proposed goals, combined with a perceived attack on traditional morphological systematists served to polarize factions in favour and against DNA barcoding. For example, Hebert et al. (2003a, 2003b, 2004a, 2004b), Hebert and Gregory (2005), Janzen et al. (2005), Proudlove and Wood (2003), Smith et al. (2006), and Tautz et al. (2003) strongly support barcoding while detractors of the method include Dunn (2003), Lipscomb et al. (2003), Prendini (2005), Seberg et al. (2003), Sperling (2003), and Will and Rubinoff (2004). Testing the barcoding concept on individual lineages of animals is one potential way to understand the benefits and pitfalls of the method.

We feel that although some of the claims were initially inflated, recent ideals promoted by the barcoding community are more tempered and realistic. For example, Hebert et al. (2003a) claimed 100% success in identifying lepidopterans in their preliminary study. More recently Hebert and Gregory (2005), Janzen et al. (2005), and Ward et al. (2005), based on a considerable amount of data, suggested that success rates of 95–97% are more realistic. Smith et al. (2006) present the first test of the concept in Diptera, applying *coxI* data to a diverse assemblage of *Belvosia* Robineau-Desvoidy, 1830 (Tachinidae). Their conclusions suggest that morphological concepts are largely supported by molecular data, but that morphology alone does not discover the true diversity hidden within some species groups. Ecological data and morphological data combined more closely to reflect the molecular species hypotheses within these cryptic species groups. Our data presented below are largely congruent with this study, but also clearly illustrate that problem groups exist where *coxI* data provide no useful signal at the species level (another example of this can be seen in Kaila & Ståhls (2006).

The data that we present below clearly illustrate some of the benefits and pitfalls of DNA barcoding. Used on its own, *coxI* data is of no more utility than reliance on any single suite of characters (e.g. male genitalia, wing venation, body colour or vestiture, etc.). However, if twinned with at least one additional dataset such as morphology, there should be no argument that *coxI* data is a superb addition to taxonomic revisions. Problems inherent in the morphological dataset are illuminated by the *coxI* data and conversely, problems with the *coxI* data are revealed using morphology. Note that this does not save time in producing taxonomic revisions as envisioned by many supporters of barcoding (e.g., Hebert and Gregory (2005)), but it does produce stronger hypotheses that clearly illustrate weak areas that require further research. The end result is time saved over the long term. Hypotheses for most species concepts included in such a competent revision (including examination of types, barcoding of 10 or more specimens from many populations of most species, etc.) will be robustly supported and thus may not need to be revisited. Focusing future efforts on taxa within the genus that clearly present problems will then be the future focus of research, rather than revising the entire group every 20–50 years as new datasets emerge.

One issue that has been oversimplified revolves around the cost of DNA Barcoding. We feel that this has never been clearly examined. The current cost estimate for DNA barcoding is \$5 per specimen (Hebert and Gregory 2005) ("including labour and sequencing"). This evidently is intended to cover only the costs of molecular work. Cost to a morphological lab includes: locating the appropriate specimen in a collection (or actually going into the field to collect it), removing tissue for DNA extraction, databasing and photographing the specimen, and packing and shipping the material to a molecular lab. This at least doubles the cost as the labour cost alone for this equates to at least 30 minutes per specimen even without consideration of field collection costs (based on actual timed trials in the first author's lab). At \$10.72 per hour (the lowest rate that we can pay a student at Agriculture and Agri-Food Canada) this costs \$5.50. Since some of the effort will require scientist involvement at a much higher rate of pay (for example, selection of the specimens to be used), the real cost is at least \$7.00 per specimen. Note that although the cost of DNA barcoding (the molecular work) is expected to fall considerably over time, the cost to morphological labs will remain constant or increase, as further field exploration is required to provide the raw material for the molecular work. This is not a complaint

about DNA barcoding, only a statement to indicate that future funding decisions should focus on dividing the resources appropriately. We feel that money going into DNA barcoding should benefit all systematics but it should be clearly earmarked to be spread equally among participants. Money should thus go towards:

1. Paying for sequencing and molecular lab development.
2. Supporting miniaturization and cost effectiveness of barcoding.
3. Supporting field collection of tissues for analysis by scientists conducting the revision.
4. Supporting natural history collection development.
5. Supporting collection databasing.
6. Supporting traditional morphological revision that is tied to DNA barcoding.

#### *Background on Pipunculidae Systematics*

Pipunculidae are a monophyletic lineage of exclusively parasitoid flies (Aczél 1948; Rafael and De Meyer 1992; Skevington and Yeates 2000; Skevington and Yeates 2001). Most species (all members of Pipunculinae and Chalarinae) are endoparasitoids of Auchenorrhyncha, usually attacking nymphs. *Nephrocerus* were recently discovered to be endoparasitoids of adult Tipulidae (Diptera) (Koenig and Young 2007; Skevington 2005). Along with some other parasitoid Diptera (such as Tachinidae), Pipunculidae have long been recognized as one of the more difficult groups to resolve taxonomically. Taxonomy for the group has evolved from a reliance on external characters (many of which are now recognized to be variable) to the present reliance on male genitalic characters. Dissection of males is usually necessary to confirm identification, a time consuming step. Even when using current keys, non-specialists often have difficulty identifying these flies, because of the need to interpret complex genitalic characters and the similarity of taxa within sibling species swarms. To add to the problem, male and female flies of one species usually look very different and association of the sexes is problematic. Even once associations are made, females are often extremely difficult to separate from closely related species. Many species are also rarely encountered (although they are possibly common in nature). This results in small numbers of specimens being available to specialists for study and the corresponding difficulty separating variation from discrete character states. Larvae are unidentifiable to species, and in many cases even to genus. This creates a problem when applied biologists attempt to use pipunculids for biological control or other research. For example, to identify pipunculid larvae recently discovered in *Oncometopia orbona* (Fabricius, 1798), Skevington et al. (2007) had to sequence parts of several genes (*cox1*, 12S and CAD) just to be able to confidently identify the specimens to genus. Species level identification will require rearing the larvae, a process that will take at least a year and has no guarantee of success.

DNA Barcoding is currently being tested on groups such as birds (Aves) (Hebert et al. 2004b), fish (Teleostei) (Ward et al. 2005), butterflies and moths (Lepidoptera) (Hebert et al. 2004b; Janzen et al. 2005; Hajibabaei et al. 2006), mayflies (Ephemeroptera) (Ball et al. 2005), spiders (Araneae) (Barrett and Hebert 2005), ants (Formicidae) (Smith et al. 2005), springtails (Collembola) (Hogg and Hebert 2004), tachinid flies (Tachinidae) (Smith et al. 2006), agromyzid flies (Agromyzidae) (Scheffer et al. 2006) and big-headed flies (Pipunculidae) (Skevington 2006). With all of these projects on groups with varying evolutionary histories and specific problems, we should be able to better evaluate barcoding in the near future. Members of some of these groups are relatively easy to identify in the field, but the real potential strength of DNA barcoding will be displayed when applied to groups like Tachinidae, Pipunculidae, Formicidae, and Ephemeroptera where there is a major taxonomic impediment and serious challenges for technology transfer even after the taxonomy has been resolved. This study aims to test the application of this technique to resolving problems within Pipunculidae. The specific goals are as follows:

1. Contrast and test the ability of *cox1* to diagnose species in two lineages of Pipunculidae that are hypothesized to have evolved at different evolutionary rates.
2. Attempt to identify unknown female pipunculids using the *cox1* tool.
3. Test the utility of *cox1* for phylogenetic inference in Pipunculidae.

4. Explore the benefits and potential pitfalls of DNA Barcoding as promoted by Hebert et al. (2003a). Can it be used as a routine part of revisionary work or on its own as a biodiversity assessment tool?

#### *Choice of groups*

Two genera of pipunculids were chosen for this study because of the availability of recent morphological revisionary work on the groups and our belief that they were likely to have contrasting evolutionary histories. The latter was supported by differences in morphological characterization of species within these genera. Australian *Clistoabdominalis* include many species groups, some endemic, and each with presumably different evolutionary pressures and histories (Skevington 2001). Nearctic *Nephrocerus* species appear to have undergone different evolutionary constraints from Palaearctic species. The latter are easily differentiated by numerous morphological characters whereas the Nearctic species are much more similar to each other. As a result, we suspected that there might have been a rapid, recent radiation of the Nearctic taxa. Australian *Clistoabdominalis* were revised recently by Skevington (2001) and revision of Nearctic *Nephrocerus* was completed in 2005 by Skevington (2005).

## Materials and Methods

#### *Specimen Acquisition*

Voucher data are available in Appendix 1. Outgroup taxa (*Chalarus* Walker, 1834) were collected into absolute ethanol in 2001 and 2005 by Skevington. Some ingroup taxa were collected into absolute alcohol; however, most were killed with cyanide and air-dried (Appendix 1). Most specimens used were collected in the last eight years but specimens as old as 24 years were successfully sequenced. For the genus *Clistoabdominalis*, 71 specimens were successfully sequenced (Appendix 1). Of these, 29 samples were preserved in absolute ethanol and stored at  $-80^{\circ}\text{C}$  and 42 specimens were pinned (see Appendix 1 for breakdown of years collected). We were unable to sequence 45 specimens, all but 4 of them dried and pinned. For the genus *Nephrocerus*, 14 specimens were successfully sequenced (Appendix 1). Of these, seven were preserved in alcohol as above and seven were dried and pinned. We were unable to sequence 27 specimens, all of them pinned, despite several attempts for some of the individuals.

An attempt was made to sample all Nearctic species of *Nephrocerus* and most species within the *helluo* and *koebeleri* groups of Australian *Clistoabdominalis*. In the end, all six Nearctic species of *Nephrocerus*, the three common European *Nephrocerus* species and 17 of 29 Australian species of *Clistoabdominalis* were sequenced. Four of six species in the *koebeleri* group and five of nine species in the *helluo* group were sequenced. Population level sampling was attempted where possible but was limited to one or two extra specimens for most taxa. *Clistoabdominalis ancylus* Skevington, 2001 was sampled most heavily to explore population variation within *cox1* (28 specimens from 12 populations in three states were sequenced) (Appendix 1). Two species from one other pipunculid genus were used to root the tree (*Chalarus* spp.). All specimens are accompanied by a label with a unique number (in the format J. Skevington Specimen # *n*). Specimens in alcohol have had the genitalia removed, macerated in lactic acid, and placed in glycerin in a microvial on a pin with a label bearing collection data and a unique number. The remainder of the specimen is maintained in the Canadian National Insect Collection DNA repository at  $-80^{\circ}\text{C}$ , except for the remains of the three Palaearctic species of *Nephrocerus* deposited at the Zoological Museum of the Finnish Museum of Natural History in Helsinki, Finland. Complete label data and a duplicate of the voucher number are in the vial with the remaining tissue. Additional voucher specimens are deposited in the following collections (standardized codons given below and in Appendix 1 follow those of Evenhuis and Samuelson (2004): Australian Museum (AMS, Sydney, New South Wales, Australia, D. Bickel), California Academy of Sciences (CAS, San Francisco, California, USA, N. Penny), Canadian National Collection of Insects (CNC, Ottawa, Ontario, Canada, J. Cum-

ming), Zoological Museum of the Finnish Museum of Natural History (MZH, Helsinki, Finland, G. Ståhls), Florida State Collection of Arthropods (FSCA, Gainesville, Florida, USA, G.J. Steck), Illinois Natural History Survey (INHS, Champaign, Illinois, USA, D.W. Webb), Queensland Museum (QM, Brisbane, Queensland, Australia, C. Burwell), Texas A-&M University Insect Collection (TAMU, College Station, Texas, USA, J. Oswald and E.G. Riley), University of Connecticut Entomology Collections (UCMS, Storrs, Connecticut, USA, J.E. O'Donnell), University of Guelph Insect Collection (DEBU, Guelph CA, S.A. Marshall), W.R. Enns Entomology Museum (UMRM, University of Missouri, Columbia, Missouri, USA, R. Sites), and University of New Hampshire Entomological Museum (DENH, Durham, New Hampshire, USA, D.S. Chandler).

#### *DNA extraction*

Nucleic acid extractions followed the tissue protocol from the QIAamp DNA Mini Kit (Qiagen). Steps 3b and 7a were omitted and full speed centrifugation was performed at 13,000 rpm. The total extraction volume (100 µl) was obtained by spinning 2 x 50 µL volumes of Buffer AE. Extractions for the Palaearctic *Nephrocerus* species were performed using the Nucleospin Tissue Kit and standard protocol (Machery-Nagel).

#### *cox1 PCR Amplification and Sequencing*

A 658 base pair fragment of the 5' end of the COI gene (now referred to as *cox1* in the 'barcoding' literature) was amplified using the primer pair LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAATCA-3') (Folmer et al. 1994). Each PCR mixture contained: 5 µl 10x *Ex Taq* PCR buffer (20 mM Mg<sup>2+</sup>), 36.75 µl RO H<sub>2</sub>O, 4 µl of dNTP mixture (2.5 mM of each dNTP, pH7~9), 0.25 µL of *TaKaRa ExTaq* Hot Start *Taq* polymerase (5 units/µL), 1 µL of each primer (5 µM), and 2.0 to 8.0µl of DNA template (the difference was made up with RO H<sub>2</sub>O as required). Amplification was achieved using the following program: 94°C (2 minutes); 34 cycles of 94°C (30 s), 49.5°C (2 minutes), 72°C (1 minute); 1 cycle of 72°C (10 minutes). Cycling conditions for the Palaearctic species of *Nephrocerus* were slightly different since they were carried out in a different lab (CK, GS): 1 cycle of 95°C (2 minutes); 35 cycles of 94°C (30 s), 45°C (30 s), 72°C (2 minutes); 1 cycle of 72°C (8 minutes), 15°C (10 minutes). The PCR products were then purified using the QIAquick PCR Purification Kit Protocol (Qiagen). All centrifugation steps were performed for 60 s and samples were left to incubate for 1–2 minutes in Buffer PE (to reduce salt concentration), and 5 minutes in Buffer EB. Purified products were sequenced in both directions using the PCR primers on an Applied Biosystems Incorporated PRISM® 3100-*Avant* Genetic Analyzer using the ABI PRISM® Big Dye® Terminator v.3.1 Cycle Sequencing Kit. The Palaearctic species of *Nephrocerus* were sequenced in both directions on an ABI 377 sequencer (Applied Biosystems) using the Big Dye® Terminator v.1.1 Cycle Sequencing Kit. The considerable overlap of strands (~90%) allowed confident assessment of nucleotide identity for all sequences analysed. All sequences were deposited in GenBank under the accession numbers DQ337628 to DQ337706 and DQ349221 (Appendix 1).

#### *Data Analysis*

No insertions or deletions occur in the dataset so alignment was unambiguous. The alignment used is lodged in the EMBL Nucleotide Sequence Database (ALIGN\_001041). To access this from the EMBL-EBI homepage at <http://www.ebi.ac.uk/embl/index.html>, select 'databases', then 'nucleotide databases', then 'EMBL-Align database' (the direct link for this is currently at <http://www.ebi.ac.uk/embl/Submission/alignment.html>). From here, log onto their ftp database (<ftp://ftp.ebi.ac.uk/pub/databases/embl/align>) and download ALIGN\_001041).

Phenetic and parsimony analyses were performed with PAUP\* (Swofford 2001). Character polarity was based on outgroup comparison (Nixon and Carpenter 1994). Two *Chalarus* species were defined as the outgroup for all analyses. Neighbour joining was used to produce the phenograms (using standard PAUP\* defaults). For parsimony analysis, the heuristic search procedure was used with stepwise-addition and 100

random replications. The heuristic search option was used with tree bisection-reconnection branch swapping, MULPARS, and random addition of taxa. Multistate characters were treated as non-additive. All individuals were analysed separately.

Evidential support for different clades was assessed using the nonparametric bootstrap (BS—1000 replicates using the same parameters described for our initial parsimony analysis) (Felsenstein 1985) and branch support (Bremer 1994). The latter was calculated with the program TreeRot 2c (Sorenson 1999). Branch support indicates the number of extra steps from the most parsimonious solution at which clade(s) fail to be resolved as successively longer trees are examined. A high numerical value for a clade indicates good support. Tree measures such as the consistency index (CI) and the retention index (RI) were used to evaluate the fit of the data to the tree.

Partitioned Bremer Supports for the total evidence *Clistoabdominalis* analyses (morphology data from Skevington (2001)) were calculated using TreeRot 2c (Sorenson, 1999). Only one molecular terminal was used to represent each putative species in order to simplify the analysis.

## Results

### General

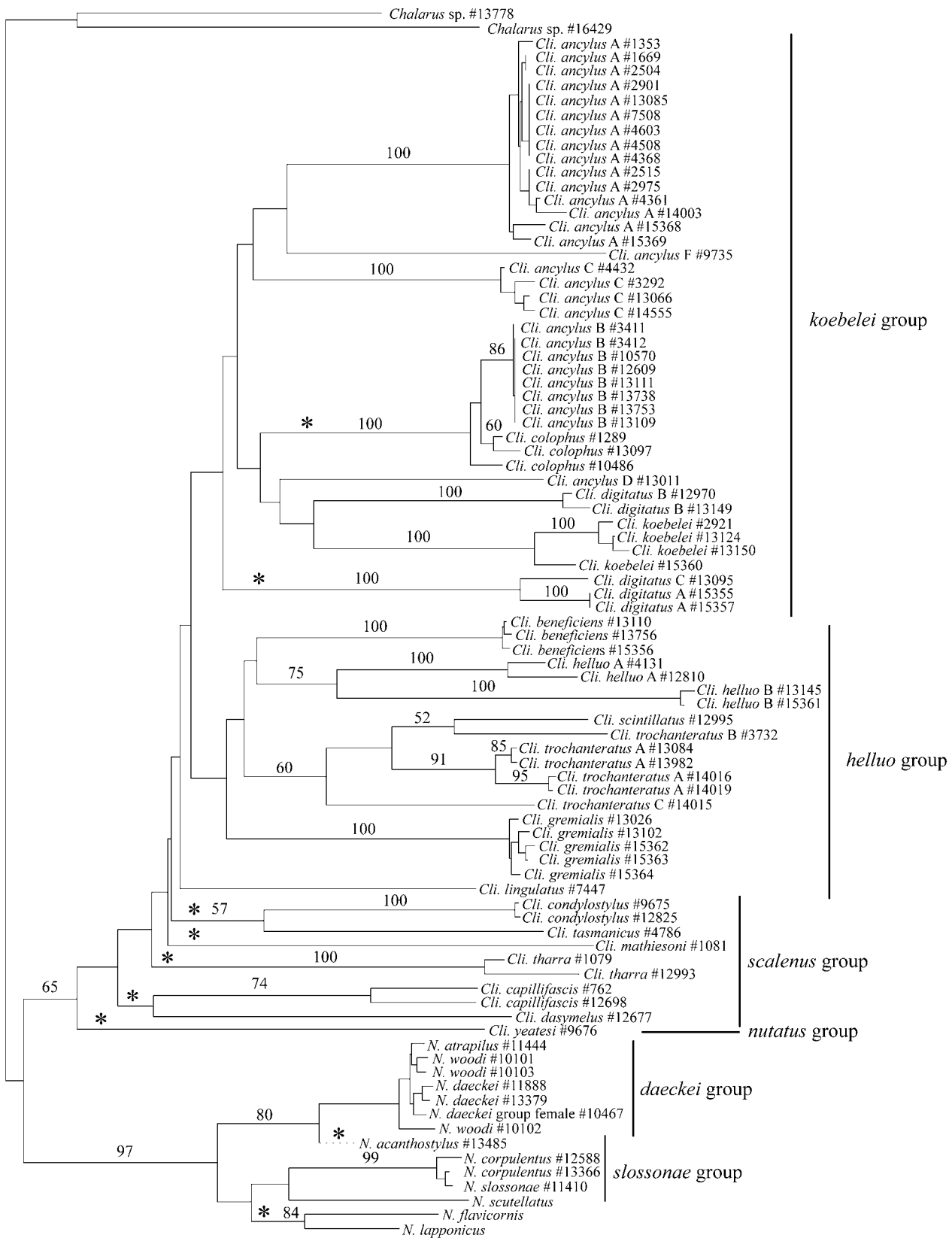
The nucleotide frequencies were typical of insect mitochondrial DNA, with an AT bias averaging 66%. The number of parsimony informative nucleotide characters for phylogenetic analysis of *cox1* was 245. A neighbour joining phenogram is presented in Figure 1. Branch supports are indicated using bootstraps. Parsimony analysis resulted in 799 equally parsimonious trees (tree length = 1398, consistency index (CI) excluding uninformative characters = 0.28, retention index (RI) = 0.78, rescaled consistency index (RC) = 0.23) (Figure 2). In this tree, putative clusters of species supported by neighbour joining analysis are supported but relationships between species are often unresolved. Successive weighting of this data produced a single most parsimonious tree (Fig. 3) that is similar in topology to the neighbour joining tree. Differences are all in the *scalenus* group with one exception: *Cli. digitatus* Skevington, 2001 A+C move into the *scalenus* group.

### *Clistoabdominalis*—Molecular Analysis

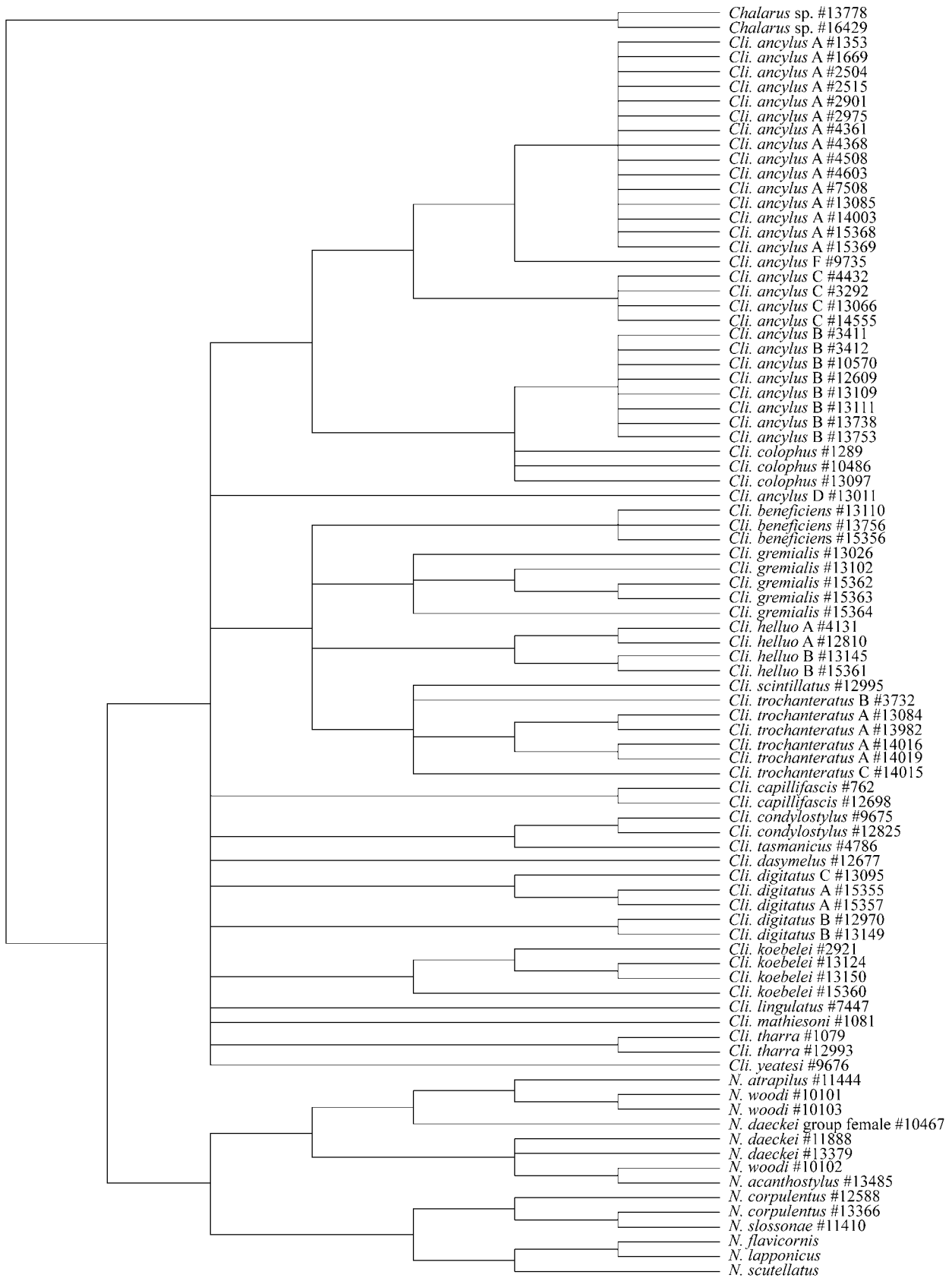
The uncorrected pairwise sequence divergences for *cox1* within *Clistoabdominalis* varied between 0.0% and 18.7% (data available from authors upon request). Most species pairs are over 9% different. For example: *Cli. ancylus* A to C (9.2–10.2%), *Cli. ancylus* A to B (9.4–10.3%), *Cli. digitatus* A to B (11.7–12.6%), *Cli. helluo* (Perkins, 1905) A to B (9.5–10.6%), *Cli. beneficiens* (Perkins, 1905) to *Cli. helluo* A and B (8.4–13.5%), and *Cli. koebelei* (Perkins, 1905) to *Cli. digitatus* A and B (9.5–12.5%). These data sample putative new species complexes discovered here as well as the well-supported species pairs or complexes proposed by Skevington (2001). Other close relatives (not necessarily sister taxa) are equally divergent and include: *Cli. tasmanicus* Skevington, 2001 to *Cli. condylostylus* Skevington, 2001 (9.2–9.5%), *Cli. tharra* Skevington, 2001 to *Cli. mathiesoni* Skevington, 2001 (14.0–15.8%), *Cli. capillifascis* Skevington, 2001 to *Cli. tharra* (13.3–15.3%), *Cli. capillifascis* to *Cli. dasymelus* Skevington, 2001 (11.6–12.3%), and *Cli. tharra* to *Cli. tasmanicus* (12.0–14.1%). The only clear exceptions of sister taxa with lower than 9% between species variation includes *Cli. colophus* Skevington, 2001 to *Cli. ancylus* B (1.1%) and *Cli. trochanteratus* (Becker, 1900) B to *Cli. scintillatus* Skevington, 2001 (5.0%).

Within species variation is more difficult to delineate. There is no variation between any of the eight specimens of *Cli. ancylus* B surveyed despite the fact that these specimens are from four populations in two states (Appendix 1). There is up to 1.2% variation within *Cli. ancylus* A. The 1.8% difference between *Cli. helluo* A specimens from Fly Hill (Queensland) and Kakadu (Northern Territory) could represent two species or variation within one species. Similarly, it is difficult to determine whether variation in the following taxa represents intraspecific variation or the presence of cryptic species: *Cli. koebelei* (2.3%), *Cli. digitatus* A and C (2.4%),

*Cli. tharra* (2.0%), *Cli. capillifascis* (3.7%). There is no morphological variation (internal or external) to separate any of these groups that vary in *cox1* sequences into putative species.

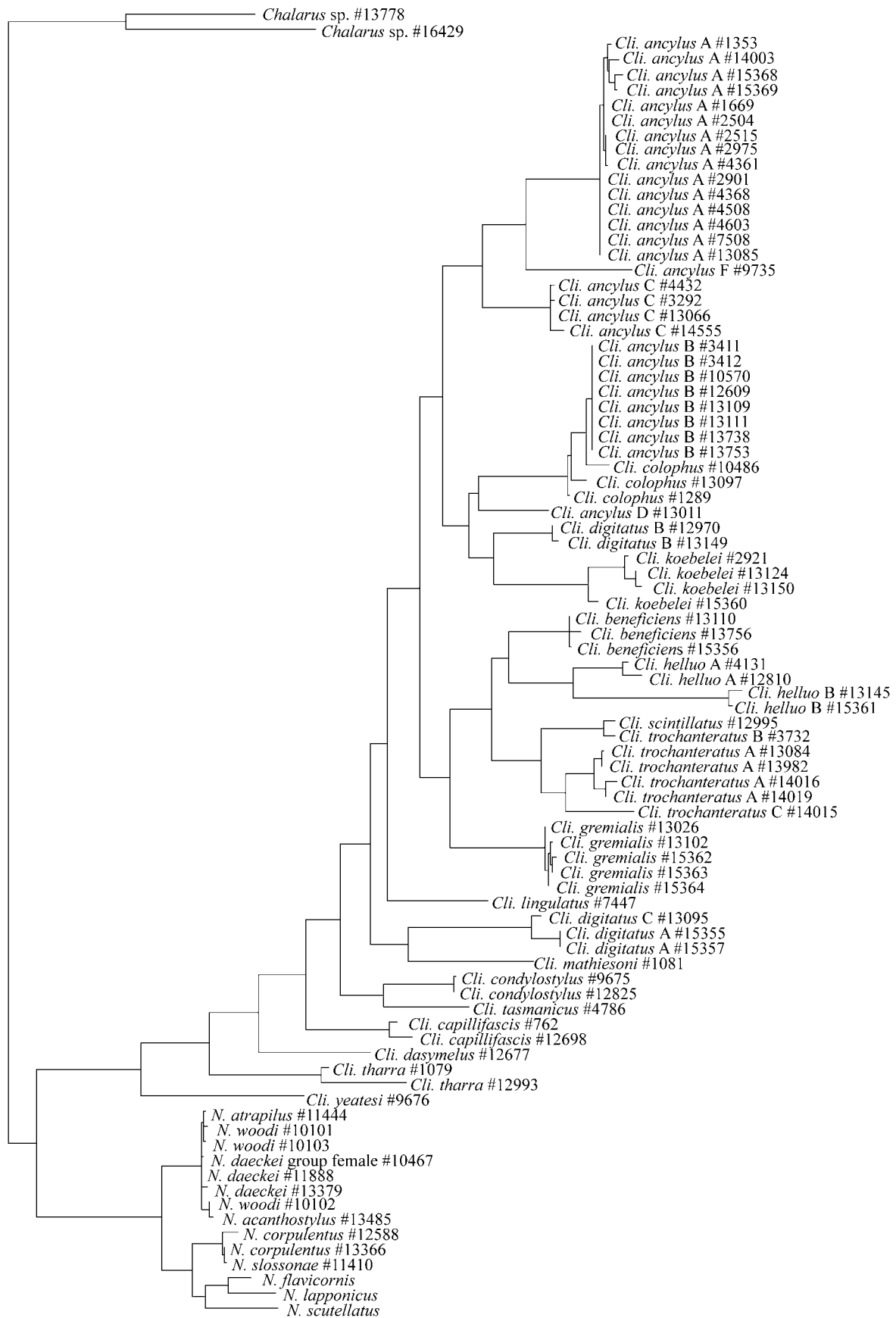


**FIGURE 1.** Neighbour joining tree of *Clistoabdominalis* and *Nephrocerus* relationships inferred from *cox1* data. Numbers above the branch are bootstrap values. Asterisks indicate branches that are incongruent with respect to previous morphological hypotheses of Skevington (2001, 2005).



**FIGURE 2.** Consensus tree of *Clistoabdominalis* and *Nephrocerus* relationships inferred from *cox1* data (from 799 most parsimonious trees).





**FIGURE 3.** Single successively weighted tree of *Clistoabdominalis* and *Nephrocercus* relationships inferred from *cox1* data.

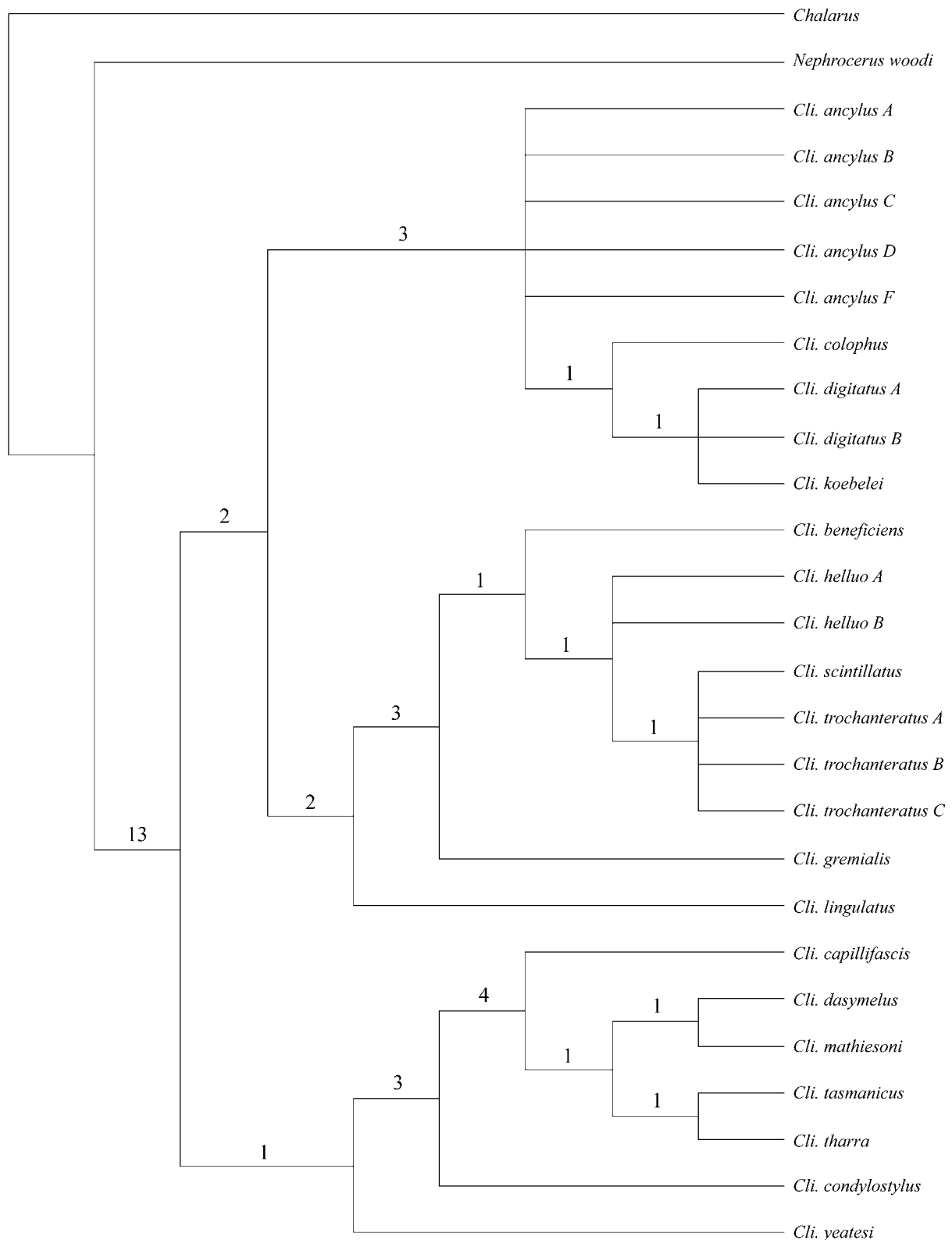
### *Clistoabdominalis*—Morphological Analysis and Total Evidence Analysis

These analyses utilized two outgroup taxa: *Chalarus* sp. and *Nephrocerus woodi* Skevington, 2005. Results of the morphological analysis are illustrated in Figure 4 (including Bremer supports; a consensus tree from two most parsimonious trees, tree length = 101, CI excluding uninformative characters = 0.71, RI = 0.89). Tree topology is identical to that shown by Skevington (2001) despite omitting 12 species from this analysis. The total evidence cladogram is illustrated in Figure 5. Both Bremer support and bootstraps are included to indicate branch supports. Partitioned Bremer support is used to show the relative contribution of each dataset to the overall branch support. Nodes with conflicting signal between the morphological and molecular datasets are clearly highlighted using this technique (one partition appears with a positive signal and the other is negative). Conflict is restricted to six nodes. Most differences between the morphological analysis and the total evidence analysis are restricted to the *scalenus* group. Skevington (2001) acknowledged that this was the weakest part of the morphological phylogenetic hypothesis and that additional data were likely to suggest changes in this part of the tree. Despite low branch supports, Skevington (2001) felt that the *helluo* group relationships were sensible and likely came close to the true phylogeny. The addition of *cox1* data largely supports that contention. The only changes are the inclusion of *Cli. colophus* within the *Cli. ancylus* complex and the sister species relationship between *Cli. helluo* and *Cli. beneficiens*. The latter move makes sense (these taxa are very similar morphologically) and was discussed as a possible scenario by Skevington (2001). The inclusion of *Cli. colophus* within the *Cli. ancylus* complex makes sense now that we perceive that there are many more species involved in what was initially proposed to be only two species.

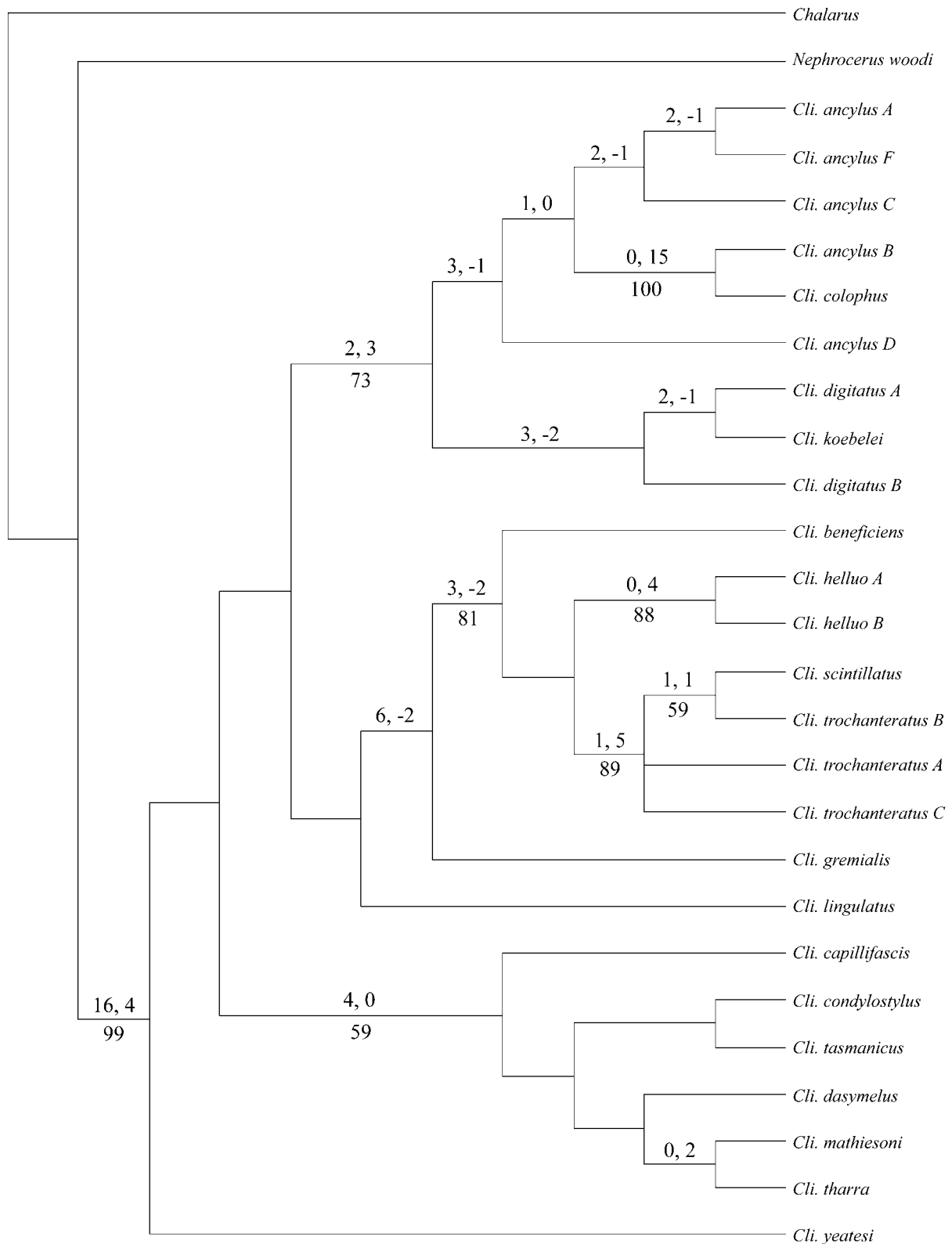
Although this total evidence phylogeny supports many aspects of the molecular phylogenies illustrated in Figures 1–3, some of the more unrealistic molecular hypotheses are neutralized. For example, the two putative species of *Cli. digitatus* remain together in a lineage with *Cli. koebelei* (in separate clades in Figs 1–3). These species are all very similar morphologically and undoubtedly related. The putative species of *Cli. ancylus* also form a lineage in the total evidence analysis (as would be expected from their morphology). Other differences occur in the *scalenus* group, where branch supports are still low and more data are certainly required to obtain a realistic hypothesis of relationships.

### *Nephrocerus*

*Nephrocerus* proved to contain much less *cox1* variation than *Clistoabdominalis*. The uncorrected pairwise sequence divergences for *cox1* within *Nephrocerus* varied between 0.2% and 7.9%. The three Palaearctic species varied from 4.0 to 7.9% from each other. Two species groups for the Nearctic taxa were proposed by Skevington (2005), the *slossonae* group (containing *N. acanthostylus* Skevington, 2005, *N. corpulentus* Skevington, 2005 and *N. slossonae* Johnson, 1915) and the *daeckei* group (containing *N. atrapilus* Skevington, 2005, *N. daeckei* Johnson, 1903 and *N. woodi*). The molecular data mostly supports this, but suggests that *N. acanthostylus* is more closely allied with the *daeckei* group. The single *N. acanthostylus* sequence is short (265 bp) so this result should be viewed with reservation. Uncorrected pairwise sequence divergences between Nearctic taxa in the two species groups was 3.9 to 5.7%. Variation within Nearctic species of the *slossonae* group was 0.2 to 0.6% and within the *daeckei* group was 0.8 to 1.9% (0.3–1.2% when *N. acanthostylus* is excluded). The Palaearctic species sampled are weakly supported as members of the *N. slossonae* group (Fig. 1) and differ from the Nearctic species by 5.4 to 6.2%. Morphologically, the Palaearctic species are more closely allied to Nearctic species of the *daeckei* group. These areas of conflict between the datasets will require further research, applying ecological data and additional genes. A useful application of ITS1 is illustrated by Smith et al. (2006), to tease apart possible hybrid events for closely related taxa within one lineage.



**FIGURE 4.** Consensus tree of *Clistoabdominalis* relationships inferred from morphological data (reduced taxon dataset to reflect the molecular taxa used in this study; using matrix from Skevington (2001)) (from two most parsimonious trees). Numbers above the branches are Bremer support values.



**FIGURE 5.** Consensus tree of *Clistoabdominalis* relationships inferred from morphological and molecular (*cox1*) data (from two most parsimonious trees; tree length = 1219, CI excluding uninformative characters = 0.33, RI = 0.42). Morphological data was taken from Skevington (2001). Numbers above the branches are partitioned Bremer support values; the first partition represents branch support conveyed from morphological data, the second represents support conveyed from *cox1* data. Numbers below the branches are bootstrap values.

## Discussion

### Overview

Morphological systematic research on Pipunculidae has been evolving since the description of the first pipunculid in 1802 by Latreille (*Pipunculus campestris*). The type series for this first described pipunculid contains three syntypes in what we now recognize as three genera (unpublished data). In those early days of pipunculid taxonomy species concepts were based on very broad similarities. Later, scientists used a suite of external features including colour, antennal shape, comparisons of wing venation, etc. to aid in species discovery. Many of these characters are now considered to be too variable to be of much use and scientists now typically circumscribe species based primarily on male genitalia. Despite the apparent improvements and evolution of taxonomy in the group, there have been few independent tests of species concepts using alternate data sets. Jervis (1992) was one of the few who combined ecological data and morphology with a revision of *Chalarus*. The ability to apply molecular data to these problems has been available for many years now but the only work that even skirted around the issue of species concepts in the family was by Skevington and Yeates (2000). This was a phylogenetic analysis but it applied 12S and 16S data to the problem and explored some of the intraspecific and interspecific mitochondrial variation in a few chosen exemplars. Very recently, Skevington (2006) applied *cox1* data to test morphological species concepts within Australasian and Oceanian species of *Collinias* Aczél, 1940. This was the first paper to apply barcoding data as part of a Diptera revision and actually follows the work presented in this paper.

This paper and Skevington (2006) are the first to truly test some of our species concepts for pipunculids and were conducted with a dual purpose in mind. It also serves to test the concept of DNA Barcoding proposed by Hebert et al. (2003a). We could have used a variety of genes to test species concepts in the family, but we really wanted to see if *cox1* would do the job as proposed. We feel that if two essentially independent datasets support the same species concepts, we can consider these concepts to be robust. When these two datasets are contradictory or uninformative, another dataset will have to be applied. We did not attempt to apply other datasets here as we simply set out to see how much of the species diversity within two lineages of Pipunculidae can be clearly described using morphology and *cox1*.

### *Clistoabdominalis*

In Skevington (2001), several potential species complexes or species with unusual disjunct ranges were noted. It was assumed that more species might be found in these lineages if other data sets (such as molecular data) were explored. Based on morphology, species complexes were suspected in what were treated as the following species: *Cli. ancylus*, *Cli. helluo*, *Cli. lingulatus* Skevington, 2001, *Cli. mathiesoni*, *Cli. nutatus* Skevington, 2001, *Cli. octiparvus* Skevington, 2001, and *Cli. trochanteratus*. Regional differences in behaviour in *Cli. helluo* (with some populations using hilltop aggregation sites and others not exhibiting this mating behaviour) were also suspicious and may be explained by the presence of more than one species in what was treated as one. As seen from the data presented above, these suspicions were well founded. In contrast to this, the *koebeleri* group posed an interesting problem. With the exception of *Cli. ancylus* and its variants, the included species were so similar that Skevington (2001) considered that they could all be one variable species. In particular, *Clistoabdominalis digitatus* and *Cli. koebeleri* are so similar that Skevington (2001) suggested that they could be conspecific. He suggested testing the species concepts of these species using molecular data. In fact, from the results shown above, molecular data support the existence of these two species and even support the existence of another cryptic species within what was treated as *Cli. digitatus*. Even more extreme is the example of *Clistoabdominalis ancylus* (also part of the *koebeleri* group) that based on the results above is likely a complex of five or more species. The following section examines the putative new species of *Clistoabdominalis* supported by *cox1* data.

### *Exploration of the Putative New Species of Clistoabdominalis*

*Cli. ancylus* complex. This 'species' was clearly marked as a problem and potential complex by Skevington (2001). For this reason, we sampled this group more heavily than any other in this study. We are now confident that this group includes five to six species (likely five), here treated as A, B, C, D, E and F. Uncorrected pairwise sequence divergences for *cox1* between these species is from 9.1 to 12.2%. Species A (*Cli. ancylus* s.s.) has 2 morphological forms, A1 and A2. These forms are sympatric and *cox1* data suggest that they are a single species (pairwise distances 0 to 1.4%). They are restricted to dry sclerophyll forest. Compared to A1, A2 has a wider subepandrial sclerite, the horn on the right surstylus is wider and the surstyli are more concave, particularly the right surstylus. This was considered to be part of normal genitalic variation by Skevington (2001). Based on the discovery of several cryptic species supported by similar amounts of morphological variation and by the molecular data, this is now less clearly variation. More gene regions will have to be investigated to find out whether the forms A1 and A2 represent different species. However, in this instance, we suspect that they are truly one species as suggested by the small amount of variation within *cox1* and based on the fact that they always co-occur in space and time. Species B is ecologically separate from species A and occurs only in rainforest. Species B also has a wider subepandrial sclerite than A1 (similar to A2). The horn on the right surstylus is the most distinctive autapomorphy – it is bent forward at 90° and is short and thick. The horn on sternite 6 is also narrower than other species in the complex. The ejaculatory ducts also appear to be slightly thicker and specimens are smaller than in all other species of the complex. Species C also appears to be restricted to rainforest; however, a specimen from Mt. Coot-tha was captured in dry sclerophyll forest. It may have come up to the hilltop where it was collected from nearby rainforest. This species is distinctive externally and has bright yellow antennae, legs and surstyli. The surstyli are proportionally larger than those in species A and B and are most similar in shape to A1. Form D is morphologically identical to Form A1 and despite being genetically distinct is almost certainly not a new species. We suspect that this is a sequencing error. The macerated genitalia are all that remain of this specimen. Tissue remaining after the initial extraction was submitted to another lab for independent analysis (to test these unexpected results) and was lost. Thus unless more material matching the *cox1* signature for Form D turn up, we will never be able to re-examine this. Specimen 9735 (species F) was an unknown female that we were hoping to place. It turns out that it is likely another new species. There are no males from this site (a hilltop surrounded by heath habitat – Mount Tempest, Morteon Island, Queensland). We have two other females, possibly of this cryptic species, from heath habitat on Bribie Island. More material, including males, is needed to find out more about this possible species. Species E was not sequenced and is not included in Figure 1. Despite this, based on the minor morphological characteristics used to separate the species above, species E is clearly different. It is another rainforest species, intermediate between species B and C and one of the main reasons that Skevington (2001) could not separate *Cli. ancylus* into its true constituents. Externally it has yellow legs with black medial bands (intermediate between C and the other species). The antennae and surstyli are yellowish brown (also intermediate). The surstyli are large as in species C but slightly different in shape from all other species (very subtle and difficult to quantify). The horn on sternite 6 is much thicker than in other species.

So, in summary, we believe that combined support from both *cox1* and morphology indicate the presence of five species in this complex, *Cli. ancylus* s.s. (including forms A1 and A2), species B, species C, species E, and species F. Of these, species B, C and E are in a state where they can be described. The senior author will do this in a separate publication. Given the very subtle differences between the morphology of these species it is easy to see why this would have been an intractable problem using morphology alone. The variation in leg colour appeared to be continuous from C (yellow) to E (yellow and brown) to B and A (brown). The surstyli and the horn on sternite 6 vary subtly between all species but appeared to do so in a continuous fashion until the molecular data were applied. After examining the molecular data, it is clear that morphological variation within species is almost non-existent (except in species A where the surstyli vary from concave to flattened on the dorsal surface). No other characters were found to vary. Skevington (2001) actually considered that his

species concepts were narrow to the point of being somewhat daring and went to great lengths to justify the complex of species including *Cli. ancylus*, *Cli. digitatus*, *Cli. koebelei* and *Cli. lambkinae* Skevington, 2001. These species are broadly sympatric, are sibling taxa and essentially differ only in the shape of the surstyli, the shape of the hook on the right surstylus and the shape of the horn on sternite 6. Colleagues questioned the authenticity of these taxa being species and it was suggested that they might be only one species with polymorphic genitalia. Skevington (2001) would never have considered breaking this already morphologically questionable group of species into an additional five or more species (four more from *Cli. ancylus* s.s. and one more from *Cli. digitatus* s.s.). He suggested that molecular data should be applied to the group to test the species concepts. Together with morphology we now have a good idea of the diversity represented within this group.

*Cli. digitatus* complex. This was not considered by Skevington (2001) to be a complex of species. Despite this, we are now confident that this group includes two to three species (likely two), here treated as A, B and C. Species A and B are clearly different based on *cox1* data (11.7–12.6% different). They can also be distinguished by morphology based on the shape of the hook on the right surstylus (this hook has two projections and has a wider, stubbier proximal projection on species A) and the size of the hairy projection on the right surstylus (shorter in species A). Form C could be treated as variation in A or as a separate entity. The molecular data are inconclusive here (2.4% different) and the morphological data are no better. The only difference detected between Species A and Form C is the very subtle difference in the hook on the right surstylus (slightly stubbier in Form C). Examination of more material would likely show that this slight morphological difference can be explained by variation.

Surprisingly, Species A and B do not cluster together in either the phenetic or parsimony analyses (Figs 1–3). These are clearly sister taxa based on morphology and we have little doubt that additional molecular data would support this hypothesis (it is supported by the total evidence analysis in Fig. 5). We believe that species B is *Cli. digitatus* (s.s.) but the type must be consulted to confirm this. The first author will describe species A in a separate publication.

*Cli. koebelei*. This may or may not represent a complex of species. Molecular data supports the existence of one to two species (likely one). Uncorrected pairwise sequence divergences for *cox1* between the *Cli. koebelei* specimens was 2.3%. No variation in male genitalia or external morphology was detected.

*Cli. helluo* complex. Skevington (2001) noted variation in the surstyli and in sternite 6 of *Cli. helluo* leading him to conclude that more than one species may be involved here. Molecular data clearly support the existence of two species, A and B (9.5–10.6% different). The two species involved in this study show no external differences. There is minor variation in size of ridge on the right surstylus and size of protuberances on sternite 6, but these differences are not quantifiable and intermediates do occur (i.e. at this point we can find no discrete differences). More molecular sampling will be required to help determine if more species are involved and if morphological variation can be broken into discrete states. Re-examination of females may provide additional data. We have no female of species A available at this time. The name-bearing specimen may belong to Species B but the type will have to be consulted and additional *cox1* sequences will have to be collected before more species are described from this complex.

*Cli. trochanteratus* complex. This is a particularly problematic species morphologically. It has the largest range of any species of *Clistoabdominalis* (Algeria, Australia, Egypt, Israel, Japan, Laos, Sudan, Thailand and Turkey) and the surstyli show subtle, but considerable variation throughout the range. This was treated as clinal variation by Kehlmaier (2005) and Skevington (2001) but both recognized that this could be a complex of species. Unpublished work by Skevington on Fijian Pipunculidae has discovered a further array of species involved in this complex. Unlike in other regions, these Fijian taxa have external differences that support their species status. Molecular data will be applied to the Fijian taxa in an attempt to avoid under representation of species diversity as has occurred in previous revisions.

Based on the limited number of *Cli. trochanteratus* populations sampled here, support exists for two to

three species. Given the large range and variation within *Cli. trochanteratus*, it is likely that many more species exist. Of the specimens examined for this study, variation is restricted to the surstyli. No other variation was detected (no other external or internal variation in males, no ovipositor variation in females). Surstyli variation is subtle, but in concert with the molecular data it suggests that there are actually three or more species present in this complex in Australia alone. Note that this taxon was described from specimens collected in Egypt. Therefore it is very likely that all forms of *Cli. trochanteratus* present in this study or in Australia might turn out to be new to science. The molecular data alone support the existence of three species, A, B and C. Species A appears to be the most widespread and sequence divergences of 1.2 to 1.4% exists within the four specimens sampled. Species B is 7.9 to 9.3% different from the other species and is almost certainly an undescribed species. Form C represents an unknown female from Cooloola National Park, Queensland. It is surprising that this did not cluster with species A (two specimens of species A are from the same site). This may represent a sequencing error and should be repeated. We attempted to re-extract and re-sequence this specimen but were unsuccessful.

More work is clearly required before the species limits of taxa within this difficult complex will be resolved. Obtaining fresh material (or even relatively recently collected material) from specimens throughout the large range of this species will be a major obstacle to progress on this lineage. Work may have to be carried out regionally.

*Cli. tharra*. As with *Cli. koebelei*, this may or may not represent a complex of species. Molecular data supports the existence of one to two species (likely one). Uncorrected pairwise sequence divergences for *cox1* between the *Cli. tharra* specimens was 2.0%. No variation in male genitalia or external morphology was detected.

*Cli. capillifascis*. As with *Cli. koebelei* and *Cli. tharra*, this may or may not represent a complex of species. Molecular data supports the existence of one to two species (likely one). Sequence divergences between the *Cli. capillifascis* specimens was 3.7%. This is quite high for intraspecific variation, but no variation in male genitalia or external morphology was detected. More population sampling using *cox1* should be attempted to explore this high level of variation.

The *cox1* results for *Clistoabdominalis* are encouraging and suggest that Skevington (2001) was overly conservative when applying morphological species concepts to the group. When used in tandem as we have done here, *cox1* data and morphological data may resolve many of the species issues in a revision, leading to strongly supported concepts. With our limited sampling focusing on previously recognized problems, we have additional support for at least eight putative undescribed species of *Clistoabdominalis*. Denser molecular sampling will almost certainly uncover several more species in this group, leading us to believe that over 40 species exist in Australia. The clear take-home message here is that neither DNA Barcoding nor morphology alone can resolve all of the problems in revising this genus, but together they resolve most of the problems and leave only a few open questions for further research to resolve. There are two considerable problems left for those of us who would like to apply the extra molecular dataset to all of our revisions. First, do we have enough high quality specimens that will yield the needed DNA? Although this problem would have been minor in the revision of Australian *Clistoabdominalis* (because Skevington collected most of the species and over half of the specimens that went into the revision since 1997), the problem may be insurmountable in some groups where fresh material is simply unavailable. The work on *Nephrocerus* discussed below serves to illustrate this problem. Second, can we afford the extra cost of applying these data. We estimate that it would cost approximately \$6,000.00 for a dataset complete enough to thoroughly treat *cox1* for Australian *Clistoabdominalis* (40 species X an average of 10 specimens per species at a cost of about \$15 per specimen. Note that this does not allow for failed sequences that cost just as much. The failure rate will presumably be lower when conducted by a specialized lab with massive throughput and experience with this gene – see Hajibabaei et al. (2005) for example). Even if these costs can be cut in half through advances in PCR and sequencing, this is still a substantial increase in cost to a relatively inexpensive discipline. Despite this, the data are worth the



cost and systematists should lobby for the funding needed to better represent their science. Otherwise, our revisions will require revision before the ink is dry. DNA Barcoding is **not** a panacea, but **in tandem** with morphology it will develop an amount of synergy that helps to move systematics that much closer to understanding the true diversity of the world and as a result will allow us to ask much more detailed questions about the evolution, biogeography and ecology of the taxa that we study. There will always be taxonomic questions remaining that require further work, but future research can focus on these specific questions rather than forcing us to embark on repeated generic revisions.

### *Clistoabdominalis* phylogeny

So how does *cox1* perform in reconstructing the phylogenetic relationships of Australian *Clistoabdominalis*? In a nutshell, considerably better than we would have expected for part of a single gene. Figure 1 illustrates the *cox1* gene tree constructed using Neighbour Joining. The main species groups recognized in morphological analyses are labelled and unexpected results are highlighted by asterisks (\*). The main structure of the cladogram is supported by the previous morphological analysis by Skevington (2001). Skevington (2001) recognized four main lineages in Australian *Clistoabdominalis* (Fig. 1). Note that although Figure 4 uses only the reduced taxon dataset represented in this paper, the topology of the resulting cladogram is congruent with the published cladogram for all Australian *Clistoabdominalis* species (Skevington, 2001). There are two obvious problems with the molecular phylogeny illustrated in Figs 1–3. The *Clistoabdominalis ancylus* and *Cli. digitatus* complexes are not monophyletic. This is not a phenetic (i.e. neighbour joining) problem as successive weighting from the parsimony analysis produces similar results (Fig. 3). It is clearly a result of weak signal. The unweighted parsimony analysis cannot resolve the relationships between the major lineages and the consensus (Fig. 2) reflects this lack of data. Combining the molecular data with morphological data in a total evidence analysis overcomes this shortcoming and results in monophyletic lineages as would be expected.

Another incongruence with Skevington's (2001) analysis is the position of *Cli. lingulatus* outside of the *helluo* group (in all analyses, Figs 1–5). It is close to the same position as hypothesized by Skevington (2001), but instead of being the sister to *helluo* group taxa it is the sister to the *helluo* and *koebelei* groups. The position of the *nutatus* group as the sister to all other Australian *Clistoabdominalis* is also only one step longer than in Skevington's original hypothesis. All of the other inconsistencies of the *cox1* data in relation to Skevington's (2001) morphological analysis are in a part of the cladogram that he had little personal confidence in (the *scalenus* group). The combined analysis (Fig. 5) indicates that relationships within the *scalenus* group possess most of the conflicting data in our hypothesis. Addition of more *scalenus* group exemplars and addition of the putatively basal *Cli. monas* (Perkins, 1905) and *Cli. collessi* (Skevington, 2001) would help with resolution in this part of the tree. It is unclear whether the molecular or morphological signal (or neither) is more likely to represent the true phylogeny of the *scalenus* group.

In summary, with this new dataset essentially supporting Skevington's (2001) overall phylogeny of Australian *Clistoabdominalis*, and with a more robust total evidence analysis, we are now more confident of this hypothesis, particularly relationships within the *helluo* group.

### *Nephrocerus*

Based on uncorrected pairwise sequence divergences for *cox1*, *Nephrocerus* have accrued fewer molecular differences than species of *Clistoabdominalis*. Despite this, many of the species could still be diagnosed by *cox1* (~4 to 8% difference between several taxa). Speciation within the two Nearctic clades of *Nephrocerus* must have been recent as molecular variation within these groups is very low. Within the *slossonae* group, *N. corpulentus* and *N. slossonae* are indistinguishable with *cox1* (0.2 to 0.8 percent different; no fixed molecular character differences). Similarly, within the *daeckeii* group, *N. atrapilus*, *N. daeckeii* and *N. woodi* are also indistinguishable with *cox1* (0.3 to 1.2 percent different; no fixed molecular character differences). *Nephro-*

*cerus acanthostylus* is only weakly differentiated from other members of the *daeckei* group. As this sequence was based on a short, unidirectional sequence, it may not be reliable. Attempts to sequence the other strand failed as did attempts to sequence other *N. acanthostylus* specimens. Until a good sequence of this species is available to repeat our findings, placement of this taxon should be considered tentative.

Molecular data support the contention that the *slossonae* group includes the common European taxa. This hypothesis is not born out by morphology. The Palaearctic taxa would be placed in the *daeckei* group based on the black bristles on the abdomen and coxae. This hypothesis should be tested using multiple genes and additional Palaearctic taxa. The repercussions of this hypothesis are interesting. It suggests that the two North American lineages of *Nephrocerus* arrived through separate colonization events, and have independently undergone recent bursts of speciation leading to morphologically distinct species, but little *cox1* differentiation. We suspected a rapid radiation within Nearctic *Nephrocerus*, but two independent, parallel bursts of speciation would be somewhat surprising. It is conceivable that some type of environmental pressure on *cox1* is preventing modification of this gene within the Nearctic Region. We believe that this is unlikely given the two independent origins of the Nearctic taxa and the fact that Palaearctic and Nearctic environmental pressures are likely to be similar. It is more parsimonious to accept that there have been two independent bursts of speciation. Again, data from additional genes would serve to test this hypothesis.

Another obvious hypothesis generated by examining the molecular data is that Skevington's (2005) species concepts for *Nephrocerus* are too narrow and the traditional taxa recognized by Johnson (*N. slossonae* and *N. daeckei*), plus perhaps *N. acanthostylus*, are the only valid Nearctic species. This is possible but unlikely. Several discrete characters (both male and female) make it easy to separate *N. slossonae* from *N. corpulentus* (circlet of bristles on apex of hind tibiae different, epandrium and surstyli shapes different, ovipositor shape distinctive). *Nephrocerus acanthostylus* is very easy to diagnose and is undoubtedly a valid species (based on unique scutellum, proepisternum and tergite 1 colouration, a specialized bristle patch on hind tibia, a distinctive hook on the left surstylus of the male, the unique shape of both surstyli, and the distinctive stubby ovipositor in the female). Species in the *daeckei* group are much harder to diagnose. Females are indistinguishable. The most distinctive species is *N. atrapilus* Skevington (based on a stubby, pointed lobe on the right surstylus). *Nephrocerus daeckei* and *N. woodi* are very similar and can be differentiated only on the basis of subtle differences in the shape of the surstylus. Despite the similarity, Skevington (2005) considered that there were more species in this group that were closely allied to these two species (including a species in Mexico known only from females and a species from the SE USA that varied so slightly he was unwilling to describe it without additional material). A large series of *N. atrapilus* reared from two species of *Tipula* Linnaeus, 1758 and a small series of *N. daeckei* reared from another host provide ecological support for the existence of these species (Koenig and Young 2006; Skevington 2005). If the differences that Skevington (2005) described can be attributed to polymorphism or variation, these series of reared specimens (n=26) would undoubtedly have exhibited this. In fact, there is no variation in the reared material.

## Summary

The important points illustrated by these datasets is that DNA Barcoding is a powerful tool that can provide positive feedback to the revisionary process (when molecular distances are large and discrete). However, negative data are uninformative. As illustrated by *Nephrocerus*, using *cox1* to refute species concepts when the molecular distances are small would be a mistake. Similarly, setting a numerical limit as to what constitutes a species is foolhardy. Although this might work for the majority of species, it will conceal true diversity in groups like *Nephrocerus* and will invent diversity that is not there as appears to be the case in *Cli. capillifascis* and *Cli. tharra*.

We thus support the use of this tool when possible in revisions but caution the cavalier use of it as a primary data source. There is a great benefit to using one gene in as many studies as possible (for diagnostic purposes). If these data were available for most pipunculid species and incorporated in up-to-date revisionary papers we have little doubt that ecological and applied studies would be routinely investigating this biologically fascinating group of flies. In the meantime, the taxonomic impediment in the group is still considerable with even recent revisions such as that conducted by Skevington (2001) far from achieving an accurate representation of the true faunal diversity. As such, we encourage extra funding for systematic research to pay for this added data and to allow systematists to describe better the true diversity of our planet. However, we want to emphasize that sponsoring molecular methods should in no means result in a decrease of financial resources available for traditional morphological work, as it has been, is and will be the primary source of taxonomic information. Note that when molecular data are applied in combination with morphology, these extra data will not accelerate the description of new species as suggested by Hebert et al. (2003a), they will actually slow the process somewhat – although this may be minimized if the work is shared between a primarily morphological and a predominantly molecular research group. Such a slow-down should not be perceived to be too much of a problem since the hypotheses proposed are certain to be stronger; it simply illustrates the need for more systematists as we continue to struggle to support this fundamental basis for all biology.

## Acknowledgements

Thanks to the curators listed in the Materials and Methods section for providing specimens for revision and then allowing us to destructively sample parts of some of these specimens. Scott Kinnee and Tonyan Tian (California Department of Food and Agriculture—CDFA) provided key guidance and advice with respect to molecular techniques. René Luna (CDFA) and Lisa Bartels (CNC) conducted most of the lab work. Funding was provided by both CDFA and Agriculture and Agri-Food Canada (AAFC). Australian specimens were collected under a Queensland National Parks collecting permit (permit number F1/000076/02/SAA).

## References

- Aczél, M.L. (1948) Grundlagen einer monographie der Dorilaiden (Diptera). Dorilaiden-Studien VI. *Acta Zoologica Lilloana*, 6, 5–168.
- Ball, S.L., Hebert, P.D.N., Burian, S.K. & Webb, J.M. (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society*, 24, 508–524.
- Barrett, R.D.H. & Hebert, P.D.N. (2005) Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*, 83, 481–491.
- Bremer, K. (1994) Branch support and tree stability. *Cladistics*, 10, 295–304.
- Caterino, M.S., Cho, S. & Sperling, F.A.H. (2000) The current state of insect molecular systematics: a thriving Tower of Babel. *Annual Review of Entomology* 45, 1–54.
- Dunn, C.P. (2003) Keeping taxonomy based in morphology. *Trends in Ecology and Evolution*, 18, 270–271.
- Evenhuis, N.L. & Samuelson, G.A. (2004) *The insect and spider collections of the World website*. Bishop Museum, Entomology. Available from: <http://hbs.bishopmuseum.org/codens/> (Date of access 6 May 2006).
- Farris, J.S., Källersjö, M., Kluge, A.G. & Bult, C. (1994) Testing significance of incongruence. *Cladistics*, 10, 315–319.
- Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783–791.
- Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.
- Gest, H. (1999) Bacterial classification and taxonomy: A "primer" for the new millenium. *Microbiology Today*, 26, 70–72.
- Hajibabaei, M., deWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M. & Hebert, P.D.N. (2005) Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1959–1967.

- Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. & Hebert, P.D.N. (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, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London (B)*, 270, 313–322.
- Hebert, P.D.N. & Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. *Systematic Biology*, 54, 852–859.
- Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H. & Hallwachs, W. (2004a) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14812–14817.
- Hebert, P.D.N., Ratnasingham, S. & deWaard, J.R. (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London, Series B*, 270, S596–599.
- Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S. & Francis, C.M. (2004b) Identification of birds through DNA barcodes. *PLoS Biology*, 2, 1657–1663.
- Hogg, I.D. & Hebert, P.D.N. (2004) Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *The Canadian Journal of Zoology*, 82, 749–754.
- Janzen, D.H., Hajibabaei, M., Burns, J.M., Hallwachs, W., Remigio, E. & Hebert, P.D.N. (2005) Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1835–1845.
- Jervis, M.A. (1992) A taxonomic revision of the pipunculid fly genus *Chalarus* Walker, with particular reference to the European fauna. *Zoological Journal of the Linnean Society*, 105, 243–352.
- Kaila, L. & Ståhls, G. (2006) DNA barcodes: Evaluating the potential of COI to differentiate closely related species of *Elachista* (Lepidoptera: Gelechioidea: Elachistidae) from Australia. *Zootaxa*, 1170, 1–26.
- Kehlmaier, C. (2005) Taxonomic revision of European Eudorylini (Insecta, Diptera, Pipunculidae). *Verhandlungen des Naturwissenschaftlichen Vereins in Hamburg*, 41, 45–353.
- Koenig, D.P. & Young, C.W. (2007) First observation of parasitic relations between big-headed flies, *Nephrocerus* Zetterstedt (Diptera: Pipunculidae) and crane flies, *Tipula* Linnaeus (Diptera: Tipulidae: Tipulinae), with larval and puparial descriptions for the genus *Nephrocerus*. *Proceedings of the Entomological Society of Washington*, 109, 52–65.
- Lipscomb, D., Platnick, N. & Wheeler, Q. (2003) The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology and Evolution*, 18, 65–66.
- Nixon, K.C. & Carpenter, J.M. (1994) On outgroups. *Cladistics*, 9, 413–426.
- Prendini, L. (2005) Comment on "Identifying spiders through DNA barcodes". *Canadian Journal of Zoology*, 83, 498–504.
- Proudlove, G. & Wood, P.J. (2003) The blind leading the blind: cryptic subterranean species and DNA taxonomy. *Trends in Ecology and Evolution*, 18, 272–273.
- Rafael, J.A. & De Meyer, M. (1992) Generic classification of the family Pipunculidae (Diptera): a cladistic analysis. *Journal of Natural History*, 26, 637–658.
- Scheffer, S.J., Lewis, M.L. & Joshi, R.C. (2006) DNA Barcoding Applied to Invasive Leafminers (Diptera: Agromyzidae) in the Philippines. *Annals of the Entomological Society of America*, 99, 204–210.
- Seberg, O., Humphries, C.J., Knapp, S., Stevenson, D.W., Petersen, G., Scharff, N. & Andersen, N.M. (2003) Shortcuts in systematics? A commentary on DNA-based taxonomy. *Trends in Ecology and Evolution*, 18, 63–65.
- Skevington, J.H. (2001) Revision of Australian *Clistoabdominalis* (Diptera: Pipunculidae). *Invertebrate Taxonomy*, 15, 695–761.
- Skevington, J.H. (2005) Revision of Nearctic *Nephrocerus* Zetterstedt (Diptera: Pipunculidae). *Zootaxa*, 977, 1–36.
- Skevington, J.H. (2006) Revision of Fijian *Collinias* Aczél (Diptera: Pipunculidae). Bishop Museum Occasional Papers: *Fiji Arthropods*. Submitted.
- Skevington, J.H., Goolsby, J.A. & Setamou, M. (2007) New Records of Pipunculidae attacking Proconiine sharpshooters – the implications for parasitism of *Homalodisca coagulata* (Auchenorrhyncha: Cicadellidae: Proconiini). *South-western Entomologist*. Submitted.
- Skevington, J.H. & Yeates, D.K. (2000) Phylogeny of the Syrphoidea (Diptera) inferred from mtDNA sequences and morphology with particular reference to classification of the Pipunculidae (Diptera). *Molecular Phylogenetics and Evolution*, 16, 212–224.
- Skevington, J.H. & Yeates, D.K. (2001) Phylogenetic classification of Eudorylini (Diptera: Pipunculidae). *Systematic Entomology*, 26, 421–452.
- Smith, I.M., Fisher, B.L., & Hebert, P.D.N. (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: The ants of Madagascar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1825–1834.
- Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W. & Hebert, P.D.N. (2006) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceed-*

- ings of the National Academy of Sciences of the United States of America*, 103, 3657–3662.
- Sorenson, M.D. (1999) TreeRot, version 2c. Boston University.
- Sperling, F. (2003) DNA Barcoding: Deus ex Machina. *Newsletter of the Biological Survey of Canada (Terrestrial Arthropods)*, 22, 50–53.
- Swofford, D.L. (2001) PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods), version 4.0b10. Sinauer Associates, Inc.
- Tautz, D., Arctander, P., Minelli, A., Thomas, R.H. & Vogler, A.P. (2003) A plea for DNA taxonomy. *Trends in Ecology and Evolution*, 18, 70–74.
- Theron, J. & Cloete, T.E. (2000) Molecular Techniques for Determining Microbial Diversity and Community Structure in Natural Environments. *Critical Reviews in Microbiology*, 26, 37–57.
- Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. & Swings, J. (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological Reviews*, 60, 407–438.
- Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. & Hebert, P.D.N. (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1847–1857.
- Will, K.W. & D. Rubinoff. (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*, 20, 47–55.

**Appendix 1. Voucher data for specimens used in this study.** Data are presented in the following order: genus, specific epithet, author, unique specimen #, sex, country, state, location, latitude, longitude, date, collection and storage methods, notes, Genbank # (deposition).

*Chalarus* sp.: JSS13778, male, Canada, Quebec, Vaudreuil Co., Summit of Mount Rigaud, 45°27'59" N, 074°19'35" W, 13.vi.2001, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337675 (CNC); JSS16429, male, United States, Alaska, Fairbanks, Lake Smith, 64°51'44" N, 147°51'50" W, 13.vi.2005, killed in alcohol and stored at -80°C C, thorax and legs used for DNA extraction, DQ789673 (CNC); *Clistoabdominalis ancylus* A: JSS1353, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 29.xi.1997, dried specimen, killed with cyanide, legs taken from dried, pinned specimens for DNA extraction, DQ337639 (UQIC); JSS1669, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 27.xi.1997, dried specimen, killed with cyanide, legs taken from dried, pinned specimens for DNA extraction, DQ337633 (UQIC); JSS2504, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 22.i.1998, dried specimen, killed with cyanide, legs taken from dried, pinned specimens for DNA extraction, DQ337635 (UQIC); JSS2515, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 22.i.1998, dried specimen, killed with cyanide, legs taken from dried, pinned specimens for DNA extraction, DQ337653 (UQIC); JSS2901, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 20.i.1998, dried specimen, killed with cyanide, legs taken from dried, pinned specimens for DNA extraction, DQ337636 (UQIC); JSS2975, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 20.i.1998, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337632 (UQIC); JSS4361, male, Australia, Queensland, Mount Tinbeerwah East of Cooroy, 26°24' S, 152°59' E, 7.xi.1998, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337652 (UQIC); JSS4368, male, Australia, Queensland, Mount Tinbeerwah East of Cooroy, 26°24' S, 152°59' E, 7.xi.1998, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337634 (UQIC); JSS4508, male, Australia, New South Wales, Warrumbungle National Park, Split Rock, 31°17'08" S, 148°59'00" E, 18.i.1999, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337654 (UQIC); JSS4603, male, Australia, Tasmania, Franklin-Gordon Wild River National Park, Donaghy's Hill, 42°11'52" S, 145°55'55" E, 7.i.1999, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337644 (UQIC); JSS7508, male, Australia, Queensland, Barakula State Forest No. 302, Summit Round Mountain, 26°22'13" S, 150°58'09" E, 8.i.2000, dried specimen, killed with cyanide, Legs taken from dried, pinned specimens for DNA extraction, DQ337638 (UQIC); JSS13085, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337646 (CNC); JSS14003, female, Australia, New South Wales, Werrikimbe National Park, 31°10'23" S, 152°09'45" E, 1–7.xii.1997, dried specimen, killing agent unknown, DQ337637 (AMS); JSS15368, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337640 (CNC); JSS15369, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337651 (CNC); *Clistoabdominalis ancylus* B: JSS3411, male, Australia, Queensland, Mount Glorious Biological Centre, main road, [27°19'54" S, 152°45'29" E], 3–10.x.1997, dried, collected into alcohol and critical point dried, legs taken from dried, pinned specimens for DNA extraction, DQ337648 (UQIC); JSS3412, female, Australia, Queensland, Mount Glorious Biological Centre, main road, [27°19'54" S, 152°45'29" E], 3–10.x.1997, dried, collected into alcohol and critical point dried, Legs taken from dried, pinned specimens for DNA extraction, DQ337629 (UQIC); JSS10570, female,

Australia, New South Wales, Port Macquarie, Sea Acres NR, 25°47' S, 152°54' E, 10-24.xii.1999, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337642 (QM); JSS12609, male, Australia, Queensland, Ravensbourne National Park, 27°22' S, 152°11' E, 28.ix.2002, dried specimen, killed with cyanide, head used for DNA extraction, DQ337643 (CNC); JSS13109, male, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337647 (CNC); JSS13111, male, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337628 (CNC); JSS13738, female, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, dried, collected into alcohol and critical point dried, head used for DNA extraction, DQ337650 (CNC); JSS13753, female, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, dried, collected into alcohol and critical point dried, DQ337631 (CNC); ***Clistoabdominalis ancylus C***: JSS3292, male, Australia, Queensland, Scrub Road, Brisbane Forest Park, 27°25' S, 152°50' E, 23-30.i.1998, dried, collected into alcohol and critical point dried, DQ337649 (UQIC); JSS4432, male, Australia, Queensland, Mount Glorious; bordering Maiala National Park, [27°19'54" S, 152°45'29" E], 1.viii.-30.ix.1996, dried, collected into alcohol and critical point dried, legs taken from dried, pinned specimens for DNA extraction, DQ349221 (UQIC); JSS13066, male, Australia, Queensland, Brisbane, Summit of Mount Coot-tha, 27°29'16" S, 152°57'02" E, 29.ix.2002, dried specimen, killed with cyanide, head used for DNA extraction, DQ337645 (CNC); JSS14555, female, Australia, New South Wales, Camden Head, 31°39' S, 152°50' E, 24.xii.1999-4.i.2000, killed in alcohol and stored at -80°C C, DQ337641 (CNC); ***Clistoabdominalis ancylus D***: JSS13011, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, dried specimen, killed with cyanide, head used for DNA extraction, DQ337630 (CNC); ***Clistoabdominalis ancylus F***: JSS9735, female, Australia, Queensland, Moreton Island, Mount Tempest summit, 27°08'37" S, 153°24'09" E, 29.x.2000, dried specimen, killed with cyanide, DQ337687 (CNC); ***Clistoabdominalis beneficiens (Perkins)***: JSS13110, male, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337657 (CNC); JSS13756, female, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, dried, collected into alcohol and critical point dried, DQ337656 (CNC); JSS15356, male, Australia, Queensland, Moreton Island, Tangalooma, 27°11' S, 153°22' E, 25.viii.2004, killed in alcohol and stored at -80°C C, DQ337655 (CNC); ***Clistoabdominalis capillifascis Skevington***: JSS762, male, Australia, Queensland, 5 km E Normanby River on Battlecamp Road NW of Cooktown, 15°17' S, 144°52' E, 6.vi.1997, dried specimen, killed with cyanide, DQ337658 (CNC); JSS12698, male, Australia, Queensland, Carnarvon National Park, summit of Sugarloaf Mountain, 24°53'55" S, 147°56'43" E, 11.x.2002, dried specimen, killed with cyanide, 3 legs and part of tergite 1 used for DNA extraction, DQ789674 (CNC); ***Clistoabdominalis colophus Skevington***: JSS1289, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 29.xi.1997, dried specimen, killed with cyanide, DQ337659 (CNC); JSS10486, female, Australia, Australian Capital Territory, Canberra, Black Mountain, 35°16' S, 149°06' E, 19-31.i.1999, dried specimen, killing agent unknown, half of thorax used for DNA extraction, DQ337661 (CNC); JSS13097, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, killed in alcohol and stored at -80°C C, DQ337660 (CNC); ***Clistoabdominalis condylostylus Skevington***: JSS9675, male, Australia, Queensland, Moreton Island, Mount Tempest summit, 27°08'37" S, 153°24'09" E, 29.x.2000, (CNC), killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337663; JSS12825, male, Australia, Queensland, Carnarvon N[ational] P[ark], summit of 'Fly Hill' near West Branch Camp, 24°58'27" S, 147°59'34" E, 12.x.2002, dried specimen, killed with cyanide, DQ337662 (CNC); ***Clistoabdominalis dasymelus Skevington***: JSS12677, male, Australia, Queensland, Carnarvon National Park, summit of Sugarloaf Mountain, 24°53'55" S, 147°56'43" E, 11.x.2002, dried

specimen, killed with cyanide, 2 legs and 1/2 of tergite 1 used for DNA extraction, DQ337664 (CNC); *Clistoabdominalis digitatus* **A**: JSS15355, male, Australia, Queensland, Moreton Island, Tangalooma, 27°11' S, 153°22' E, 25.viii.2004, killed in alcohol and stored at -80°C C, DQ337667 (CNC); JSS15357, male, Australia, Queensland, Moreton Island, Tangalooma, 27°11' S, 153°22' E, 25.viii.2004, killed in alcohol and stored at -80°C C, DQ337666 (CNC); *Clistoabdominalis digitatus* **B**: JSS12970, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, dried specimen, killed with cyanide, head used for DNA extraction, DQ337669 (CNC); JSS13149, male, Australia, Queensland, Barakula State Forest No. 302, Summit of Round Mountain, 26°22'13" S, 150°58'09" E, 5.xii.2000, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337665 (CNC); *Clistoabdominalis digitatus* **C**: JSS13095, male, Australia, Queensland, Brisbane Forest Park, Scrub Creek, 27°25'41" S, 152°50'18" E, 28.ix.-15.x.2002, killed in alcohol and stored at -80°C C, DQ337668 (CNC); *Clistoabdominalis gremialis* **Skevington**: JSS13026, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, dried specimen, killed with cyanide, thorax and legs used for DNA extraction, DQ337674 (CNC); JSS13102, male, Australia, Queensland, Great Sandy National Park, Cooloola Section, 26°03' S, 153°01' E, 6-8.x.2002, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337670 (CNC); JSS15362, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337673 (CNC); , JSS15363, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337672 (CNC); JSS15364, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337671 (CNC); *Clistoabdominalis helluo* **A**: JSS4131, male, Australia, Northern Territory, Kakadu National Park, Mirrai Lookout, 12°52'01" S, 132°42'12" E, 27.ix.1998, dried specimen, killed with cyanide, DQ337679 (CNC); JSS12810, male, Australia, Queensland, Carnarvon N[atational] P[ark], summit of 'Fly Hill' near West Branch Camp, 24°58'27" S, 147°59'34" E, 12.x.2002, dried specimen, killed with cyanide, DQ337678 (CNC); *Clistoabdominalis helluo* **B**: JSS13145, female, Australia, Queensland, Bribe Island, QDPI Fisheries Site, 27°03' S, 153°11' E, 13-20.iii.1998, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337676 (CNC); JSS15361, female, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337677 (CNC); *Clistoabdominalis koebelei* (**Perkins**): JSS2921, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Mount Moffatt Summit, 25°03'35" S, 148°02'38" E, 20.i.1998, dried specimen, killed with cyanide, DQ337682 (CNC); JSS13124, male, Australia, Queensland, Barakula State Forest No. 302, Round Mountain, 26°22'13" S, 150°58'09" E, 5.xii.2000, killed in alcohol and stored at -80°C C, DQ337683 (CNC); JSS13150, male, Australia, Queensland, Barakula State Forest No. 302, Summit of Round Mountain, 26°22'13" S, 150°58'09" E, 5.xii.2000, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337680 (CNC); JSS15360, male, Australia, Queensland, Brisbane, Mount Coot-tha, 27°29'16" S, 152°57'02" E, 21.viii.2004, killed in alcohol and stored at -80°C C, DQ337681 (CNC); *Clistoabdominalis lingulatus* **Skevington**: JSS7447, male, Australia, Western Australia, M[oun]t Magnet, M[oun]t Waramboo, 28°01'55" S, 117°49'23" E, 14.xii.1999, dried specimen, killed with cyanide, DQ337684 (CNC); *Clistoabdominalis mathiesoni* **Skevington**: JSS1081, male, Australia, Queensland, Moreton Island, Mount Tempest, 27°09' S, 153°24' E, 20.ix.1997, dried specimen, killed with cyanide, DQ337685 (CNC); *Clistoabdominalis scintillatus* **Skevington**: JSS12995, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, dried specimen, killed with cyanide, 5 legs used for DNA extraction, DQ789675 (CNC); *Clistoabdominalis tasmanicus* **Skevington**: JSS4786, female, Australia, Tasmania, Cradle Mountain National Park, 41°37'38" S, 145°56'44" E, 22.xii.1998-8.i.1999, dried, collected into alcohol and critical point dried, DQ337688 (CNC); *Clistoabdominalis tharra* **Skevington**: JSS1079, male, Australia, Queensland, Moreton Island, Mount Tempest, 27°09' S, 153°24' E, 20.ix.1997, dried speci-



men, killed with cyanide, DQ337690 (CNC); JSS12993, male, Australia, Queensland, Carnarvon National Park, Mount Moffatt Section, Summit of Mount Moffatt, 25°03'35" S, 148°02'38" E, 13.x.2002, dried specimen, killed with cyanide, 3 legs used for DNA extraction, DQ337689 (CNC); *Clistoabdominalis trochanteratus* **A**: JSS13084, male, Australia, Queensland, Carnarvon National Park, summit of 'Fly Hill' near West Branch Camp, 24°58'27" S, 147°59'34" E, 12.x.2002, killed in alcohol and stored at -80°C C, half of thorax used for DNA extraction, DQ337691 (CNC); JSS13982, female, Australia, Queensland, 5 km S Mica Creek, 20°52' S, 139°28' E, 2.iii.2002, dried specimen, killing agent unknown, DQ337694 (QM); JSS14016, female, Australia, Queensland, [Great Sandy National Park,] Cooloola National Park, 26°03' S, 153°01' E, 6–8.x.2002, killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337692 (CNC); (Becker), JSS14019, male, Australia, Queensland, [Great Sandy National Park,] Cooloola National Park, 26°03' S, 153°01' E, 6–8.x.2002, killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337693 (CNC); *Clistoabdominalis trochanteratus* **B**: JSS3732, male, Australia, Queensland, Moreton Island, Mount Tempest, Summit, 27°09' S, 153°24' E, 22.iii.1998, dried specimen, killed with cyanide, DQ337695 (CNC); *Clistoabdominalis trochanteratus* **C**: JSS14015, female, Australia, Queensland, [Great Sandy National Park,] Cooloola National Park, 26°03' S, 153°01' E, 6–8.x.2002, killed in alcohol and stored at -80°C C, DQ337686 (CNC); *Clistoabdominalis yeatesi* **Skevington**: JSS9676, male, Australia, Queensland, Moreton Island, Mount Tempest summit, 27°08'37" S, 153°24'09" E, 29.x.2000, killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337696 (CNC); *Nephrocerus acanthostylus* **Skevington**: JSS13485, male, United States, Minnesota, Becker Co., 6 mi[les] SW Detroit Lakes, [46°49' N, 95°51' W], 5.vi.1981, dried specimen, killing agent unknown, 3 legs and 2 tergites used for DNA extraction, DQ789676 (CAS); *Nephrocerus atrapilus* **Skevington**: JSS11444, male, United States, Texas, Brazos Co., College Station, Lick Creek Park, [30°33' N, 96°14' W], 28–30.iii.2000, dried specimen, killing agent unknown, 3 legs used for DNA extraction, DQ337697 (DEBU); *Nephrocerus corpulentus* **Skevington**: JSS12588, female, United States, North Carolina, Gr[eat] Sm[oky] M[oun]t[ains] N[at]ional P[ark], Rainbow Falls, [35°40' N, 83°28' W], 28.v.1999, dried specimen, killing agent unknown, 3 legs used for DNA extraction, DQ337700 (CNC); JSS13366, male, United States, Missouri, Boone Co., 5 mi[les] E Ashland, Ashland Wildlife Area, [38°45' N, 92°12' W], 20–28.v.1992, dried specimen, killing agent unknown, 3 legs, tergites 2 and 3 removed for DNA extraction, DQ337698 (UMRM); *Nephrocerus daeckei* **Johnson**: JSS11888, male, United States, Pennsylvania, Berks Co., 5 km S Pine Grove, [40°33' N, 76°27' W], 16.vi.1997, dried specimen, killing agent unknown, 3 legs used for DNA extraction, DQ337699 (CNC); JSS13379, male, United States, Connecticut, [Tolland Co.], Storrs, TLS window, [41°48' N, 72°15' W], 1.vii.1996, dried specimen, killing agent unknown, 3 legs used for DNA extraction, DQ789677 (UCMS); *Nephrocerus daeckei* **group**: JSS10467, female, United States, North Carolina, Swain Co., Great Smoky Mountains National Park, Big Cove Road, Site #4, 35°30.635' N, 83°17.874' W, 19–26.v.2001, killed in alcohol and stored at -80°C C, head used for DNA extraction, DQ337702 (CNC); *Nephrocerus flavicornis* **Zetterstedt**: MZH\_Y33, female, Finland, Savonia australis, Kangaslampi, [61°56' N, 26°52' E], 21.vi–14.vii.2003, Malaise trap with 90% alcohol, stored at -20°C, 3 legs used for DNA extraction, DQ789678 (MZH); *Nephrocerus lapponicus* **Zetterstedt**: MZH\_Y65, female, Finland, Tavastia australis, Lammi, Evo, [61°13' N, 25°09' E], 9–23.vi.2003, Malaise trap with <70% alcohol, whole specimen used for DNA extraction, DQ789679 (MZH); *Nephrocerus scutellatus* **Macquart**: CK1, male, Germany, Saxony, Dresden, Dresdner Heide, Prießnitztal, 160m asl., 51°05' N, 13°46' E, 3.vi.2005, killed in alcohol, stored at room temperature, hind leg, wing, antennae and genitalia retained, DQ789680 (MZH); *Nephrocerus slossonae* **Johnson**: JSS11410, male, Canada, Ontario, Lambton County, Pinery Provincial Park, Burley Campground, [43°14' N, 81°53' W], 15–19.vi.1995, dried specimen, killing agent unknown, tergites 1 and 2 used for DNA extraction, DQ337701 (DEBU); *Nephrocerus woodi* **Skevington**: JSS10101, male, Canada, Quebec, Co. Vaudreuil, Summit of Mount Rigaud, 45°27'59" N, 074°19'35" W, 13.vi.2001, killed in alcohol and stored at -80°C C, DQ337705 (CNC); JSS10102, male, Canada, Quebec, Co. Vaudreuil, Summit of Mount Rigaud, 45°27'59" N, 074°19'35" W, 13.vi.2001, killed in alcohol and stored at -

80°C C, head removed and used for DNA extraction, DQ337704 (CNC); JSS10103, male, Canada, Quebec, Vaudreuil Co., Summit of Mount Rigaud, 45°27'59" N, 074°19'35" W, 13.vi.2001, killed in alcohol and stored at -80°C C, DQ337703 (CNC).