Winter 2008

Rapid identification of New England freshwater copepods using a novel genetic barcode

Elisah B. Allan
University of New Hampshire, Durham

RAPID IDENTIFICATION OF NEW ENGLAND FRESHWATER COPEPODS USING A NOVEL GENETIC BARCODE

BY

ELISHA B. ALLAN
B.S. Biology: Evolution, Ecology, and Behavior, University of New Hampshire, 2006

THESIS

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Master of Science
in
Zoology

December, 2008

This thesis has been examined and approved.

Thesis Co-Director, Dr. James F. Haney
Professor of Zoology

Thesis Co-Director, Dr. Marianne K. Litvaitis
Professor of Zoology

Dr. Larry G. Harris, Professor of Zoology

December 14, 2018
Date

DEDICATION

I would like to dedicate this thesis to all of my teachers: my formal teachers and professors, my parents Richard and Carol, my grandparents Paul, Catherine, and Eleanor, my brother Josh, my friends, Todd and all the other people who took time to share their knowledge and creativity with me. I would not be at this place in my life without your time, patience, and inspiration.

ACKNOWLEDGEMENTS

My sincerest gratitude to my thesis advisers, Dr. Marianne Litvaitis and Dr. James Haney, for all they have taught me and for all of their encouragement, support, and time. This project could not have become what it is without their guidance.

I would like to thank my committee member, Dr. Larry Harris, for all of his support and advice, specifically his input into my research goals and experimental design. I would like to extend my gratitude to the staff of the United States Environmental Protection Agency (US EPA) Region 1 Regional Environmental Monitoring and Assessment Program (REMAP), specifically Hilary Snook, and all of the members of the University of New Hampshire Center for Freshwater Biology for sample collection and support. I greatly appreciate the taxonomic help provided by Dr. Janet Reid (Virginia Museum of Natural History). This research was funded by the US EPA Region 1 through the New England Interstate Water Pollution Control Commission. I am grateful for Dr. Kate Rawlinson's assistance with DNA protocols.

Additionally, I would like to thank the faculty and other graduate students at UNH for their help and encouragement. I especially want to thank Dr. Jessica Bolker for her advice, editing, and friendship. Thanks also to Dr. Marcela Bolaños for being my role model for graduate level research.

I deeply appreciate all of the love, support, and encouragement offered by my whole family and all of my friends, especially my parents, Richard and Carol, my brother Josh, and Todd. I would like to thank my dad, Josh, and Todd for

helping me sample and teaching me different computer programs. Additionally, I appreciate all of the time Todd kept me company while I worked in the laboratory.

Table of Contents

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

ABSTRACT
RAPID IDENTIFICATION OF NEW ENGLAND FRESHWATER COPEPODS USING A NOVEL GENETIC BARCODE

by

Elisha B. Allan

University of New Hampshire, December, 2008

Identification of freshwater calanoid and cyclopoid copepods is limited to adults of certain sexes because morphological keys are mostly based on mature reproductive structures, necessitating an alternate method. Genetic barcodes are an additional tool for distinguishing species using variation in short segments of DNA. I tested the utility of the 28S rDNA D3 expansion segment as a barcode for identifying five species of calanoids and five species of cyclopoids from multiple lakes from New England. Neighbor-joining trees grouped all conspecifics together with high bootstrap support, except for *Leptodiaptomus minutus.* Comparisons of intra- vs. interspecific variation revealed a barcode gap for both calanoids and cyclopoids. Fifty characteristic attributes (CAs) were identified that separate specimens from ordinal to specific levels. Overall, the barcode shows promise as an alternate identification tool for freshwater calanoids and cyclopoids and future research should evaluate the barcode for more species over a wider geographic range.

INTRODUCTION

According to Dayrat (2005), most investigations into speciation, ecosystems diversity and management, or conservation actions depend on accurate species identifications. Correct species identifications are the provenance of taxonomy, which incorporates morphological, behavioral, geographic, and increasingly molecular data to develop hypotheses about species groups and species boundaries. Traditionally, morphological keys have been developed to determine species based on the known taxonomy. However, many keys are limited to a specific sex in dimorphic species or to certain life stages in cases where species undergo ontogenic changes. This deficiency in morphological keys has created the need for an additional identification tool.

Genetic barcodes may provide a supplementary approach for species identifications using molecular rather than morphological data. Such barcodes are short segments of DNA that are used to differentiate unknown samples to species based on variations in nucleotide sequences. This approach has been used extensively for identification of organisms that cannot be distinguished easily using morphology, e. g., bacteria and viruses (Blaxter, 2004). Among metazoans, DNA barcoding may allow for species identifications that are difficult when based on morphology. Additionally, species assignments may be possible for individuals of sexually dimorphic species or for life stages that are not included in morphological keys. A case where identifications would be greatly aided by genetic barcodes are the freshwater copepods. Calanoida and

Cyclopoida are two orders of crustaceans commonly studied in lake ecology because of their ubiquitous presence and ecological importance in lakes (Boxshall & Defaye, 2008; Wetzel, 1975). Understanding the role of small herbivores and predators such as copepods requires that all life history stages of individual species be accurately identified. Copepods are of great importance because they have been used as indicators of lake health (Yan et al., 2004; Binks et al., 2005; Gerten & Adrian, 2002). Yet, morphological keys for the freshwater copepods are limited mostly to mature reproductive structures that exclude certain sexes (because of dimorphisms) and immature specimens. Additionally, the adult morphology of closely related taxa is highly similar, making species distinctions difficult for inexperienced researchers. In this thesis, I propose that genetic barcodes may be a valuable identification tool for freshwater calanoids and cyclopoids.

In the first chapter of my thesis, I describe the diversity of freshwater calanoids and cyclopoids and their roles within the lake ecosystem, demonstrating the importance of correct species identifications. The second chapter focuses on DNA barcoding and discusses the use of ribosomal DNA as an alternative genetic barcode for freshwater copepods. The last chapter contains the findings of my study. It consists of an evaluation of the D3 expansion segment of the nuclear large ribosomal subunit (28S rDNA) as an alternate DNA barcode. I used five calanoid and five cyclopoid species (a total of 177 adult individuals) and six immature copepods from one calanoid species and two cyclopoid species in my study. The results not only demonstrate the potential

use of this gene region to identify sexes and life stages previously unidentifiable, but also provide insight into how these barcodes should be used to distinguish species.

Hypotheses and Experimental Design

The overall goal of my thesis was to evaluate the utility of the D3 expansion segment of the 28S rDNA gene as an alternate barcode to the traditionally used CO-I gene for the identification of freshwater cyclopoid and calanoid copepods. The following goals were developed to test the effectiveness of the barcode:

- To determine the range of values for intraspecific versus interspecific variation for five species each of freshwater calanoid and cyclopoid copepods.
- To compare intraspecific genetic variation of the D3 expansion segment within and between populations.
- To determine if the barcode can identify all life stages: eggs through adults.

The following hypotheses were tested:

- Hypothesis 1. Intraspecific variation will be less than interspecific variation.
- Hypothesis 2. Intraspecific variation within and between populations will be smaller than interspecific variation, and conversely interspecific variation within and between populations will be greater than intraspecific variation. Therefore the barcode will identify conspecifics across populations.

- Hypothesis 3. DNA can be extracted, amplified, and sequenced from all life history stages and the sequences will assign all immature specimens to the correct species.

To determine whether intraspecific variation among all species will be less than interspecific variation among all species (Hypothesis 1), the proposed barcode region was evaluated for at least fifteen individuals for each of five species of calanoids and five species of cyclopoids (excluding *Tropocyclops* spp). Nucleotide sequence divergences were compared for all individuals within each species (intraspecific variation) and between all individuals of different species using pairwise comparisons. Percent similarity for each comparison was calculated and evaluated by gene trees and represented via a frequency histogram.

Hypothesis 2 was tested by including copepods of the same species from at least three different populations (lakes). Genetic distances between members of the same and different populations were compared to determine variation between populations and to evaluate any potential geographic trend.

I determined whether the barcode region could determine species for all life history stages of copepods (Hypothesis 3) by extracting and sequencing DNA from eggs, nauplii larvae, and copepodids. Eggs were removed from identified females. DNA from some eggs was sequenced, whereas some eggs were allowed to further develop to provide nauplii. Sequences derived from eggs and from larval stages were included in a gene tree with sequences derived from

adults to determine whether immature stages would group with adults of their identified species.

CHAPTER I

FRESHWATER CALANOID AND CYCLOPOID COPEPODS: IMPORTANCE AND CHALLENGES WITH SPECIES IDENTIFICATIONS

Copepods are a diverse group of over ten thousand species (Cole, 1994). They are members of the phylum Arthropoda and the subphylum Crustacea. Their body plan consists of a metasome (dorsally fused cephalothorax) and a urosome (a segmented abdomen); which fulfills the three section body plan of Arthropods (Fig. 1; Wetzel, 1975; Cole, 1994). Copepods have two pairs of antennae, three pairs of mouthparts, six swimming legs, and attached to the last abdominal segment are two caudal rami with setae (Fig. 1; Wetzel, 1975; Cole, 1994). Copepods evolved in the marine environment, but some families have colonized freshwater and diversified into many different genera and species (Boxshall & Jaume, 2000).

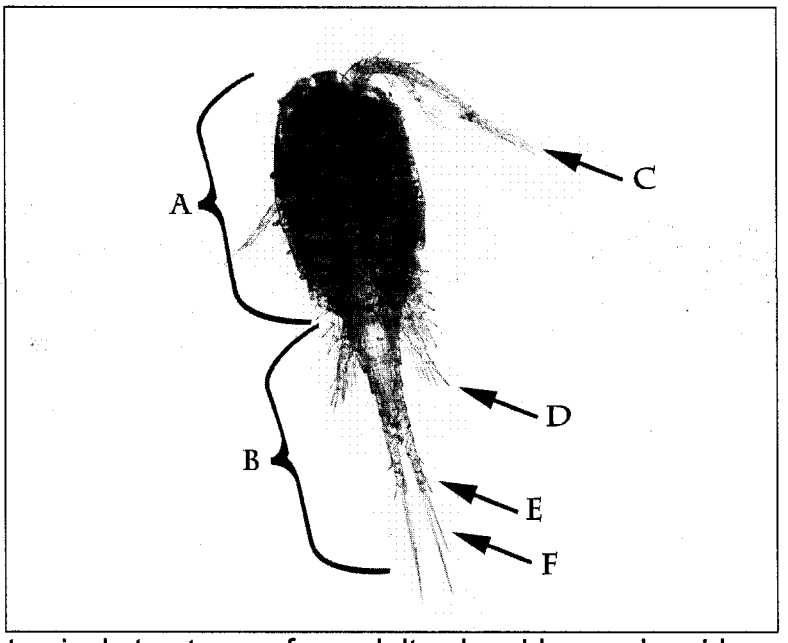

Figure 1. Anatomical structures of an adult calanoid or cyclopoid copepod. Female adult cyclopoid pictured. **A.** Metasome **B**. urosome **C.** Two pairs of antennae **D.** Six pairs of swimming legs **E.** Caudal rami **F.** Caudal setae

Freshwater planktonic copepods from the orders Calanoida and Cyclopoida inhabit a wide range of habitats of all size ranges, including ground water and hot springs (Boxshall & Defaye, 2008). Within New England, copepods are found in lakes ranging in size, trophic levels, and of different animal and plant composition. Most calanoid species are planktonic (residing in the deeper regions of a lake and avoiding the littoral region) and most cyclopoid species are littoral or benthic, although a few species of each can be found occupying the alternate region (Wetzel, 1975). All copepods are considered macro-zooplankton (body size >50 μm), and each species' size range greatly influences their food preferences.

The feeding habits of copepods place them in a central position within lake food webs. Calanoid and cyclopoid copepods link bacteria, phytoplankton

(primary producers), and smaller zooplankton (herbivores, omnivores) with larger zooplankton (omnivores and carnivores) and planktivorous fish (Fig. 2; Cole, 1994; Wetzel, 1975). The microbial food loop is affected by copepods as they prey on ciliates at rates increasing with eutrophication (Burns & Schallenberg, 2001). Freshwater copepods can influence terrestrial animals, such as the endangered larval flatwood salamander, *Ambyostoma cingulatum* COPE, whose stomach content contained 16% comprised of cyclopoid copepods (Whiles et al., 2004). Additionally, some freshwater copepods affect other organisms' life cycles as intermediate hosts of parasites.

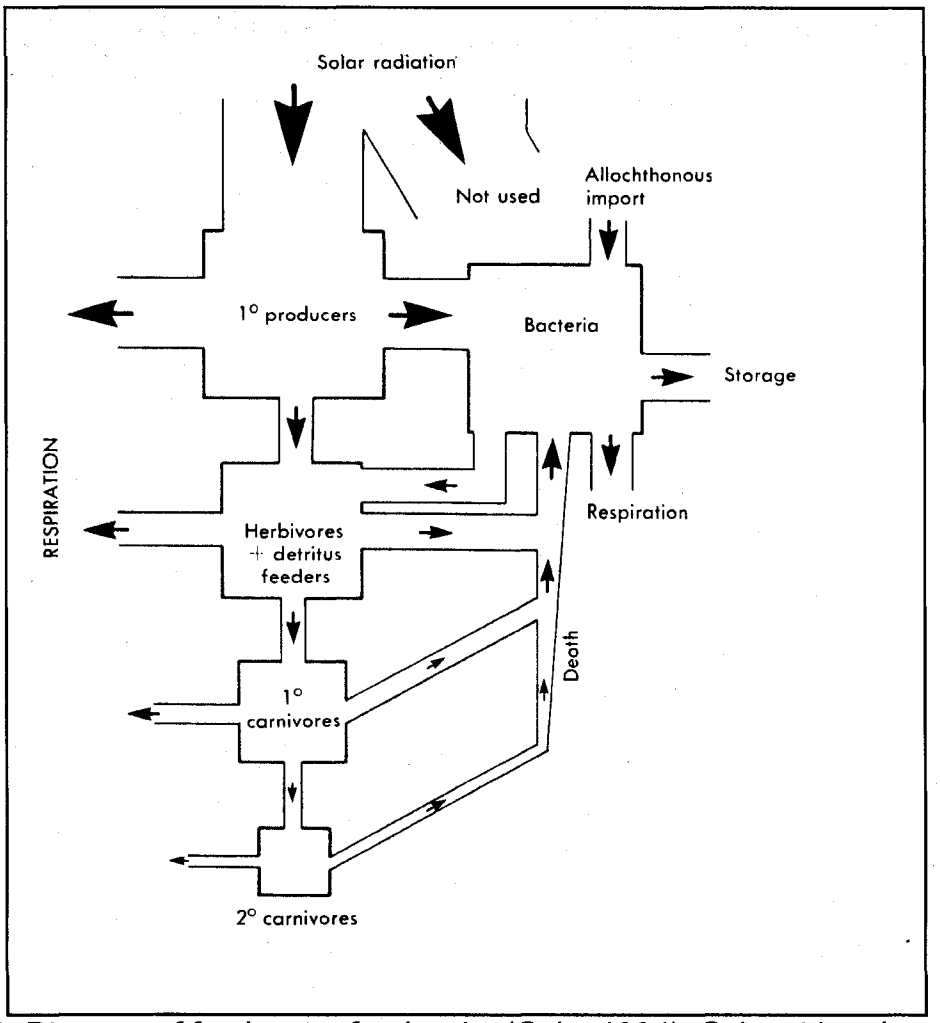

Figure 2. Diagram of freshwater food webs (Cole, 1994). Calanoid and cyclopoid copepods can occupy the herbivore, primary carnivore, and secondary carnivore levels within the food web.

Cyclopoid copepods influence human health as parasite vectors and in pest control. Species of *Mesocyclops* and *Tropocyclops* are intermediate hosts to the nematode parasite, *Dracunculus medinensis* LINNAEUS which is a health problem in West Africa and India (Boxshall & Defaye, 2008). They can also be

host to *Diphyllobothrium latum* LINNAEUS, a fish tapeworm (Boxshall & Defaye, 2008). In addition to carrying parasites, *Mesocyclops* species prey on mosquito larvae and have been used for mosquito control especially in areas with high incidence of malaria and dengue fever (Boxshall & Defaye, 2008). Although success rates have varied, copepod mosquito control has been reliable in Vietnam (Boxshall & Defaye, 2008). In addition to hosting parasites, some freshwater copepods are parasites themselves, influencing the life of other organisms within their lake system. *Eucyclops bathanalicola* BOXSHALL & STRONG is a parasite of the gastropod, *Bathanalia straeleni* LELOUP, representing the rare account of copepods parasitizing freshwater invertebrates (Boxshall & Strong, 2006).

The influence of copepods goes beyond interactions with other metazoans and extends to the phytoplankton community. Copepod grazing rates alter the ratio of nitrogen and phosphorous by recycling nutrients back into the lake after feeding (Andersen & Hessen, 1991). The direct link of copepods to lake processes is caused not only by their involvement in lake cycles, but by the way they react to alterations in lake conditions.

As lake trophic levels shift, copepod species compositions may change, as certain species prefer particular ranges in nutrient levels (Maier, 1996). Copepod species can also be used to indicate the acidity level in lakes, such as *Tropocyclops extensus* KIEFER and *Mesocyclops edax* FORBES which will quickly disappear as a lake becomes acidic (Waiseng et al., 2003). Certain species of copepods can re-colonize lakes recovering from acidification and heavy-metal

toxin loading before other zooplankton populations can rebound (Yan et al., 2004; Binks et al., 2005).

Copepods respond to the effects of global warming because their population dynamics change in response to temperature changes. Thermophilic copepods have longer growth periods and larger population sizes as the water warms, but copepod species that are adapted to a wide range of temperature conditions remain unaffected (Gerten & Adrian, 2002). Studies of freshwater copepods provide information on lake conditions and health, and on many organismal interactions; especially because freshwater copepods are ubiquitous worldwide.

There are approximately 2, 814 freshwater species of copepods. The two dominant families are the calanoid Diaptomidae and the cyclopoid Cyclopidae; followed by the calanoid Temoridae and Centropagidae (Boxshall & Halsey, 2004). Copepod species are mostly (90.9%) endemic to a single region (Fig. 3; Boxshall & Defaye, 2008). Cyclopidae comprise one third of all species found within the Nearctic region (North America), although only one genus is endemic (Boxshall & Defaye, 2008).

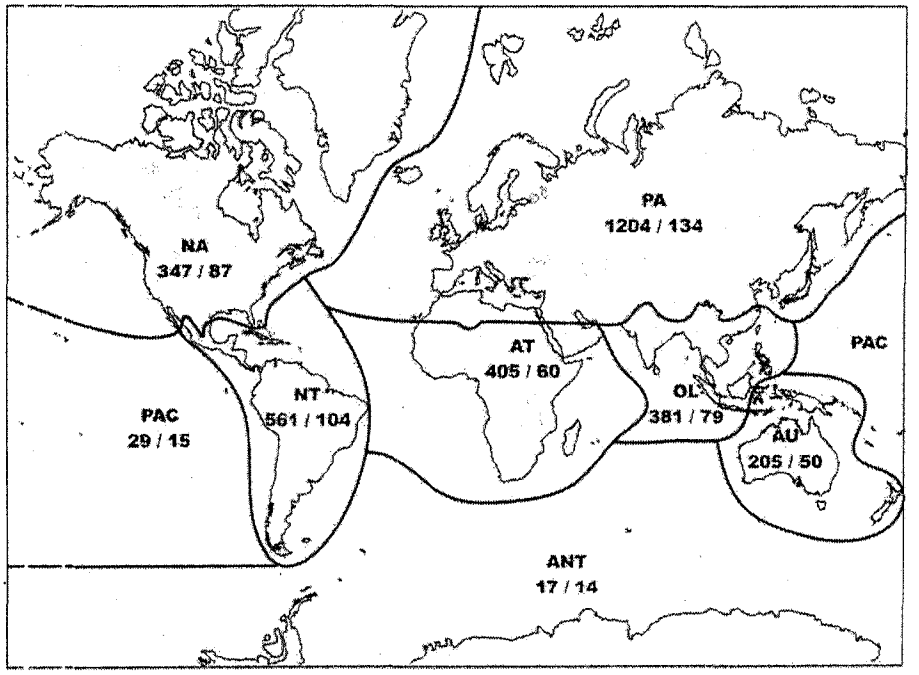

Figure 3. Copepod distribution worldwide (Boxshall & Defaye, 2008). Species number/Genus number. Geographic regions: **PA:** Palearctic **NA:** Nearctic **NT:** Neotropical **AT:** Afrotropical **OL:** Oriental **AU:** Australasian **PAC:** Pacific Oceanic Islands **ANT:** Antarctic

Temoridae are located exclusively in the Northern Hemisphere and the only freshwater genera are *Epischura* and *Heterocope* (Boxshall & Halsey, 2004). There are 30 species worldwide; within the Nearctic region there are 18 species in three genera and none are endemic (Fig. 3; Boxshall & Defaye, 2008).

Diaptomidae consists of about 410 species, making it the largest family within the Calanoida (Boxshall & Defaye, 2008). They are the most widespread geographically and inhabit every continent (Boxshall & Halsey, 2004). Within the Nearctic region, 18% of all copepods are from four diaptomid genera, consisting of 77 species, all endemic (Fig. 3; Boxshall & Defaye, 2008). Endemic genera of North America include: *Acanthodiaptomus, Aglaodiaptomus, Arctodiaptomus, Diaptomus, Eudiaptomus, Hesperodiaptomus, Leptodiaptomus,*

Mastigodiaptomus, Mixodiaptomus, Nordodiaptomus, Onchyodiaptomus, and *Skistodiaptomus* (Boxshall & Halsey, 2004).

Many taxa located within the Nearctic region have distributions affected by Pleistocene glaciers (Fig. 3; Boxshall & Defaye, 2008). The range of *Skistodiaptomus* spp. within the northeastern United States shows varying distribution affected by glacial events and some environmental conditions (Fig. 4; Thum & Stemberger, 2006). The range limits of *S. oregonensis* LILLJEBORG (New York, Vermont, and one location in Maine) and *S. pygmeus* PEARCE (throughout New Hampshire and Maine) occur because of spatial distribution in glacial waterways and is independent of environmental variables (Thum & Stemberger, 2006). *Skistodiaptomus pallidus* LILLJEBORG resides in New York and Southern Connecticut, and is limited by high productivity, high pH, and high nitrogen conditions (Thum & Stemberger, 2006). This species is in direct competition with *S. pygmeus* which prevents the co-occurrence of the two in the same lake (Thum & Stemberger, 2006).

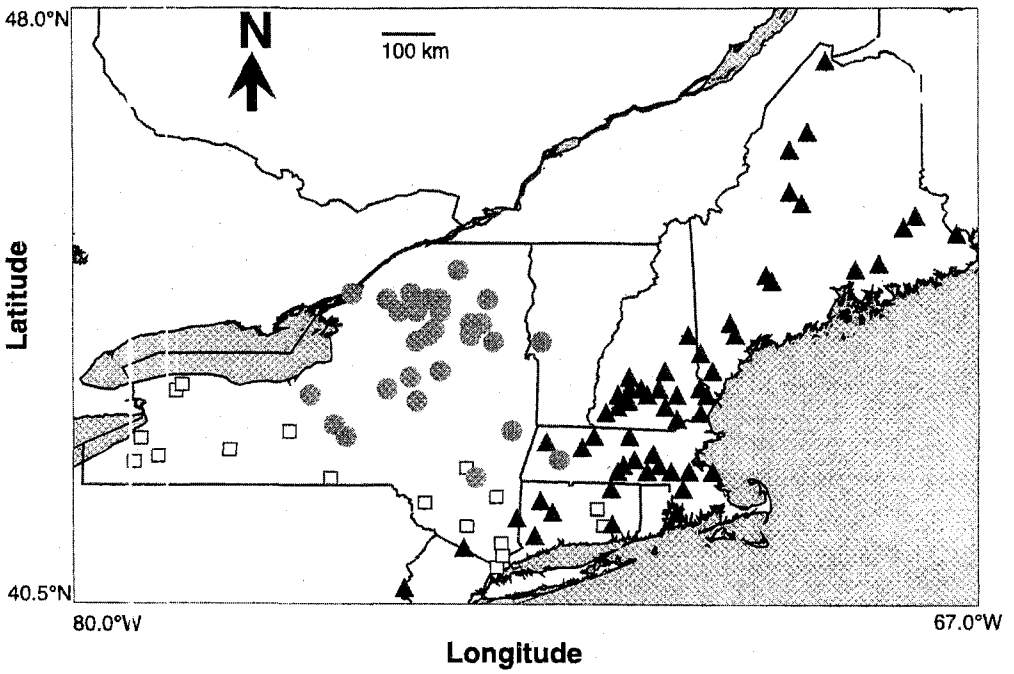

Figure. 4. Map of *Skistodiaptomus* distribution within the Northern United States (Thum & Stemberger, 2006). Shaded circles: *Skistodiaptomus oregonensis* Solid triangles: *Skistodiaptomus pygmeus* Open squares: *Skistodiaptomus pallidus*

One main difficulty in studying copepods involves correct species identifications. The taxonomy of copepods is often scrutinized because of the highly similar morphology of some species and the wide range of variation within species (Fryer, 1954). Diaptomid identifications are challenging because many genera are not separated based on synapomorphies and because most keys are based on male features, making it hard to distinguish females at the generic level (Boxshall & Halsey, 2004). Within *Cyclops* there is a lot of variation that is not used for new species descriptions instead taxonomically distinguishing characters are based on slightly differing features (Lowndes, 1929 in Fryer, 1954). Cyclopoid copepods within the '*vernalis*' complex show a wide range of different forms caused by environmental factors, which makes determining group membership highly error-prone (Fryer, 1954). This same type of variation has

been observed in diaptomids, specifically in two different lake populations of *Diaptomus shoshone* FORBES from lakes in Alberta, Canada. These conspecific from different populations are of very different sizes, which may correspond to available prey or different water temperatures, making them appear morphologically distinct (Anderson, 1967).

Correctly identifying copepods is important to studies because some adult copepods have highly similar morphology but very different behavior. For example, *Leptodiaptomus minutus* LILLJEBORG is similar in structure to *Skistodiaptomus oregonensis*, but the former is an acid-sensitive species and the latter an acid-tolerant species (Binks et al., 2005). This can be further complicated when multiple species with similar morphology occupy the same lake.

The coexistence of multiple copepod species within a lake can make species composition studies more difficult. Cohabitation is hypothetically caused by seasonal separation, vertical separation due to lake stratification and food types, or body size differences affecting the food items on which copepods prey (Wetzel, 1975). In a study of shallow Saskatchewan ponds in Canada, diaptomid species found in the same pond had different reproductive cycles and body lengths (Hammer & Sawchyn, 1968). Sandercock (1967) found that all three factors (size differences, seasonal separation, and different vertical distributions) allowed for the coexistence of three diaptomid species in Clark Lake, Ontario. Seasonal separation is caused by different reproductive cycles and length and can be influenced by the presence of diapause (resting) stages.

Calanoid and cyclopoid copepod are dioecious and are either univoltine or bivoltine (Wetzel, 1975). Mating occurs when the male spermatophore is transferred to the female genital pore; however fertilization of eggs could take place immediately or a few months after copulation (Wetzel, 1975). Calanoid copepods produce a single egg sac and cyclopoids produce two egg sacs, both of which contain multiple eggs (Wetzel, 1975). Clutch size for both orders of copepods depends on the species and time of year (Wetzel, 1975).

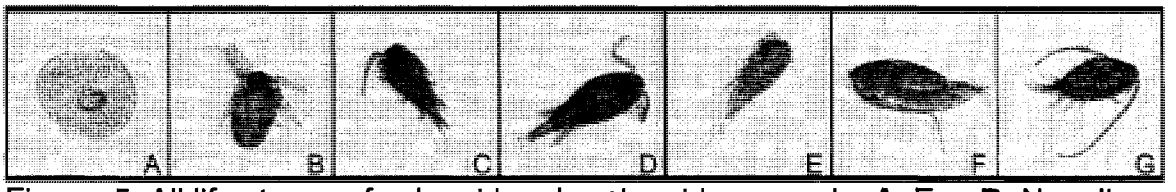

Figure. 5. All life stages of calanoid and cyclopoid copepods. **A.** Egg **B.** Nauplius **C.** Copepodid **D.** Adult male cyclopoid **E.** Adult female cyclopoid **F.** Adult male calanoid **G.** Adult female calanoid

Free-swimming larvae, called nauplii, are released from the egg sac at the time of hatching, and then proceed through six naupliar stages in the calanoids and five naupliar stages in the cyclopoids before reaching the copepodid stage (Fig. 5; Wetzel, 1975; Cole, 1994). The copepod will undergo six copepodid stages, the last of which results in a mature adult copepod (Wetzel, 1975). Cyclopoid copepods often slow down their metabolism and sink to the lake bottom in diapause at the copepodid stages C4 and C5 (Santer & Lampert, 1995) Calanoid copepods can drop their egg sacs directly into lake sediment in the form of diapause eggs. These strategies may protect calanoid and cyclopoid populations during unfavorable conditions such as droughts or times of high fish predation (Hairston et al., 1990; Santer & Lampert, 1995). Eggs remain viable in

the sediment until environmental cues signal the embryos to hatch (Hairston & Munns, 1984; Hairston et al., 1990; Dowell, 1997; DeStasio, 2004).

Production of diapause eggs not only ensures copepod hatching after unfavorable abiotic conditions, but it can allow for copepod eggs to survive predation. Bartholmé et al. (2005) tested the viability of subitaneous and diapause eggs of calanoids (*Eudiaptomus gracilis* SARS and *E. graciloides* LILLJEBORG) and the cyclopoid, *Macrocyclops albidus* JURINE, after passing through fish guts. They found the diapause eggs were slightly more viable than the subitaneous eggs in *E. graciloides*, but both egg types could successfully hatch after passing through fish guts. This can have implications in evolutionary studies of fitness and studies of competition within copepods and between calanoids and cyclopoids and other zooplankton (Bartholmé et al., 2005). In addition to studies of current species, studies on previous copepod populations are possible because they are preserved in lake sediments when in diapause.

Isolating diapause eggs from sediment makes possible species determinations of copepods that may no longer exist in the lake. Most methods for identifying diapause eggs rely on hatching the copepods which often has low success rates and can take many years (Lohner et al., 1990). Identification of cyclopoid copepods in the copepodid stage can be difficult when specimens in the sediment are partially degraded.

The entire life cycle of freshwater copepods from egg to adult is relatively long compared to other zooplankton. Generally spring and summer generations develop in four weeks and fall generations develop in six weeks (Wetzel, 1975).

Diaptomus stagnalis in a Chicago lake took approximately 30 days to develop from eggs to adults (Brewer, 1964). *Aglaodiaptomus clavipes* SCHACHT was reported to develop in about 68 days (Gehrs & Robertson, 1975). In tropical waters, *Tropocylops prasinus* and *Mesocyclops longisetus* have developmental times ranging from 20 to 46 days depending on temperature (Melão & Rocha, 2004). Abdullahi (1990) reported the total life span of *Mesocyclops viridis* JURINE at 20°C as approximately 240 d for females and 160 d for males. The relatively long life cycle of copepods often results in the presence of only immature copepods of a species in a lake at time of sampling.

The main taxonomic characters used in copepod identification are reproductive structures, which may be significant to speciation events by creating reproductive barriers. However, the focus on mature structures excludes the inclusion of immature life history stages and often certain sexes from identification. Since at time of sampling, adults of certain species may not be present, species composition data may become skewed by the under-representation of certain species. Additionally, the lack of keys to both sexes necessitates sampling of specific individuals to allow identification. To overcome the challenges of determining copepod species morphologically, an alternate identification tool is needed and may be genetic barcoding.

CHAPTER II

GENETIC BARCODES AND RIBOSOMAL DNA AS A FRESHWATER CALANOID AND CYCLOPOID BARCODE

Researchers are proposing genetic barcodes to distinguish species as an alternate identification tool to morphological keys. Genetic barcodes (synonymous with molecular tags and DNA signature sequences) use variations in short segments of DNA (~300-600 base pairs) to distinguish between species (Waugh, 2007; Hebert et al., 2003; Floyd et al., 2002). The term barcoding can be misleading by implying that, like a grocery store USP barcode, there is only one unique sequence for each species when there is often variation within conspecific sequences (Moritz & Cicero, 2004). However, genetic barcoding is based on the premise that variation within a species (intra-specific) is much less than the variation between species (inter-specific). Nucleotide sequences should provide enough unique sequences to barcode all of earth's taxa because in a sequence of 15 base pairs lies the possibility of one billion different genetic barcodes (Hebert et al., 2003). The barcodes are established from known taxonomic species then these gene segments can be matched with unknown specimens to identify them to species.

Barcodes are developed from freshly collected species that are carefully identified and determined to be the same as the holotype (Tautz et al., 2003). Barcodes cannot replace taxonomy; it can only be used as an identification tool. An organism's taxonomy is established using many different characters which

provide information to create species hypotheses which can be tested. Whereas a genetic barcode relies only a short nucleotide segment and therefore cannot provide enough information on species boundaries to establish new species (Will & Rubinoff, 2004; Dayrat, 2005; Bhadury et al., 2006). Ultimately, the major limitation of DNA barcoding is the correct species identification for the specimen on which the barcode is based (Hebert & Gregory, 2005), which becomes less problematic when the identification is based on well-established taxonomy. The barcode must be accompanied by a voucher specimen, which allows for examination of the organism if any identifications are questioned or new species arise (Will & Rubinoff, 2004; Waugh, 2007). Since DNA extraction of smaller organisms often requires the use of the entire individual, photographic vouchers may be used instead of a preserved specimen (Tautz et al., 2003).

Barcoding is widely used for species identification of bacteria and fungi; groups that are often impossible to identify using traditional taxonomic methods (Blaxter, 2004). Since genetic barcodes do not use morphological characters, partial specimens can still be identified, as well as immature stages or sexes that are not represented in morphological keys (Lord, 2005; Steinke et al., 2005; Mortiz & Cicero, 2004; Will & Rubinoff, 2004; Tautz et al., 2003; Hebert et al., 2003). Calanoids and cyclopoids are an ideal test organism for genetic barcoding because they demonstrate sexual dimorphisms, ontogenic life history shifts, and have high morphological similarity between sister taxa.

There are different approaches for determining the ability of barcodes to distinguish species. Distance methods look for barcode gaps (separation of intra-

and interspecific variation in a population) and are used to create Neighbor-joining (NJ) trees (Hebert et al., 2004a; Meyer & Paulay, 2005; Wiemers & Fielder, 2007). A barcode gap is created when intraspecific variation is lower than interspecific variation with no overlap between them (Fig. 6).

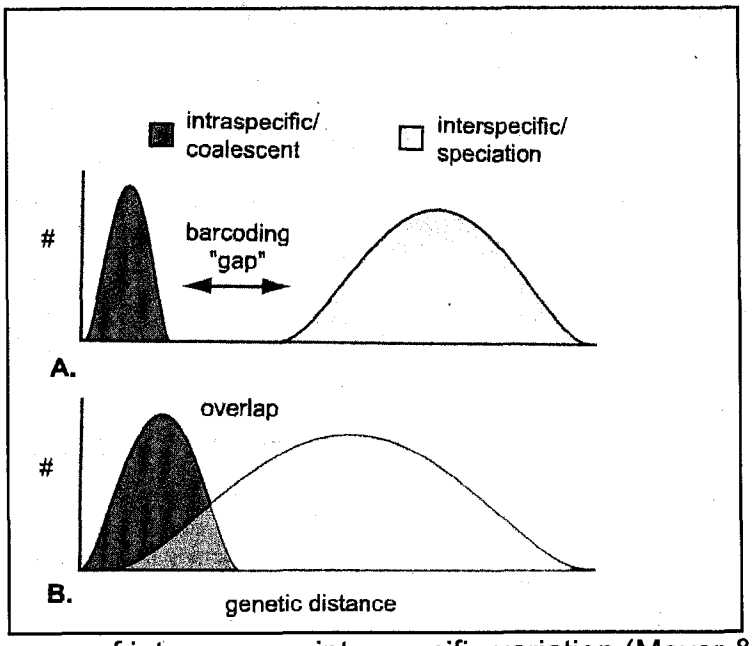

Figure 6. Frequency of intra- versus interspecific variation (Meyer & Paulay, 2005). **A.** An informative barcode where a barcode gap separate intra- and interspecific variation is observed **B.** A non-informative barcode where intra- and interspecific variation overlap.

Gene trees group sequences from conspecifics together while excluding sequences from different species. However, gene tree building is prone to systematic errors when the sequence data do not match the models of sequence evolution used in the tree-building algorithms (Swofford et al., 1996; DeSalle et al., 2005). Wiemers and Fiedler (2007) found that NJ tree profiles resolve species more accurately than barcode gaps, but that the trees still failed to correctly identify 17% of species due to a lack of reciprocal monophyly.

Errors in distance methods can be attributed to false positives and false negatives. False positives include an individual into a group of conspecifics to which it does not belong. This occurs if species have formed genetically distinct populations because of limited gene flow or in allopatric populations with interrupted gene flow (Wiemers & Fiedler, 2007). False positives may raise questions of cryptic species which would need to be investigated by taxonomic methods (Wiemers & Fiedler, 2007). False negatives are the exclusion of an individual from its conspecifics. This can occur when there is not enough sequence variation to distinguish between closely related species, again demonstrating the need for integrative approaches to identifying species (Wiemers & Fiedler, 2007). False positives and negatives can either open or close a barcode gap erroneously. Insufficient sampling often creates an artificial "barcode gap" that does not exist with more expansive sampling (Wiemers & Fielder, 2007; Moritz & Cicero, 2004).

DNA barcoding can be evaluated with distance methods (as described above) or by qualitative data. Characteristic attributes (CAs) are nucleotides at a specific position in the DNA sequence that are unique to a taxon (Rach et al., 2008). CAs can be pure or private meaning they are found in all individuals of the taxon or 80% of all individuals of the taxon respectively (Rach et al, 2008). Additionally, CAs can be simple when it consists of a nucleotide at only one position or compound if nucleotides at multiple positions are combined as one CA (Rach et al, 2008). Simple, pure CAs are the most conservative for species identifications. These characters provide qualitative information and in the

absence of information will not result in a species determination (DeSalle et al., 2005). This may make them a more accurate test of a barcodes ability to distinguish species because it will not rely on intra- versus interspecific variation but on specific characters found within taxa.

DNA barcodes are most likely to fail where they would be most useful, i.e., in instances of closely related species that have very little evolutionary divergence, highly similar morphologies, paraphyly, or polyphyly (Moritz & Cicero, 2004; Meyer & Paulay, 2005). This can be caused by incorrect species hypotheses, hybridizations or introgressions, incomplete lineage sorting, or speciation of peripheral isolates (Funk & Omland, 2003). These problems can be resolved by taxonomic investigation into the species boundaries and by testing the ability of the barcode to distinguish species that are closely related or that have a wide range of variations. Sample size and geographic area sampled need to be representative of the variation within the study group, including that of closely related sister taxa (Moritz & Cicero, 2004). These potential sampling errors pose a threat not only to barcode gap analysis but to other methods of barcode assessment as well.

A limitation to barcoding is that there is no way to predict the length of a sequence that needs to be used as a barcode because rates of molecular evolution vary between different gene segments and different species (Hebert et al., 2003). Barcode lengths need to include enough variation to distinguish between closely related species but should be short enough to reduce sequencing cost and to allow the possibility of DNA extraction from older or ill-

preserved samples. Hajibabaei et al. (2006) obtained sequences of "mini-barcodes" (407 bp, 221 bp, 135 bp, and 134 bp) from 84-98% of samples between 1-21 years old that were either ethanol preserved or oven-dried; but only obtained sequences from 24-39% of sequences when using a full length barcode. However, when sequences are greater than 400 bp, forward and reverse sequencing needs to be used which doubles barcoding costs.

The success of a barcode initiative is based on correctly selecting a gene segment which evolves at a rate that reflects speciation events of the study organism. The DNA region chosen for barcoding needs to be conserved enough within a group of organisms to allow for reliable amplification, yet variable enough to make species distinctions possible (Floyd et al., 2002).

Hebert et al. (2003) proposed using a 650 base pair region at the 5'-end of the cyctochrome *c* oxidase subunit I (CO-I) gene as a universal barcode for all organisms in every phylum. Cytochrome *c* oxidase is a large transmembrane protein that is highly conserved across species that use oxidative phosphorylation for metabolism (Waugh, 2007). This gene region was chosen as a DNA barcode because its functional constraint meant it evolved at a rate that allowed for the design of universal primers (Hebert et al., 2003). At the same time, CO-I evolution is rapid enough within species thus yielding enough variation to separate closely related taxa (Hebert et al., 2004a; Cox & Hebert, 2001).

Like all other mitochondrial genes, CO-I lacks introns (Saccone et al., 1999) thus facilitating sequence alignments (Waugh, 2007). On the other hand, the limitations of using mitochondrial DNA (mtDNA) as a barcode include the

retention of ancestral polymorphisms, its maternal mode of inheritance, and potential transfer of mitochondrial genes into the nuclear genome, creating pseudogenes (Moritz & Cicero, 2004; Roe & Sperling, 2007; Tautz et al., 2003). Nuclear transfer of mtDNA results in nuclear pseudogenes (Numts, nuclear mitochondrial pseudogenes), which can co-amplify. Numts may evolve at different rates than target genes and can lead to incorrect conclusions about speciation events (Funk & Omland, 2003; Benasasson et al., 2001). However, there currently are methods to detect Numts and to prevent against their inclusion (Funk & Omland, 2003; Benasasson et al., 2001).

The major disadvantage of CO-I as a universal barcode lies in its inability to reliably amplify all members of a taxon. Specifically, CO-I is not informative for anthozoans, fungi, polyclad worms, plants, or amphibians (Erpenbeck et al., 2005; Litvaitis, pers. comm., Marshall, 2005; Vences et al., 2005). Roe and Sperling (2007) point out that many barcoding studies do not use the universal primers for CO-I (Folmer, 1994), but rather employ taxon specific primers or multiple primers. Lee (2004) argues that no one gene will be able to differentiate between every species, which is echoed by other researchers who emphasize that multiple genes should be considered to determine the best barcode for each organismal group (Nielsen & Matz, 2006; Will & Rubinoff, 2004).

Although the universal CO-I primers do amplify marine copepods (Bucklin et al., 1999), they do not provide a CO-I product in all freshwater copepod species (Thum & Derry, 2008; Guarneri, 1996; M. Litvaitis, unpublished data; K. Rawlinson, unpublished data). These inconsistencies are due to variations at the

wobble bases which prevent universal primers from annealing to the target regions (Palumbi, 1996).

Alternative genes or gene segments have been suggested as DNA barcodes, specifically the nuclear ribosomal genes (Floyd et al., 2002; Markmann & Tautz, 2005; Tautz et al., 2003).

Ribosomal DNA is under constrained evolution because it is functionally critical to the accuracy of translation during protein synthesis (Lafontaine & Tollervey, 2001). The secondary structure of rDNA forms regions with stems (helices) that lead into expansion segments (loop structures). Sequence divergence is highest in the expansion segments (Fig. 7; Hassouna et al., 1984). The large rDNA gene is characterized by 12 expansions segments (D1-D12) (Hillis & Dixon, 1991; Hassouna et al., 1984).

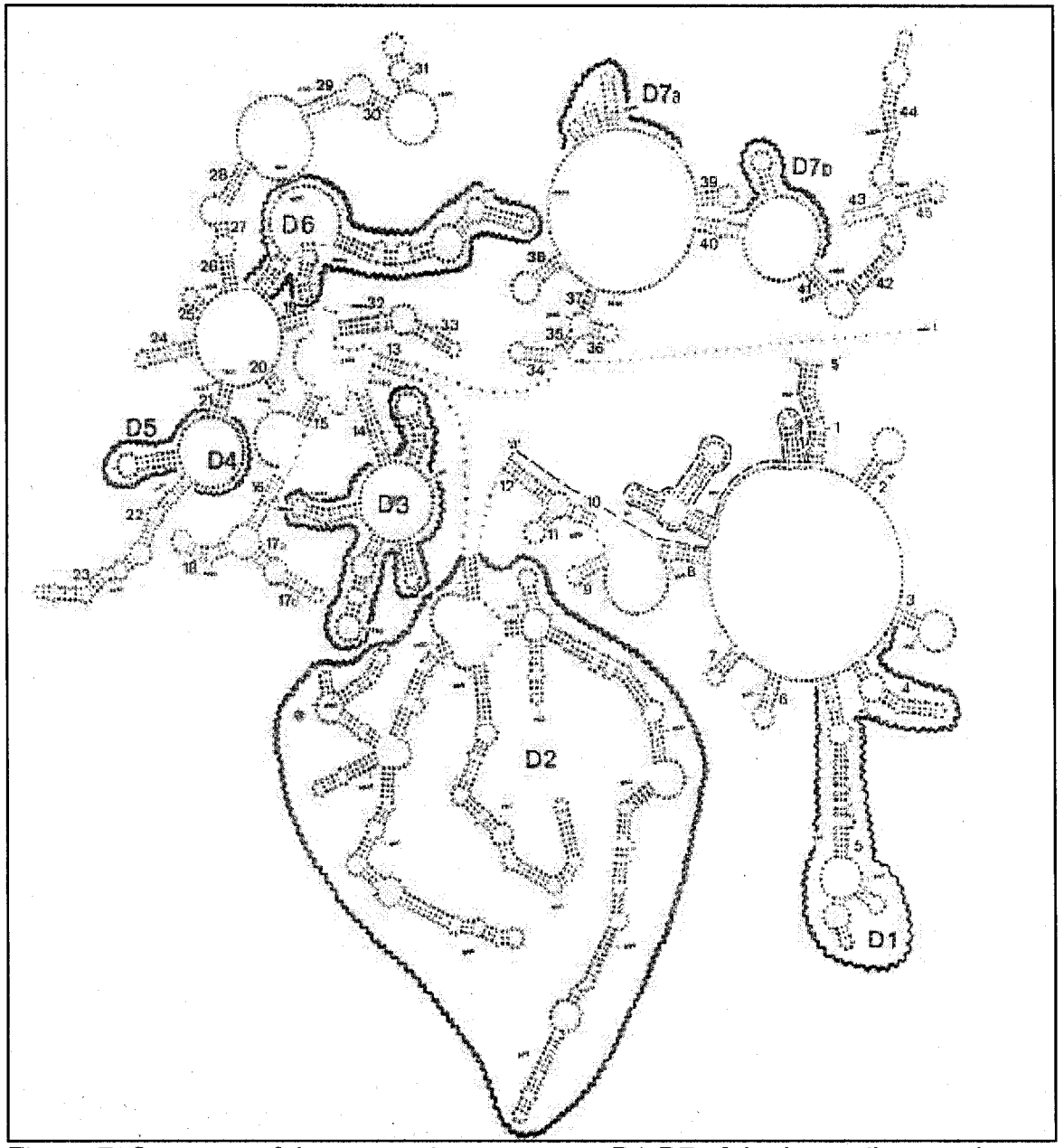

Figure 7. Structure of the expansion segments D1-D7 of the large ribosomal subunit of *Mus musculus* LINNAEUS (Michot et al., 1984).

Expansion segments are evolutionarily non-conserved, do not interfere with the function of the mature rRNA product, and increase the size of eukaryotic rDNA relative to that of prokaryotes (Clark et al., 1984). The slower evolving regions (helices) allow for the design of universal primers (Tautz et al., 2003),

whereas the more rapidly evolving expansion segments show enough variation to differentiate between closely related taxa (Litvaitis et al., 1994).

Multiple copies of rDNA are found in the genome, facilitating DNA extraction and amplification even from a small amount of tissue (Long & Dawid, 1980; Markmann & Tautz, 2005). Although both the large (28S) and small (18S) subunits of rDNA are transcribed together and either can be used for species distinctions, the 28S subunit has more expansion segments (Markmann & Tautz, 2005).

The D3 expansion segment of the 28S rDNA has been used to distinguish species in other invertebrate groups (Litvaitis et al., 1994; Markmann & Tautz, 2005). Markmann & Tautz (2005) found about 74% sequence conservation in the D3-D5 expansion segments, and a few highly variable regions which provided enough information for distinguishing between closely related taxa. Therefore, I tested the effectiveness of the D3 expansion segment of the 28S rDNA as a genetic barcode to identify species of freshwater calanoid and cyclopoid copepods at all life history stages.

CHAPTER III

RAPID IDENTIFICATION OF NEW ENGLAND FRESHWATER COPEPODS USING A NOVEL GENETIC BARCODE

Introduction

Genetic barcodes rely on nucleotide variations in short segments of DNA (about 300-600 bp) to distinguish among species of a particular taxon (Waugh, 2007; Hebert et al., 2003; Floyd et al., 2002). They have the potential to expedite and verify routine species determinations, especially of samples that are difficult to identify using traditional morphology-based approaches (Hebert et al., 2003). The use of genetic barcodes for delineating species boundaries and identifying species is currently under debate (Lee, 2004; DeSalle, 2005; Meyer & Paulay, 2005; Hebert et al., 2004a; Wiemers & Fiedler, 2007). Despite this fact, molecular approaches to species identifications may be useful in cases where traditional morphology-based identifications are difficult or even impossible.

Current species identifications of freshwater calanoid and cyclopoid copepods are based mostly on the morphology of mature reproductive structures. These structures may exhibit only slight differences among sister taxa and often rely on one adult sexual form only (e. g., male calanoids and

female cyclopoids), necessitating the presence of specific individuals in a sample. As a result, studies of copepod species composition in lake communities may be skewed due to misidentifications. Not only are adult copepods difficult to distinguish but most morphology-based keys exclude immature stages. The relatively long life cycle of copepods (e. g., 68 days for *Aglaodiaptomus clavipes* SCHACHT; Gehrs & Robertson, 1975) often result in the collection of only immature specimens at time of sampling, again producing incomplete data sets.

Furthermore, copepods have been used as indicator species for the health of lake systems (Gerten & Adrian, 2002; Waiseng et al., 2003; Yan et al., 2004) in which case correct species identifications are critical. For example, Binks et al. (2005) have shown that species of similar morphologies (*Leptodiaptomus minutus* LILLJEBORG vs. *Skistodiaptomus oregonensis* LILLJEBORG) actually have different ecological requirement. Specifically, *L. minutus* is an acid-sensitive species, whereas *S. oregonensis* exhibits acid tolerance (Binks et al., 2005).

The original barcode proposed by Hebert et al. (2003) focuses on about 650 base pairs of the 5'-end of the cytochrome oxidase I (CO-I) gene. Since then, several studies have questioned the universal utility of CO-I and alternate genes have been proposed (Markmann & Tautz, 2005; Vences et al., 2005; Roe & Sperling, 2007). A pilot study revealed that CO-I may not be an appropriate choice for freshwater copepods (M. Litvaitis and K. Rawlinson, unpublished data), a conclusion upheld by Thum & Derry (2008) and Guarnieri (1996). Hence, our primary objective was to test the utility of the D3 expansion segment of the 28S rDNA (~350 bp) gene for species identifications in two orders of freshwater

copepods. This segment has been used successfully to distinguish species in other invertebrate groups (Litvaitis et al., 1994; Markmann & Tautz, 2005).

A major criticism of using molecular barcodes centers on sample size and thus, the range of variation that should be represented to evaluate the barcode (Moritz & Cicero, 2004). In an attempt to offset this short-coming, we included at least three populations of each species with five individuals representing each population (with some exceptions). This approach provides a more thorough test of intraspecific variation within and among populations than previous barcode studies. Because all samples were collected from New England (Northeast United States) lakes, we maximized the geographic range by using samples from lakes that were most distant from each other. However, this sampling strategy still represents a limited range for *Epischura* spp. FORBES, the Diaptomidae, and the Cyclopoida which are present in other geographic regions (Boxshall & Defaye, 2008). Additionally, immature copepods (eggs, nauplii) that were offspring of identified mothers were included in the analysis to confirm the utility of the genetic barcode for all life stages.

Materials and Methods

Sample Collection and Taxonomic Identification

Five species of Calanoida, representing four genera and two families, and five species of Cyclopoida distributed among four genera and one family, were collected from 27 lakes or ponds throughout New England (Table 1). To assure that we captured the intra-specific range of variation, five specimens of each

species were used from at least three populations (except for *Tropocyclops* spp). All samples were collected by vertical tow net in 2006 and 2007 either by the Environmental Protection Agency (EPA) Region 1 as part of the Regional Environmental Monitoring and Assessment program (REMAP) or by members of the University of New Hampshire Center for Freshwater Biology (UNH CFB). Samples were preserved in =95% undenatured ethanol and stored at room temperature.

Table 1. Taxonomic and locality information for samples used in this study.

Taxon		No of Indiv.	Lake	Location	Georeferences	Year Collected
Calanoida						
Diaptomidae	*Aglaodiaptomus spatulocrenatus* PEARCE	5	Granite Lake	Munsonville, NH	43.0215269, -72.1423662	2006
		5	Lake Machias	T5 ND BPP, ME	45.1271600, -68.0067100	2007
		5	Great East Lake	Acton, ME	43.5818360, -70.9461220	2006
		3	Goose Pond	Tyringham, MA	42.2830070, -73.1960350	2007
	Leptodiaptomus minutus LILLJEBORG	5	Lake Hayward	East Haddam, CT	41.5210794, -72.3289410	2006
		5	Granite Lake	Munsonville, NH	43.0215269, -72.1423662	2006
		5	Alexander Lake	Dayville, CT	41.8605197, -73.1960350	2006
		5	Great East Lake	Acton, ME	43.5818360, -70.9461220	2006 & 2007
	Skistodiaptomus pygmeus PEARCE	5	Lake Whalom	Lunenburg, MA	42.5742210, -71.7403380	2006
		3	Goose Pond	Tyringham, MA	42.2830070, -73.1960350	2007
		5	Watchoag Pond	Charlestown, RI	41.3842638, -71.6901586	2006
		5	Norton Reservoir	Norton,MA	41.9844420, -71.1935130	2007
		6	Baboosic Lake	Amherst, NH	42.8843000, -71.5790340	2006 & 2007
		2	Mendum's Pond	Dover, NH	43.1733330, -71.0641670	2006
Temoridae	*Epischura lacustris* S.A. FORBES	5	Silver Lake	Hollis, NH	42.7576600, -71.5983300	2006
		5	Granite Lake	Munsonville, NH	43.0215269, -72.1423662	2006
		5	Otter Pond	T8R5 WELS, ME	46.3747460, -68.3561350	2007
	Epischura nordenskioldi LILLJEBORG	5	Lake Mattawa	Orange, MA	42.5665070, -72.3221270	2006
		5	Lake Hayward	East Haddam, CT	41.5210794, -72.3289410	2006
		5	Alexander Lake	Dayville, CT	41.8605197, -73.1960350	2006
Cyclopoida						
Cyclopidae	*Cyclops scutifer* SARS	1	Figure Eight Pond	North Augusta, ME	44.3934445, -69.8245264	2006
		5	Long Lake	St. Agatha, ME	47.2130000, -68.1456200	2007
		5	4th Debs	T1R11 WELS, ME	45.7510900, -69.0826000	2007
		6	Island Pond	T15 R9 WELS, ME	46.9533800, -68.8417100	2007
	Diacyclops thomasi S.A. FORBES	5	East Twin Lake	Salisbury, CT	42.0264448, -73.3849578	2006
		5	Lake Whalom	Lunenburg, MA	42.5742210, -71.7403380	2006
		5	Beach Pond	Exeter,RI	41.5790000, -71.4729000	2007
		5	Long Pond	Plymouth, MA	41.8593840, -70.6048130	2006
	Mesocyclops edax S.A. FORBES	5	Lake Hayward	East Haddam, CT	41.5210794, -72.3289410	2006
		5	Harvey Lake	Northwood, NH	43.2151000, -71.2067000	2007
		5	Pleasant Lake	Topsfield, ME	45.3594480, -67.9195050	2007
		5	Forest Lake/Nelson Pond	S. Woodbury, VT	44.4084683, -72.4417809	2006
		5	Lake Whalom	Lunenburg, MA	42.5742210, -71.7403380	2006
		5	Keoka Lake	Waterford, ME	44.1798324, -70.7060482	2006
		2	Silver Lake	Barnard, VT	43.7292857, -72.6113220	2006
		2	Figure Eight Pond	North Augusta, ME	44.3934445, -69.8245264	2006
		7	Old Durham Reservoir	Durham, NH	43.1850320, -70.9299290	2007
	Tropocyclops jerseyensis KIEFER	2	Silver Lake	Hollis, NH	42.7576600, -71.5983300	2006
	Tropocyclops prasinus FISCHER	5	Silver Lake	Barnard, VT	43.7292857, -72.6113220	2006

Individual adult calanoids and female adult cyclopoids were identified based on morphology (Hudson et al., 2003; Lesko et al., 2003; Reid, 1991; UNH CFB On-line Identification Key, 2007; Wilson, 1959; Yeatman, 1959). Digital images were acquired of the whole organism and of taxonomically informative structures. Immature stages (eggs, nauplii) of known parentage were obtained by collecting eggs from identified mature females and nauplii hatched from these eggs were cultured in 5 ml filtered well water at 20°C. Specimens were fed every 3 d with 1 X 10^5 cells ml^{-1} of *Nannochloropsis* spp. Only one cyclopoid nauplius reached the copepodid stage and no calanoid nauplii survived. All samples, including eggs, nauplii, copepodid, and adults were placed individually into microfuge tubes containing 5% Chelex (Sigma-Aldrich, St. Louis, MO) and stored at -20°C until DNA extraction.

Molecular Protocols

Following the addition of proteinase K (Fermentas Life Sciences, Glen Burnie, MD) to each thawed sample (1U per 500 µl of Chelex), the samples were incubated overnight at 55° C. Five microliters of genomic DNA was used in 50 µl PCR reactions, containing sterile water, 1x Taq buffer or GoTaq colorless buffer, 1.5 mM $MgCl_2$, 0.2 mM nucleotides, 1 µM of each primer, and 0.3 U/µl GoTaq Polymerase (all reagents from Promega Inc. Madison, WI). The primers target the D3 expansion segment of the 28S rDNA, a fragment of about 350 base pairs (Litvaitis et al., 1994). PCR conditions were as follows: 2 min pre-

PCR incubation at 96° C, followed by 30 cycles of 96°C for 30 sec, 50°C for 1 min, and 72°C for 2 min with a final extension at 72°C for 4 min.

Two different protocols were used for amplicon purification. Amplicons were visualized on a 1% SeaKem (Lonza, Rockland, Maine) agarose gel and DNA was recovered using Costar Spin-X columns (Corning Inc, Corning, New York). Alternatively, amplified DNA was gel-purified in a 1% SeaPlaque (Lonza, Rockland, Maine) agarose gel, followed by excision of the PCR products from the gel and by agarase digestion (Litvaitis & Litvaitis, 1996). Because PCR products were less than 400 bp in length, amplicons were sequenced using the forward primer only (Geneway, Hayward, CA).

Sequence Analysis

Chromatographs were edited using FinchTV v.1.3.1 (Geospiza Inc). Sequences have been submitted to NCBI GenBank and are available under accession numbers EU598277-EU598447. Sequences were aligned using ClustalW (Thompson et al., 1994) as implemented in Geneious Pro v. 3.8.5 (Biomatters Ltd, 2005-2008), and improved by eye. Pairwise comparisons were calculated between all species using PAUP* (Swofford, 2002).

Two neighbor-joining trees (NJ trees; Saitou & Nei, 1987) employing the logarithmic determinant (LogDet)/paralinear model (Lake, 1994; Lockhart et al., 1994), were calculated in PAUP * (Swofford, 2002) and re-sampled with 1000 bootstrap replicates (Felsenstein, 1985). Both trees were rooted using three individuals of a more distantly related crustacean order: two specimens of *Daphnia dubia* HERRICK and one specimen of *D. cawtaba* COKER. The first NJ tree

was based on all adult copepod samples with the exclusion of multiple identical sequences of conspecifics of the same population. The second NJ tree included all samples of the first tree plus immature specimens of known parentage. Frequencies of intra- and interspecific variations were determined for each order and intra- and interpopulation variations were determined for each species. According to Rach et al. (2008), character states that are unique to species can be used as characteristic attributes (CAs) to separate taxa. We identified pure, simple CAs (all organisms within the taxa have a certain nucleotide at a single position) at the ordinal, familial, and generic levels from a ClustalW alignment of consensus sequences of each species.

Results

A total of 90 adult calanoid and 87 adult cyclopoid specimens were isolated from lake samples and sequenced (for species, see Table 1). These samples represented five individuals of each species from three populations each (for exceptions to number of individuals or populations, see Table 1). In addition, one calanoid egg, two calanoid nauplii, two cyclopoid eggs and one cyclopoid nauplius of known parentage were included.

A neighbor-joining tree (NJ) of adult calanoid and cyclopoid samples revealed a clear separation with 100% bootstrap support of the two orders (Fig. 8). This was also reflected in a mean inter-ordinal variation of 19.7% $\pm$ 0.0108 SD (Table 3). Furthermore all genera with the exception of *Leptodiaptomus* were supported with high bootstrap values (85-100%). All specimens clustered within

their respective species, indicating complete congruence between morphological and molecular identifications.

Within Calanoida, the two species of *Epischura* (Temoridae) were clearly distinguished from each other (Fig. 8). Temoridae formed a sister clade to Diaptomidae with 100% bootstrap support (Fig. 8), and a mean inter-familial variation of 7.36% $\pm$ 0.802% SD (Table 3). Within Cyclopoida, separation of *Tropocyclops jerseyensis* and *T. prasinus* was supported by 100% bootstrap value (Fig. 8). The remaining three species *(Mesocyclops edax, Diacyclops thomasi, Cyclops scutifer)* all formed monophyletic clades with high bootstrap support (Fig. 8).

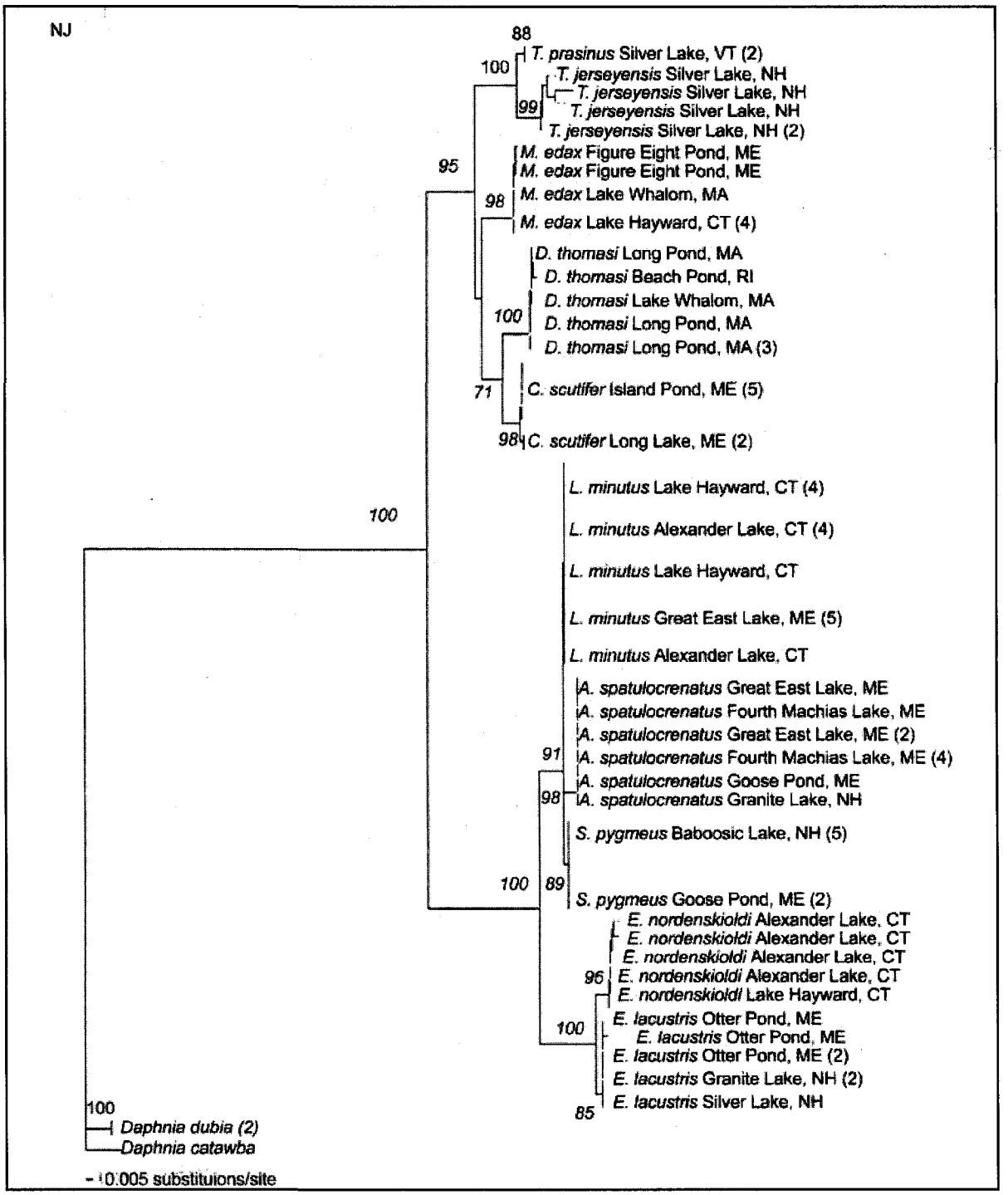

Figure 8. Neighbor-joining tree of 75 adult calanoid and cyclopoid copepods. Numbers at nodes represent bootstrap values. Number in parentheses denotes number of individuals with identical sequences from the same population.

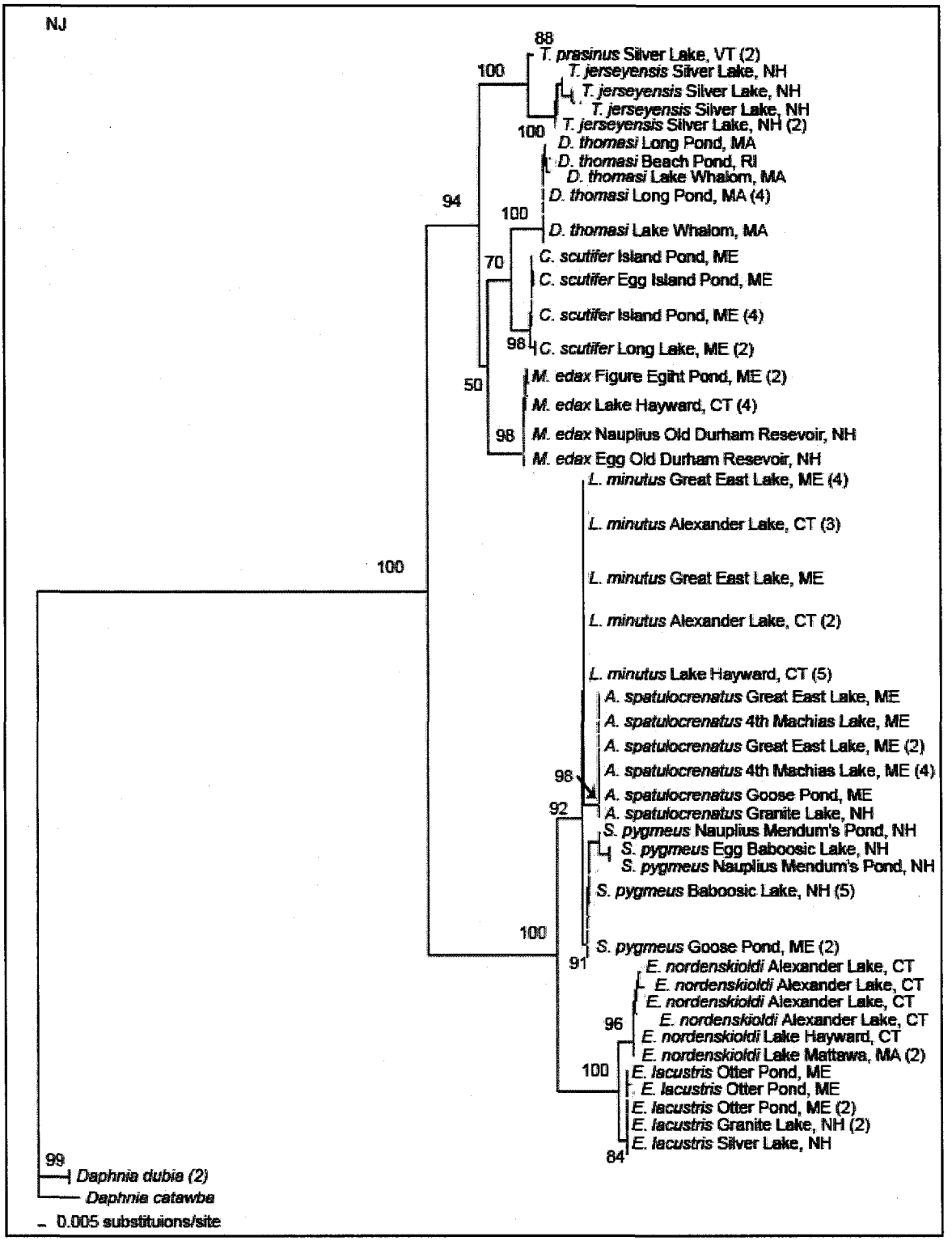

Figure 9. Neighbor-joining tree of 75 adult calanoid and cyclopoid copepods plus six immature stages of known parentage. Numbers at nodes represent bootstrap values. Number in parentheses denotes number of individuals with identical sequences from the same population.

Table 2. Intraspecific variation for all species included in this study.

Species	Mean	SE	SD	Range	No Individuals per Comparison
Calanoida					
Aglaodiaptomus spatulocrenatus	0.000	0.000	0.000	0.000-0.000	17
Leptodiaptomus minutus	0.000	0.000	0.000	0.000-0.000	20
Skistodiaptomus pygmeus	0.000	0.000	0.000	0.000-0.000	23
Epischura lacustris	0.000411	0.000103	0.00105	0.000-0.00308	15
Epischura nordenskioldi	0.000	0.000	0.000	0.000-0.000	15
Cyclopoida					
Cyclops scutifer	0.000943	0.000157	0.00172	0.000-0.00409	16
Diacyclops thomasi	0.000	0.000	0.000	0.000-0.000	20
Mesocyclops edax	0.000	0.000	0.000	0.000-0.000	39
Tropocyclops jerseyensis	0.00137	0.000558	0.00176	0.000-0.00342	5
Tropocyclops prasinus	0.000	NA	NA	0.000-0.000	2

Table 3. Interspecific variation among species of the same family, inter-familial variation within Calanoida, and inter-ordinal variation between Calanoida and Cyclopoida.

Comparison	Mean	SE	SD	Range
Family Diaptomidae				
Leptodiaptomus minutus vs. *Skistodiaptomus pygmeus*	0.00399	0.000	0.000	0.00399-0.00399
Aglaodiaptomus spatulocrenatus vs. *Leptodiaptomus minutus*	0.0108	0.00000220	0.0000405	0.0108-0.0109
Aglaodiaptomus spatulocrenatus vs. *Skistodiaptomus pygmeus*	0.0149	0.0000109	0.000216	0.0108-0.0151
Family Temoridae				
Epischura lacustris vs. *E. nordenskioldi*	0.0159	0.0000609	0.000909	0.0156-0.0207
Family Cyclopidae				
Tropocyclops jerseyensis vs. *T. prasinus*	0.0257	0.00105	0.00332	0.0239-0.0322
Cyclops scutifer vs. *Tropocyclops jerseyensis*	0.0766	0.000550	0.00492	0.0725-0.0899
Cyclops scutifer vs. *Tropocyclops prasinus*	0.0634	0.000494	0.00280	0.0604-0.0695
Mesocyclops edax vs. *Tropocyclops jerseyensis*	0.0835	0.000349	0.00488	0.0741-0.0939
Mesocyclops edax vs. *Tropocyclops prasinus*	0.0722	0.000289	0.00255	0.0679-0.0820
Mesocyclops edax vs. *Cyclops scutifer*	0.0701	0.0000916	0.00229	0.0692-0.0855
Diacyclops thomasi vs. *Tropocyclops jerseyensis*	0.0985	0.000616	0.00616	0.0944-0.110
Diacyclops thomasi vs. *Tropocyclops prasinus*	0.0751	0.000400	0.00253	0.0726-0.0778
Diacyclops thomasi vs. *Cyclops scutifer*	0.0362	0.0000866	0.00155	0.0356-0.0448
Diacyclops thomasi vs. *Mesocyclops edax*	0.0464	0.0000414	0.00116	0.0454-0.0536
Inter-familial				
Temoridae and Diaptomidae	0.0736	0.000189	0.00802	0.0156-0.0976
Inter-ordinal				
Calanoida and Cyclopoida	0.197	0.000126	0.0108	0.0699-0.259

A NJ tree consisting of adult specimens and of immature stages of *Cyclops scutifer, Mesocyclops edax,* and *Skistodiaptomus pygmeus* unequivocally clustered eggs and nauplii with their correct species (Fig. 9). The addition of immature specimens did not change the topology of the tree. Furthermore, value of bootstrap support for the three species containing immature samples either remained the same *(C. scutifer, M. edax)* or increased slightly *(S. pygmeus).*

A frequency diagram comparing intra- and interspecific variation within Calanoida revealed a small gap between 0.00308-0.00399 (Fig. 10). With the exception of *Epischura lacustris*, intraspecific variation was 0 for all calanoids (Table 2). Mean inter-specific variation within the order ranged from 0.4% to 1.6% with the largest value found between the two congeneric species of *Epischura* (1.6% $\pm$ 0.091% SD; Table 3). A distinct barcoding gap (0.00409-0.0239) was found in a comparison of intra- vs. interspecific variation within Cyclopoida (Fig. 11). Mean interspecific variation ranged from 2.57% to 8.35% within the order (Table 3). The congeneric *Tropocyclops* were separated by the lowest interspecific variation (2.57% $\pm$ 0.332% SD; Table 3). Finally, no geographic structuring among populations within lakes was evident.

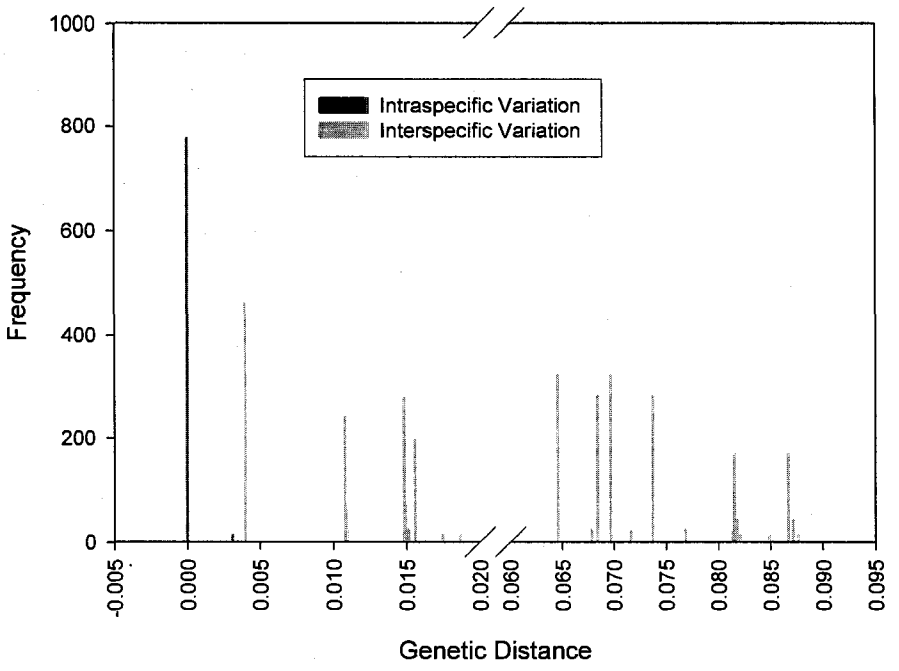

Figure 10. Frequency diagram of intra- versus interspecific variation in freshwater Calanoida.

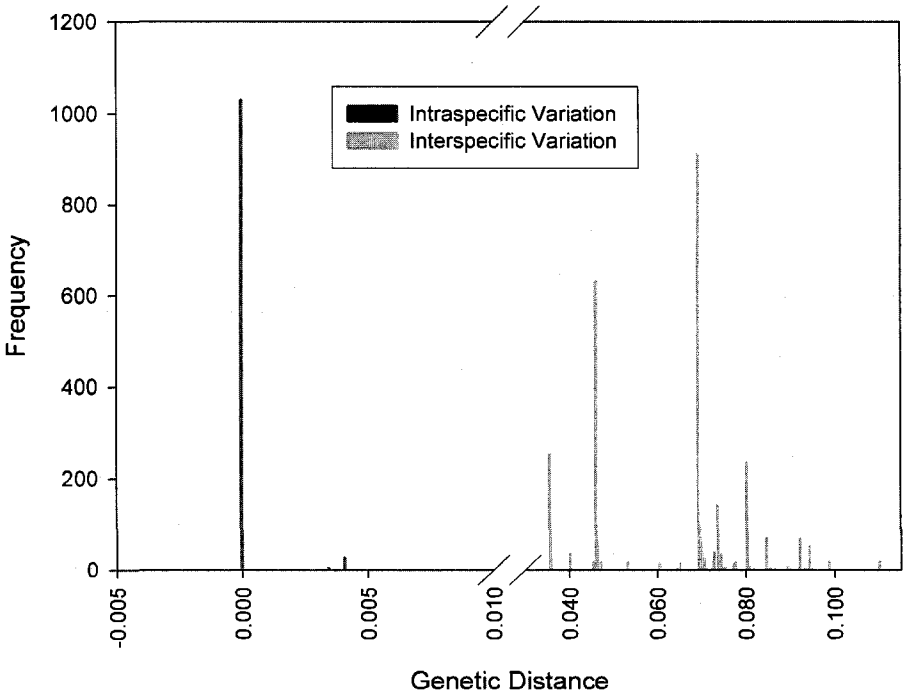

Figure 11. Frequency diagram of intra- versus interspecific variation in freshwater Cyclopoida.

Calanoida and Cyclopoida clearly separated at the following 24 order-unique CAs (numbering follows the 28S rDNA gene of *Mus musculus;* first base Calanoida, second base Cyclopoida): 1178, 1305: T→G; 1195: M→G; 1211, 1292, 1310, 1463: A→G; 1217: T→A; 1226, 1248, 1286: C→G; 1227: A→T; 1229: C→T; 1247: C or G→T; 1282, 1315: C→A ; 1306, 1314: G→C; 1307: G→A; 1311, 1326, 1366: T→C; 1302: T or deletion→C; 1324: R→ C (Table 4). Within Calanoida, Diaptomidae and Temoridae separate at the following family specific CAs (first base, Diaptomidae, second base Temoridae): 1192: T→C; 1193, 1293: A→T; 1195: C→A; 1247, 1313: C→G; 1250, 1281: G→C; 1269, 1320, 1360, 1381: C→T; 1302: T→deletion; 1376: G→A. Generic separation within Diaptomidae occurred at positions 1249 and 1300 (Table 4). Specific separation was possible for *Epischura lacustris* and *E. nordenskioldi* at position 1216.

Within Cyclopoida, *Tropocyclops* was distinguishable from the remaining genera at the following positions (*Tropocyclops* first base, other cyclopoids second base): 1194: G→A; 1200; C→ T; 1321, 1463: A→ G; 1325: A→ C; 1366: T→ C (Table 4). Generic separation was possible for *Mesocyclops edax* from *Cyclops scutifer* and *Diacyclops thomasi* at 1214: A→G; 1230: T→C; 1231, 1371: C→T; 1300: T→A; 1386: G→A, and for *C. scutifer* from *D. thomasi* at positions 1240 A→G and 1274: T→C. A species-specific CA was found for *Tropocyclops prasinus* and *T. jerseyensis* at position 1249 (Table 4).

Table 4. Characteristic attributes (CAs) of freshwater copepods. CAs that separate orders are in bold, CAs that separate families are italicized, and generic CAs are underlined.

Taxon	**1 1 7 8**	**1 1 7 9**	**1 1 9 2**	**1 1 9 3**	**1 1 9 4**	**1 1 9 5**	**1 2 0 0**	**1 2 1 1**	**1 2 1 4**	**1 2 1 6**	**1 2 1 7**	**1 2 2 6**	**1 2 2 7**	**1 2 2 9**	**1 2 3 0**	**1 2 3 1**	**1 2 4 0**	**1 2 4 7**	**1 2 4 8**	**1 2 4 9**	**1 2 5 0**	**1 2 6 9**	**1 2 7 4**	**1 2 8 1**	**1 2 8 2**
Calanoida																									
A. spatulocrenatus	**T**	T	*T*	*A*	G	*C*	C	**A**	A	T	**T**	**C**	**A**	**C**	T	C	A	*C*	**G**	T	*G*	*C*	T	*G*	**A**
L. minutus	**T**	T	*T*	*A*	G	*C*	C	**A**	A	T	**T**	**C**	**A**	**C**	T	C	A	*C*	**G**	C	*G*	*C*	T	*G*	**A**
S. pygmeus	**T**	T	*T*	*A*	G	*C*	C	**A**	A	T	**T**	**C**	**A**	**C**	T	C	A	*C*	**G**	C	*G*	*C*	T	*G*	**A**
E .lacustris	**T**	T	*C*	*T*	G	*A*	C	**A**	A	T	**T**	**C**	**A**	**C**	T	C	A	*G*	**G**	C	*C*	*T*	T	*C*	**A**
E. nordenskioldi	**T**	T	*C*	*T*	G	*A*	C	**A**	A	C	**T**	**C**	**A**	**C**	T	C	A	*G*	**G**	C	*C*	*T*	T	*C*	**A**
Cyclopoida																									
C. scutifer	**G**	C	*C*	*A*	A	*G*	T	**G**	G	C	**A**	**G**	**C**	**T**	C	T	A	*T*	**C**	C	*C*	*C*	T	*G*	**C**
D. thomasi	**G**	C	*C*	*A*	A	*G*	T	**G**	G	C	**A**	**G**	**C**	**T**	C	T	G	*T*	**C**	C	*C*	*C*	C	*G*	**C**
M. edax	**G**	C	*C*	*A*	A	*G*	T	**G**	A	C	**A**	**G**	**C**	**T**	T	C	G	*T*	**C**	C	*C*	*C*	C	*G*	**C**
T. jerseyensis	**G**	T	*C*	*A*	G	*G*	C	**G**	G	C	**A**	**G**	**C**	**T**	C	C	A	*T*	**C**	T	*C*	*C*	T	*G*	**C**
T. prasinus	**G**	T	*C*	*A*	G	*G*	C	**G**	G	C	**A**	**G**	**C**	**T**	C	C	A	*T*	**C**	C	*C*	*C*	T	*G*	**C**

Table 4. Continued.

Taxon	1286	1292	1293	1299	1300	1302	1305	1307	1310	1311	1313	1314	1315	1320	1321	1324	1325	1326	1360	1366	1371	1376	1381	1386	1463
Calanoida																									
A. spatulocrenatus	**C**	**A**	*A*	C	T	*T*	**T**	**G**	**A**	**T**	*C*	**G**	**C**	*C*	A	**R**	A	**T**	*C*	T	T	*G*	*C*	A	A
L. minutus	**C**	**A**	*A*	C	A	*T*	**T**	**G**	**A**	**T**	*C*	**G**	**C**	*C*	A	**R**	A	**T**	*C*	T	T	*G*	*C*	A	A
S. pygmeus	**C**	**A**	*A*	C	A	*T*	**T**	**G**	**A**	**T**	*C*	**G**	**C**	*C*	A	**R**	A	**T**	*C*	T	T	*G*	*C*	A	A
E .lacustris	**C**	**A**	*T*	T	A	-	**T**	**G**	**A**	**T**	*G*	**G**	**C**	*T*	A	**R**	A	**T**	*T*	T	T	*A*	*T*	A	A
E. nordenskioldi	**C**	**A**	*T*	T	A	-	**T**	**G**	**A**	**T**	*G*	**G**	**C**	*T*	A	**R**	A	**T**	*T*	T	T	*A*	*T*	A	A
Cyclopoida																									
C. scutifer	**G**	**G**	*A*	C	A	*C*	**G**	**A**	**G**	**C**	*G*	**C**	**A**	*C*	A	**C**	A	**C**	*C*	T	T	*G*	*C*	A	A
D. thomasi	**G**	**G**	*A*	T	A	*C*	**G**	**A**	**G**	**C**	*G*	**C**	**A**	*C*	A	**C**	A	**C**	*C*	T	T	*G*	*C*	A	A
M. edax	**G**	**G**	*A*	C	T	*C*	**G**	**A**	**G**	**C**	*G*	**C**	**A**	*C*	A	**C**	A	**C**	*C*	T	C	*G*	*C*	G	A
T. jerseyensis	**G**	**G**	*A*	C	T	*C*	**G**	**A**	**G**	**C**	*G*	**C**	**A**	*C*	G	**C**	C	**C**	*C*	C	C	*G*	*C*	G	G
T. prasinus	**G**	**G**	*A*	C	T	*C*	**G**	**A**	**G**	**C**	*G*	**C**	**A**	*C*	G	**C**	C	**C**	*C*	C	C	*G*	*C*	G	G

Discussion

In a revolutionary approach to species identifications, Hebert et al. (2003) proposed the use of 650 bp of the mitochondrial CO-I gene as an alternative to traditional morphology-based identification methods. Since then numerous studies, both in support and in opposition of the use of CO-I as a DNA barcode have been published (Lee, 2004; DeSalle, 2005; Meyer & Paulay, 2005; Hebert et al., 2004a; Hebert et al., 2004b; Wiemers & Fiedler, 2007). Although many objections to DNA barcoding focus on its validity in species discovery and delineation of species boundaries, there also is a methodological obstacle that has been limiting many attempts at using CO-I for barcoding. Specifically, for many taxa the universal primers designed by Folmer et al. (1994) do not find target sequences during the annealing step of PCR and hence, no products are amplified. In the absence of truly universal primers for CO-I of all metazoans, the promised ease and speed of DNA barcoding clearly is not existent, and circumventing the problem with re-designing primers (for some taxa even multiple re-designed primers) defies the purpose of the approach.

Although DNA barcoding using CO-I has been successfully employed with marine copepods (Bucklin et al., 1999; Hill et al., 2001; Bucklin et al., 2003), the highly variable third codon position prevents successful amplification in freshwater calanoids and cyclopoids. Therefore, we tested the utility of the D3 expansion segment of the 28S rDNA gene as an alternative barcode for the identification of freshwater copepods in New England lakes. Previous studies using the D1-D2 (Sonnenberg et al., 2007) and the D3-D5 expansion segments

(Markman & Tautz, 2005) have shown the superior performance of ribosomal sequences for species identifications. Our results clearly confirm these earlier findings (Fig. 8). Furthermore, we also have shown that species assignments were possible for immature stages (eggs, nauplii) and for partial specimens (antennae) (Fig. 9).

For both orders, we determined a barcoding gap between intra- and interspecific variation (Figs. 10 and 11), an indication that the D3 expansion segment is able to differentiate species. Bucklin et al. (2003) found intra-specific variations of CO-I sequences ranged between 1% and 2% for most marine calanoids examined in their study. We found that intra-specific variation for all species was always less than 1%. This is due to the more conserved rate of evolution of ribosomal genes. Thum (2004) also noted rapid evolution of CO-I in diaptomid copepods and provided evidence that relationships within this family can be resolved using 18S rDNA. However, the 18S rDNA is too conserved for species-specific resolution (Thum, 2004).

In a recent study, Thum & Derry (2008) compared mitochondrial sequences of four morphospecies belonging to the Diaptomidae. One of these, *Leptodiaptomus minutus* corresponds to one of our species, and two others *(Skistodiaptomus pallidus, S. reighardi* MARSH*)* are congeners of *S. pygmeus,* included in our study. Similarly to our findings, Thum & Derry (2008) recorded low genetic divergence (1.1%) among *L. minutus* haplotypes. *Skistodiaptomus pallidus* on the other hand, exhibited an extremely high intraspecific variation of 16.7% among four deeply branching, geographically isolated clades. The authors

suggest that each clade may actually represent a cryptic species, a fact further supported by intra-clade variations ranging from 0.1 to 1.5% (Thum & Derry, 2008). In comparison, intra-specific variation using the D3 expansion segment was zero for *L. minutus* and *S. pygmeus,* and we found no evidence of sequence divergences in geographically disparate populations of either species. We recognize that our study was geographically limited to New England and inclusion of additional samples may increase intra-specific variation in our study species. However, this comparison demonstrates the need for the careful selection of a gene or gene segment to distinguish taxa at specific levels, i. e., the D3 segment is appropriate for species-level but not for population-level separations, whereas CO-I which evolves more rapidly is useful for population-level distinctions for species whose DNA can be reliably amplified by CO-I primers.

Interestingly, we recovered the lowest interspecific variation (~0.4%) between two species of diaptomid copepods, *Leptodiaptomus minutus* and *Skistodiaptomus pygmeus* (Table 3). A value that is considerably lower than the intra-specific variation among the different clades of *S. pallidus* (Thum & Derry, 2008). Therefore, it appears that CO-I evolution in freshwater copepods is accelerated. This observation, combined with the failure of the universal primers used for the amplification of this gene (Folmer et al., 1994), supports our preference for the use of the more conserved D3 expansion segment for species identifications.

The identification of characteristic attributes may ultimately form the basis for species separations using restriction enzyme digests (Billiones et al., 2004) or multiplex species-specific PCR (Bucklin et al., 1999; Hill et al., 2001). However, we recognize that the D3 expansion segment may not be of sufficient length to provide size fragments and priming sites that unequivocally allow for the resolution of all species. We therefore, recommend that a longer segment, possibly D3-D5, be used in future studies. The length of this segment will increase sequencing cost as forward and reverse reactions will need to be run, but this cost will then be equal with barcoding cost for CO-I. Additionally, this segment may not be able to readily amplify partially degraded sequences from ill-preserved or older specimens as readily as the D3 segment alone, but may be more accurate at distinguishing species among closely related taxa. This portion of the 28S rDNA gene already has been employed successfully for the separation of benthic meiofauna (Markman & Tautz, 2005).

Rapid and accurate species identifications in lakes are crucial for an improved understanding of species interactions and their trophic relationships. Knowledge of species composition assists lake managers in their monitoring programs and may allow for the potential identification of bioindicator species. In this way, barcoding can be a tool for lake environmental studies. It may be possible to use genetic barcodes to determine shifts in copepod populations over time by extracting copepod DNA from lake sediment (Bisset et al., 2005). Another possibility involve the use of species specific primers to detect the presence of certain species (especially parasitic or invasive taxa) in lakes by

extracting DNA from water samples (Ficetola et al., 2008). Another important implication for copepod barcoding is to determine species of copepods from gut contents of fish predators and with the development of algal and bacterial barcodes it may be possible to determine species-specific food selection of copepods. A similar study has been conducted determining the gut contents of Antarctic krill, *Euphausia superba* DANA (Passmore et al., 2006). We recognize that our study was limited to New England, however, our results are of value because lake ecologists often rely on regional rather than global sampling. Hence, genetic barcoding using the D3 expansion segment has the potential to become an additional valuable tool in limnological studies.

LIST OF REFERENCES

Abdullahi, B.A., 1990. The effect of temperature on reproduction in three species of cyclopoid copepods. Hydrobiologia 196: 101-109.

Andersen, T. & D.O. Hessen, 1991. Carbon, nitrogen, and phosphorous content of freshwater zooplankton. Limnology and Oceanography 36: 807-814.

Anderson, R.S., 1967. Diaptomid copepods from two mountain ponds in Alberta. Canadian Journal of Zoology 45: 1043-1047.

Bartholmé, S., L. Samchyshyna, B. Santer & W. Lampert, 2005. Subitaneous eggs of freshwater copepods pass through fish guts: survival, hatchability, and potential ecological implications. Limnology and Oceanography 50: 923-929.

Benasasson, D., D. Zhang, D.L. Hartl & G.M. Hewitt, 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. TRENDS in Ecology and Evolution 16: 314-321.

Bhadury, P., M.C. Austen, D.T. Bilton, P.J.D. Lambshead, A.D. Rogers & G.R. Smerdon, 2006. Development and evaluation of a DNA-barcoding approach for rapid identification of nematodes. Marine Ecology Progress Series 320: 1-9.

Billiones, R., M. Brehm, J. Klee & K. Schwenk, 2004. Genetic identification of *Hyalodaphnia* species and interspecific hybrids. Hydrobiologia 526: 43-53.

Binks, J.A., S.E. Arnott & W.G. Sprules, 2005. Local factors and colonist dispersal influence crustacean zooplankton recovery from cultural acidification. Ecological Applications 15: 2025-2036.

Bisset, A., J.A.E. Gibson, S.N. Jarman, K.M. Swadling & L. Cromer, 2005. Isolation, amplification, and identification of ancient copepod DNA from lake sediments. Limnology and Oceanography: Methods 3: 533-542.

Blaxter, M.L., 2004. The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society B 359: 669-679.

Boxshall, G.A. & D. Defaye, 2008. Global diversity of copepods (Crustacea: Copepoda) in freshwater. Hydrobiologia 595: 195-207.

Boxshall, G.A. & D. Jaume, 2000. Making waves: the repeated colonization of freshwater water by copepod crustaceans. Advances in Ecological Research 31: 61-79.

Boxshall, G. A. & S. H. Halsey, 2004. An Introduction to Copepod Diversity. The London Series. London, The Ray Society. 166 part 1 and 2.

Boxshall, G.A. & E. E. Strong, 2006. An extraordinary shirt in life habit within a genus of cyclopoid copepods in Lake Tanganyika. Zoological Journal of the Linnean Society 146: 275-285.

Brewer, R.H., 1964. The phenology of *Diaptomus stagnalis* (Copepoda: Calanoida): the development and the hatching of the egg stage. Physiological Zoology 37: 1-20.

Bucklin, A., M. Guarnieri, R.S. Hill, A.M. Bentley & S. Kaartvedt, 1999. Taxonomic and systematic assessment of planktonic copepods using mitochondrial CO-I sequence variation and competitive, species-specific PCR. Hydrobiologia 401: 239-254.

Bucklin, A., B. W. Frost, J. Bradford-Grieve, L. D. Allen & N. J. Copley, 2003. Molecular systematics and phylogenetic assessment of 34 calanoid copepod species of the Calanoidae and Clausocalanoida. Marine Biology 142: 333-343.

Burns & Schallenberg, 2001. Calanoid copepods versus cladocerans: Consumer effects on protozoa in lakes of different trophic status. Limnology and Oceanography 46; 1558-1565.

Clark, C.G., B.W. Tague, V.C. Ware & S.A. Gerbi, 1984. *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications. Nucleic Acids Research 12: 6197-6220.

Cole, G.A., 1994. Textbook of Limnology. 4th edition. Waveland Press, Inc, Long Grove, IL: 44-46, 94.

Cox, A.J. & P.D.N. Hebert, 2001. Colonization, extinction, and phylogeographic patterning in a freshwater crustacean. Molecular Ecology 10: 371-386.

Dayrat, B., 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society 85: 407-415.

DeSalle, R., M.G. Egan & M. Siddall, 2005. The unholy trinity: taxonomy, species delimitation, and DNA barcoding. Philosophical Transactions of the Royal Society B 360: 1905-1916.

DeStasio, B.T., 2004. Diapause in calanoid copepods: within-clutch hatching patterns. Journal of Limnology 63: 26-31.

Dowell, K.M., 1997. Evidence of diapause in the freshwater copepod *Skistodiaptomus pallidus*. American Midland Naturalist 137: 362-368.

Erpenbeck, D., J.N.A. Hooper & G. Wörheide, 2006. CO-I phylogenies in diploblasts and the 'Barcoding of Life'-are we sequencing a suboptimal partition? Molecular Ecology Notes 6: 550-553.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Ficetola, G.F., C. Miaud, F. Pompanon & P. Taberlet, 2008. Species detection using environmental DNA from water samples. Biology Letters 4: 423-425.

Floyd, R., E. Abebe, A. Papert & M. Blaxter, 2002. Molecular barcodes for soil nematode identification. Molecular Ecology 11: 839-850.

Folmer, O., M. Black, W. Hoech, R. Lutz & R. Vrijenhoek, 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294-299.

Fryer, G., 1954. Contributions to our knowledge of the biology and systematics of the freshwater Copepoda. Hydrobiologia 16: 64-77.

Funk, D.J. & K.E. Omland, 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 34:397-423.

Gehrs, C.W. & A. Robertson. 1975. Use of life tables in analyzing the dynamics of copepod populations. Ecology 56: 665-672.

Gerten, D. & R. Adrian, 2002. Species-specific changes in the phenology and peak abundance of freshwater copepods in response to warm summers. Freshwater Biology 47: 2163-2173.

Guarnieri, M., 1996. A preliminary study of genus variation within and among populations of *Diaptomus leptopus* (Copepoda: Calanoida): a DNA sequence analysis of the mitochondrial cytochrome oxidase I gene. MS Thesis, Concordia University.

Hairston, N.G. Jr. & W.R. Munns Jr., 1984. The timing of copepod diapause as an evolutionary stable strategy. The American Naturalist 123: 733-751.

Hairston, N.G. Jr., T.A. Dillon & B.T. DeStasio Jr., 1990. A field test for the cues of diapause in a freshwater copepod. Ecology 71: 2218-2223.

Hajibabaei, M., M.A. Smith, D.H. Janzen, J.J. Rodriguez, J.B. Whitefield & P.D.N. Hebert, 2006. A minimalist barcode can identify a specimen whose DNA is degraded. Molecular Ecology Notes 6: 959-964.

Hammer, U.T. & W.W. Sawchyn, 1968. Seasonal success and cogeneric associations of *Diaptomus* spp. (Copepoda) in some Saskatchewan ponds. Limnology and Oceanography 13: 476-484.

Hassouna, N., B. Michot & J.-P. Bachellerie, 1984. The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Research 12: 3563-3583.

Hebert, P.D.N., A. Cywinksa, S.L. Ball & J.R. deWaard, 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society London B 270: 313-321.

Hebert, P.D.N. & T.R. Gregory, 2005. The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852-859.

Hebert, P.D.N., M.Y. Stoeckle, T.S. Zemlak & C.M. Francis, 2004a. Identification of birds through DNA barcodes. Public Library of Science Biology 2:1657-1662.

Hebert, P.D.N., E.H. Penton, J.M.Burns, D.H. Janzen & W. Hallwachs. 2004b. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Science 101:14812-14817.

Hill, R. S., L. D. Allen & A. Bucklin, 2001. Multiplexed species-specific PCR protocol to discriminate four N. Atlantic *Calanus* species, with a mtCO-I gene tree for ten *Calanus* species. Marine Biology 139: 279-287.

Hillis, D.M. & M.T. Dixon, 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. The Quarterly Review of Biology 66: 411-446.

Hudson, P. L., L. T. Lesko, J. W. Reid & M. A. Chriscinske, 2003. Cyclopoid copepods of the Laurantian Great Lakes. Great Lakes Science Center Home Page http://www.glsc.usgs.gov/greatlakescopepods/Key.asp?GROUP=Cyclopoid

Jamieson, C.D. & B. Santer, 2003. Maternal aging in the univoltine freshwater copepod *Cyclops kolensis:* variation in egg sizes, egg development times, and naupliar development times. Hydrobiologia 510:75-81.

Lafontaine, D.L. & D. Tollervey, 2001. The function and synthesis of ribosomes. Nature Reviews Molecular Cell Biology 2:514-520.

Lake, J.A., 1994. Reconstructing evolutionary trees from DNA protein sequences: paralinear distances. Proceedings of the National Academy of Science USA 91: 1455-1459.

Lee, M.S.Y., 2004. The molecularisation of taxonomy. Invertebrate Systematics 18:1-6.

Lesko, L. T., P. L. Hudson & M. A. Chriscinske, 2003. Calanoid copepods of the Laurentian Great Lakes. Great Lakes Science Center Home Page http://glsc.usgs.gov/greatlakescopepods/Key.asp?GROUP=Calanoid

Litvaitis, M.K. & J.A. Litvaitis, 1996. Using mitochondrial DNA to inventory distribution of remnant populations of New England cottontails. Wildlife Society Bulletin 24: 725-730.

Litvaitis, M.K., G. Nunn, W.K. Thomas & T.D. Kocher, 1994. A molecular approach for the identification of meiofaunal turbellarians (Platyhelminthes, Turbellaria). Marine Biology 120: 437-442.

Lockhart, P.J., M.A. Steel, M.D. Hendy & D. Penny, 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Molecular Biology and Evolution 11: 605-612.

Lohner, L.M., N.G. Hairston JR. & W.R. Schaffner, 1990. A method for distinguishing subitaneous and diapausing eggs in preserved samples of the Calanoid Copepod genus Diaptomus. Limnology and Oceanography 35: 763-767.

Long, E.O. & I. B. Dawid, 1980. Repeated genes in eukaryotes. Annual Review of Biochemistry 49: 727-764.

Lord, A., 2005. DNA barcoding a useful tool for taxonomists. Nature 435: 17.

Lowndes, A.G., 1929. *Cyclops americanus* MARSH. A discussion and description of its specific characteristics and its occurrence in Europe. International Review of Hydrobiologia 19: 12-20.

Maier, G., 1996. Copepod communities in lakes of varying trophic degree. Archiv für Hydrobiologie 136: 455-465.

Markmann, M. & D. Tautz, 2005. Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences. Philosophical Transactions of the Royal Society B 360: 1917-1924.

Marshall, E., 2005. Will DNA bar codes breathe life into classification? Science 307: 1037.

Melão, M.D.G.G. & O. Rocha, 2004. Life history, biomass and production of two planktonic cyclopoid copepods in a shallow subtropical reservoir. Journal of Planktonic Research 26: 909-923.

Meyer, C.P. & G. Paulay, 2005. DNA barcoding: error rates based on comprehensive sampling. Public Library of Science 3: 2229-2238.

Michot, B., N. Hassouna & J-P. Bachellerie, 1984. Secondary structure of mouse 28S rRNA and general model for the folding of the large rRNA in eukaryotes. Nucleic Acids Research 12: 4259-4279.

Moritz, C. & C. Cicero, 2004. DNA barcoding: promise and pitfalls. Public Library of Science Biology 2: 1529-1531.

Nielsen, R. & M. Matz, 2006. Statistical approaches for DNA barcoding. Systematic Biology 55: 162-169.

Palumbi, S.R., 1996. Nucleic acid II: the polymerase chain reaction. In Hillis, D.M., C. Moritz, & B.K. Mable (eds), Molecular Systematics. 2nd ed. Sinauer Associates, Inc., Sunderland MA: 494-495.

Passmore, A.J., S.N. Jarman, K.M. Swadling, S. Kawaguchi, A. McMinn & S.N. Nicol, 2006. DNA as a dietary biomarker in Antarctic krill, *Euphausia superba*. Marine Biotechnology 8: 686-696.

Rach, J., R. DeSalle, I.N. Sarkar, B. Schierwater & H. Hadrys, 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proceedings of the Royal Society B 275: 237-247.

Reid, J.W., 1991. Some species of *Tropocyclops* (Crustacea, Copepoda) from Brazil, with a key to the American species. Bijdragen tot de Dierkunde 61: 3-15.

Roe, A.D. & F.A.H. Sperling, 2007. Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications in DNA barcoding. Molecular Phylogenetics and Evolution 44: 325-245.

Saccone, C., C. DeGiorgi, C. Gissi, G. Pesole & A. Reyes, 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Genes 238: 195-209.

Saitou, N. & M. Nei, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4: 406-425.

Sandercock, G.A., 1967. A study of selected mechanism for the coexistence of *Diaptomus* spp. in Clark Lake, Ontario. Limnology and Oceanography 12: 97-112.

Santer, B. & W. Lampert, 1995. Summer diapause in cyclopoid copepods: adaptive response to a food bottleneck? Journal of American Ecology 64: 600-613.

Sonnenberg, R., A. W. Nolte & D. Tautz, 2007. An evaluation of LSU rDNA D1-D2 sequences for their use in species identification. Frontiers in Zoology 4: doi:10.1186/1742-9994-4-6

Steinke, D., M. Vences, W. Salzburger & A. Meyer, 2005. Tax1: a software tool for DNA barcoding using distance methods. Philosophical Transactions of the Royal Society B 360: 1975-1980.

Swofford, D. L., 2002. Phylogenetic analysis using parsimony (*and other methods), v. 4. Sinauer Associates, Sunderland, Massachusetts.

Swofford, D.L., G.J. Olsen, P.J. Waddell & D.M. Hillis, 1996. Phylogenetic Inference. In Hillis, D.M., C. Moritz, & B.K. Mable (eds), Molecular Systematics. 2nd ed. Sinauer Associates, Inc., Sunderland MA: 494-495.

Tautz, D., P. Arctander, A. Minelli, R.H. Thomas & A.P. Vogler, 2003. A plea for DNA taxonomy. TRENDS in Ecology and Evolution 18: 70-74.

Thompson, J.D., D.G. Higgins & T.J. Gibson, 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673-4680.

Thum, R. A., 2004. Using 18S rDNA to resolve diaptomid copepod (Copepoda: Calanoida: Diaptomidae) phylogeny: an example with the North American genera. Hydrobiologia 519: 135-141.

Thum, R.A. & A. M. Derry, 2008. Taxonomic implications for diaptomid copeods based on contrasting patterns of mitochondrial DNA sequence divergences in four morphospecies. Hydrobiologia 614: 197-207.

Thum, R.A. & R. S. Stemberger, 2006. Pure spatial and spatially structured environmental variables explain *Skistodiaptomus* copepod range limits in the northeastern USA. Canadian Journal of Fisheries and Aquatic Sciences 63: 1397-1404.

University of New Hampshire Center for Freshwater Biology, 2007. Image-Based Key to the Zooplankton of the Northeast (USA) v.2.0. University of New Hampshire Center for Freshwater Biology http://cfb.unh.edu/CFBkey/html/index.html

Vences, M., M. Thomas, R.M. Bonnet & D.R. Vieites, 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philosophical Transactions of the Royal Society B 360: 1859-1868.

Waiseng, B., N.D. Yan & A.K. Schartau, 2003. Littoral microcrustacean (Cladocera and Copepoda) indicators of acidification in Canadian shield lakes. Ambio 32: 208-213.

Waugh, J. 2007. DNA barcoding in animal species: progress, potential, and pitfalls. BioEssays 29: 188-197.

Wetzel, R.G., 1975. Limnology. W.B. Saunders Company, Philadelphia PA: 458-465.

Whiles, M.R., J.B. Jensen, J. G. Palis & W. G. Dyer, 2004. Diets of larval flatwoods salamanders, *Ambyostoma cingulatum,* from Florida and South Carolina. Journal of Herpetology. 38: 208-214.

Wiemers, M. & K. Fiedler, 2007. Does the DNA barcoding gap exist?-a case study in blue butterflies (Lepidoptera:Lycaenidae). Frontiers in Zoology 4: doi:10.1186/1742-9994-4-8.

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.

Wilson, M. S., 1959. Free-living Copepoda: Calanoida. In Edmondson, W. T. (ed), Freshwater Biology. John Wiley and Sons, Inc., New York: 738-794.

Yan, N.D., R. Girard, J.H. Heneberry, W.B. Keller, J.M. Gunn & P.J. Dillon, 2004. Recovery of copepod, but not cladoceran, zooplankton from severe and chronic effects of multiple stressors. Ecology Letters 7: 452-460.

Yeatman, H. C., 1959. Free-living Copepoda: Cyclopoida. In Edmondson, W. T. (ed), Freshwater Biology. John Wiley and Sons, Inc., New York: 795-814.