*Bulletin of Entomological Research* (2016) **106**, 34–46
doi:10.1017/S0007485315000486


# Morphological identification and COI barcodes of adult flies help determine species identities of chironomid larvae (Diptera, Chironomidae)

**A.J. Failla[1], A.A. Vasquez[1], P. Hudson[2], M. Fujimoto[1] and J.L. Ram[1]***

[1]Department of Physiology, Wayne State University School of Medicine, 5374 Scott Hall, 540 E. Canfield St., Detroit, MI 48201: [2]Great Lakes Science Center, U.S. Geological Survey, 1451 Green Road, Ann Arbor, MI 48105, USA

## Abstract

Establishing reliable methods for the identification of benthic chironomid communities is important due to their significant contribution to biomass, ecology and the aquatic food web. Immature larval specimens are more difficult to identify to species level by traditional morphological methods than their fully developed adult counterparts, and few keys are available to identify the larval species. In order to develop molecular criteria to identify species of chironomid larvae, larval and adult chironomids from Western Lake Erie were subjected to both molecular and morphological taxonomic analysis. Mitochondrial cytochrome c oxidase I (COI) barcode sequences of 33 adults that were identified to species level by morphological methods were grouped with COI sequences of 189 larvae in a neighbor-joining taxon-ID tree. Most of these larvae could be identified only to genus level by morphological taxonomy (only 22 of the 189 sequenced larvae could be identified to species level). The taxon-ID tree of larval sequences had 45 operational taxonomic units (OTUs, defined as clusters with >97% identity or individual sequences differing from nearest neighbors by >3%; supported by analysis of all larval pairwise differences), of which seven could be identified to species or 'species group' level by larval morphology. Reference sequences from the GenBank and BOLD databases assigned six larval OTUs with presumptive species level identifications and confirmed one previously assigned species level identification. Sequences from morphologically identified adults in the present study grouped with and further classified the identity of 13 larval OTUs. The use of morphological identification and subsequent DNA barcoding of adult chironomids proved to be beneficial in revealing possible species level identifications of larval specimens. Sequence data from this study also contribute to currently inadequate public databases relevant to the Great Lakes region, while the neighbor-joining analysis reported here describes the application and confirmation of a useful tool that can accelerate identification and bioassesment of chironomid communities.

**Keywords:** Chironomidae, DNA barcoding, species-level identification, morphology, Great Lakes

(Accepted 13 May 2015; First published online 15 June 2015)

*Author for correspondence
Phone: 1-313-577-1558
Fax: 1-267-345-9684
E-mail: jeffram@gmail.com

## Introduction

Chironomids represent a dominant group of benthic macro-invertebrate populations and have been observed as one of the principal groups of aquatic organisms both in terms of number and distribution in sampling studies (Carr & Hiltunen, 1965). Chironomids are ecologically important due to their contribution to the food web, nutrient cycling, and pollutant accumulation; however, the adaptive ability of chironomids and ease of both natural and anthropogenic-mediated transport warrant a concern for their potential role as invasive pests, especially in recently disturbed environments (Failla *et al*., 2015). The terrestrial adult stage is short-lived and often characterized by swarms of mating adults, sometimes presenting a great nuisance to humans in environments where emergences occur (Ali *et al*., 1985, Tabaru *et al*., 1987, Iwakuma, 1992, Hirabayashi & Okino, 1998). The ecological roles of chironomids have made sampling and subsequent species-level identification an important and useful biological tool for monitoring lake health (Langdon *et al*., 2006).

The identification of chironomid larvae to species level represents a challenging task for taxonomists. Identifying characteristics are more difficult to distinguish among the immature features of larval specimens, as compared with mature adults (Oliver, 1971). Adults possess more developed and specific features and thus are more amenable in establishing species level identifications (Ekrem *et al*., 2007). Larval keys usually identify chironomids only to genus level (Ram *et al*., 2014), and few chironomid keys exist that enable identification to species level. Identification of larvae to species level is usually accomplished by rearing larvae and collecting the pupae or adults (Inoue *et al*., 2008). The fully developed pupae or adults are then morphologically identified in order to assign species level identifications for the corresponding larvae.

DNA barcoding has been instrumental in facilitating identification of cryptic larval chironomid species (Sharley *et al*., 2004, Pfenninger *et al*., 2007, Carew *et al*., 2011). Studies combining the use of adult and larval DNA sequences have aided the species level identification of larvae within specific genera such as *Cricotopus* (Sinclair & Gresens, 2008), *Cladopelma* (Carew *et al*., 2005), *Procladius* (Carew *et al*., 2011) and *Corynoneura* (Silva & Wiedenbrug, 2014). However, these techniques have not been applied in describing the composition of diverse communities of chironomid larvae in the Great Lakes.

The present study specifically addresses the assessment of a benthic community from the standpoint of aquatic species monitoring and identification with its application in the Western Lake Erie region. Studies of larval and adult mitochondrial cytochrome c oxidase I (COI) genes have shown that DNA barcodes are an excellent tool provided that a comprehensive DNA barcode library at the species level is available. Such DNA libraries should contain a set of sequences that have been obtained from diverse larvae or adults that have been morphologically identified by expert taxonomists (Ekrem *et al*., 2007).

Although the public databases are quantitatively abundant, only a small number of sequences are useful in the Great Lakes region. Expansion of the reference databases in regions where chironomids are of interest would make species level identifications from molecular analysis more accessible and consistent. The present study uses newly determined sequences of taxonomically identifiable adult specimens to enrich the chironomid sequence database for Western Lake Erie and to improve identification of the diverse larval community in this region.

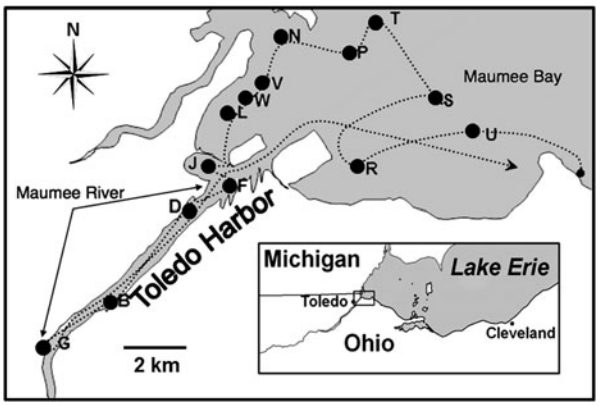


Fig. 1. Map of 14 benthic collection sites from 2012. Benthic communities of the Maumee Bay and Maumee River of the Western Lake Erie region (inset) were sampled at the lettered locations.

## Materials and methods

### *Collection of larvae*

Benthic organisms were collected from the Maumee Bay area of Western Lake Erie and the Toledo Harbor region of the Maumee River in the spring and summer of 2012, at 14 sites illustrated in fig. 1. On 2 and 30 May, 12 June, 12 July, 6 and 28 August, 2012 a total of 128 benthic samples were collected from 14 different sites . Benthic grab samples were obtained via a hand-operated bottom dredge (AMS, Ben Meadows, Janesville, WI), as previously described (Ram *et al*., 2014). Samples were collected from most sites on each date, washed on a 0.5 mm sieve to remove fine sediments while retaining organisms, and stored temporarily on ice in the field in 80% ethanol obtained by adding a fourfold volume of 100% ethanol (Fisher Science, Pittsburgh, PA). Upon delivery to the laboratory, samples were washed again on a 0.5 mm sieve and stored in 90% ethanol at 4 °C until shipment to EcoAnalysts, Inc (Moscow, ID) for sorting and morphological identification.

### *Morphological taxonomy of larvae*

EcoAnalysts performed morphological identification of all organisms in each sample, sorting different taxa into separate vials. These animals included 2410 chironomid larvae that were mostly classified only to genus level by morphological features. Only a small proportion could be identified to species level, as described in results. Thousands of non-chironomid specimens were also observed and will be reported elsewhere.

### *DNA sequencing of larvae*

Sorted, identified chironomid larvae were returned to the laboratory for molecular analysis. All specimens of rare taxa (those identified in each collection by Ecoanalysts fewer than five times) were sequenced. Among more common taxa, at least five specimens of each taxon randomly chosen from

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

among the available specimens were sequenced from each collection. Due to particular interest in *Cricotopus* spp., known agricultural pests, all *Cricotopus* larvae that were collected were sequenced. The result of this selection, designed to obtain multiple representative sequences, whenever possible, from every taxon identified by Ecoanalysts, was that a total of 189 larvae out of the 2410 larvae collected were chosen for sequencing.

The full body, anterior aspect, and posterior aspect of chironomid larvae chosen for sequencing were photographed, and then sterile dissecting methods were used to obtain a small piece of tissue from the mid-portion of each selected specimen. Each dissected piece was placed in 30 µl of 100% ethanol in individual wells of a 96-well micro-plate and sent to the Canadian Center for DNA Barcoding (CCDB; Biodiversity Institute of Ontario, University of Guelph, Ontario, Canada) for sequencing of up to 658 base pairs of the mitochondrial COI gene using forward and reverse primers HCO2198 and LCO1490 (Folmer *et al.*, 1994, Hebert *et al.*, 2003*a*). The anterior and posterior ends of each larva were retained as voucher specimens.

### *Collection of adults*

Adults were collected from the Maumee Bay region, including from the hull and sides of the boat while the benthic collections already described were in process. Briefly, the flies were trapped via a hand-held vacuum cleaner and subsequently emptied into a series of 50 ml collection tubes containing isopropyl alcohol (pilot studies indicated that DNA was as readily obtained from specimens preserved in isopropyl as in ethyl alcohol). Adult flies were then sorted, and each individual was placed in its own vial containing 85% ethanol. Thirty-nine undamaged adults of diverse macroscopic characteristics were chosen for morphological and molecular analysis.

### *DNA sequencing of adult chironomids*

For the first set of 20 specimens, two legs were detached from each adult chironomid and preserved in 90% ethanol for DNA analysis. DNA isolation was performed according to a Qiagen DNA spin-column_protocol (https://www.qiagen.com/us/resources/resourcedetail?id=21b4511b-4aaa-470a-a141-191ed91c54be&lang=en). Isolated DNA was amplified by polymerase chain reaction (PCR) using COI forward and reverse primers, HCO2198 and LCO1490 (Folmer *et al.*, 1994, Hebert *et al.*, 2003*a*). The PCR product was purified and diluted with sterile PCR water to a concentration appropriate for sequencing. Genewiz (South Plainfield, NJ) sequenced the purified COI product in the forward and reverse directions. A consensus sequence was determined using DNAbaser software (DNA Baser Sequence Assembler v4.x 2014, Heracle BioSoft SRL, www.DnaBaser.com), evaluating chromatograms and aligning sequences from both directions. For the remaining 19 adults, two legs were removed from each fly and placed in a microplate that was sent to CCDB for COI sequencing, based on the protocol described above for larvae. Altogether, 37 of the 39 specimens belonged to the Chironomidae. The remaining two specimens, a culicid (*Anopheles* species) and a chaoborid (*Chaborus punctipennis*) were excluded from the analysis.

### *Morphological identification of adults*

Morphological identification of the 39 specimens (minus the two legs used for molecular taxonomy) was completed at the Great Lakes Science Center (Ann Arbor, MI). The tissues were digested, and the remnants were mounted on slides to identify key morphological features. Initially, a specimen's size, color and shape were noted and then a pair of legs, wings and one antenna were mounted on a glass slide while the rest of the body was cleared of muscle tissue and then mounted. Keys by Townes (1945) and Cranston *et al.* (1989) were used to key adult specimens to genus. Species within a genus were identified using Townes (1945), Dendy & Sublette (1959), Roback (1971), (Epler, 1988), Heyn (1992), Saether (2009), and Saether (2011). When available, at least two specimens of each species were mounted and identified; additional specimens were studied for confirmation based on size, color and shape and returned to their labeled vials pending any possible need to mount and reconfirm based on DNA analysis.

### *Initial database search*

All of our larvae and adult sequences were initially screened with the Barcode of Life Data Systems (BOLD) species level identification engine (http://www.boldsystems.org/index.php/IDS_OpenIdEngine) to determine if there were matches that differed by <3% in the BOLD database. If no match was found, the sequence was then subjected to a BLAST search of the National Center for Biotechnology Information (NCBI) GenBank database to confirm that no match existed in either database. If the result from BOLD included published results that also existed in the GenBank database, no further search was done. If a sequence had a database match differing by <3%, representative sequences from the matches were used as reference sequences in subsequent analysis.

### *Neighbor-joining analysis of larvae, adults and reference sequences*

Neighbor-joining analysis was performed using MEGA software (Tamura *et al.*, 2011). Pairwise differences of all larval sequences were calculated and analyzed graphically to determine natural groupings of sequences to define operational taxonomic units (OTUs) of sequences. COI barcode relationships were determined by constructing neighbor-joining trees and calculating pairwise differences using a maximum composite likelihood algorithm (Tamura *et al.*, 2011).

In addition to the reference sequences chosen from the results of BOLD and GenBank searches, a database of useful species-level reference sequences was developed by downloading all sequences that registered as 'chironomidae' in the GenBank database and selecting quality sequences from among them. After aligning the potential reference sequences using Clustal W in MEGA, sequences that were too short or of poor quality (e.g., with multiple 'N's, <75% contiguous, etc.) were removed from the analysis.

COI barcode relationships of larval sequences were inferred from these quality reference sequences, comparing 619 nucleotide positions. Since the analysis of average pairwise differences supported defining OTUs as clusters having no more than 3% pairwise differences within the OTU (see below), reference sequences that differed from the larval OTUs by more than 3% were removed and redundant reference sequences were eliminated. Subsequent neighbor-joining analysis defined OTUs as clusters having >97% identity. Any genetic grouping that is described in this paper as being the same species or OTU adhered to this standard. Sequences obtained from chironomid adults were then added to the

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

Table 1. Species level morphological identification of adult and larval chironomids.

| Item | Adults | Larvae |
|---|---|---|
| Total number of specimens examined by taxonomists | 39 (39 sequenced) | 2410 (189 sequenced) |
| Total number of specimens (# of chironomids) identified to genus level | 39 (37 chironomids) | 2410 (2410 chironomids) |
| Total number of chironomid specimens (species) identified to species or species group level; *list of species* | 33 (15)<br>*Ablabesmyia annulata* (Say, 1823)<br>*Axarus festivus* (Say, 1823)<br>*Chironomus crassicaudatus* Malloch, 1915<br>*Chironomus decorus* (Johannsen, 1905)<br>*Cladopelma viridulum* (Linnaeus, 1767)<br>*Coelotanypus scapularis* (Loew, 1866)<br>*Cryptochironomus fulvus* (Johannsen, 1905)<br>*Cryptochironomus digitatus* (Malloch, 1915)<br>*Dicrotendipes lucifer* (complex) (Johannsen, 1907)<br>*Glyptotendipes senilis* (Johannsen, 1937)<br>*Glyptotendipes meridionalis* Dendy & Sublette, 1959<br>*Procladius bellus* (Loew, 1866)<br>*Procladius denticulatus* Sublette, 1964<br>*Stictochironomus devinctus* (Say, 1829)<br>*Tanypus stellatus* Coquillett, 1902 | 73 (6); 22 were sequenced<br>*Dicrotendipes simpsoni* Epler, 1987<br>*Dicrotendipes modestus* (Say, 1823)<br>*Polypedilum halterale* (Coquillett, 1901)<br>*Polypedilum scalaenum* (Schrank, 1803)<br>*Cricotopus bicinctus* (Meigen, 1818)<br>*Ablabesmyia annulata* (Say, 1823) |
| Percentage of chironomid specimens identified to species level | 89% (33/37) | 3% (11.6% of sequenced) |

analysis, creating a final tree with sequences from larval specimens, taxonomically identified adult specimens, and quality reference chironomid sequences from GenBank and BOLD .

Genus and species names, their authorities and years, and their family and subfamily identities were verified by reference to http://zipcodezoo.com/Key/ and through use of the NCBI taxonomy browser (http://www.ncbi.nlm.nih.gov/taxonomy/).

## Results and conclusions

The numbers of larval and adult specimens that were classified by morphological criteria are summarized in table 1, along with a subset of specimens for which COI barcodes were sequenced. Altogether, 2410 larvae were identified morphologically by EcoAnalysts, revealing 23 genera, among which only 6 genera had specimens that could be identified to species level. Of the 2410 specimens, 189 were selected for sequencing, including 22 (11.6%) specimens that had been identified by EcoAnalysts to species (table 1). Out of 39 adult specimens subjected to morphological and molecular analysis, 37 were adult males, along with two other insect specimens. Of the 37 adult chironomid specimens 33 (89%) were identified to species level, comprising the 15 different species listed in table 1. COI barcode sequences were obtained from all 39 adult specimens. All of these newly identified sequences have been submitted to NCBI for inclusion in the GenBank database as accession numbers KP954634–KP954653 (adults), KR085203–KR085223 (adults) and KR085224–KR085412 (larvae).

### *Pairwise differences of larval sequences*

A histogram of pairwise differences of 189 larval sequences illustrated in fig. 2A shows structure that helped define OTUs for this study. Graphic analysis reveals 3 major peaks: (1) pairwise differences of <3% that we used in subsequent analysis to define OTUs; (2) a second peak between 3 and 6% pairwise differences. The pairs within this peak include four different genera (*Cryptochironomus*, *Procladius*, *Microchironomus*, and *Dicrotendipes*) for which the members of each pair had the same genus but whose species taxa had been designated only as 'sp.' by EcoAnalysts; and (3) pairwise differences that were >11%, peaking at ∼20% as the most frequent pairwise difference. Pairs within this peak included both intragenus (e.g., *Chironomus*, 16%; *Coelotanypus*, 14%; *Cryptochironomus*, 14%; and *Polypedilum* 11%) and intergenus (e.g., *Glyptotendipes* and *Procladius*, 20%) differences.

### *Reference sequences*

Out of over 2000 species level 'chironomidae' COI sequences downloaded from GenBank and BOLD, deletion of short or poor quality sequences left 1447 as reference database sequences for COI barcode analysis. Of these 1447 sequences, 11 non-redundant sequences differed by <3% from one or more larval and adult sequences and were used as reference sequences in constructing neighbor joining taxon-ID trees. Altogether, 250 larval, adult and reference database sequences were compared at 619 nucleotide positions. The histogram of pairwise differences retained its three-peak structure with the inclusion of the adult and reference database sequences (fig. 2B).

### *Neighbor-joining analysis larvae, adults and reference sequences*

A neighbor-joining taxon-ID tree of all larval sequences contained 45 OTUs (fig. 3). Seven of these (15.6%) were supported by reference sequences from GenBank as shown in detail (fig. 4). OTUs in fig. 4 were identified as follows: *Chironomus entis/plumosus* (OTU 2), *C. quinnitukqut/decorus* (OTU 13), *Micropsectra insignilobus* (OTU 22), *Paratanytarsus natvigi* (OTU 23), *Paratanytarsus grimmii* (OTU 33), *Cricotopus sylvestris* (OTU 35) and *Cricotopus bicinctus* (OTU 36). OTUs 2, 23, 33, 35 and 36 were identified reliably based on close relationships to reference sequences from GenBank. However, OTU 13 grouped with

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

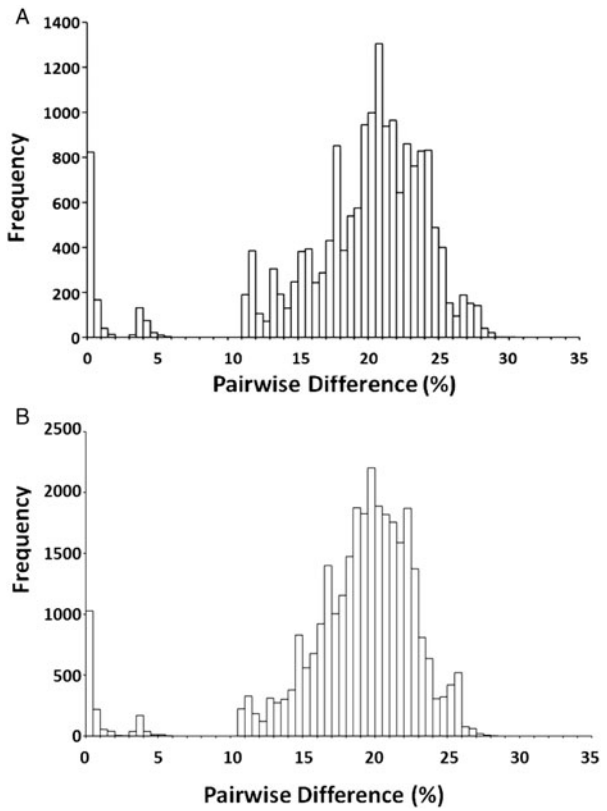


Fig. 2. Histograms of pairwise distance values of (A) larval sequences and (B) all sequences (larvae, reference database and adults) combined.

both an adult and a reference sequence, the identification of which differed from each other and will be discussed below. For OTU 22, the genus of the reference sequences from GenBank differed from the genus of the larva morphologically identified by EcoAnalysts, a discrepancy that will also be discussed below. In OTU 2, two reference sequences enabled identification of both larvae and adults that had previously only been identified to genus level.

Figure 5 shows 13 larval OTUs out of 45 (29%) that could be compared with sequences of identified adult chironomids, from which they differed by <3%. OTUs 4, 11, 15, 16, 37, 38, 39, 40 and 45 were identified reliably (*Cryptochironmus digitatus*, *C. fulvus*, *Cladopelma viridulum*, *Glyptotendipes meridionalis*, *Procladius bellus*, *Ablabesmyia annulata*, *Tanypus stellatus*, *Coelotanypus scapularis*, *Procladius denticulatus*, respectively). OTUs 1, 13, 21 and 25 (*Chironomus crassicaudatus*, *C. decorus*, *Dicrotendipes lucifer*, *Stictochironmus devinctus*, respectively) contain discrepancies when compared with adult and reference sequences. Of the 15 different chironomid species identified as adults (table 1), only *Axarus festivus* failed to have a corresponding sequence among the larvae. In addition, a *Glyptotendipes senilis* adult was taxonomically identified but not included in the phylogeny due to poor sequence quality and length. While 19 out of the 45 OTUs (42%) are identified, 26 out of the 45 still have no reference sequences or identified adults from which to assign the larvae with putative species identifications.

Based on COI sequence relationships, OTU 22 (fig. 4) shows a specimen classified by EcoAnalysts as a *Cladotanytarsus* sp. larva that is most likely to actually be *M. insignilobus*. In OTU 22, the claimed *Cladotanytarsus* larva falls within a large *Micropsectra* reference sequence clade, being most closely related to *M. insignilobus* as indicated by fig. 6. Furthermore, the sequence of this specimen differs from four other larvae also identified by EcoAnalysts as *Cladotanytarsus* sp. by a pairwise difference of ~20%. Contamination of this sample during sequencing by *Microspectra* DNA from another specimen is highly unlikely because no other *Micropsectra* larvae were collected or handled during the present study.

## Discussion

This study improves the reference databases of COI barcodes for chironomid larval identification by conducting a quality review of existing database sequences of chironomid COI barcodes and determining additional sequences from newly collected morphologically identifiable adult chironomids. This study also provides support, through its analysis of pairwise differences in COI barcodes, for using 97% identity as a natural amount of within-species variation defining chironomid OTUs. Sequences provide specific, reliably generated data for classifying specimens. Nevertheless, as will be discussed, the need for referencing the sequences to potentially ambiguous morphological identifications and to databases that may vary in sequence quality and taxonomic reliability means that ambiguities, inconsistencies, and errors may still occur, and care must be taken in using sequence data for identification.

### *Improvement of chironomid reference databases*

A search for high quality reference sequences in GenBank to identify 45 OTUs of chironomid larvae specimens from Maumee River and the Maumee Bay area of Western Lake Erie revealed that only 15.5% of sequenced larvae OTUs had corresponding reference sequences in GenBank. In order to achieve even that degree of identification this study reviewed over 2000 chironomid sequences in GenBank to assure selection of only high quality sequences of sufficient length and adequate annotation. This study adds 33 new reference sequences to the public database, based on careful taxonomic identification and COI barcoding of adult specimens. Addition of the sequences from taxonomically identified adult specimens to these databases aided in identification of OTUs 1, 4, 11, 15, 16, 21, 25, 37, 38, 39, 40 and 45, which would not have been possible prior to this study. Nevertheless, more work needs to be done, as 26 of the 45 OTUs in the tree have yet to be identified. In addition, we have yet to collect larvae with sequences corresponding to the adult specimens of *A. festivus*. Absence of these corresponding larvae in our dataset could indicate that *A. festivus* larvae occupy different or harder substrate habitats than those we sampled. In any case, addition of *A. festivus* to the GenBank database will make future identifications of *Axarus* larvae possible.

### *Relationship of taxa to pairwise differences*

Data in fig. 2 show gaps in the distribution at around 3% difference and between 6 and 10%. A threshold of 3% for species differentiation has been used for various animal groups

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

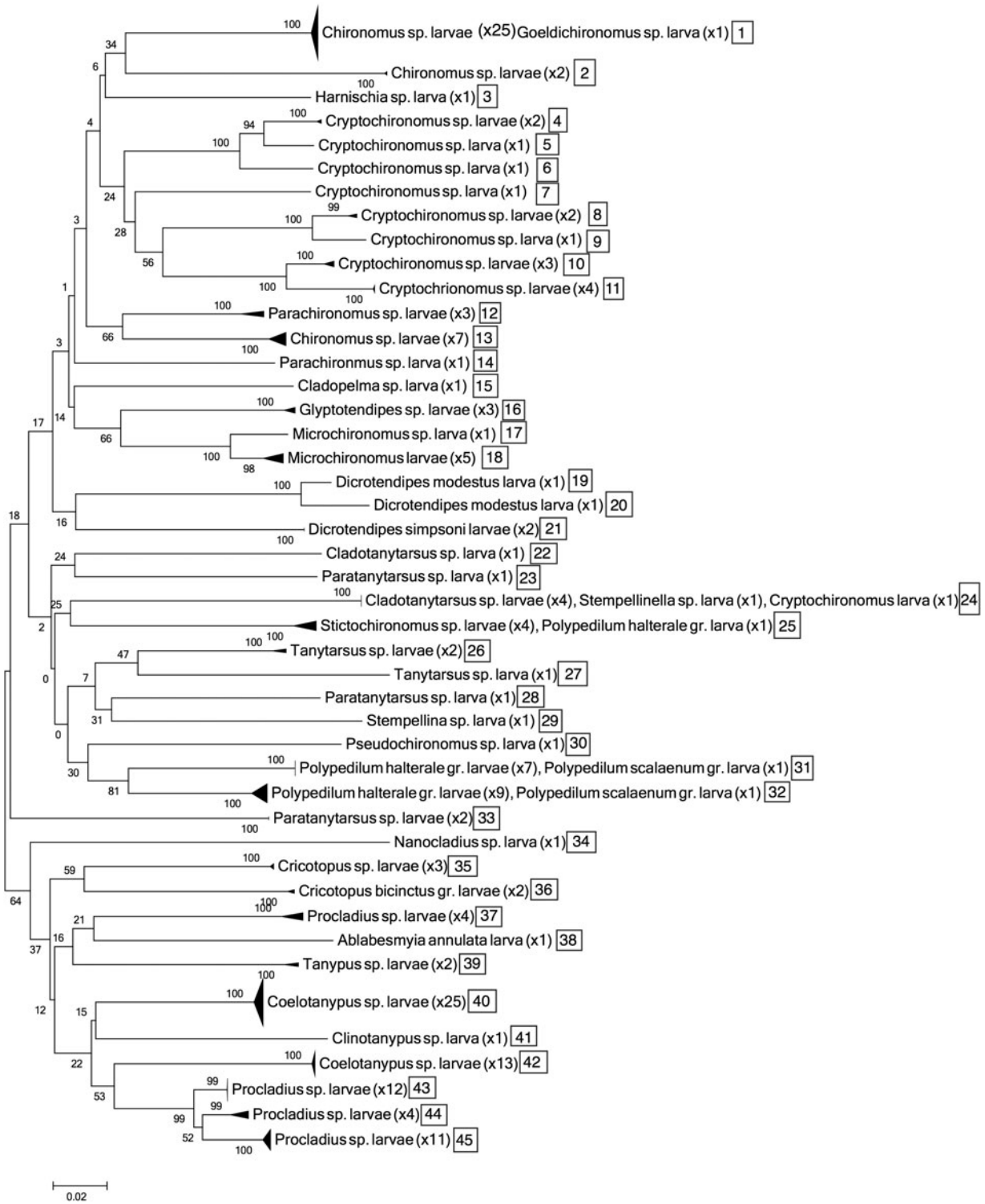


Fig. 3. Condensed neighbor – joining tree with maximum composite likelihood algorithm depicting 45 distinct larval operational taxonomic units (OTUs) based on mitochondrial cytochrome c oxidase I (COI) DNA sequences. Values at nodes correspond to a bootstrap 1000 test. Numbers in parentheses indicate the number of sequences within each branch. Numbers in blocks to the right of each branch correspond to the OTU numbers referred to throughout the rest of the paper. The analysis is based on 619 nucleotide positions in 189 larval sequences. The identifications are according to the highest taxonomic level achievable by EcoAnalysts. The scale represents fractional difference of sequence according to the length of the branch. For condensed branches, triangle height represents the number of sequences in the OTU, while width indicates the distance value corresponding to the lowest branch point.

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

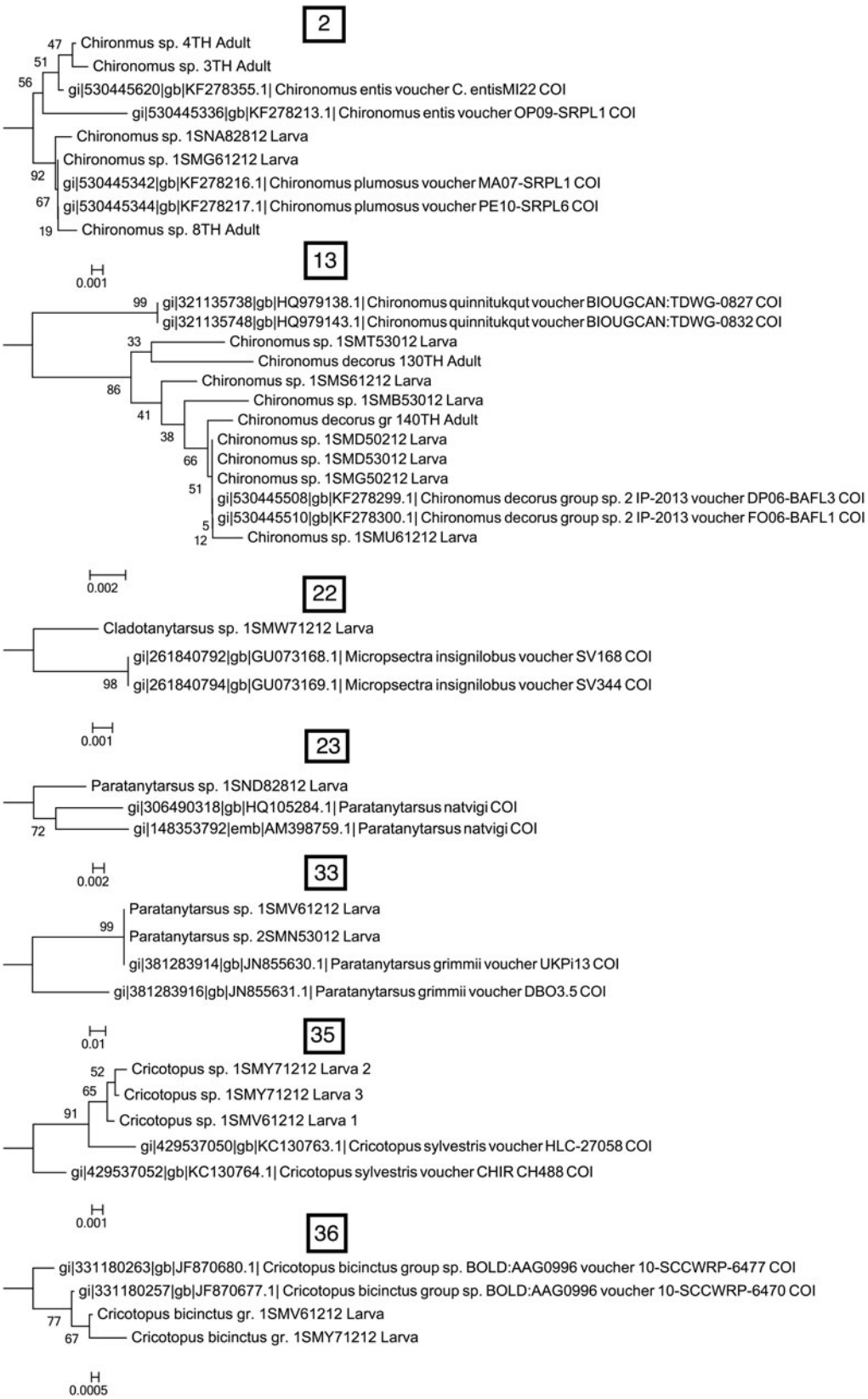


Fig. 4. GenBank reference sequences identified seven larval operational taxonomic units (OTUs) to species level. Each subtree has its own scale in the bottom left corner. Numbers in boxes indicate the clade number that corresponds to the larval tree (fig. 3). For OTU 33, only the UKpi13 reference sequence of *Paratanytarsus grimmii* was within a 3% distance. The DBO3.5 reference sequence is included to indicate the scale.

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

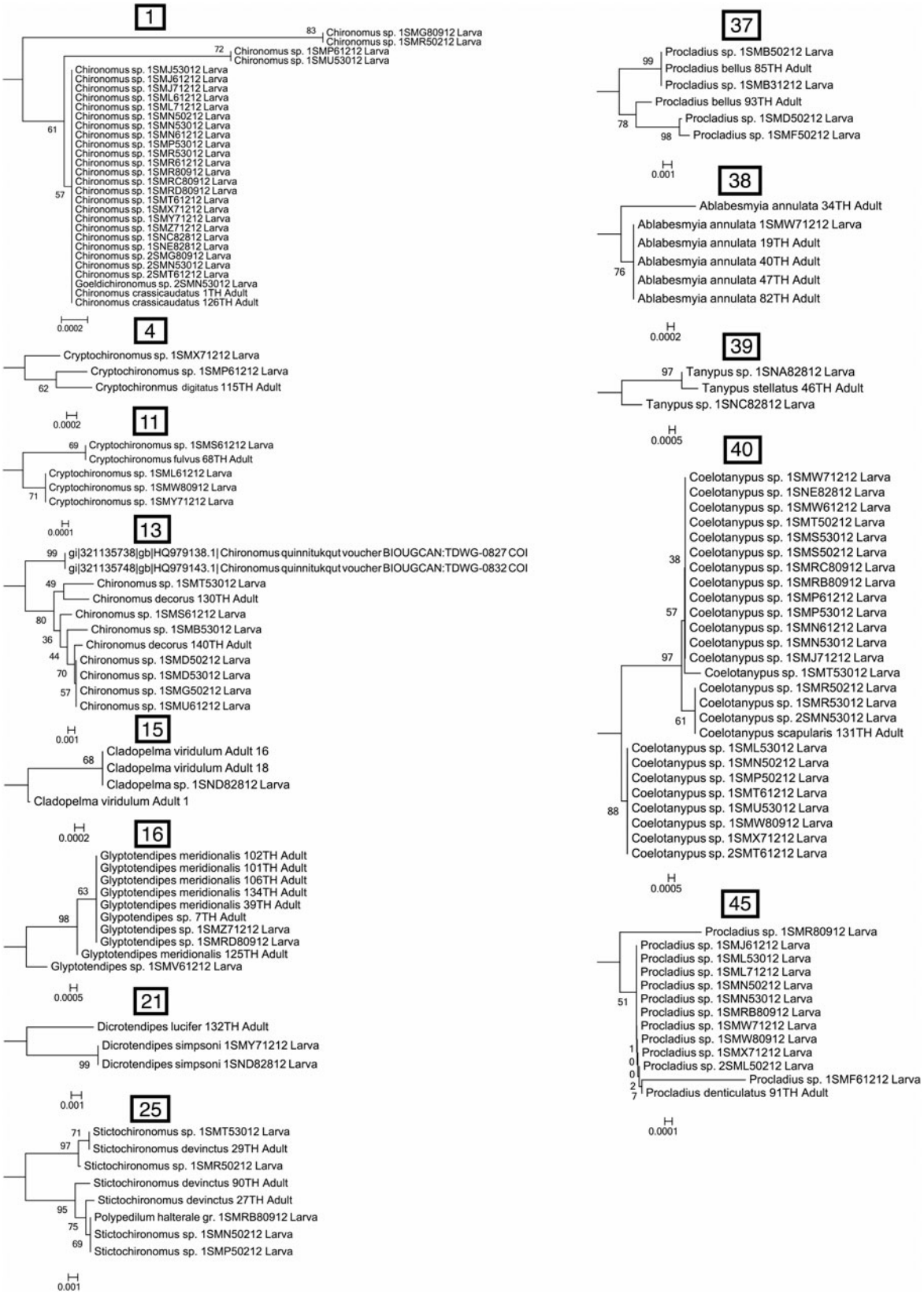


Fig. 5. Larval operational taxonomic units (OTUs) identified by adult sequences. Each subtree has its own scale in the bottom left corner. Numbers in boxes indicate the clade number that corresponds to the larval tree (fig. 3).

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

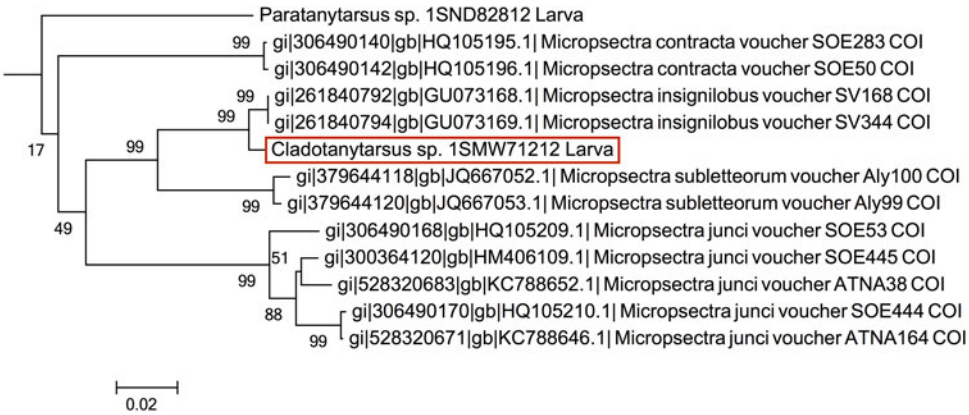


Fig. 6. A subtree extracted from the neighbor-joining taxon-ID tree constructed with GenBank reference and larval sequences before redundant sequences were removed. The relationship between the GenBank reference sequences and the sequenced larva in operational taxonomic unit (OTU) 22 that had been classified by EcoAnalysts as *Cladotanytarsus* sp. suggests that this specimen was most likely misclassified during morphological analysis, as its position within a large *Micropsectra* clade is evident.

(Hebert *et al.*, 2003*a*, Hebert *et al.*, 2003*b*, Sinclair & Gresens, 2008). In some cases, different species appear within the same COI cluster (e.g., OTU 24 and 25). Although these specimens remain within the confines of the 3% threshold designated by molecular morphology, the taxonomic designations do not agree. Evolutionary processes such as hybridization or rapid speciation where divergent mutations have not yet accumulated may explain this phenomenon (Michailova & Fischer, 1986, Ekrem *et al.*, 2007, Proviz, 2008, Abbott *et al.*, 2013). The second peak in pairwise differences, between 3 and 6%, suggests that some chironomid groups may have a greater within species variation. The question also arises: could the specimens in the region between the two larger histogram peaks represent instances of incipient speciation? In most cases, in our data pairwise differences >11% clearly represent different species and usually different genera. However, cases where apparently the same species has pairwise differences this large might also represent cryptic species (Anderson *et al.*, 2013), revealed by large intraspecific pairwise differences. *Polypedilum halterale*, represented by OTU 31 and 32, which differ by 11%, is one such example.

### *Ambiguities, inconsistencies and possible errors*

Despite the care with which larvae and adults were sequenced and the adult specimens were identified, the results shown here exhibit several ambiguities, inconsistencies and possible database or identification errors. Examples of ambiguities in identifying larvae are OTU 13 and 21 (figs. 4 and 5). In OTU 13, a reference sequence and a sequenced adult were classified as different species within the genus *Chironomus*. For some *Chironomus* species, the COI barcode and morphological identification alone may be inadequate for establishing species level identifications (Ekrem *et al.*, 2010, Proulx *et al.*, 2013) For example, *C. quinnitukqut* is a part of the *C. decorus* group, and species within this group are often separated based on karyotype analysis (Martin *et al.*, 2011). In OTU 21, *Dicrotendipes simpsoni* represents a species within the *Dicrotendipes lucifer* complex. Similarly, difficulty in assigning species level identifications to chironomids within the genus *Cricotopus* exists, specifically within the *C. sylvestris* species group (Gresens *et al.*, 2012). Assignment of species level identifications within these types of difficult genera will be ambiguous unless additional markers are utilized, techniques such as karyotyping are incorporated, or morphological keys are improved.

Ambiguity is also present in OTU 2 (fig. 4) and OTU 24 (fig. 3). For OTU 2, identification as *C. entis* is evident, yet examination of a karyotype could place it as a closely related species, such as *Chironomus borokensis* Kerkis, Filippova, Shobanov, Gunderina & Kiknadze, 1988 (Proviz & Bazova, 2013). Both of these taxa belong to the *C. plumosus* group, which contains sibling species differentiated by karyotype (Kiknadze *et al.*, 2000, Golygina & Kiknadze, 2012). For OTU 24, six larval specimens with identical sequences were classified as *Cladotanytasus* sp. (4 specimens), *Stempellinella* sp. (1 specimen), and *Cryptochironomus* sp. (1 specimen). Since this OTU is fully within a clade in which identification of most other specimens are in the tribe Tanytarsini, and *Cryptochironomus* is not classified as a member of this tribe, the *Cryptochironomus* identification is likely to be in error. Whether the correct designation of OTU 24 is *Cladotanytarsus* or *Stempellinella* remains ambiguous.

In some cases, the assignment of a larval taxon by qualified taxonomists (e.g., those who work for EcoAnalysts) even at the genus level is inconsistent with identifications derived from molecular data. In the present study, the morphological identification of some larval specimens in OTUs 1, 13, 21, 22 and 25 did not agree with the identification determined from GenBank reference or adult chironomid COI sequences. Morphological identification may be compromised by condition and maturity of specimens, preservation and inadequate reference materials available to taxonomists. It is also possible that after identification a specimen might get cross-contaminated by DNA from another specimen or a specimen or vial might be mislabeled, leading to such inconsistencies. Avoiding errors is essential to use this process effectively. To as great an extent possible, we consulted expert taxonomists and sterilized dissecting equipment with ethanol and flame between specimens-to-minimize the risk of sample-to-sample cross-contamination prior to subsequent barcoding.

Since mistakes in reference databases could also lead to ambiguities or incorrect identification, we consider here several instances in which reference databases may have errors. For example, in fig. 4, OTU 33 has two *Paratanytarsus* sp. larval

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

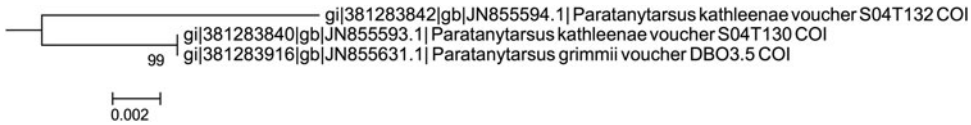


Fig. 7. The neighbor-joining taxon-ID tree with reference sequences only, which was constructed as an intermediate step in obtaining the final tree.


Fig. 8. *Paratanytarsus* sequence relationships for a part of the uncut reference sequence tree.

sequences grouping with a 100% identical *P. grimmii* reference sequence; however, another *P. grimmii* reference sequence (label: DBO3.5) is >3% divergent and is suspected to be incorrectly labeled as *P. grimmii*. Taxon-ID trees of quality reference sequences constructed without the inclusion of adult sequences (figs. 7 and 8) reveals that the highly divergent '*Paratanytarsus grimmii*' sequence is the only one out of 23 very closely related sequences with this designation, leading us to suspect that this identification is likely incorrect (fig. 7). Instead, the correct identity of this reference sequence is more likely to be *P. kathleenae*, in agreement with the other 22 specimens in this clade (fig. 8).

Despite the existence of such ambiguities, inconsistencies and possible errors, the clades of the taxon-ID tree of larvae in fig. 3 mostly show excellent congruence with previous morphological taxonomic classification to the family, subfamily, or tribe levels. Thus, the clades of OTUs 1–21 and 30–33 contain genera that are all classified as subfamily Chironominae in the tribe Chironomini; genera of OTUs 37–45 are all members of the subfamily Tanypodinae; and genera of OTUs 34–36 are all members of the subfamily Orthocladiinae. Only OTUs 22–29 represent an exception to this general congruence of molecular clades with known subfamilies and tribes: while the majority of these genera are classified as subfamily Chironominae of the tribe Tanytarsini, exceptions are one specimen in OTU 24 and 5 specimens of OTU 25, whose genera are classified as members of the tribe Chironomini. The mix of tribes within this clade may simply indicate the difficulty of determining larval morphology or it may reflect identification errors.

### *Significance of identifying chironomid larvae*

Species level identification of chironomid larvae is useful due to the importance of larvae in aquatic food webs (Oliver, 1971, Armitage *et al*., 1995) and the negative impacts of some species as pests (Ali, 1996, Broza *et al*., 2003) and potential invaders (Failla *et al*., 2015). For example, OTU 33 (fig. 4) confirms the presence of *P. grimmii*, a parthenogenic nuisance species known for colonizing water treatment ponds and their associated equipment, such as granular activated carbon absorbers (Langton *et al*., 1988, Olsen *et al*., 2009). Also, OTUs 35 and 36 identify the presence of two species of *Cricotopus* that are both known to be colonizers and detrimental pests of rice fields in California (Clement *et al*., 1977).

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

In addition to identifying known nuisance species, this method has the potential to identify the presence of new species that are either previously undetected or invasive. For example, OTU 22 indicates the presence of *M. insignilobus*, a species associated with waters of low organic content that has a very limited record of detection in the Great Lakes, but whose distribution is well described in Northern Europe (Saether, 1979, Ilyashuk & Ilyashuk, 2001). To our knowledge, no previous records exist of this species in Lake Erie, although new species of the *Micropsectra* genus, such as *M. subletteorum*, a sister species of *M. insignilobus*, have recently been described in Minnesota via similar use of molecular and morphological methods (Anderson *et al.*, 2013).

### *Future research needs*

Despite the possibility of ambiguities, inconsistencies and reference database errors, we recommend the molecular barcode methods used in this study to identify chironomid larvae in future studies. Disagreements in molecular identifications exist, suggesting that developing a more comprehensive library of diverse genetic markers and employing additional identification techniques, such as karyotyping, may resolve some issues. Because some OTUs are only identified by one reference sequence, confirming their identity with barcodes of replicate reference specimens would be beneficial. In addition, improving the quality of existing databases is needed. To some degree, errors were avoided by sorting out sequences that had low quality scores or many ambiguous bases (N's, for example). COI is able to provide presumptive species-level identifications in many cases and in general is considered accurate and reliable (Silva *et al.*, 2013, White *et al.*, 2013). Nevertheless, COI does not provide as great a resolution as CAD or *Cyt b* genes (Ekrem *et al.*, 2010, Carew *et al.*, 2011). The use and establishment of other DNA markers in the future could contribute significantly to the reference data and improve the field of DNA-based taxonomic identification.

The larvae and adults in this study were collected from just one small portion of Lake Erie and the Maumee River. For studies throughout the Great Lakes, reference databases ought to be established for specimens elsewhere in the region since reference sequences are likely to be region-specific. While we have expanded the number of reference sequences available for this relatively little studied portion of Western Lake Erie, if this study were performed in Scandinavia, a heavily studied region, the number of reference sequences that would match sequences of environmentally collected midge sequences would likely be higher. Establishing barcodes and analyzing the phylogeny is important in regions where chironomids have not been heavily studied in this manner (Bergsten *et al.*, 2012).

In order to improve the results obtained from the methods of chironomid identification used in this study, a more comprehensive collection and identification of adult flies within the Western Lake Erie region needs to be done. For future studies, obtaining a larger sample size and a more diverse assortment of adult flies from the region in question may increase the number of identifiable larval OTUs. Establishing quality reference sequences that are supported by professional taxonomists is integral to utilize this process. As more quality sequences are submitted to public databases species identifications based on molecular taxonomy will be more accessible.

Because chironomid larvae contribute significant biomass and diversity to aquatic ecology, it is important to have reliable methods of species level identification. The use of barcodes from adult midges collected from Western Lake Erie helped to resolve the species level identity of several larval clades collected in benthic samples. The present study validated the use of adults for further resolution of larval species identification. Our study enhances existing work in regions where chironomid populations are prominent and allows species level identifications to be more reliable, accurate and accessible. Establishing a comprehensive reference database of multiple DNA barcodes using reliable, cross-referenced adults identified by expert taxonomists, as was done in this study could potentially resolve problems of species level taxonomy of larvae in the family Chironomidae.

## Acknowledgements

This project was supported by grants from the Environmental Protection Agency (Grant number GL00E00808-0) and funds from Wayne State University. We thank Tjorborn Ekrem for his advice and communication, as well as helpful reviewers for aiding in improving the presentation of our research. This article is contribution 1945 of the U.S. Geological Survey Great Lakes Science Center.

## References

**Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M. G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W. & Zinner, D.** (2013) Hybridization and speciation. *Journal of Evolutionary Biology* **26**, 229–246.

**Ali, A.** (1996) A concise review of chironomid midges (Diptera: Chironomidae) as pests and their management. *Journal of Vector Ecology* **21**, 105–121.

**Ali, A., Stanley, B.H. & Majori, G.** (1985) Daily abundance patterns of pestiferous Chironomidae (Diptera) in an urban lakefront in central Florida. *Environmental Entomology* **14**, 780–784.

**Anderson, A.M., Stur, E. & Ekrem, T.** (2013) Molecular and morphological methods reveal cryptic diversity and three new species of Nearctic *Micropsectra* (Diptera: Chironomidae). *Freshwater Science* **32**, 892–921.

**Armitage, P.D., Cranston, P.S. & Pinder, L.C.V.** (1995) *The Chironomidae: Biology and Ecology of Non-Biting Midges*. London, Chapman and Hall.

**Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., Ribera, I., Nilsson, A.N., Barraclough, T.G. & Vogler, A.P.** (2012) The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology* **61**, 851–869.

**Broza, M., Halpern, M., Gahanma, L. & Inbar, M.** (2003) Nuisance chironomids in waste water stabilization ponds: monitoring and action threshold assessment based on public complaints. *Journal of Vector Ecology* **28**, 31–36.

**Carew, M.E., Pettigrove, V. & Hoffmann, A.A.** (2005) The utility of DNA markers in classical taxonomy: using cytochrome oxidase I markers to differentiate australian *Cladopelma* (Diptera: Chironomidae) midges. *Annals of the Entomological Society of America* **98**, 587–594.

**Carew, M.E., Marshall, S.E. & Hoffmann, A.A.** (2011) A combination of molecular and morphological approaches

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

resolves species in the taxonomically difficult genus *Procladius* Skuse (Diptera: Chironomidae) despite high intraspecific morphological variation. *Bulletin of Entomological Research* **101**, 505–519.

**Carr, J.F. & Hiltunen, J.K.** (1965) Changes in the bottom fauna of Western Lake Erie from 1930 to 1961. *Limnology and Oceanography* **10**, 551–569.

**Clement, S., Grigarick, A. & Way, M.** (1977) The colonization of California rice paddies by chironomid midges. *Journal of Applied Ecology* **14**, 379–389.

**Cranston, P.S., Dillon, M.E., Pinder, L.C.V. & Reiss, F.** (1989) The adult males of Chironominae (Diptera, Chironomidae) of the Holarctic region – keys and diagnoses. *Entomologica Scandinavica* **34**, 353–502.

**Dendy, J.S. & Sublette, J.E.** (1959) The Chironomidae (=Tendipedidae: Diptera) of Alabama with descriptions of six new species. *Annals of the Entomological Society of America* **52**, 506–519.

**Ekrem, T., Willassen, E. & Stur, E.** (2007) A comprehensive DNA sequence library is essential for identification with DNA barcodes. *Molecular Phylogenetics and Evolution* **43**, 530–542.

**Ekrem, T., Willassen, E. & Stur, E.** (2010) Phylogenetic utility of five genes for dipteran phylogeny: a test case in the Chironomidae leads to generic synonymies. *Molecular Phylogenetic and Evolution* **57**, 561–571.

**Epler, J.H.** (1988) Biosystematics of the genus *Dicrotendipes* Kieffer, 1913 (Diptera: Chironomidae: Chironominae) of the world. *Memoirs of the American Entomological Society* **36**, 1–214.

**Failla, A.J., Vasquez, A.A., Fujimoto, M. & Ram, J.L.** (2015) The ecological, economic and public health impacts of nuisance chironomids and their potential as aquatic invaders. *Aquatic invasions* **10**, 1–15.

**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.

**Golygina, V.V. & Kiknadze, II.** (2012) A revision of chromosome II (CD) mapping in *Chironomus plumosus* (Linnaeus, 1758) group (Diptera, Chironomidae). *Comparitive Cytogenetics* **6**, 249–266.

**Gresens, S., Stur, E. & Ekrem, T.** (2012) Phenotypic and genetic variation within the *Cricotopus sylvestris* species-group (Diptera, Chironomidae), across a Nearctic – Palaearctic gradient. *Fauna norvegica* **31**, 137–149.

**Hebert, P.D., Cywinska, A., Ball, S.L. & deWaard, J.R.** (2003*a*) Biological identifications through DNA barcodes. *Proceedings of the Biological Sciences* **270**, 313–321.

**Hebert, P.D.N., Ratnasingham, S. & de Waard, J.R.** (2003*b*) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, S96–S99.

**Heyn, M.** (1992) A review of the systematic position of the North American species of the genus *Glyptotendipes*. *Netherland Journal of Aquatic Ecology* **26**, 129–137.

**Hirabayashi, K. & Okino, T.** (1998) Massive flights of chironomid midge nuisance insects around a hypereutrophic lake in Japan: a questionnaire survey of tourists. *Journal of the Kansas Entomological Society* **71**, 439–446.

**Ilyashuk, B. & Ilyashuk, E.** (2001) Response of alpine chironomid communities (Lake Chuna, Kola Peninsula, northwestern Russia) to atmospheric contamination. *Journal of Paleolimnology* **25**, 467–475.

**Inoue, E., Kimura, G., Denda, M., Tokioka, T., Tsushima, K., Amano, K. & Hirabayashi, K.** (2008) Utility of larval rearing for assessment of lotic chironomid assemblages: a case study in a riffle/pool section of the middle reaches of the Shinano River, Japan. *Boletim do Museu Municipal do Funchal* **13**, 119–126.

**Iwakuma, T.** (1992) Emergence of Chironomidae from the shallow eutrophic Lake Kasumigaura, Japan. *Hydrobiologia* **245**, 27–40.

**Kiknadze, I.I., Butler, M.G., Golygina, V.V., Martin, J., Wulker, W.F., Sublette, J.E. & Sublette, M.F.** (2000) Intercontinental karyotypic differentiation of *Chironomus entis* Shobanov, a Holarctic member of the *C. plumosus* group (Diptera, Chironomidae). *Genome* **43**, 857–873.

**Langton, P., Cranston, P. & Armitage, P.** (1988) The parthenogenetic midge of water supply systems, *Paratanytarsus grimmii* (Schneider) (Diptera: Chironomidae). *Bulletin of Entomological Research* **78**, 317–328.

**Langdon, P.G., Ruiz, Z., Brodersen, K.P. & Foster, I.D.L.** (2006) Assessing lake eutrophication using chironomids: understanding the nature of community response in different lake types. *Freshwater Biology* **51**, 562–577.

**Martin, J., Sublette, J.E. & Caldwell, B.A.** (2011) Description of *Chironomus* quinnitukqut, n. sp., closely related to the *C. decorus* group in North America, with characterization of an additional larval form from halobiontic habitats. *Zootaxa* **2716**, 29–41.

**Michailova, P. & Fischer, J.** (1986) Speciation within the plumosus group of the genus *Chironomus* Meigen (Diptera, Chironomidae). *Journal of Zoological Systematics and Evolutionary Research* **24**, 207–222.

**Oliver, D.** (1971) Life history of the Chironomidae. *Annual Review of Entomology* **16**, 211–230.

**Olsen, A., Leadbeater, B.S., Callow, M.E., Holden, J.B. & Bale, J. S.** (2009) The origin and population dynamics of annually re-occurring *Paratanytarsus grimmii* (Diptera: Chironomidae) colonising granular activated carbon (GAC) adsorbers used in potable water treatment. *Bulletin of Entomological Research* **99**, 643–651.

**Pfenninger, M., Nowak, C., Kley, C., Steinke, D. & Streit, B.** (2007) Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic Chironomus (Diptera) species. *Molecular Ecology* **16**, 1957–1968.

**Proulx, I., Martin, J., Carew, M. & Hare, L.** (2013) Using various lines of evidence to identify *Chironomus* species (Diptera: Chironomidae) in eastern Canadain lakes. *Zootaxa* **3741**, 401–458.

**Proviz, V.I.** (2008) Speciation and chromosomal evolution of the Baikalian endemic chironomids of the genus *Sergentia* Kief. (Diptera, Chironomidae): Karyotype divergence and chromosomal polymorphism in the populations of eurybathic species *Sergentia flavodentata* Tshern. and littoral species *Sergentia baicalensis* Tshern. *Russian Journal of Genetics* **44**, 1037–1048.

**Proviz, V.I. & Bazova, N.V.** (2013) Karyotypic features of *Chironomus entis* and *Chironomus borokensis* (Diptera, Chironomidae) from Lake Kotokel (Lake Baikal Basin). *Entomological Review* **93**, 35–44.

**Ram, J.L., Banno, F., Gala, R.R., Gizicki, J.P. & Kashian, D.R.** (2014) Estimating sampling effort for early detection of non-indigenous benthic species in the Toledo harbor region of Lake Erie. *Management of Biological Invasions* **5**, 209–216.

**Roback, S.S.** (1971) The adults of the subfamily Tanypodinae (=Pelopiinae) in North America (Diptera: Chironomidae). *Academy of Natural Sciences of Philadelphia Monographs* **17**, 1–410.

**Saether, O.A.** (1979) Chironomid communities as water quality indicators. *Holarctic Ecology* **2**, 65–74.

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press

**Saether, O.A.** (2009) *Cryptochironomus* Kieffer from Lake Winnipeg, Canada, with a review of Nearctic species (Diptera: Chironomidae). *Zootaxa* **2208**, 1–24.

**Saether, O.A.** (2011) *Glyptotendipes* Kieffer and *Demeijerea* Kruseman from Lake Winnipeg, Manitoba, Canada, with the description of four new species (Diptera: Chironomidae). *Zootaxa* **2760**, 39–52.

**Sharley, D.J., Pettigrove, V. & Parsons, Y.M.** (2004) Molecular identification of *Chironomus* spp. (Diptera) for biomonitoring of aquatic ecosystems. *Australian Journal of Entomology* **43**, 359–365.

**Silva, F.L. & Wiedenbrug, S.** (2014) Integrating DNA barcodes and morphology for species delimitation in the *Corynoneura* group (Diptera: Chironomidae: Orthocladiinae). *Bulletin of Entomological Research* **104**, 65–78.

**Silva, F.L.d., Ekrem, T. & Fonseca-Gessner, A.A.** (2013) DNA barcodes for species delimitation in Chironomidae (Diptera): a case study on the genus *Labrundinia*. *The Canadian Entomologist* **145**, 589–602.

**Sinclair, C.S. & Gresens, S.E.** (2008) Discrimination of *Cricotopus* species (Diptera: Chironomidae) by DNA barcoding. *Bulletin of Entomological Research* **98**, 555–563.

**Tabaru, Y., Moriya, K. & Ali, A.** (1987) Nuisance midges (Diptera, Chironomidae) and their control in Japan. *Journal of the American Mosquito Control Association* **3**, 45–49.

**Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S.** (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731–2739.

**Townes, H.K. Jr.**, (1945) The Nearctic species of Tendipedini [Diptera, Tendipedidae (=Chironomidae)]. *American Midland Naturalist* **34**, 1–206.

**White, B.P., Pilgrim, E.M., Boykin, L.M., Stein, E.D. & Mazor, R. D.** (2013) Comparison of four species-delimitation methods applied to a DNA barcode data set of insect larvae for use in routine bioassessment. *Freshwater Science* **33**, 338–348.

https://doi.org/10.1017/S0007485315000486 Published online by Cambridge University Press