

*HARMOTHOE IMBRICATA*: SPECIES COMPLEX OR COMPLEX SPECIES?

By  
Angela Gastaldi, B.S.

A Thesis Submitted in Partial Fulfillment of the Requirement  
for the Degree of

Master of Science  
in  
Marine Biology

University of Alaska Fairbanks  
May 2019

APPROVED:

J. Andres Lopez, Committee Co-Chair  
Sarah Hardy, Committee Co-Chair  
Amanda Kelley, Committee Member  
Derek Sikes, Committee Member  
Matthew Wooller, Chair  
*Department of Marine Biology*  
Bradley Moran, Dean  
*College of Fisheries and Ocean Sciences*  
Michael Castellini, *Dean of the Graduate School*

## ABSTRACT

Accurate estimates of species diversity are constrained by cryptic species complexes, in which multiple closely related species are grouped under a single species name due to the absence of clear morphological differences. Cryptic diversity is known to be prevalent in polychaete worms, a mostly marine group commonly known as bristle worms. A recent survey of polychaete diversity discovered that the widespread scale-worm *Harmothoe imbricata* comprises multiple distinct mitochondrial lineages based on analysis of the Cytochrome c oxidase I (COI) gene, which is often referred to as the ‘barcoding’ gene. Analyses based solely on DNA sequences from COI may overestimate the number of lineages comprising a cryptic species complex, so it has been recommended that cryptic species investigations incorporate nuclear gene sequences. The goal of this study was to determine whether the incorporation of DNA sequences from the nuclear genome corroborates the designation of *H. imbricata* as a cryptic species complex. I sequenced segments of COI and five nuclear genes: ITS1, ITS2, H3, and portions of the 18S and 28S genes of *H. imbricata* and analyzed them using distance measures, maximum likelihood, and Bayesian inference. I compared phylogenetic trees produced from mitochondrial and nuclear DNA sequences, as well as from a combined mitochondrial/nuclear dataset. *Harmothoe imbricata* was found to include five mitochondrial lineages, whereas the nuclear sequences only supported four well-defined lineages. These results corroborate previous reports showing COI-based cryptic species investigations find more lineages than nuclear DNA based investigations. These results provide additional lines of evidence that *H. imbricata* is a cryptic species complex. These divergent lineages likely arose after being separated during the last glacial maximum but they are now found in sympatry. A thorough morphological study of *H. imbricata* populations may reveal phenotypic differences correlated with the genetic lineages identified here.



## Table of Contents

	Page
Abstract.....	i
Table of Contents.....	iii
List of Figures.....	iv
List of Tables.....	iv
General Introduction.....	1
Chapter 1.....	10
Abstract .....	10
Introduction .....	11
Materials and Methods .....	13
Results.....	17
Discussion.....	20
General Conclusion.....	26
Figures and Tables.....	29
References.....	39

## List of Figures

	Page
Figure 1: Map of sample distribution .....	29
Figure 2: Phylogenetic tree of the <i>Harmothoe imbricata</i> species complex (ML).....	30
Figure 3: Phylogenetic tree of the <i>Harmothoe imbricata</i> species complex (BI) .....	31
Figure 4: Phylogenetic tree of the <i>Harmothoe imbricata</i> species complex (BI/ML).....	32
Figure 5: Haplotype network of ITS1 and ITS2 regions.....	33

## List of Tables

	Page
Table 1: List of samples used in analyses.....	34
Table 2: COI and ITS K2P values between <i>H. imbricata</i> lineages and outgroup .....	38

## GENERAL INTRODUCTION

Biodiversity plays a vital role in ecosystem functioning and resilience, as different species within a community perform various important ecological roles (Duarte 2000). Global biodiversity is in decline and a baseline assessment of extant species is needed to better detect and predict the consequences of future diversity loss (Butchart et al. 2010). Current estimates of global biodiversity vary greatly. Approximately 1.8 million species have been described and this number increases as efforts to find and describe new species continues (Roskov et al. 2018). Estimates of global biodiversity range from less than 2 million (Costello et al. 2012) to over 6 billion species (Larsen et al. 2017). Estimates of marine biodiversity are similarly variable; approximately 250,000 marine species have been described, while estimates of total species richness range from 300,000 (Costello et al. 2012) to 10 million (Grassle et al. 1992).

Estimates of species richness are hindered by cryptic species complexes, which occur when multiple species are considered to be a single species based on morphological similarity. Cryptic species have been identified across a wide array of taxa and biogeographic regions (Pfenninger and Schwenk 2007), although cryptic speciation may be more prevalent in tropical rainforests and marine biomes (Bickford et al. 2007). Cryptic diversity is common in polychaete worms (phylum Annelida; e.g. Schmidt and Westheide 2000, Bleidorn et al. 2006). Many polychaetes have been described as having cosmopolitan distributions, yet many of these wide ranging species are likely cryptic species complexes (Nygren 2014). A recent effort to document polychaete diversity in Canadian waters examined nearly 2000 specimens representing 142 morphologically identified species. Thirty-four of these morphospecies were composed of multiple genetically distinct lineages, suggesting that they may represent cryptic species complexes (Carr 2010). Additionally, nearly all the species previously described as cosmopolitan

contained multiple lineages with enough genetic divergence to support classification as provisional species (Carr et al. 2011). In this study, I further investigate patterns of genetic variation of one such polychaete.

### *Polychaete diversity*

Polychaetes are an abundant, diverse, and ecologically important class of annelid worms found in all marine habitats, especially benthic environments. Over 10,000 species of polychaetes have been described (Hutchings and Fauchald 2000) from more than 80 families (Fauchald and Rouse 1997). Polychaetes show a variety of feeding types including herbivory, carnivory, suspension feeding, and deposit feeding. Multiple feeding types are commonly documented within a single family, genus, or even species (Jumars et al. 2015). Polychaete reproductive strategies are similarly diverse (Rouse and Pleijel 2006). While the majority of polychaetes are broadcast spawners, brooding and encapsulation also occur; developmental modes include planktotrophic, lecithotrophic and direct development (Wilson 1991). Due to the wide variety of feeding and reproductive strategies exhibited within the class, polychaetes perform many functions in the ecosystems they inhabit (Snelgrove 1997).

Polychaete systematics has been based primarily on external morphology (Fauchald 1974, Fauchald and Rouse 1997). Among annelids, polychaetes have been distinguished from Clitellata (i.e. earthworms and leeches), the other major annelid group, since the middle of the 19<sup>th</sup> century (Grube 1850). Clitellates are distinguishable by their clitellum, a cocoon-forming segment that is used in reproduction, in addition to having reduced or non-existent chaetae. Polychaeta are distinguishable by their parapodia containing chitinous chaetae and are divided into two groups: Errantia and Sedentaria. This division was initially based on a difference in

lifestyle and not on evolutionary relationships (Fauchald 1974). Errantia grouped the free-living and generally predatory forms, while Sedentaria comprised sessile, often tube-dwelling forms.

The first major morphology-based cladistics analysis of polychaetes disregarded the Errantia/Sedentaria classifications and instead divided Polychaeta into two clades: Palpata and Scolecida (Rouse and Fauchald 1997). Palpata was further subdivided into two clades: Canalipalpata, which included mainly sessile and burrowing deposit feeders, and Aciculata, which mainly included active foragers and predators, and closely resembled the group previously referred to as Errantia. Scolecida was characterized by the lack of head appendages and palps and mostly included burrowers and deposit feeders. Echiura (spoon worms) and Sipuncula (peanut worms) were considered sister phyla to Annelida (Rouse and Fauchald 1997).

Evidence from genetics has yielded new insights into polychaete and annelid systematics and evolution over the past several decades. Early molecular phylogenetic analyses generally relied on single gene sequences, such as 18S rRNA and elongation factor 1 $\alpha$ , and included only a small representation of annelid diversity (ex: Winnepennickx et al. 1995, McHugh 1997, Winnepennickx et al. 1998). Although these analyses proved inadequate for resolving relationships among polychaete families, they did provide the first line of evidence that Echiura and Clitellata are nested within Polychaeta (McHugh 2000). Sampling and sequencing efforts increased over time producing support for relationships within many annelid families, but the relationships among families were still poorly supported. Additionally, Annelida was consistently found to be paraphyletic with respect to other non-annelid lophotrochozoans. Analysis of a large-scale transcriptome dataset split Annelida into two large groups, which were reclassified as Errantia and Sedentaria because the family groups placed within each clade closely resembled their original members (Struck et al. 2011).

High-throughput sequencing techniques invented and refined over the past few decades have greatly reduced cost and increased rate of sequencing DNA. Two large molecular datasets were produced using next-generation sequencing (Weigert et al. 2014; Struck et al. 2015). Combined, these datasets consisted of 679 orthologous genes with over 189,000 amino acid positions from 80 annelid species. Weigert and Bleidorn (2016) used these data to produce a comprehensive phylogeny of annelids which suggests that Annelida is a large, monophyletic group composed of the sister clades Errantia and Sedentaria, as well as five basal groups. Most of the groups previously referred to as polychaetes fall within Errantia and Sedentaria, with a few exceptions. Clitellata and Echiura are firmly nested within Sedentaria, while Sipuncula is a basal group. The other four basal groups, Magelonidae, Chaetopteridae, Oweniidae, and Amphinomidae, were all previously considered families within Polychaeta. Although these analyses show that Polychaeta is a paraphyletic group, the term remains in use to describe marine segmented worms.

Aphroditiformia is a particularly diverse and successful group of polychaetes that are distinguished by two rows of scales, also referred to as elytra, along their dorsal surface, giving them the common name of scale-worms. Over two-thirds of all scale-worm species are placed in the family Polynoidae (WoRMS 2018). All scale-worm genera were initially grouped in the family Aphroditidae, but over time different groups of species were transferred to separate families based on diagnostic traits; the remaining genera were reclassified as family Polynoidae, which lack unifying characteristics and are unlikely to represent a monophyletic group (Fauchald and Rouse 1997). A recent combined molecular and morphological cladistics analysis recovered Polynoidae as monophyletic, although this analysis only included representatives from 8 of the 20 recognized subfamilies within this group (Gonzalez 2018). This study also revealed the genus

*Harmothoe* to be paraphyletic with respect to other genera within the family Polynoidae, including: *Eunoe*, *Gattyana*, *Lepidonotus*, and *Bylgides*. *Harmothoe* recently underwent a major taxonomic revision, which resulted in the delineation of 28 valid species from a starting set of 126 described species (Barnich and Fiege 2009).

#### *Harmothoe imbricata*- *A cryptic species complex*?

The overarching goal of this study was to examine cryptic diversity within the polychaete scale-worm *Harmothoe imbricata* uncovered by earlier work based on mitochondrial COI gene sequences (Carr et al. 2011, Hardy et al. 2011). This species is widely distributed across the Arctic, North Pacific, North Atlantic, North Sea, Mediterranean Sea, and Baltic Sea. It inhabits the intertidal zone down to 300m, free-living or living commensally with tube dwelling polychaetes. Despite extensive research attention, *H. imbricata* remains one of the most taxonomically problematic polychaetes. It was originally described by Carl Linnaeus in 1776, but his description was not detailed enough to distinguish between multiple species of *Harmothoe*. The type specimen was thought to reside in Sweden, but the type material is deemed lost (Barnich and Fiege 2009). Based on a redescription by Malmgren in 1866, a neotype was designated in 2009 from a sample collected off Iceland which was deposited in the Swedish Museum of Natural History in Stockholm, Sweden (Barnich and Fiege 2009).

A recent barcoding effort aimed at documenting polychaete distribution and diversity in Canadian waters revealed high levels of genetic diversity within *H. imbricata* (Carr 2010). Based on mitochondrial cytochrome c oxidase subunit 1 (COI) DNA sequences alone, *H. imbricata* was classified as a cryptic species complex containing six provisional species (i.e. unconfirmed species), which were named *H. imbricata* CMC01 – *H. imbricata* CMC06. The mean Kimura 2-

paramater (K2P) pairwise distance between specimens, over the full set of *H. imbricata* specimens sequenced, was 11.6%, while that between the most divergent lineages identified exceeded 16% (Hardy et al. 2011). These values are well above expected and commonly seen K2P pairwise distances generally seen within species (Hebert et al. 2003b). Multiple provisional species occurred in sympatry, and some regions in the central Canadian Arctic contained five of the six provisional species. Subsequent analyses incorporated additional samples from regions surrounding Alaska, USA, and led to a new phylogenetic analysis which reduced the number of provisional species from six to five (C. Carr personal correspondence). Carr's (2010) sequences appear in GenBank with the updated provisional species designations (i.e. *Harmothoe imbricata* CMC01- *Harmothoe imbricata* CMC05).

I used DNA sequences from nuclear loci to test the hypothesis that *H. imbricata* is a cryptic species complex. The gene regions targeted were the two nuclear ribosomal internal transcribed spacers (ITS1 and ITS2), a partial sequence of the nuclear ribosomal 18S gene (also known as the nuclear small ribosomal subunit), a partial sequence of the nuclear ribosomal 28S gene (also known as the nuclear large ribosomal subunit), and Histone-3 (H3), in addition to COI. The selection of genes targeted in this study was constrained by the availability of amplification primers that match target sequences in the species of interest. COI and the ITS regions are broadly used in molecular systematics studies because primers with broad taxonomic targets were developed early. H3, 18S, and 28S have been sequenced in closely related taxa, so primers that work well with the species examined here were available.

The nuclear ribosomal 18S rRNA gene is a slowly evolving gene, which makes it a good candidate for studying distant evolutionary relationships (Hillis and Dixon 1991, Halanych and

Janosik 2006). The nuclear ribosomal 28S rRNA gene has regions that show more variability than the 18S gene and other regions that are more conserved, making it applicable for a broader range of phylogenetic questions than 18S (Halanych and Janosik 2006). The internal transcribed spacers are regions of the ribosomal DNA coding array that are spliced and discarded following transcription (Lafontaine and Tollervey 2001). Spacer sequences are able to accumulate substitutions more quickly than the surrounding coding regions presumably because they do not encode a functioning RNA product, making them useful for inferring phylogenies of closely related taxa (Hillis and Dixon 1991). The ITS regions are often used in conjunction with COI to infer phylogenies of closely related species and uncover cryptic species complexes (examples in Annelids: Gustafsson et al. 2009, Liu et al. 2017; examples in other taxa: Schuchert 2014, Calderon et al. 2006). Histone-3 is a protein coding gene that functions in DNA-binding and gene regulation (Bhasin et al. 2006). This gene has been shown to be fairly conserved at the amino acid level, but quite variable in the third codon position (Brown et al. 1999). H3 is often included in investigations of cryptic species in polychaetes (ex: Glasby et al. 2013, Nygren et al. 2009, Colgan et al. 2006, Achurra and Erseus 2013).

The COI gene is the most commonly sequenced mitochondrial gene in both vertebrates and invertebrate animals, partially because of its status as the ‘barcoding gene’ (Hebert et al. 2003b). This protein coding gene is useful for multiple levels of phylogenetic analyses; the first and second codon positions evolve fairly slowly due to constraints on COI critical function. In contrast, synonymous substitutions at the third codon position accumulate at the fast rate that characterizes mitochondrial DNA molecular evolution. These differences lead to the first and second codon position being useful for inferring genus and family level phylogenetic

relationships, while the third codon position is useful for studying closely related species (Halanych and Janosik 2006).

DNA barcoding based on the sequencing of the COI gene (Hebert et al. 2003a) can be used to identify species and infer relationships between populations (e.g. Bleidorn et al. 2006; Ruiz-García et al. 2015; Titus and Daly 2015). DNA barcoding is often used to reveal cryptic species complexes (e.g. Hebert et al. 2004, Carr et al. 2011), but relying solely on COI sequences can lead to overestimates of the number of lineages within cryptic species complexes (Liu et al. 2017). Although mitochondrial DNA (mtDNA) data has long been the standard in phylogeographic studies (Avice et al. 2016), delimiting species based solely on COI gene sequences has been widely criticized (e.g. Will and Rubinoff 2004, Collins and Cruickshank 2013). The biology and evolutionary pressures of mitochondrial genomes differ from the nuclear genome. Mitochondria are haploid, inherited maternally, and do not undergo recombination. The mitochondrial genome accumulates mutations approximately 10 times faster than the nuclear genome (Brown et al. 1979). These differences can cause significant bias when inferring evolutionary histories (Ballard and Whitlock 2004), so an accurate representation of the true evolutionary history of a group of organisms requires inclusion of nuclear DNA markers that evolve at an appropriate rate in addition to mtDNA (Hurst and Jiggins 2005, Rubinoff and Holland 2005).

I analyzed new and previously published nuclear and mitochondrial DNA sequences to test the hypothesis that *Harmothoe imbricata* is a cryptic species complex. I compare the phylogenies produced from COI, nuclear gene sequences, and a concatenated mitochondrial/nuclear dataset using distance measures, maximum likelihood, and Bayesian

phylogenetic inference methods. My results show that *Harmothoe imbricata* consists of multiple lineages, but the number of lineages and confidence of phylogenies differs depending on inference tool and gene regions analyzed.

## ***Chapter 1: Harmothoe imbricata: species complex or complex species?***

### **ABSTRACT**

The polychaete scale-worm *Harmothoe imbricata* was recently deemed a cryptic species complex composed of five morphologically identical yet genetically distinct lineages based on mitochondrial COI DNA sequences. It has been recommended that investigations of cryptic species include the analysis of nuclear genes, as mitochondrial genes alone are more likely than nuclear genes to overestimate the number of lineages in a cryptic species complex.

Consequently, I sequenced five regions from the nuclear genome of *H. imbricata*: ITS1, ITS2, H3, and portions of the 18S and 28S genes in addition to the mitochondrial COI gene. Multiple sequence alignments of the COI sequences, the concatenated nuclear sequences, and the concatenated mitochondrial/nuclear sequences were analyzed using distance measures, maximum likelihood and Bayesian inference. The resulting phylogenetic trees were examined to identify areas of disagreement between the genomes and to determine the degree of support for cryptic species identified solely on the basis of COI evidence. Haplotype networks were used to visualize patterns of nucleotide variation in the ITS regions. The *H. imbricata* individuals examined were found to include the five mitochondrial lineages previously reported, however nuclear gene sequence variation only supports the existence of four distinct nuclear lineages. For both datasets, the degree of confidence varied depending on the analysis. This work provides further evidence that *H. imbricata* is a cryptic species complex and that COI DNA sequences support more lineages than inferred from nuclear DNA. The morphology of this species should be further investigated to identify phenotypic variation that correlates with the genetic boundaries apparent in the DNA sequences.

## INTRODUCTION

Estimates of marine species diversity are impeded by cryptic species complexes, which occur when multiple separately evolving lineages are morphologically indistinguishable and thus considered a single species (Bickford et al. 2007, Struck et al. 2018). Although cryptic species are common in all taxa and biogeographic regions (Pfenninger and Schwenk 2007), marine biomes likely harbor a significant number of cryptic species (Bickford et al. 2007). Polychaete worm diversity (phylum Annelida) has been shown to harbor a significant amount of cryptic species (e.g. Schmidt and Westheide 2000, Bleidorn et al. 2006).

With over 10,000 species described (Hutchings and Fauchald 2000), polychaetes are a diverse and ecologically important group of marine invertebrates. Scale-worms, polychaete suborder Aphroditiformia (Fauchald 1977), are a particularly successful and diverse group of polychaetes that are distinguished by the two rows of scales along their dorsal surface. The aphroditiform family Polynoidae contains over two-thirds of all scale worm species (WoRMS 2018), although this family is likely not monophyletic (Fauchald and Rouse 1997). *Harmothoe*, a speciose genus of polynoids, recently underwent a major taxonomic revision that reduced the number of valid species from 126 to 28 (Barnich and Fiege 2009). This genus was also recently shown to be paraphyletic with respect to multiple other genera of polynoids (Gonzalez 2018).

*Harmothoe imbricata* is a widespread polynoid that inhabits the intertidal zone down to 300m and is found in the Arctic, North Pacific, North Atlantic, North Sea, Mediterranean Sea, and Baltic Sea. Despite being well-studied in regards to reproductive biology (Daly 1972; Daly 1974; Garwood 1980; Garwood 1981; Garwood and Olive 1982) and behavior (Gaudron and Bentley 2002; Gaudron et al. 2007), *H. imbricata* remains one of the most taxonomically

problematic polychaetes (Barnich and Fiege 2009). A recent study used the barcoding gene, cytochrome c oxidase subunit 1 gene (COI), to assess polychaete diversity and distribution patterns in Canadian waters and found high levels of genetic diversity within *H. imbricata* (Carr 2010). Based on COI DNA sequences alone, *H. imbricata* was deemed a cryptic species complex consisting of six provisional species, which were given code names *H. imbricata CMC01* - *H. imbricata CMC06* (Carr et al. 2011). The mean K2P distance among all lineages was 11.6%, while the more deeply divergent lineages were over 16% (Hardy et al. 2011). Additional samples were collected from regions surrounding Alaska, USA, and further analyses reduced the number of provisional species from six to five (C. Carr personal correspondence). Some of the provisional species were found in multiple ocean basins, and in many regions multiple provisional species were found to occur in sympatry. It has been speculated that these lineages diverged in various glacial refugia during the Pleistocene glaciation, and the current overlapping distribution is the result of reproductive isolation following secondary contact (Hardy et al. 2011).

DNA barcoding (Hebert et al. 2003a) is often used to reveal cryptic species complexes (e.g. Hebert et al. 2004, Carr et al. 2011) but COI sequences alone are more likely than nuclear DNA to overestimate the number of lineages within cryptic species complexes (Liu et al. 2017). Nuclear DNA should be included in cryptic species investigations to generate more complete representation of the evolutionary history of the populations (Ballard and Whitlock 2004, Hurst and Jiggins 2005, Rubinoff and Holland 2005). Proposed cryptic species in *H. imbricata* were based solely on mtDNA; here we incorporate additional sequence data from nuclear genes to test the hypothesis that *H. imbricata* is a cryptic species complex.

## MATERIALS AND METHODS

### *Sampling*

A total of 193 specimens collected in 12 regions primarily surrounding Alaska and Canada, in addition to a small number of samples from Western Russia and Scandinavia, were examined in this study. The samples span a wide area, with the best represented regions being the Chukchi Sea, the Gulf of Alaska, and British Columbia (see Figure 1), although much of the described distribution of *Harmothoe imbricata* is unrepresented, including: the Western and Eastern coasts of the United States, the Western Canadian Arctic, the Beaufort Sea, Greenland, Iceland, and most of Europe and Russia. Of the total sample set, 38 individuals are new collections from the Chukchi Sea made during the Arctic Shelf Growth, Advection, Respiration and Deposition (ASGARD) and the Arctic Marine Biodiversity Observing Network (AMBON) cruises in June and August 2017, respectively. The samples from Scandinavia and Svalbard were obtained as a tissue grant from the Gothenburg Natural History Museum in Göteborg, Sweden and were part of the materials examined by Nygren et al. (2011). The samples from previous phylogenetic investigations of *H. imbricata* (Carr et al. 2011, and Hardy et al. 2011) were obtained from the Canadian Center for DNA Barcoding at the Biodiversity Institute of Ontario. A variety of other species of scale-worms were collected during ASGARD to use as potential outgroups. Additional samples were collected on the research vessels Oscar Dyson and Oshoru Maru from the Chukchi and Bering Seas in September 2007.

Field collected specimens were frozen in RNAlater or molecular grade ethanol. DNA was extracted from preserved tissue samples using a Gentra Puregene Tissue Kit, following the protocol listed for tissue (<https://www.qiagen.com/us/>). Samples provided by the Canadian Center for DNA Barcoding were obtained as DNA extracts.

### *Amplification and sequencing*

Six gene fragments, including one from the mitochondrial genome, were targeted for sequencing. The mitochondrial gene, Cytochrome c oxidase subunit I (COI), was amplified using primers LCO1490 and HCO2198 and reaction conditions described in Folmer et al. (1994). Internal transcribed spacers 1 and 2 (ITS1 and ITS2) were amplified and sequenced using primers described in Nygren et al. (2009). The D9 region of the 28S ribosomal DNA was amplified using primers described in Brown et al. (1999). H3 was amplified using primers from Colgan et al. (2000). Species-specific primers targeting a fragment of the 18S were designed using PRIMER3 (Untergasser et al. 2012) based on an initial set of 18S *H. imbricata* sequences produced with an unpublished primer set (H. Wiklund personal communication). The 18S species-specific primer sequences were: Forward- 5'-ACCTATCAAATGTCGATGGTAAGTG-3'; Reverse- 5'-ACCGAGGTCCTATTCTATTATTCCA-3'.

Reaction conditions for amplification of 28S-D9, H3, and 18S were: 1X GoTaq reaction buffer, 0.025 U/uL GoTaq DNA polymerase (Promega), 0.8mM dNTP's (0.2mM each), 3.5 mM MgCl<sub>2</sub>, 1uM each primer, 1 uL of template DNA in 25uL reactions incubated under the following temperature profile: 95°C for 3 min., 35 cycles of 95°C for 30 s, 50°C for 45 s and 72°C for 30 s, and a final elongation step of 72°C for 5 min. Reaction products were visualized on 2% agarose gels stained with 1X GelRed (Biotium). PCR products were treated with ExoSAP-IT (Thermo Fisher Scientific) to break down residual primers and unincorporated nucleotides. Purified amplification products were used as templates in Sanger sequencing reactions followed by automated sequence determination from Sanger reaction products. Sequencing reactions and sequence determinations were performed by Eurofins Genomics ([www.eurofinsgenomics.com](http://www.eurofinsgenomics.com)).

### *Data analyses*

Forward and reverse sequences from each sample were reviewed, aligned and edited using CodonCode Aligner v. 7.1.2. Sequences were inspected and edited for base-calling artifacts and low quality output. Sequences were also checked for evidence of heterozygosity. Low quality segments at the starts and ends of sequences were removed. All targeted fragments were sequenced in forward and reverse directions. After editing for quality, sequences were imported to Mesquite v. 3.31 (Maddison and Maddison 2018) to create multiple sequence alignments (MSA) for each gene using MUSCLE (Edgar 2004). For each gene, sequences that were significantly shorter than the others were removed, and remaining sequences were trimmed to match start and end positions. A full concatenated dataset of nuclear loci was produced from individuals sequenced at all loci. Another concatenated dataset was produced for ITS1 and ITS2. When available, COI sequences from *H. imbricata* individuals included in my sample set were downloaded from GenBank; a MSA was assembled from a combination of newly determined sequences and downloaded COI sequences. A final concatenated dataset was produced that included all 6 gene regions (combined mitochondrial and nuclear) for samples that had been sequenced for all loci.

Distance matrices were calculated in PAUP 4.0 (Swofford 2002) for the ITS1/ITS2 concatenation and the COI dataset using the Kimura 2-parameter (K2P) model with gaps being ignored to compare divergences among and within mitochondrial lineages. A neighbor-joining tree was calculated for the COI sequences to determine the mitochondrial lineage affinities of newly sequenced specimens following the methods of Carr et al. (2011). The resulting tree contained five distinct mitochondrial lineages. In this study, the mitochondrial lineages corresponding to Carr et al.'s (2011) CMC-types are referred to as HI1-HI5 to highlight that they

represent mitochondrial lineages rather than species. Provisional species designations of samples used in the previous study can be found in Table 1. Average K2P distances among and within mitochondrial lineages were calculated to determine if the patterns of divergence seen in COI correlated with those observed in the ITS regions.

Maximum likelihood (ML) and Bayesian inference (BI) were used to infer phylogenies. Three datasets were used for the ML and Bayesian analyses: the nuclear concatenation (i.e., ITS1, ITS2, H3, 18S, and 28S comprising 2302 nucleotide sites, with 25 sites being parsimony informative), the COI fragment, which spanned 575 sites, and a combined nuclear/mitochondrial concatenation (previously described nuclear concatenation combined with COI sequences). Nuclear and mitochondrial loci were analyzed separately and together so the inferred phylogenies could be compared. *Harmothoe extenuata* was used as an outgroup, as preliminary investigations found it to be the closest relative to *H. imbricata* of all species available. PartitionFinder2 (Lanfear et al. 2016, Lanfear et al. 2012) was used to find the best-fitting partitioning schemes and models of nucleotide evolution for these two datasets.

Bayesian phylogenetic inference was performed using MrBayes 3.2.6 (Ronquist and Huelsenbeck 2003) as implemented in CIPRES Science Gateway V. 3.3 (Miller et al. 2010). All Bayesian analyses were run for 50 million generations, with two independent runs of four chains each (one cold and three hot), a 25% relative burnin, and tree sampling every 1000 generations. BIC and AIC model selection criteria identified HKY and GTR+I, respectively, as best fits to the nuclear concatenation, both with a single partition spanning all sites. We performed Bayesian analyses using each of these two models. Similarly, model selection criteria identified two different models as best fits for the COI alignment. The AICc and BIC indicated the GTR+G model with a single partition spanning all sites, while the AIC supported the GTR+I for a first

and second codon positions and the GTR+I+G for third codon positions. We performed analyses under each model to evaluate sensitivity of results to model selection. Convergence was diagnosed for each analysis using Effective Sample Sizes (ESS) of key parameters as calculated by functions in the R package ‘RWTY’ (Warren et al. 2017).

Maximum likelihood phylogenetic inference and bootstrapping were performed using RAxML-HPC2 version 8.2.10 (Stamatakis 2014) on XSEDE (Towns et al. 2014) as implemented by CIPRES Science Gateway V. 3.3 (Miller et al. 2010). Rapid bootstrapping was conducted and allowed to continue until halted by RAxML using the GTR + G model, since RAxML advises against using the estimation of invariable sites option (+I).

The resulting tree files were visualized using the online portal Interactive Tree of Life (iTOL; Letunic and Bork 2016). Results generated by each method (distance, ML and BI) were compared to determine areas of agreement and disagreement between analyses and datasets.

Haplotype networks were used to visualize the extent of allele variation and degree of correspondence between allele relationships and putative species designations within *H. imbricata*. Haplotype networks of ITS1 and ITS2 were created in PopART (Leigh and Bryant 2015) using the median joining algorithm (Bandelt et al. 1999).

## RESULTS

Sequences were obtained from segments of 18S (90 individuals; 529 sites, 0 sites parsimony informative) 28S (78 individuals; 596 sites, 1 site parsimony informative), ITS1 (149 individuals; 472 sites, 13 sites parsimony informative), ITS2 (100 individuals; 400 sites, 5 sites parsimony informative), H3 (81 individuals; 333 sites, 7 sites parsimony informative), and COI (30 newly determined sequences combined with 72 publicly available sequences; 575 sites, 144

sites parsimony informative). COI, ITS, and H3 showed the most divergence (see Table 2); 18S and 28S showed little to no variation within the ingroup. Seventeen individuals were heterozygous at a single site in the H3 sequence and those sites were labeled as ambiguous for downstream analysis. Sequences for all targeted gene regions were obtained from 55 individuals. A concatenation of sequences from these specimens was used in subsequent phylogenetic analyses. This subset included the following numbers of specimens from each mitochondrial lineage: HI1, 24; HI2, 13; HI3, 6; HI4, 6; and HI5, 5. The concatenated nuclear dataset spanned 2302 nucleotide sites.

### *Phylogenetic analyses*

The K2P distances of both the concatenated ITS1/ITS2 and COI sequences from the five *H. imbricata* mitochondrial lineages and the outgroup *H. extenuata* can be found in Table 2. The K2P distances within the putative cryptic lineages ranged from 0.31 to 2.04% and 0 to 0.16% for COI and ITS, respectively. The K2P distances between putative cryptic lineages for COI and ITS ranged from 4.04 to 18.11% and 0.10 to 1.92%, respectively. For both regions, the greatest levels of divergence were seen between individuals that carried the HI3 and HI5 mitochondrial lineages. The range of interspecific pairwise K2P distances (i.e. between *H. extenuata* and all *H. imbricata* mitochondrial lineages) for COI and ITS were 16.32-18.75% and 7.41-8.63%, respectively.

Both Bayesian inference and maximum likelihood analyses of the COI sequences produced phylogenies that split *Harmothoe imbricata* into five clades, but the topologies differed between the analyses (See Figures 2 and 3). ML analysis of the mitochondrial dataset suggested

that the HI4 and HI5 mitochondrial lineages are sister taxa, although the support was weak (bootstrap (BS) support value of 57%), while BI places the HI5 mitochondrial lineage as being the sister taxon to a clade that includes all other mitochondrial lineages (PP- 100), and the HI4 lineage as sister taxa to the HI1/HI2/HI3 clade, with a posterior probability (PP) of 88. In both analyses, HI3 is a well-supported clade within the larger HI1/HI2/HI3 clade (BS- 88, PP- 97). HI2 is a poorly supported clade according to BI (PP- 50), and does not form a clade according to ML. HI1 forms a well-supported clade according to BI (PP- 96), but ML found lower support for this clade (BS- 61).

Although phylogenetic trees produced based on the analyses of the nuclear concatenation grouped individuals by mitochondrial lineages, not all mitochondrial lineages formed monophyletic groups, and the inferred distances between clades were substantially shorter than those based on the COI sequences (See Figures 2 and 3). Maximum likelihood and Bayesian inference produced similar topologies, but the levels of support differed for some clades. Both analyses of the nuclear concatenation split *Harmothoe imbricata* into four clades: one unique to individuals carrying the HI5 mitochondrial lineage (BS- 74, PP- 63), a second unique to HI4 carriers (BS- 81, PP- 80), and a third clade that included HI1, HI2, and HI3 carriers (BS- 66, PP- 99). Nested within this third clade, HI3 carriers form their own well-supported clade (BS- 88, PP- 100). The topology shows HI5 being the sister taxon to the remaining clades, and HI4 being the sister taxon to the HI1/HI2/HI3 clade. Analyses of the combined nuclear and mitochondrial concatenation produced phylogenetic trees most similar to trees produced using COI sequences alone, but like the nuclear DNA-only analysis, found support for four, rather than five clades (see Figure 4). These analyses do not add new evidence to support the mitochondrial lineages HI1 and HI2 as distinct clades.

### *Haplotype/allele networks*

We identified eleven ITS1 haplotypes and six ITS2 haplotypes. Median joining haplotype networks are depicted in Figure 5. For ITS1, the most common haplotype (42% of samples) was shared between HI1, HI2, and HI3 mitochondrial type carriers. Three ITS1 haplotypes were unique to HI1 carriers and they differed from the most common haplotype by 0.7-2.7% uncorrected proportion of site differences. Another haplotype was unique to HI4 carriers and was 0.7% different compared to the most common haplotype. Two haplotypes were unique to HI5 carriers and they differed from the haplotype unique to HI4 carriers by 2.7% and 3.4%. Two haplotypes were unique to HI3 carriers, and they split off between the main haplotype and the haplotypes unique to HI1 carriers. For ITS2, the main haplotype was shared between HI1, HI2, and HI3 carriers. Two haplotypes differed from the main haplotype by 1% each, and both were unique to HI1 carriers. Two other haplotypes differed from the main haplotype by 2% and 3%, and both were unique to HI5 carriers. The haplotype that is most divergent from the main haplotype was unique to HI4 carriers and varied from the most divergent haplotype unique to HI5 carriers by an additional 2%.

## **DISCUSSION**

Cryptic species diversity has gained research attention over the past two decades as it has become evident that a significant amount of biodiversity is hidden within cryptic species complexes. This increase in research attention is in large part due to increased access to genetic evidence to aid in the discovery of cryptic species, yet a large scale literature review of 606 studies investigating cryptic diversity found that 35.5% of these studies relied on only a single

molecular marker, most often the mitochondrial COI gene (Struck et al. 2018). Although high COI divergences do suggest the possibility of hidden biodiversity, it has been shown that analyses based on COI data are more likely to overestimate the number of lineages in a cryptic species complex (Liu et al. 2017), likely due to the fast rate of evolution of mtDNA. Additionally, deeply divergent COI sequences can be maintained in sympatry without reproductive isolation (Giska et al. 2015), so cryptic species investigations should also incorporate other forms of evidence, such as nuclear sequence data and phenotypic data (Ballard and Whitlock 2004, Rubinoff and Holland 2005). A recent survey of Canadian polychaetes found a surprising amount of genetic diversity in the mitochondrial COI gene of the polynoid *Harmothoe imbricata* and deemed it a cryptic species complex (Carr et al. 2011). My study aimed to determine whether the analysis of multiple nuclear genes corroborated the designation of *H. imbricata* as a cryptic species complex.

The new COI sequences generated here agree with previous findings that *Harmothoe imbricata* comprises 5 distinct and substantially divergent mitochondrial lineages based on a K2P distance analysis. Bayesian inference of COI sequence data similarly supported the designation of five mitochondrial lineages, while maximum likelihood bootstrapping offered less support for the mitochondrial lineages HI1 and HI2 as distinct clades. Both Bayesian inference and maximum likelihood analyses of the nuclear concatenation offered support that *H. imbricata* is composed of four lineages: HI3, HI4, and HI5 each form distinct clades and individuals from the HI1 and HI2 mitochondrial lineages form a single clade. The combined nuclear and mitochondrial analyses produced phylogenies similar to those based solely on COI sequences but like the nuclear analyses, support only four clades – mitochondrial lineages HI3, HI4, and HI5 as distinct clades, with the fourth paraphyletic clade composed of HI1 and HI2.

Alternative phylogenetic inference methods revealed broadly similar patterns; all analyses resolved HI3, HI4, and HI5 as distinct clades. Some analyses resolved HI1 and HI2 as distinct clades (although the support was generally low), while others grouped the two into a single clade. It is likely that these clades originated recently and have not had enough time to accumulate substantial differences in the genes examined. The haplotype networks provide a visualization of the close relationship among the clades of *H. imbricata*. For both ITS1 and ITS2, the most common haplotype was shared among multiple mitochondrial lineages. Although some individuals from each mitochondrial lineage represented unique haplotypes, the degree of separation in the ITS regions were only a few nucleotide differences at most, highlighting the very short history of these variants. These results further support observations that phylogenies based on COI divergences alone find more lineages in a cryptic species complex than nuclear DNA (e.g. Liu et al. 2017).

An interesting discrepancy was found when comparing COI and ITS divergences (see Table 2); according to COI, the mitochondrial lineages HI and HI5 were more divergent from each other than HI was to the outgroup, *H. extenuata* (17.8% vs. 16.39%), but based on ITS, it is much less divergent (1.5% vs. 8.3%). The similar level of COI differentiation between clades of *H. imbricata* and *H. extenuata* may indicate that COI has reached substitution saturation, consequently limiting its value as a source of phylogenetic information for this deep split. COI should show greater divergence than ITS because the mitochondrial genome accumulates mutations approximately ten times faster than the nuclear genome (Brown et al. 1979), but some studies of cryptic species complexes have found much higher rates of ITS differentiation between lineages. In the case of *Paranaitis wahlbergi*, a phyllodocid polychaete found in both Arctic and boreal waters, the Arctic conspecifics attained smaller maximum body sizes and

produced larger eggs (Kato and Pleijel 2003). When ITS and COI K2P distances were compared between Arctic and boreal populations, ITS distances were greater than COI distances (ITS1-22.2%, ITS2-16.04%, COI-14.6%). These differences supported the separation of the Arctic and boreal populations as distinct species (Nygren et al. 2009). Another example is from the phyllodocid *Eumida sanguinea*, which was determined to be a species complex containing 10 provisional species based on combined COI and ITS data. The ITS and COI distances were similar between some provisional species pairs, although generally the COI distances were slightly higher than the ITS distances, and in a few cases the ITS distances were greater (Nygren and Pleijel 2011). There is extensive variation in the relationship between COI and ITS divergences within cryptic species complexes and the findings reported here adds support to the importance of including evidence from both nuclear and mitochondrial loci in cryptic species investigations.

It has been hypothesized that cryptic diversity can develop through three mechanisms: recent divergence, niche evolution, or morphological convergence (Fiser et al. 2017). The molecular evidence shows that all lineages of *H. imbricata* are closely related, so morphological convergence is unlikely to be the mechanism leading to cryptic species in these taxa. It seems plausible that the mechanism causing cryptic diversity in the *Harmothoe imbricata* complex is either recent divergence, niche evolution, or a combination of the two. In this study, multiple mitochondrial lineages of *H. imbricata* occurred in sympatry in almost all sampling regions, with the greatest genetic diversity seen in the Arctic. Individuals carrying the HI1 mitochondrial lineage were found to occur in the Pacific, Arctic, and Atlantic Oceans. *Harmothoe imbricata* lineages may have diverged in multiple glacial refugia when much of the species' northern range was ice-covered during the Pleistocene glaciation, then remained reproductively isolated when

they came into secondary contact after ice sheets retreated approximately 9,000 years ago (Hardy et al. 2011). The divergence could be recent enough that morphological differences have not had time to accumulate. The isolated groups of *H. imbricata* could have experienced different environmental conditions and selection pressures could have led to physiological differences between the populations. These physiological differences could lead to the different populations using separate ecological niches, thereby limiting breeding between the groups. Alternatively, there could be a behavioral or physiological difference that evolved during their separation that prevents reproduction between the lineages.

There are multiple mechanisms that could be isolating reproduction between these divergent lineages. Although most polychaetes exhibit the reproductive strategy of broadcast spawning, in which animals release gametes into the water column to be fertilized externally and receive no parental care, *H. imbricata* employs a complex reproductive strategy that involves brooding of developing eggs under the elytra (Wilson 1991). Although individuals are generally solitary and antagonistic towards conspecifics, mature females emit a pheromone after gamete maturation that attracts male conspecifics and mating pairs form (Daly 1972, Watson et al. 2000). Prior to spawning, females undergo an elongation of their nephridial papillae, which turn upwards between the parapodia and release eggs into the space between their body and scales. Males position themselves dorsally on the female and shed their sperm, which is moved over the eggs by use of ciliary action under the female's elytra (Segrove 1938), leading to fertilization. It is possible that the pheromones of the different lineages have changed and males are either not attracted to, or maintain antagonistic behaviors towards, females from different lineages, thereby preventing spawning.

Alternatively, there may be gamete incompatibilities, with either the sperm not binding to the egg or the egg not accepting the sperm (e.g. Styan et al. 2008, Palumbi 1999). This species has been bred successfully in captivity (e.g. Daly 1972); a captive breeding experiment could help determine whether the divergent lineages of *H. imbricata* are willing and capable of producing viable offspring. An in-depth morphological comparison should also be conducted among the lineages of *H. imbricata* to check for subtle morphological variation, as it is not uncommon for morphological comparisons of cryptic species complexes to lead to the discovery of inconspicuous differences among lineages (Saez and Lozano 2005). Another important goal would be to obtain DNA from the neotype of this species; if it is decided that the divergent lineages of *H. imbricata* deserve separate species names, the name '*Harmothoe imbricata*' must remain with the lineage to which the neotype belongs.

## GENERAL CONCLUSION

Molecular evidence has revealed that many cosmopolitan species are cryptic species complexes composed of multiple genetically distinct, yet morphologically indistinguishable lineages, potentially leading to underestimates of global biodiversity (Bickford et al. 2007, Struck et al. 2018). Many studies aimed at uncovering cryptic diversity use the mitochondrial COI gene, commonly referred to as the barcoding gene, to describe cryptic species complexes. Although COI divergence does suggest the possibility of a cryptic species complex, investigations of cryptic species complexes should also incorporate nuclear sequence data (Ballard and Whitlock 2004, Rubinoff and Holland 2005). The polychaete scale-worm *Harmothoe imbricata* was recently suggested to be a cryptic species complex based on the observation of large sequence divergences in COI (Carr 2010). This study aimed to determine whether analysis of nuclear DNA data supported or rejected the designation of *H. imbricata* as a species complex. The analysis of multiple nuclear genes suggest that *H. imbricata* is a species complex composed of four divergent lineages, rather than the five lineages previously described. This finding corroborates previous studies that have shown that COI is likely to find more lineages than nuclear DNA in a cryptic species complex (e.g. Liu et al. 2017).

Although it is generally agreed upon that species are an indispensable unit of biology, the definition of a species is a much more complicated issue. Over 20 species concepts exist, some of which are conflicting and incompatible (Mayden 1997). Some of the most commonly used species concepts include the biological species concept (BSC), the morphological species concept (MSC), the ecological species concept (ESC), and the genetic species concept (GSC; as summarized by Mayden 1997). According to the BSC, species are groups of organisms that are

able to naturally produce reproductively viable offspring. The MSC relies on distinct, visible differences in morphology to differentiate species; it is often assumed that phenotypic differences are indicative of reproductive isolation. According to the ESC, different species use a particular ecological niche, leading to reproductive isolation. The GSC defines species as interbreeding, reproductively isolated (i.e. genetically distinct) groups of individuals.

While these concepts are not necessarily contradictory in theory, they can be in reality. For example, if two subsets within a group of closely related individuals opportunistically take advantage of slightly different habitats, they could be considered different species based on the ESC. If both of those groups employ broadcast spawning concurrently and produce viable offspring, they would not be considered separate species based on BSC. Additionally, the ongoing reproductive success would limit genetic differentiation between individuals using different habitats, so these groups would not be separate species based on GSC. Additionally, there are instances where large amounts of genetic diversity are discovered within a morphologically identical group (i.e. cryptic diversity). Alternatively, there are cases in which researchers did not find genetic variation that correlated with appreciable morphological differences, as was seen with the color polymorphs of *H. imbricata* found in Scandinavia (Nygren et al. 2011). These are examples of how the outcome of species delimitation can be variable depending on the species concept used. Regardless of the differences between the various species concepts, the unifying concept amongst them is that species represent separately evolving metapopulations (De Queiroz 2007). There is a need to place entities into categories (i.e. species A and species B); a challenge with this is that speciation is an ongoing process, and many groups of closely related organisms are still undergoing the process of differentiation and speciation, so it is not always simple to define species as A and B.

The “provisional species” of *H. imbricata* can have extensive to little or no differentiation, depending on the gene examined. This speaks to the important issue that gene trees and species trees are not necessarily synonymous, and inferring phylogenetic relationships based on a single gene is not appropriate. A multilocus approach provides a more representative depiction of evolutionary history, but even this approach only uses a small proportion of the genome. Additionally, the genes targeted in this study were chosen due to availability of primers. Little to no variation was seen in the regions of the 28S and 18S genes that were sequenced; these genes are better used for studying more distant relationships. The internal transcribed spacers displayed variation that made them useful for inferring the phylogeny of *H. imbricata*, but there was not enough differentiation among lineages to produce a well-supported phylogeny, even when combined with the other genes. In order to gain a better understanding of genetic divergence within this group, it would be good to obtain genome-wide sequence data (Struck et al. 2018). A small scale genomic scan using double digest restriction associated DNA (ddRAD) sequencing is in progress (Gastaldi in progress) using recently acquired samples from the Chukchi Sea. Unfortunately, those samples only represent the HI1, HI4, and HI5 mitochondrial lineages, but preliminary results point to a significant distinction between all three of these mitochondrial lineages.

## FIGURES AND TABLES

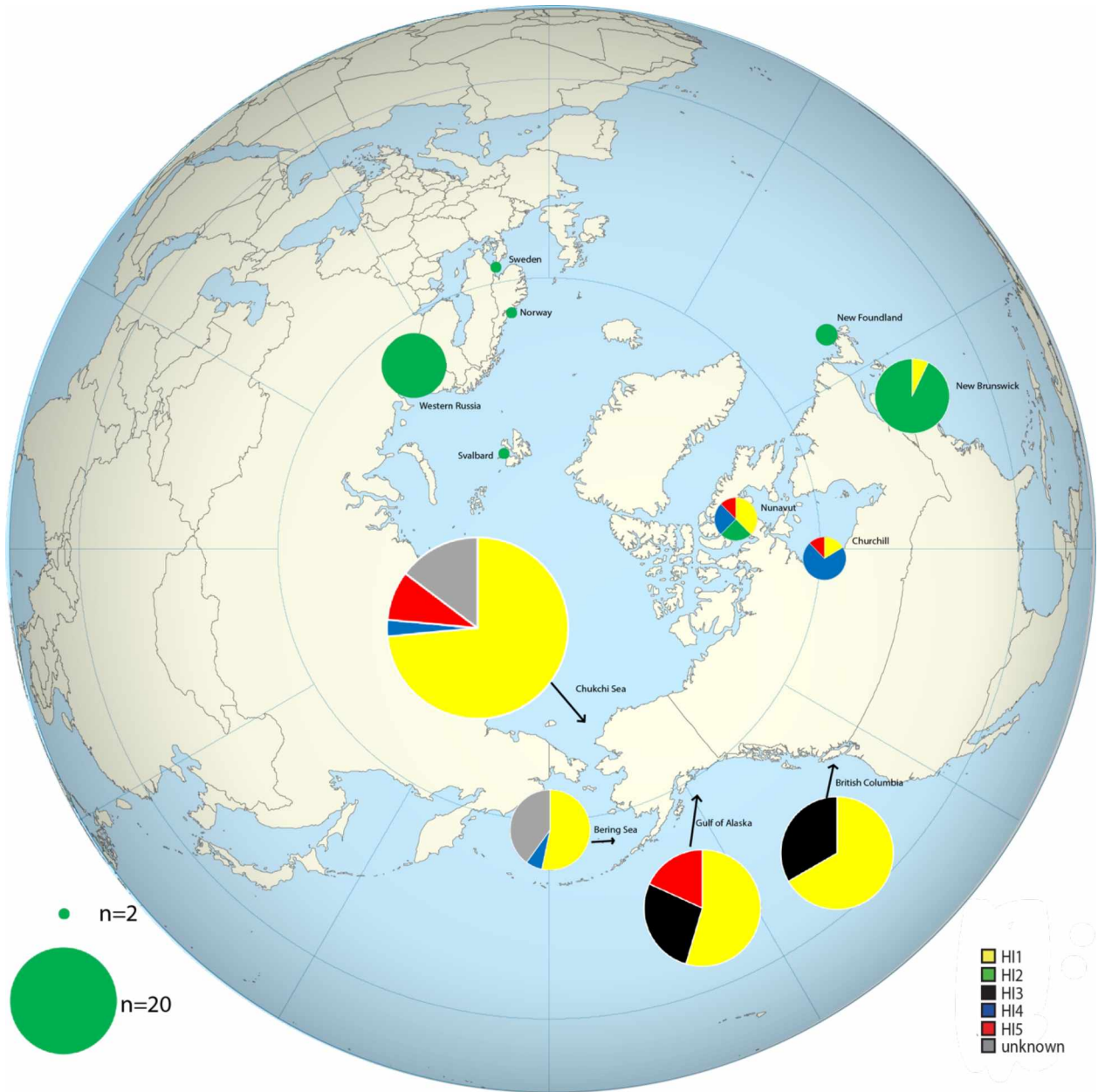


Figure 1: Map of sample distribution. Circles represent approximate sampling region. Size of circle is relative to the number of *Harmothoe imbricata* samples from that region and colors represent the proportion of each mitochondrial lineages from each region.

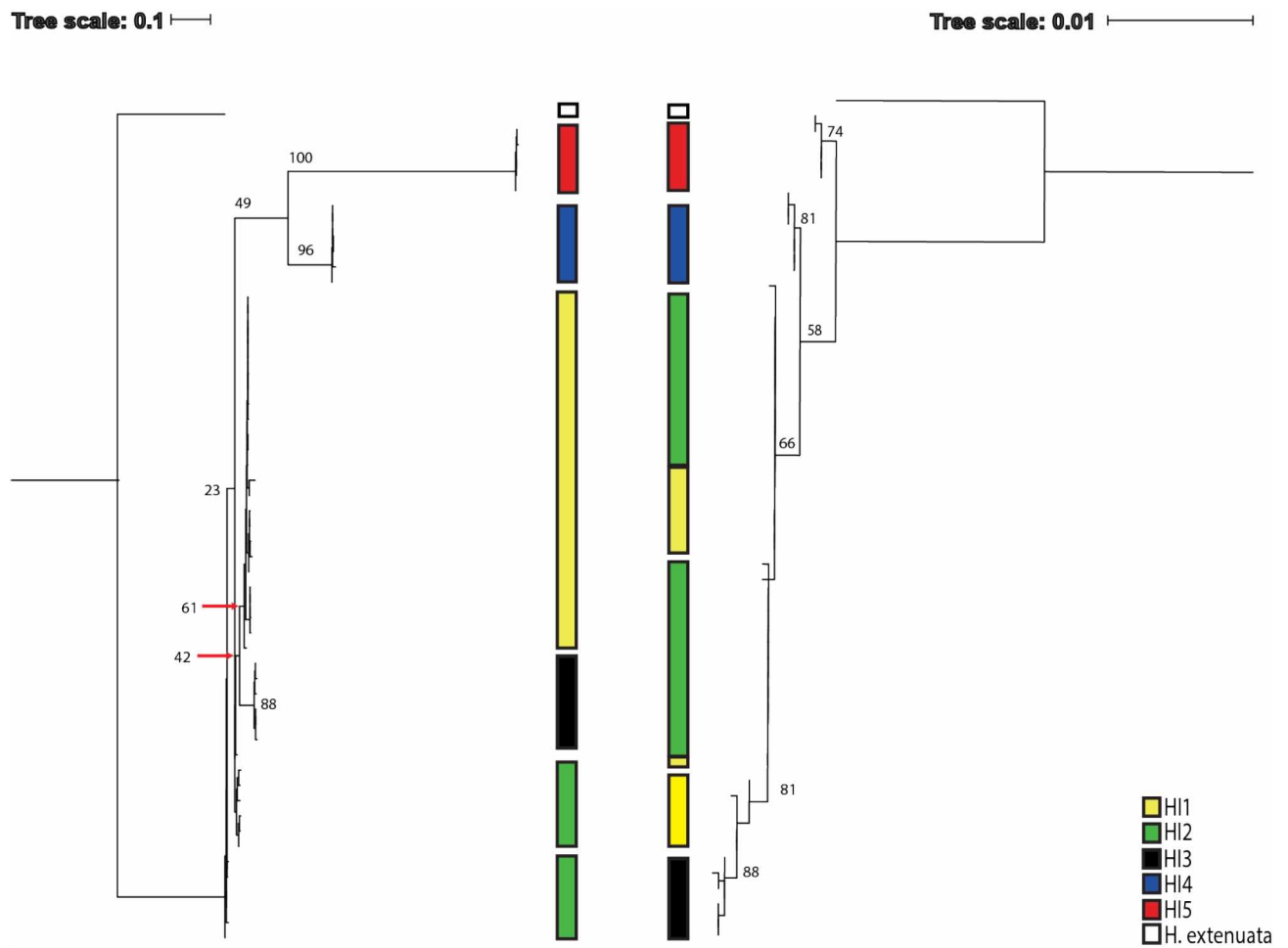


Figure 2: Phylogenetic tree of the *Harmothoe imbricata* species complex as determined by maximum likelihood analysis based on COI (left) and the nuclear concatenation (right). Numbers next to nodes represent bootstrap values. Both trees were rooted at the outgroup *H. extenuata* (note that the scale of the two trees are different).

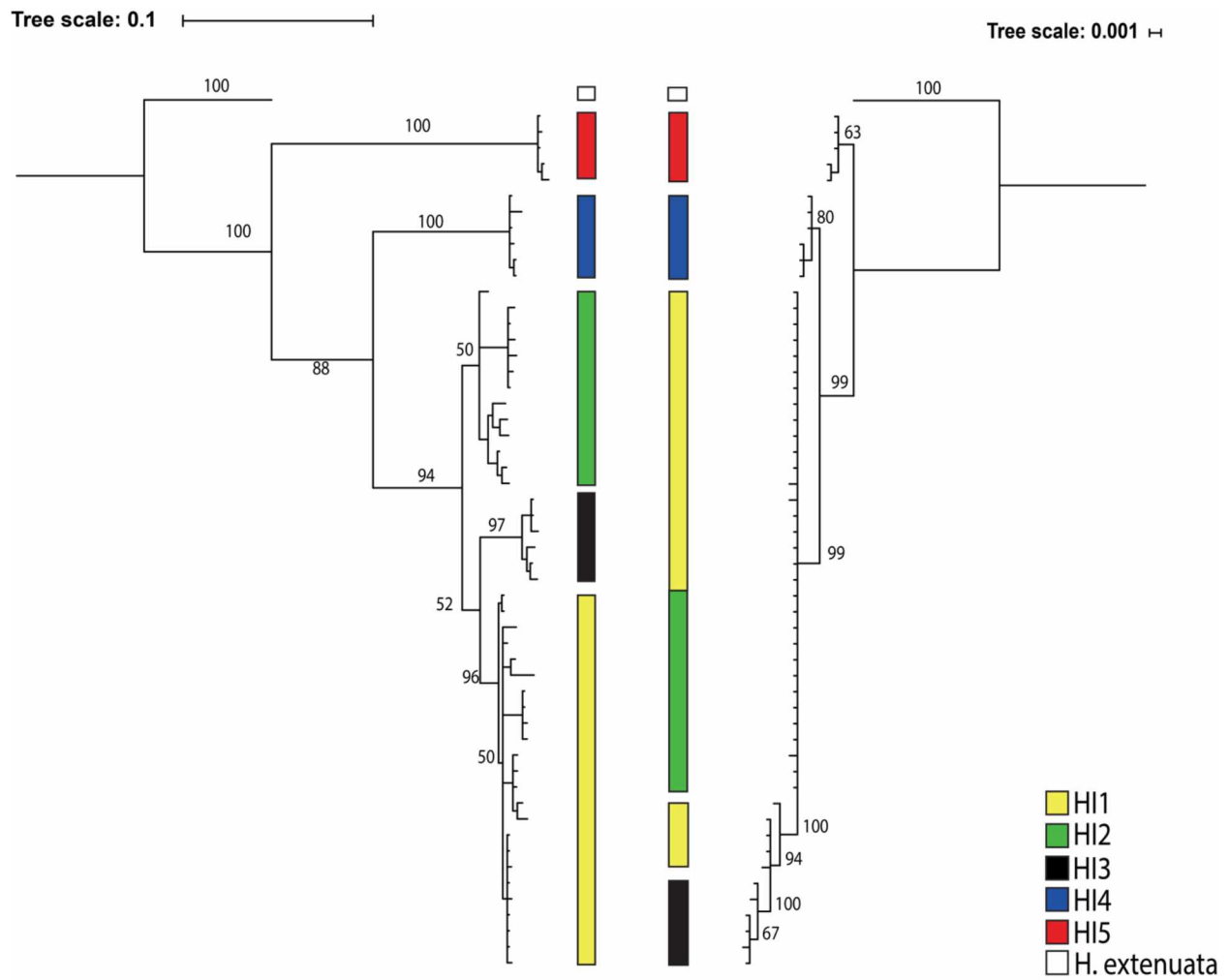


Figure 3: Phylogenetic tree of the *Harmothoe imbricata* species complex as determined by Bayesian inference based on COI (left) and the nuclear concatenation (right). Numbers next to nodes represent posterior probabilities. Both trees were rooted at the outgroup *H. extenuata* (note that the scale of the two trees are different).

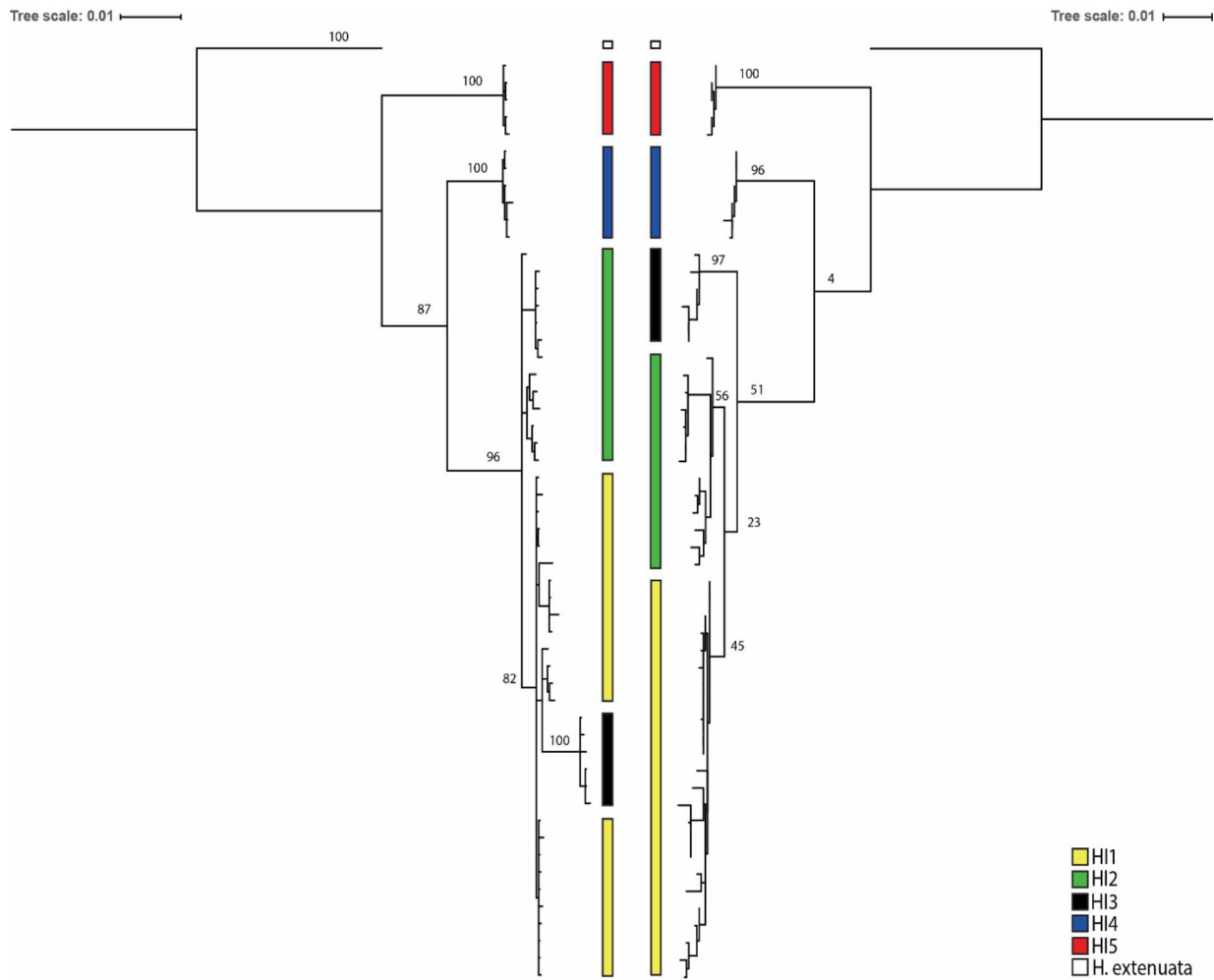


Figure 4: Phylogenetic tree of the *Harmothoe imbricata* species complex based on a concatenation of mitochondrial and nuclear sequences as determined by Bayesian inference based (left) and maximum likelihood analysis (right). Numbers next to nodes represent posterior probabilities (left) and bootstrap values (right). Both trees were rooted at the outgroup *H. extenuata*.

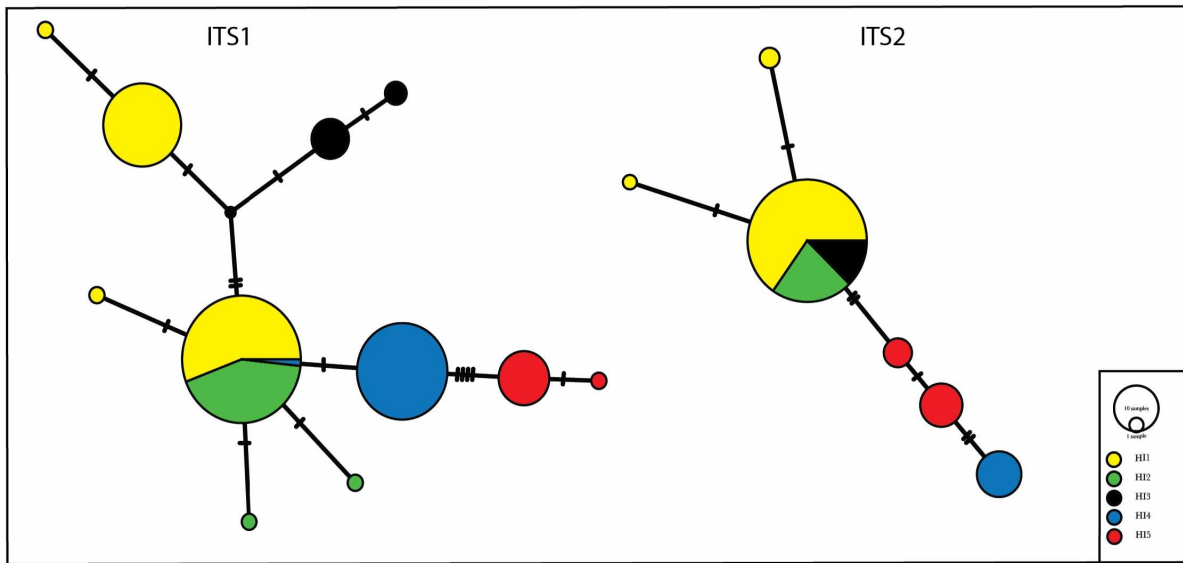


Figure 5: Haplotype network showing distribution of mutations in the ITS1 and ITS2 regions. Each circle represents a distinct haplotype and circle size represents the number of individuals per haplotype. Hatch marks represent single mutation and color represents mitochondrial lineage.

Table 1: List of samples used in analyses in addition to their: provisional species designations (as designated by Carr et al. 2011), sample identification, collection coordinates, mitochondrial lineage affinity, GenBank accession numbers, and voucher repository.

Species/Provisional species	Sample ID	Lat.	Long.	Mit. lineage	COI	ITS1	ITS2	H3	18S	28S	Voucher info
<i>Harmothoe extenuata</i>	ML1_5_HI1	70.52	164.13	N/A	MK390769	MK356573	MK376352	MK372666	MK386471	MK399303	UAM:Inv:21020
<i>Harmothoe imbricata</i> CMC01	08PROBE0087	58.79	-94.22	HI1	HQ023526*	MK356575	MK376353	MK372668	MK386480	MK399305	BIOUG<CAN>:08PROBE-0087
<i>Harmothoe imbricata</i> CMC01	08PROBE0248	58.75	-94.33	HI1	HQ023537*	MK356579	MK376354	MK372669	MK386481	MK399306	BIOUG<CAN>:08PROBE-0248
<i>Harmothoe imbricata</i> CMC01	11BIOAK0812	59.47	-151.55	HI1	MF121160*	MK356581	MK376355	MK372670	MK386482	MK399307	BIOUG<CAN>:11BIOAK-0812
<i>Harmothoe imbricata</i> CMC01	BAMPOL0015	48.82	-125.16	HI1	HM473408*	MK356592	MK376360	MK372672	MK386484	MK399311	BIOUG<CAN>:BAMPOL0015
<i>Harmothoe imbricata</i> CMC01	BAMPOL0022	48.82	-125.16	HI1	HM473411*	MK356595	MK376361	MK372673	MK386485	MK399312	BIOUG<CAN>:BAMPOL0022
<i>Harmothoe imbricata</i>	BBL_7_HI1	69.92	166.06	HI1	MK390776	MK356603	MK376367	MK372676	MK386474	MK399316	UAM:Inv:21031
<i>Harmothoe imbricata</i>	BBL_7_HI2	69.92	166.06	HI1	MK390777	MK356604	MK376368	MK372677	MK386475	MK399317	UAM:Inv:21032
<i>Harmothoe imbricata</i>	BBL_8_HI1	70.01	166.47	HI1	MK390780	MK356605	MK376370	MK372678	MK386476	MK399318	UAM:Inv:21035
<i>Harmothoe imbricata</i>	BBL_9_HI1	70.10	166.91	HI1	MK390781	MK356606	MK376371	MK372679	MK386477	MK399319	UAM:Inv:21036
<i>Harmothoe imbricata</i>	BP2010171	53.69	-132.17	HI1	HQ932555*	MK356607	MK376373	MK372680	MK386488	MK399320	BIOUG<CAN>:BP2010-171
<i>Harmothoe imbricata</i>	CBW5_HI1	64.17	-171.51	HI1	MK390784	MK356610	MK376376	MK372682	MK386491	MK399321	UAM:Inv:21009

<i>Harmothoe imbricata</i> CMC01	HUNTSPOL0382	44.99	-67.03	HI1	HQ024041*	MK356611	MK376377	MK372683	MK386492	MK399322	BIOUG<CAN>:HUNTSPOL0382
<i>Harmothoe imbricata</i>	ML1_5_HI2	70.52	164.13	HI1	MK390785	MK356612	MK376378	MK372684	MK386549	MK399323	UAM:Inv:21021
<i>Harmothoe imbricata</i>	ML5_1_HI1	70.65	166.00	HI1	MK390792	MK356617	MK376384	MK372688	MK386553	MK399325	UAM:Inv:21026
<i>Harmothoe imbricata</i>	ML6_10_HI1	71.19	160.27	HI1	MK390795	MK356619	MK376386	MK372690	MK386556	MK399326	UAM:Inv:21043
<i>Harmothoe imbricata</i>	ML6_10_HI2	71.19	160.27	HI1	MK390796	MK356620	MK376387	MK372691	MK386557	MK399327	UAM:Inv:21044
<i>Harmothoe imbricata</i>	ML6_9_HI1	71.07	160.74	HI1	MK390798	MK356621	MK376388	MK372692	MK386558	MK399328	UAM:Inv:21040
<i>Harmothoe imbricata</i>	ML6_9_HI2	71.07	160.74	HI1	MK390799	MK356622	MK376389	MK372693	MK386559	MK399329	UAM:Inv:21041
<i>Harmothoe imbricata</i> CMC01	NUNAV0176	69.37	-81.79	HI1	HQ024337*	MK356623	MK376391	MK372694	MK386493	MK399330	BIOUG<CAN>:NUNAV-0176
<i>Harmothoe imbricata</i> CMC01	OD141	65.12	-168.07	HI1	HM473758*	MK356628	MK376395	MK372696	MK386497	MK399332	BIOUG<CAN>:OD141
<i>Harmothoe imbricata</i> CMC01	OD142	65.12	-168.07	HI1	HM473759*	MK356629	MK376396	MK372697	MK386498	MK399333	BIOUG<CAN>:OD142
<i>Harmothoe imbricata</i> CMC01	OD154	69.97	-165.76	HI1	HM473764*	MK356633	MK376400	MK372699	MK386502	MK399335	BIOUG<CAN>:OD154
<i>Harmothoe imbricata</i> CMC01	OM19	66.18	-168.87	HI1	HM473765*	MK356634	MK376401	MK372700	MK386503	MK399336	BIOUG<CAN>:OM19
<i>Harmothoe imbricata</i> CMC01	PPA018	58.81	-94.21	HI1	HQ938293*	MK356635	MK376402	MK372701	MK386504	MK399337	BIOUG<CAN>:PPA018
<i>Harmothoe imbricata</i> CMC02	09BBPOL001	48.44	-54.00	HI2	KM612146*	MK356637	MK376403	MK372702	MK386505	MK399338	BIOUG<CAN>:09BBPOL-001
<i>Harmothoe imbricata</i> CMC02	09BBPOL003	48.44	-54.00	HI2	KM612232*	MK356638	MK376404	MK372703	MK386506	MK399339	BIOUG<CAN>:09BBPOL-003
<i>Harmothoe imbricata</i>	GNM 13107	78.66	20.99	HI2	GQ478931*	MK356639	MK376405	MK372704	MK386543	MK399341	GNM Polychaeta- 13107

<i>Harmothoe imbricata</i>	GNM 13119	63.44	10.35	HI2	GQ478943*	MK356641	MK376407	MK372706	MK386545	MK399342	GNM Polychaeta- 13119
<i>Harmothoe imbricata</i>	GNM 13160	58.86	11.20	HI2	GQ478984*	MK356643	MK376408	MK372708	MK386547	MK399344	GNM Polychaeta- 13160
<i>Harmothoe imbricata</i>	GNM 13163	56.91	11.90	HI2	GQ478987*	MK356644	MK376409	MK372709	MK386548	MK399345	GNM Polychaeta- 13163
<i>Harmothoe imbricata</i> CMC02	HUNTSPOL0039	45.08	-67.07	HI2	HQ024048*	MK356650	MK376410	MK372710	MK386507	MK399346	BIOUG<CAN>:HUNTSPOL0039
<i>Harmothoe imbricata</i> CMC02	HUNTSPOL0151	45.07	-67.04	HI2	HQ024051*	MK356652	MK376411	MK372711	MK386508	MK399347	BIOUG<CAN>:HUNTSPOL0151
<i>Harmothoe imbricata</i> CMC02	HUNTSPOL0153	45.07	-67.04	HI2	HQ024052*	MK356653	MK376412	MK372712	MK386509	MK399348	BIOUG<CAN>:HUNTSPOL0153
<i>Harmothoe imbricata</i> CMC02	NUNAV0188	69.37	-81.79	HI2	HQ024351*	MK356657	MK376413	MK372714	MK386510	MK399349	BIOUG<CAN>:NUNAV-0188
<i>Harmothoe imbricata</i> CMC02	WS0008	66.55	33.11	HI2	GU670829*	MK356659	MK376415	MK372716	MK386512	MK399351	BIOUG<CAN>:WS0008
<i>Harmothoe imbricata</i> CMC02	WS0078	66.92	33.10	HI2	GU672604*	MK356660	MK376416	MK372717	MK386513	MK399352	BIOUG<CAN>:WS0078
<i>Harmothoe imbricata</i> CMC02	WS0138	66.92	33.10	HI2	GU672560*	MK356662	MK376417	MK372718	MK386514	MK399353	BIOUG<CAN>:WS0138
<i>Harmothoe imbricata</i> CMC03	11BIOAK1062	59.46	-151.71	HI3	MF120994*	MK356665	MK376419	MK372719	MK386515	MK399355	BIOUG<CAN>:11BIOAK-1062
<i>Harmothoe imbricata</i> CMC03	11BIOAK1066	59.46	-151.71	HI3	MF121300*	MK356666	MK376420	MK372720	MK386516	MK399356	BIOUG<CAN>:11BIOAK-1066
<i>Harmothoe imbricata</i> CMC03	11BIOAK1638	59.49	-151.65	HI3	MF121424*	MK356668	MK376422	MK372722	MK386518	MK399357	BIOUG<CAN>:11BIOAK-1638
<i>Harmothoe imbricata</i> CMC03	BAMPOL0345	48.86	-125.16	HI3	HM473426*	MK356669	MK376423	MK372723	MK386519	MK399358	BIOUG<CAN>:BAMPOL0345
<i>Harmothoe imbricata</i> CMC03	BAMPOL0455	49.69	-124.87	HI3	HM473427*	MK356670	MK376424	MK372724	MK386520	MK399359	BIOUG<CAN>:BAMPOL0455
<i>Harmothoe imbricata</i> CMC03	BP2010413	53.67	-132.38	HI3	MF121449*	MK356672	MK376427	MK372726	MK386523	MK399361	BIOUG<CAN>:BP2010-413

<i>Harmothoe imbricata</i> CMC04	08PROBE0031	58.79	-94.22	HI4	HQ023558*	MK356681	MK376428	MK372727	MK386524	MK399362	BIOUG<CAN>:08PROBE-0031
<i>Harmothoe imbricata</i> CMC04	08PROBE0033	58.79	-94.22	HI4	HQ023560*	MK356683	MK376429	MK372728	MK386525	MK399363	BIOUG<CAN>:08PROBE-0033
<i>Harmothoe imbricata</i> CMC04	08PROBE0538	58.77	-94.29	HI4	HQ023550*	MK356691	MK376430	MK372729	MK386526	MK399364	BIOUG<CAN>:08PROBE-0538
<i>Harmothoe imbricata</i>	BBL_10_HI1	70.19	167.36	HI4	MK390771	MK356701	MK376431	MK372730	MK386472	MK399365	UAM:Inv:21039
<i>Harmothoe imbricata</i> CMC04	NUNAV0203	69.37	-81.79	HI4	HQ024353*	MK356702	MK376432	MK372731	MK386527	MK399366	BIOUG<CAN>:NUNAV-0203
<i>Harmothoe imbricata</i> CMC04	NUNAV0239	69.37	-81.79	HI4	HQ024354*	MK356703	MK376433	MK372732	MK386528	MK399367	BIOUG<CAN>:NUNAV-0239
<i>Harmothoe imbricata</i> CMC05	08PROBE0202	58.85	-93.76	HI5	HQ023571*	MK356709	MK376438	MK372736	MK386533	MK399369	BIOUG<CAN>:08PROBE-0202
<i>Harmothoe imbricata</i> CMC05	09PROBE01171	58.85	-93.81	HI5	GU672401*	MK356710	MK376439	MK372737	MK386534	MK399370	BIOUG<CAN>:09PROBE-01171
<i>Harmothoe imbricata</i> CMC05	09PROBE01179	58.80	-94.04	HI5	GU672098*	MK356711	MK376440	MK372738	MK386535	MK399371	BIOUG<CAN>:09PROBE-01179
<i>Harmothoe imbricata</i> CMC05	09PROBE08212	58.83	-93.82	HI5	GU672129*	MK356712	MK376441	MK372739	MK386536	MK399372	BIOUG<CAN>:09PROBE-08212
<i>Harmothoe imbricata</i> CMC05	HLC30563	74.68	-94.86	HI5	HQ024355*	MK356719	MK376449	MK372745	MK386542	MK399376	BIOUG<CAN>:HLC-30563

Table 2: Average K2P p-values as percentages and standard deviations between and within *Harmothoe imbricata* mitochondrial lineages and *H. extenuata*, the outgroup, for COI (upper number) and ITS (lower number).

	HI1	HI2	HI3	HI4	HI5
<b>HI1</b>	<b>1.140 ± 0.715</b> <b>0.160 ± 0.205</b>				
<b>HI2</b>	4.039 ± 0.743 0.095 ± 0.174	<b>2.042 ± 1.027</b> <b>0 ± 0</b>			
<b>HI3</b>	4.812 ± 0.385 0.566 ± 0.140	5.661 ± 0.631 0.62 ± 0.001	<b>0.730 ± 0.361</b> <b>0.076 ± 0.074</b>		
<b>HI4</b>	12.122 ± 0.263 0.783 ± 0.117	12.401 ± 0.333 0.717 ± 0	11.768 ± 0.231 1.196 ± 0.069	<b>0.383 ± 0.323</b> <b>0 ± 0</b>	
<b>HI5</b>	17.822 ± 0.267 1.507 ± 0.118	17.748 ± 0.274 1.441 ± 0	18.105 ± 0.274 1.921 ± 0.069	16.562 ± 0.216 1.009 ± 0.000	<b>0.313 ± 0.214</b> <b>0 ± 0</b>
<b>H. extenuata (Outgroup)</b>	16.391 ± 0.282 8.309 ± 0.071	16.321 ± 0.324 8.284 ± 0	16.812 ± 0.18 8.629 ± 0.076	17.768 ± 0.071 8.284 ± 0	18.754 ± 0.148 7.407 ± 0

## REFERENCES

- Avise JC, Bowen BW, Ayala FJ (2016) In the light of evolution X: Comparative phylogeography. *Proc Natl Acad Sci U S A* 113:7957–7961. doi:10.1073/pnas.1604338113
- Achurra A, Erseus C (2013) DNA barcoding and species delimitation: the *Stylodrilus heringianus* case (Annelida: Clitellata: Lumbriculidae). *Invertebr Syst* 27:118–128
- Ballard J, Whitlock M (2004) The incomplete natural history of mitochondria. *Mol Ecol* 13:729–744. doi:10.1046/j.1365-294X.2003.02063.x
- Bandelt H, Forster P, Rohl A (1999) Median-Joining Networks for Inferring Intraspecific Phylogenies. *Mol Biol* 16:37–48. doi:10.1093/oxfordjournals.molbev.a026036
- Barnich R, Fiege D (2009) Revision of the genus *Harmothoe* Kingberg, 1856 (Polychaeta: Polynoidae) in the Northeast Atlantic. *Zootaxa* 2104:1–76
- Bhasin M, Reinherz EL, Reche PA (2006) Recognition and Classification of Histones Using. *J Comput Biol* 13:102–112
- Bleidorn C, Kruse I, Albrecht S, Bartolomaeus T (2006) Mitochondrial sequence data expose the putative cosmopolitan polychaete *Scoloplos armiger* (Annelida, Orbiniidae) as a species complex. *BMC Evol Biol* 6:47. doi:10.1186/1471-2148-6-47
- Bickford D, Lohman DJ, Sodhi NS, et al. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155. doi:10.1016/j.tree.2006.11.004
- Brown S, Rouse G, Hutchings P, Colgan D (1999) Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships. *Aust J Zool* 47:499. doi:10.1071/ZO99026
- Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* 76:1967–1971. doi:10.1146/annurev.es.18.110187.001413

- Butchart SHM, Walpole M, Collen B, et al. (2010) Global biodiversity: Indicators of recent declines. *Science* (80- ) 328:1164–1168. doi:10.1126/science.1187512
- Calderon I, Garrabou J, Aurelle D (2006) Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *J Exp Mar Bio Ecol* 336:184–197. doi:10.1016/j.jembe.2006.05.006
- Carr CM (2010) *The Polychaeta of Canada: Exploring diversity and distribution patterns using DNA barcoding*. The University of Guelph
- Carr CM, Hardy SM, Brown TM, et al. (2011) A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes. *PLoS One* 6. doi:10.1371/journal.pone.0022232
- Colgan DJ, Ponder WF, Egglar PE (2000) Gastropod evolutionary rates and phylogenetic relationships assessed using partial 28s rDNA and histone H3 sequences. *Zool Scr* 29:29–63. doi:10.1046/j.1463-6409.2000.00021.x
- Colgan DJ, Hutchings PA, Braune M (2006) A multigene framework for polychaete phylogenetic studies. *Org Divers Evol*. doi:10.1016/j.ode.2005.11.002
- Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Mol Ecol Resour* 13:969–975. doi:10.1111/1755-0998.12046
- Costello MJ, Wilson S, Houlding B (2012) Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Syst Biol* 61:871–883. doi:10.1093/sysbio/syr080
- Daly JM (1972) The maturation and breeding biology of *Harmothoe imbricata*. *Mar Biol* 12:53–66
- Daly JM (1974) Gametogenesis in *Harmothoe imbricata* (Polychaeta: Polynoidae). *Mar Biol* 25:35–40

- De Queiroz K (2007) Species concepts and species delimitation. *Syst Biol* 56:879–886.  
doi:10.1080/10635150701701083
- Duarte CM (2000) Marine biodiversity and ecosystem services: an elusive link. *J Exp Mar Bio Ecol* 250:117–131. doi:10.1016/S0022-0981(00)00194-5
- Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797
- Fauchald K (1974) Polychaete phylogeny: a problem in protostome phylogeny. *Syst Biol* 23:493–506
- Fauchald K (1977) The polychaete worms. Definitions and keys to the orders, families and genera. *Natural History Museum of Los Angeles Sci ser* 28 1-188
- Fauchald K, Rouse G (1997) Polychaete systematics: Past and present. *Zoologica Scr* 26:71–138. doi:10.1111/j.1463-6409.1997.tb00411.x
- Folmer O, Black M, Hoeh W, et al. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299
- Garwood PR (1980) The role of temperature and daylength in the control of the reproductive cycle of *Harmothoe imbricata* (Polychaeta: Polynoidae). *J Exp Mar Bio Ecol* 47:35–53. doi:10.1016/0022-0981(80)90136-7
- Garwood PR (1981) Observations of the cytology of the developing female germ cell in the polychaete *Harmothoe imbricata*. *Int J Invertebr Reprod* 3:333–345. doi:10.1080/01651269.1981.10553408
- Garwood PR, Olive PJW (1982) The influence of photoperiod on oocyte growth and in the control of the reproductive cycle of the polychaete *Harmothoe imbricata*. *Int J Invertebr Reprod* 5:161–165

- Gaudron SM, Bentley MG (2002) Control of reproductive behaviour in the scale worm *Harmothoe imbricata*. *Invertebr Reprod Dev* 41:109–118
- Gaudron SM, Watson GJ, Bentley MG (2007) Induction of pairing in male scale worm *Harmothoe imbricata* (Polychaeta: Polynoidae) by chemical signals. *J Mar Biol Ass UK* 87:1115–1116. doi:10.1017/S002531540705581
- Giska I, Sechi P, Babik W (2015) Deeply divergent sympatric mitochondrial lineages of the earthworm *Lumbricus rubellus* are not reproductively isolated. *BMC Evol Biol* 15:1–13. doi:10.1186/s12862-015-0488-9
- Glasby JC, Wei N V, Gibb KS (2013) Cryptic species of Nereididae (Annelida: Polychaeta) on Australian coral reefs. *Invertebr Syst* 27:245–264. doi:10.1071/IS12031
- Gonzalez BC, Martínez A, Borda E, et al. (2018) Phylogeny and systematics of Aphroditiformia. *Cladistics* 34:225–259. doi:10.1111/cla.12202
- Grassle JF, Maciolek NJ (1992) Deep-Sea Species Richness: Regional and Local Diversity Estimates from Quantitative Bottom Samples. *The* 139:313–341
- Gustafsson DR, Price DA, Erseus C (2009) Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Mol Phylogenet Evol* 51:182–189. doi:10.1016/j.ympev.2008.12.016
- Halanych KM, Janosik AM (2006) A review of molecular markers used for Annelid phylogenetics. *Integr Comp Biol* 46:533–543. doi:10.1093/icb/icj052
- Hardy SM, Carr CM, Hardman M, et al. (2011) Biodiversity and phylogeography of Arctic marine fauna: Insights from molecular tools. *Mar Biodivers*. doi:10.1007/s12526-010-0056-x
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proc R Soc B Biol Sci* 270:313–321. doi:10.1098/rspb.2002.2218

- Hebert PDN, Ratnasingham S, de Waard JR (2003b) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc R Soc B Biol Sci* 270:S96–S99. doi:10.1098/rsbl.2003.0025
- Hebert PDN, Penton EH, Burns JM, et al. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci* 101:14812–14817. doi:10.1073/pnas.0406166101
- Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Q. Rev. Biol.* 66:411–453
- Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc Biol Sci* 272:1525–1534. doi:10.1098/rspb.2005.3056
- Hutchings, PA, Fauchald, K (2000) Class Polychaeta: Definition and general description. In: Beesley, PL, Ross, GJB, Glasby, CJ (eds), *Polychaetes and Allies: the Southern Synthesis*, CSIRO Publishing, Melbourne, pp 1-3
- Jumars PA, Dorgan KM, Lindsay SM (2015) Diet of Worms Emended: An Update of Polychaete Feeding Guilds. *Ann Rev Mar Sci* 7:497–520. doi:10.1146/annurev-marine-010814-020007
- Kato T, Pleijel F (2003) A revision of *Paranaitis* Southern, 1914 (Polychaeta: Phyllodocidae). *Zool J Linn Soc* 138:379–429. doi:10.1046/j.1096-3642.2003.00069.
- Lafontaine DLJ, Tollervey D (2001) The function and synthesis of ribosomes. *Nat Rev Mol Cell Biol* 2:514–520. doi:10.1038/35080045
- Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol Biol Evol* 29:1695–1701. doi:10.1093/molbev/mss020

- Lanfear R, Frandsen PB, Wright AM, et al. (2016) Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol* 34:772–773. doi:10.1093/molbev/msw260
- Larsen BB, Miller EC, Rhodes MK, Wiens JJ (2017) Inordinate fondness multiplied and redistributed: the number of species on Earth and the new pie of life. *Rev Lit Arts Am* 92:230–265. doi:10.1086/693564
- Leigh JW, Bryant D (2015) POPART: Full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116. doi:10.1111/2041-210X.12410
- Liu Y, Fend SV, Martinsson S, Erseus C (2017) Extensive cryptic diversity in the cosmopolitan sludge worm *Limnodrilus hoffmeisteri* (Clitellata, Naididae). *Org Divers Evol* 17:477–495. doi:10.1007/s13127-016-0317-z
- Maddison WP, Maddison DR (2018) Mesquite: A modular system for evolutionary analysis. Version 3.40 <http://mesquiteproject.org>
- Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: *Species: The Units of Biodiversity*. pp 381–423
- McHugh D (1997) Molecular evidence that echiurans and pogonophorans are derived annelids. *Proc Natl Acad Sci* 94:8006–8009. doi:10.1073/pnas.94.15.8006
- McHugh D (2000) Molecular phylogeny of the Annelida. *Can J Zool* 78:1873–1884
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gatew Comput Environ Work GCE 2010. doi:10.1109/GCE.2010.5676129
- Nygren A, Eklof J, Pleijel F (2009) Arctic-boreal sibling species of *Paranaitis* (Polychaeta, Phyllodocidae). *Mar Biol Res* 5:315–327. doi:10.1080/17451000802441301

- Nygren A, Norlinder E, Panova M, Pleijel F (2011) Colour polymorphism in the polychaete *Harmothoe imbricata*. *Mar Biol Res* 7:54–62. doi:10.1080/17451001003713555
- Nygren A, Pleijel F (2011) From one to ten in a single stroke - resolving the European *Eumida sanguinea* (Phyllodocidae, Annelida) species complex. *Mol Phylogenet Evol* 58:132–141. doi:10.1016/j.ympev.2010.10.010
- Nygren A (2014) Cryptic polychaete diversity: A review. *Zool Scr* 43:172–183. doi:10.1111/zsc.12044
- Palumbi SR (1999) All Males Are Not Created Equal: Fertility Differences Depend on Gamete Recognition Polymorphisms in Sea Urchins. *Proc Natl Acad Sci* 96:12632–12637. doi:10.1073/pnas.96.22.12632
- Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol Biol* 7:1–6. doi:10.1186/1471-2148-7-121
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574. doi:10.1093/bioinformatics/btg180
- Roskov Y, Abucay L, Orrell T, Nicolson D, Bailly N, Kirk PM, Bourgoin T, DeWalt RE, Decock W, De Wever A, Nieukerken E, Zarucchi J, Penev L, eds. (2018). *Species 2000 & ITIS Catalogue of Life, 2018 Annual Checklist*. Digital resource at [www.catalogueoflife.org/annual-checklist/2018](http://www.catalogueoflife.org/annual-checklist/2018). *Species 2000: Naturalis, Leiden, the Netherlands*. ISSN 2405-884X.
- Rouse G, Pleijel F (eds) (2006) *Reproductive biology and phylogeny of Annelida. Reproductive Biology and Phylogeny*. Science Publishers, Enfield, NH. 688
- Rouse G, Fauchald K (1997) Cladistics and polychaetes. *Zool Scr* 26:139–204. doi:10.1111/j.1463-6409.1997.tb00412.x

- Rubinoff D, Holland BS (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Syst Biol* 54:952–961. doi:10.1080/10635150500234674
- Ruiz-Garcia M, Luengas-Villamil K, Leguizamon N, et al. (2015) Molecular phylogenetics and phylogeography of all the *Saimiri* taxa (Cebidae, Primates) inferred from mt COI and COII gene sequences. *Primates*. doi:10.1007/s10329-014-0452-0
- Schmidt H, Westheide W (2000) Are the meiofaunal polychaetes *Hesionides arenaria* and *Stygocapitella subterranea* true cosmopolitan species? - Results of RAPD-PCR investigations. *Zool Scr* 29:17–27. doi:10.1046/j.1463-6409.2000.00026.x
- Schuchert P (2014) High genetic diversity in the hydroid *Plumularia setacea*: A multitude of cryptic species or extensive population subdivision? *Mol Phylogenet Evol* 76:1–9. doi:10.1016/j.ympev.2014.02.020
- Segrove F (1938) An account of surface ciliation in some polychaete worms. *Proc Zool Soc London* 108:85–107
- Snelgrove PVR (1997) The Importance of Marine Sediment Biodiversity in Ecosystem Processes. *Ambio* 26:578–583
- Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. Doi:10.1093/bioinformatics/btu033
- Struck TH, Paul C, Hill N, et al. (2011) Phylogenomic analyses unravel annelid evolution. *Nature* 471:95–98. doi:10.1038/nature09864
- Struck TH, Golombek A, Weigert A, et al. (2015) The Evolution of Annelids Reveals Two Adaptive Routes to the Interstitial Realm. *Curr Biol* 25:1993–1999. doi:10.1016/j.cub.2015.06.007

- Struck TH, Feder JL, Bendiksbj M, et al. (2018) Finding Evolutionary Processes Hidden in Cryptic Species. *Trends Ecol Evol* 33:153–163. doi:10.1016/j.tree.2017.11.007
- Styan CA, Kupriyanova E, Havenhand JN (2008) Barriers to cross-fertilization between populations of a widely dispersed polychaete species are unlikely to have arisen through gametic compatibility arms-races. *Evolution (N Y)* 62:3041–3055. doi:10.1111/j.1558-5646.2008.00521.x
- Swofford DL (2002) Phylogenetic Analysis Using Parsimony. *Options* 42:294–307. doi:10.1007/BF02198856
- Titus BM, Daly M (2015) Fine-scale phylogeography reveals cryptic biodiversity in Pederson's cleaner shrimp, *Ancylomenes pedersoni* (Crustacea: Caridea: Palaemonidae), along the Florida Reef Tract. *Mar Ecol*. doi:10.1111/maec.12237
- Towns J, Cockerill T, Dahan M, et al. (2014) XSEDE: Accelerating Scientific Discovery. *Comput Sci Eng* 16:62–74
- Untergasser A, Cutcutache I, Koressaar T, et al. (2012) Primer3- new capabilities and interfaces. *Nucleic Acids Res* 40:1–12. doi:10.1093/nar/gks596
- Warren DL, Geneva AJ, Lanfear R, Rosenberg M (2017) RWTY (R We There Yet): An R package for examining convergence of Bayesian phylogenetic analyses. *Mol Biol Evol* 34:1016–1020. doi:10.1093/molbev/msw279
- Watson GJ, Langford FM, Gaudron SM, Bentley MG (2000) Factors influencing spawning and pairing in the scale worm *Harmothoe imbricata*. *Biol Bull* 199:50–58
- Weigert A, Helm C, Meyer M, et al. (2014) Illuminating the base of the Annelid tree using transcriptomics. *Mol Biol Evol* 31:1391–1401. doi:10.1093/molbev/msu080
- Weigert A, Bleidorn C (2016) Current status of annelid phylogeny. *Org Divers Evol*. doi:10.1007/s13127-016-0265-7

- Will KW, Rubinoff D (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20:47–55. doi:10.1111/j.1096-0031.2003.00008.x
- Wilson HW (1991) Sexual reproduction modes in polychaetes: classification and diversity. *Bull Mar Sci* 48:500–516
- Winnepenninckx B, Thierry B, De Wachter R (1995) Phylogeny of Protostome Worm Derived from 18S rRNA Sequences. *Mol Biol Evol* 12:641–649
- Winnepenninckx BMH, Van Peer YDE, Backeljau T (1998) Metazoan relationships on the basis of 18S rRNA sequences: A few years later. *Am Zool* 38:888–906. doi:10.1093/icb/38.6.888
- WoRMS Editorial Board (2018). World Register of Marine Species. Available from <http://www.marinespecies.org> at VLIZ. Accessed 2018-08-10. doi:10.14284/170