Diversity and evolution of annelids with emphasis on Spionidae

by

Viktoria Bogantes Aguilar

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 8th, 2020

Keywords: taxonomy, COI, Antarctica, Iceland, Phylogenomics, Spionidae, biodiversity

Approved by

Kenneth M. Halanych, Chair, Professor of Biology
Jonathan W. Armbruster, Professor of Biological Sciences and Curator of Fishes
Leslie R. Goertzen, Associate Professor of Biological Sciences
Kelly M. Dorgan, Assistant Professor of Marine Sciences, University of South Alabama

Abstract

Annelida is a diverse phylum with over 21,000 species described that occupy a variety of environments including marine, terrestrial, and fresh water. This group includes earthworms, leeches, and marine segmented worms. Although, there are some exceptions in which segmentation is highly modified (e.g., siboglinids), incomplete (e.g., echiurids), or uncertain (e.g., sipunculans). Sequencing data and phylogenomic analyses in the last decades provided a well-supported annelid backbone tree, but many areas of the annelid tree are still poorly resolved. The goal of this dissertation is to integrate and develop methods to study morphological and molecular diversity in annelids.

This study contains four research chapters. The first two chapters explore the use of an integrative taxonomic approach that involves a combination of morphological data, COI barcoding, and whole mitochondrial genomes to assess diversity and biogeography of particular annelid groups. More specifically, I examined morphological and genetic diversity of the scale worm *Polyeunoa laevis* (Polynoidae), and its association with the soft coral *Thouarella* (Primnoidae) from the Southern Ocean and found that for both taxonomic groups biodiversity in the Southern Ocean appears to be underestimated. Following the integrative approach, I studied the diversity and phylogenetic relationships of *Laonice* (Spionidae) from Iceland, a group of spionid-annelids with a higher diversity in deep waters. Although this study was hampered by quality of preservation of animals, several species that were previously known to science were recovered (e.g., *Laonice blakei*, *Laonice sarsi*, *Laonice* cf. *norgensis*, and *Laonice cirrata*) and

one new species was described (*Laonice plumisetosa*). Moreover, this study provides a preliminary phylogeny for *Laonice*.

The third research chapter provides a description of the development of musculature, nervous system, and cilitiation patterns during larval formation of *Pseudopolydora paucibranchiata* using confocal laser scanning microscopy. In contrast to other planktotrophic annelid larvae, *P. paucibranchiata* shows a simultaneous development of muscles associated with the body wall and digestive system. Ciliation is extensive and includes multiple ciliary cells around the head, stomodeum and gut, and on the pygidium. Interestingly, no apical tuft is distinguished. The location of the first serotonergic cells could not be determined, as early larval stages showed no serotonergic activity. Thus, to study the development of the nervous system in spionid larvae, other markers will certainly be required.

Lastly, phylogenetic relationships of Spionidae are explored using a combination of WGS data and previously collected transcriptome data. Taxon sampling includes 21 Operational Taxonomic Units (OTUs) representing 17 spionid genera, as well as 1 trochochaetid, and 2 sabellarid annelids as the outgroups. Maximum likelihood phylogenetic recovered two major groups with high support. Clade 1 includes the *Polydora*-complex, *Pygospio*, *Spio*, *Scolecolepides*, *Scolelepis*, *Dispio*, *Lindaspio* and *Rhynchospio*. Clade 2 includes the *Prionospio*-complex, *Aonides*, *Spiophanes*, and *Laonice*. Importantly, *Trochochaeta* was recovered as the sister taxon to all spionid taxa, directly contrasting previous phylogenetic studies which suggested Spionidae represented a paraphyletic group. In addition to supporting Spionidae as monophyletic using WGS datasets, this study provides a well-resolved backbone within the family for the evolutionary relationships of 17 spionid genera, yielding a dataset which is highly reproducible and can be readily appropriated for future studies.

Acknowledgments

I want to thank my advisor Ken Halanych for giving me the opportunity to join his lab, and for being such a supportive advisor. While working with Ken I had the opportunity to be involved in international collaborations, as well as attending many scientific meetings. Working in his lab has been an amazing experience and I am very grateful for it. I also want to thank my committee members: Leslie Goertzen, Jonathan Armbruster and Kelly Dorgan, for their time, advices, and especially their motivation after every committee meeting.

I have also been extremely fortunate to have amazing lab mates, and I want to thank all the past and current members of the Molette Lab for their support in the lab and life in general: Michael Tassia, Pamela Branock, Nathan Whelan, Matthew Galaska, Yuanning Li, Susan Rashid, Elisa Acosta, Kyle David, Caitlin Redak, Oluchi Oyekwe, Yu Sun.

To my favorite labmate and partner Damien Waits thank you for all the support, for the good times, and for holding me during the difficult times, I would have not done it without your help.

I want to thank my Auburn friends Katherine Mincey, Stephen Sefick, and Juanita Rodriguez for their support in many ways.

I also thank Karin Meissner for being my spionid mentor, for her guidance and patience, and for helping me to get more familiar with the group. I also thank her for receiving me in her lab, and for all the worms contributed to this dissertation!

I want to thank all the people that sent me their spionids: Jason Williams, Andrew Davis, Mauricio Shimabukuru, Cinthya Santos, Allan Carrillo, Liron Goren, William Walton and Scott Rikard. I also thank Christina Zakas for sharing *Streblospio* genome data.

I want to thank Kevin Kocot for allowing me to use his SEM, for sharing his deep sea spionids with me, and one very important transcriptome, in addition to all his useful suggestions with phylogenomics. I also thank Nicole Garrison for her help with discussions on phylogenetic methods.

This study was funded by numerous funding sources: WormNet II, Link Foundation, Smithsonian Marine Station at Fort Pierce, Auburn University, Auburn University Museum of Natural History, IceAGE, and National Science Foundation.

Many thanks to my family, for always supporting me to keep studying even when I told them I was going to be working with worms. Thank you for the endless support.

This dissertation is dedicated to Damien and Elliot.

Table of Contents

List of Tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations

ACC	Antarctic circumpolar current
AMOVA	Analysis of molecular variance
APF	Antarctic polar front
AUMNH	Auburn University Museum of Natural History
BI	Bayesian Inference
CLSM	Confocal laser scanning microscopy
COI	Cytochrome c oxidase subunit I
DIC	Differential interference contrast
IceAGE	Icelandic marine Animals: Genetics and Ecology
mtDNA	Mitochondrial DNA
OG	Orthology group
PCR	Polymerase chain reaction
SEM	Scanning electron microscopy

Chapter 1: Introduction

1.1 General introduction and background

Annelida, or segmented worms, is a diverse phylum that includes earthworms, leeches, and several marine worms, and currently comprises approximately 21,000 species described (Purschke et al. 2020). Annelids can be found worldwide, occupying a variety of environments, including marine, freshwater, and terrestrial habitats. They range in size from less than 0.5 mm to more than 3 m and have a variety of feeding modes including carnivores, deposit feeders, filter feeders and herbivores (Jumars et al. 2014).

Annelids are one of the few groups besides arthropods and chordates with true segmentation (Seaver 2003); although, there are numerous exceptions in which segmentation is highly modified (e.g. siboglinids), incomplete (e.g. echiurids, and several meiofaunal taxa), or uncertain (e.g. Sipuncula), these exceptions have led to problems in recognizing such groups as annelids (Halanych et al. 2002). The generalized annelid body form is composed of the head with a prostomium and peristomium, a segmented trunk (except for cases previously mentioned), and a pygidium. In most marine representatives the trunk has lateral appendages that bear bundles of chaetae. However, some annelids are highly specialized, displaying different body and larval forms (Rouse 1999).

Traditionally, annelids were classified into Polychaeta (marine segmented worms) and Clitellata (earthworms and leeches). However, phylogenetic analyses in the last decades have changed views of annelid systematics. Former phyla including Pogonophora and Vestimentifera (now Siboglinidae), Echiura, and more recently Sipuncula are now part of annelids (McHugh 1997; Halanych et al. 2002; Weigert et al. 2014). Molecular and morphological data suggested that the groups mentioned above including clitellates to be nested within Polychaeta, indicating

that “Polychaeta” is paraphyletic and now is recognized as a synonym of annelids (Rouse and Fauchald 1997; Struck et al. 2007; Struck et al. 2011; Weigert et al. 2014). At this time, the term polychaete represents an informal name and not a valid taxonomic rank. Recently, sequencing data and phylogenomic analyses provided a well-supported annelid backbone tree (Struck et al. 2007; Struck et al. 2011; Weigert et al 2014; Andrade et al. 2015; Struck et al. 2015). Current classification of annelids consists of two main clades, Errantia and Sedentaria, and together form Pleistoannelida. Only a few lineages fall outside Pleistoannelida, and they constitute the sister taxa to all other annelids (Struck et al. 2011; Weigert et al 2014).

Although higher-level annelid phylogeny is coming into focus, many areas of the annelid tree are poorly resolved. One main gap is the lack of research on Spionidae phylogeny. Spionids represent one of the largest groups within annelids, and currently 590 species and 38 genera are recognized (Blake et al. 2020). They can be recognized based on the presence of two long palps and shape of the anterior end (Fig. 1). They are commonly found in shallow benthic sediments but are also present in the deep ocean. In addition, some species bore into calcareous substrates and are sometimes considered pests in commercially important mollusks (Simon et al. 2009; Sato-Okoshi et al., 2017; Martinelli et al. 2020). Spionids have over 10 different reproductive modes that includes sexual and asexual reproduction, and it is the only annelid group known to have diverse methods of larval development (e.g. planktotrophic, lecithotrophic) within the same species, a term known as poecilogony (Wilson 1991).

The most recent phylogenetic analysis (Blake and Arnofsky 1999) is based on a parsimony analysis using 38 characters based on reproduction, larval development, and adult morphology spanning 36 taxa. This analysis recovered spionids as a paraphyletic group, with other recognized families being nested within spionids like Trochochaetidae, Poecilochaetidae,

and Uncispionidae (Blake and Arnofsky 1999). Given their diversity and global distribution, spionids represent a valuable group to test previous hypotheses on the evolution of species complexes, evolutionary transitions of their diverse reproductive processes, and life history traits.

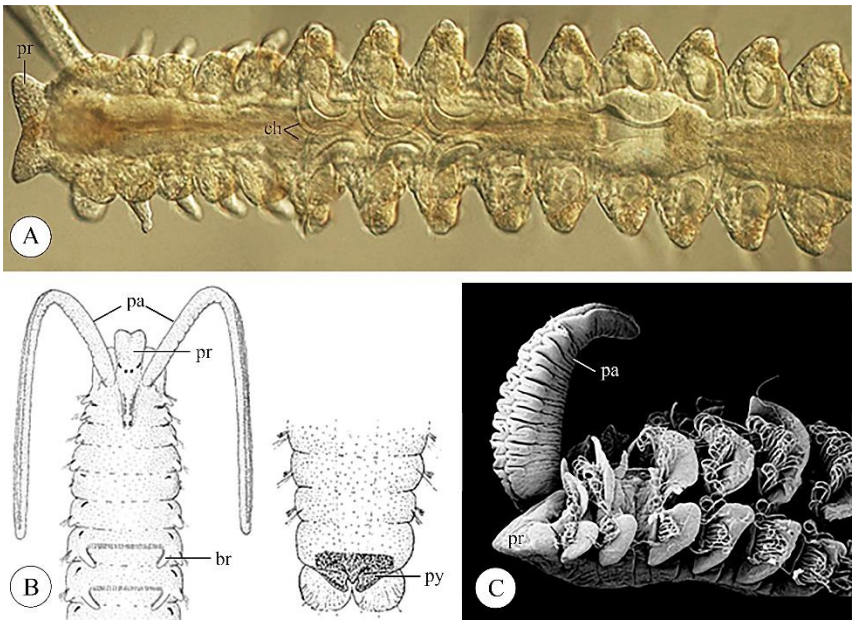

Figure 1. General morphology and anatomy of spionids. A. *Spiophanes duplex* juvenile (modified from Radashevsky 2012). B. *Polydora quadrilobata*, anterior and posterior region (modified from Blake 1969). C. *Laonice cirrata*, anterior region (modified from Radashevsky 2012) br: branchia; ch: chaeta; pa: palp; pr: prostomium; py: pygidium.

However, there are many challenges when working with this group; most spionids are small (less than 1 mm to 5 cm), fragile worms that fragment very easily when handled. Their identification is based on morphological characters that specimens frequently lose at the time of collection (e.g. palps, branchiae; Fig. 1), which is exacerbated by several collection methods (e.g. trawls, dredges, grabs) that tend to retrieve incomplete and damaged specimens. In fact, it is not uncommon that many species descriptions are based only on anterior fragments (Bogantes et al., 2018; Sikorski and Pavlova 2016; Meißner et al. 2014). The lack of descriptions of complete specimens make spionid taxonomy complicated because characters from posterior regions cannot

be described, and morphologically similar taxa are hard to separate even to the expert eye (Radashevsky et al. 2014).

Given the challenges of working with this group, to better understand Spionidae diversity current and future studies should include, when possible, integrative approaches. For example, molecular barcodes can diminish some of the taxonomic difficulties, but such barcodes need to be available and reliable. Additionally, molecular barcodes usually involve sequencing one or a few markers (e.g. COI, 18S, 28S) and under an integrative framework they have proven to be a valuable tool for assisting work with problematic taxa such as species complexes, invasive species, deep sea organisms with limited sampling, unrecognized diversity etc (Meiβner et al. 2014; Abe et al. 2016; Sato-Okoshi et al. 2017). Molecular barcodes usually involve an initial amplification of the fragment of interest with PCR followed by Sanger sequencing. One main advantage of this method is that high molecular weight DNA is not required, thus facilitating its usage when working with difficult DNA templates. However, molecular barcoding sequencing has limited resolution for resolving deep level phylogenies.

Advances of high throughput sequencing have facilitated large scale sequencing for non-model organisms, increasing the power to resolve phylogenies. Numerous phylogenetic studies have used techniques to reduce or partition the complexity of genomes such as transcriptome sequencing (Borner et al. 2014; Kocot et al. 2020; Stiller 2020), and target capture (Lemmon et al. 2012; Hedtk et al. 2013; McCormack et al. 2013) allowing sequencing of selected subsets of the genome, while reducing sequencing and computational costs involved with *de novo* whole genome sequencing (WGS). Fortunately, newer platforms (e.g. Hiseq X Ten, NovaSeq 6000) have reduced sequencing costs, making them comparable to approaches that only sequence a portion of the genome (Zhang et al. 2019). It is expected that continued decline of sequencing

costs (Wetterstrand 2020) should facilitate implementation of WGS for phylogenomic studies in a broad variety of taxa including spionids.

A goal of this research is to integrate and develop methods to study morphological and molecular diversity and evolution in annelids with an emphasis on Spionidae. Interestingly, this is one of the most abundant and diverse families of annelids, although before this study no molecular phylogenetic work has been conducted for the family. Moreover, data generated from this research will significantly increase future genomic resources for Spionidae.

References

Abe H, Kondoh T, Sato-Okoshi W. First report of the morphology and rDNA sequences of two *Pseudopolydora* species (Annelida: Spionidae) from Japan. *Zoological Science*. 2016;33:650–8.

Andrade, S.C.S., Novo, M., Kawauchi, G.Y., Worsaae, K., Pleijel, F., Giribet, G. & Rouse, G.W. 2015. Articulating Barchiannelids: phylogenomics and annelid relationships, with emphasis on meiofaunal taxa. *Molecular Biology and Evolution*, 32: 2860-2875.

Blake, J.A. & Arnofsky, P.L. 1999. Reproduction and larval development of the spioniform Polychaeta with application to systematics and phylogeny. *Hydrobiologia* 402: 57–106.

Blake, J.A., Macioleck, N.J. & Meiβner, K. 2020. Spionidae Grube, 1850. In: Handbook of Zoology. Annelida. Volume 2: Pleistoannelida, Sedentaria II. Purschke, G., Böggemann, M., Westheide, W. (eds.) DeGruyter, Berlin.

Bogantes V.E, Halanych, K.M, Meißner, K. Diversity and phylogenetic relationships of North Atlantic *Laonice* Malmgren, 1867 (Spionidae, Annelida) including the description of a novel species. *Marine Biodiversity*, 2018; 48:737–749.

Borner, J. Peter, R., Schill, R.O., Ebersberger, I. & Burmester, T. 2014. A transcriptome approach to ecdysozoan phylogeny. *Molecular Phylogenetics and Evolution*, 80: 79-87.

Eilertsen, M.H., Kongsrud, J.A., Alvestad, T., Stiller, J., Rouse, G.W. & Rapp H.T. Do ampharetids take sedimented steps between vents and seeps? Phylogeny and habitat-use of Ampharetidae (Annelida, Terebelliformia) in chemosynthesis-based ecosystems. *BMC Evolutionary Biology*, 2017;17(1):222.

Halanych, K.M., Dahlgren, T.G. & McHugh, D. 2002. Unsegmented annelids? Possible origins of four lophotrochozoan worm taxa. *Integrative and Comparative Biology*, 42(3):678-684. doi:10.1093/icb/42.3.678

Hedtk, S.M., Morgan, M.J., Cannatella, D.C. & Hillis, D.M. 2013. Targeted Enrichment: Maximizing Orthologous Gene Comparisons across Deep Evolutionary Time. *Plos one* 8(7):e67908.

Jumars P.A., Dorgan K.M. & Lindsay S.M. 2014. Diet of worms emended: an update of polychaete feeding guilds. *Annual Review of Marine Science*, 7: 497-520.

Kocot, K.M., Poustka, A.J., Stöger, I., Halanych, K.M. & Schrödl, M. 2020. New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. *Scientific Reports* 10:101.

Lemmon, A.R., Emme, S.A., Lemmon, E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61:727–744.

Li, Y., Kocot, K.M., Whelan, N.V., Santos, S.R., Waits, D.S., Thornhill, D.J. & Halanych, K.M. 2017. Phylogenomics of tubeworms (Siboglinidae, Annelida) and comparative performance of different reconstruction methods. *Zoologica Scripta* 46: 200-213.

Martinelli, J.C., Lopes, H.M., Hauser, L., Jimenez-Hidalgo, I. King, T.L., Padilla-Gamiño, J.L., Rawson, P., Spencer, L.H., Williams, J.D. & Wood, C.L. 2020 Confirmation of the shell-boring oyster parasite *Polydora websteri* (Polychaeta: Spionidae) in Washington State, USA. *Science Reports*, 10: 3961.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS ONE*, 8 e54848.

McHugh, D. 1997. Molecular evidence that echiurans and pogonophorans are derived annelids. *Proceedings of the National Academy of Sciences of the United States of America*, 94: 8006–8009.

Meißner, K., Bick, A., Guggolz, T. & Götting, M. 2014. Spionidae (Polychaeta: Canalipalpata: Spionida) from seamounts in the NE Atlantic. *Zootaxa* 3786:201–245.

Purschke, G., Böggemann, M., & Westheide, W. 2020. Handbook of Zoology: Annelida, 2: Annelida Basal Groups and Pleistoannelida, Sedentaria II. De Gruyter, Berlin.

Radashevsky, V.I, Neretina, T.V., Pankova, V.V. & Nuzhdin, S.V. 2014. Molecular identity, morphology and taxonomy of the *Rhynchospio glutaea* complex with a key to *Rhynchospio* species (Annelida, Spionidae). Syst Biodivers 12:37–41.

Rouse, G.W, Pleijel, F. 2001. Polychaetes. Oxford University Press, Oxford.

Rouse, G.W. (1999). Trochophore concepts: ciliary bands and the evolution of larvae in spiralian Metazoa. *Biological Journal of the Linnean Society*, *66*(4), 411-464.

Rouse, G.W., & Fauchald, K. (1997). Cladistics and polychaetes. *Zoologica Scripta*, *26*(2), 139-204.

Sato-Okoshi W, Abe H, Nishitani G, Simon CA. And then there was one: *Polydora uncinata* and *Polydora hoplura* (Annelida: Spionidae), the problematic polydorid pest species represent a single species. *Journal of the Marine Biological Association of the United Kingdom*. 2017;97:1675–84.

Seaver, E.C. 2003. Segmentation: mono- or polyphyletic? *The International Journal of Developmental Biology*, 47: 583–595.

Sigvaldadóttir, E., Mackie, A.S. & Pleijel, F. 1997. Generic interrelationships within the Spionidae (Annelida: Polychaeta). *Zoological Journal of the Linnaean Society*, 119: 473–500.

Sikorski, A.V. & Pavlova, L. 2016. Three new species of *Laonice* (Polychaete: Spionidae) from West and Southwest Africa. *Zootaxa* 4097:353–368.

Simon, C.A. 2009 *Pseudopolydora* species associated with mollusc shells on the south coast of South Africa, with the description of *Ps. dayii*, sp nov. *Journal of the Marine Biological Association of the United Kingdom*. 89:681–7.

Stiller, J., Tilic, E., Rousset, V., Pleijel, F. & Rouse, G.W. 2020. Spaghetti to a Tree: A Robust Phylogeny for Terebelliformia (Annelida) Based on Transcriptomes, Molecular and Morphological Data. *Biology*, *9*, 73.

Struck, T.H., Golombek, A., Weigert, A., Franke, F.A., Westheide, W., Purschke, G., Bleidorn, C. & Halanych, K.M. 2015. The evolution of annelids reveals two adaptive routes to the interstitial realm. *Current Biology*, 25: 1993–1999.

Struck, T.H., Paul, C., Hill, N., Hösel, C., Kube, M., Lieb, B., Meyer, A., Tiedemann, R., Purschke, G. & Bleidorn, C. 2011. Phylogenomic analyses unravel annelid evolution. Nature 471: 95-98.

Struck, T.H., Schult, N., Kusen, T., Hickman, E., Bleidorn, C., McHugh, D. & Halanych, K.M. 2007. Annelid phylogeny and the status of Sipuncula and Echiura. *BMC Evolutionary Biology*, 7, 57.

Weigert, A., Helm, C., Meyer, M., Nickel, B., Arendt, D., Hausdorf, B., Santos, S.R., Halanych, K.M., Purschke, G., Bleidorns, C. & Struck, T.H. 2014. Illuminating the base of the annelid tree using transcriptomics. *Molecular Biology and Evolution*, 31: 1391–1401.

Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata. Accessed May-10-2020.

Wilson, W.H. 1991. Sexual reproductive modes in polychaetes: classification and diversity. *Bulletin of Marine Science*, 48: 500–516.

Zhang, F., Ding, Y. H., Zhu, C. D., Zhou, X., Orr, M. C., Scheu, S., & Luan X.Y. 2019. Phylogenomics from low-coverage whole-genome sequencing. Methods in Ecology and Evolution, 10, 507–517.

Chapter 2: Unrecognized diversity of a scale worm, *Polyeunoa laevis* (Annelida: Polynoidae), that feeds on soft coral

2.1 Abstract

A goal of systematics is to develop a classification based on phylogeny that helps biologists see with evolutionary history of the morphological and genetic diversity of organismal lineages However, morphological and genetic diversity may not always be concordant, leading to challenges in systematics. The scale worm *Polyeunoa laevis* has been hypothesized to represent a species complex based on morphology, although there is little knowledge of its genetic diversity. Commonly found in Antarctic waters and usually associated with gorgonian corals (especially *Thouarella*), this taxon is also reported from the southwest Atlantic, Magellanic and sub-Antarctic regions. We employ an integrative taxonomic approach to examine the traditional morphological characters used for scale worm identification in combination with COI mitochondrial gene data and whole mtDNA genomes. Moreover, we consider *P. laevis*'s association with *Thouarella* by examining data from the mMutS gene, a soft-coral phylogenetic marker. Analyses for *P. laevis* recovered 3 clades, two in Antarctic waters and one from the Argentina-Indian Ocean. Interestingly, genetic and morphological results show differences between specimens from South Argentina and the Antarctic region, suggesting that open ocean barriers might have limited gene flow from these regions. Bayesian phylogenetic analyses for *Thouarella* resulted in at least 12 lineages, although some of the lineages consist of only a single individual. Our results show different evolutionary histories for both species, confirming that association between these scale worms and their hosts is not restrictive. For both taxonomic groups, biodiversity in the Southern Ocean appears to be underestimated.

2.2 Introduction

The Southern Ocean consists of approximately eight percent of global ocean surface area, and waters surrounding Antarctica are estimated to sustain around five percent of all marine diversity (Barnes & Peck, 2008; Poulin et al., 2014). Several studies on a wide variety of taxa including isopods, molluscs, pycnogonids, and echinoderms (Brandt et al., 2007; Galaska, Sands, Santos, Mahon, & Halanych, 2017a; Held & Leese, 2007; Hunter & Halanych, 2008; Mahon, Thornhill, Norenburg, & Halanych, 2010; Munilla & Soler Membrives, 2009) have shown high levels of endemism and species radiations, which may have been driven by the glacial and oceanographic history of the Southern Ocean (e.g., Antarctic Circumpolar Current and Antarctic Polar Front) (Baird, Miller, & Stark, 2011; Halanych & Mahon, 2018; Thatje, Hillenbrand, & Larter, 2005). Additionally, molecular studies have revealed the presence of previously unrecognized diversity in numerous Antarctic taxonomic groups (Allcock et al., 2011; Brokeland & Rapauch, 2008; Hunter & Halanych, 2008; Janosik & Halanych, 2010; Linse, Cope, Lörz, & Sands, 2007; Mahon et al., 2010). More recently, sampling of deep sea Antarctic polychaetes recovered cryptic diversity in 50% of species sampled based on comparison of mitochondrial DNA (Brasier et al., 2016), suggesting that diversity for this group is highly underestimated. Systematic studies on annelids from Antarctica have been mostly dependent on morphology (Blake, 2015; Parapar & Moreira, 2008), but including molecular data will aid to improve understanding of species boundaries.

The scale worm *Polyeunoa laevis* McIntosh, 1885 (Polynoidae) offers an example of a species whose taxonomy may not match recognized morphological variation and genetic diversity. This species was first described from Prince Edward Island in the Indian Ocean but was also reported during the same expedition in the Strait of Magellan (McIntosh 1885). Records

of *P. laevis* also include the Southern Ocean, sub-Antarctic waters, and southwest Atlantic (Barnich, Gambi, & Fiege, 2012; Pettibone, 1969; Stiller, 1996). In addition to their wide distribution, previous studies (Barnich et al., 2012; McIntosh, 1885) have shown that morphological characters used to identify *P. laevis* show considerable variation. Presence of cephalic peaks, coloration (Figure 1 A-C), and number and arrangements of pairs of elytra have created confusion when identifying closely related species of scale worms with similar morphology. Further complicating identification, the original descriptions of *Polyeunoa* species were ambiguous to the point that past workers have stated that *P. laevis* could not be identified based on information in the literature (see Barnich et al. 2012). To help with this situation, Barnich et al. (2012) proposed a redescription for *P. laevis*, identifying this species based on the tip of the neuropodial acicular lobe not being extended to supra-acicular process and most of the neurochaetae being unidentate (Figure 1 D-F). Their description facilitated the identification of *P. laevis*, although they hypothesized that this taxon might represent a species complex (i.e. two or more species previously classified as one species) (Barnich et al. 2012) of *P. laevis* and undescribed species.

Polyeunoa laevis is considered a feeding specialist, feeding primarily on gorgonian cnidarians (Stiller, 1996) including *Thouarella*, a widespread soft-coral species commonly called the bottlebrush corals. Even though this genus can be very common, especially in Antarctic and sub-Antarctic waters, species identification and delimitation is problematic (Taylor, Cairns, Agnew, & Rogers, 2013) as morphological features are variable. The association between *P. laevis* and *Thouarella* species has been documented in several studies (Barnich et al., 2012; Pettibone, 1969; Stiller, 1996). Hartmann-Schröeder (1989) reported finding a worm on every *Thouarella* collected in her study, and they proposed that this association afforded the worms

protection from predators. However, no study has examined whether species-level diversity of *Polyeunoa* corresponds to species-level diversity of *Thouarella*.

Given the distribution of *P. laevis*, we examined the following questions: 1) Does current taxonomy accurately reflect morphological diversity? 2) Are the morphological features described in the literature useful for characterizing diversity? 3) Lastly, given the feeding association, do *P. laevis* and *Thouarella* have similar biogeographic patterns in Antarctica? To this end we employ an integrative taxonomic approach that examines the traditional morphological characters used for scale worm identification in combination with molecular markers. Moreover, we consider *P. laevis*'s relationship to *Thouarella* species in the Southern Ocean.

2.3 Materials and Methods

2.3.1 *Sample collection*

Specimens were collected during four expeditions (i.e. *RVIB Nathaniel B. Palmer* NBP12-10, *R/V Laurence M. Gould* LMG13-12, LMG04-14, and LMG06-05) to South America and Antarctica (Figure 2a). *Polyeunoa laevis* was collected from 1 station in South America and 31 stations from the Antarctic waters including the Antarctic Peninsula and Weddell, Bellingshausen, Amundsen, and Ross Seas (Figure 2a; Table 1; Supporting Information Table S1). *Thouarella* specimens were collected from 32 stations in Antarctic waters (Figure 2a; Table 1; Supporting Information Table S2). Samples were obtained with a Blake trawl. Samples were frozen after collection at -80 °C, preserved in ~95% ethanol, or 4% formalin.

2.3.2 *Morphological examination*

A total of 111 *Polyeunoa* worms were initially examined individually under a Leica MDG41 dissecting scope. Identifications were based on the morphological characters outlined in Barnich et al. (2012), including number of elytral insertions, presence of cephalic peaks, subacicular process, and type of chaetae (i.e., unidentate or slightly bidentate). Most worms were disassociated from their soft coral host because *Polyeunoa* tended to fall off of soft coral when collected by dredging and because of the specimen preservation workflow used on the ship. However, 16 of the 111 sampled worms were retrieved directly from known *Thouarella* specimens sampled for this study.

A total of 82 *Thouarella* individuals were identified based on descriptions from Taylor et al. (2013). We documented sclerite and polyp morphology of a representative subset of species using scanning electron microscopy (SEM). Individuals examined on SEM were chosen based on preservation quality of morphological features while also trying to analyze at least one individual per clade (see below). Selected specimens were washed in 30% ethanol and then transferred to successively increasing dilutions up to 100% ethanol before being air dried. Dry specimens were mounted and sputter coated using gold prior to analysis on a Zeiss EVO 50 SEM. After placing morphological examinations in the context of molecular phylogenetic results (see below), we determined that morphological characters alone were not sufficient for species-level identification of *Thouarella*. Thus, species examined here are identified by genetic/phylogenetic similarity to specimens identified and sequenced in other studies, or in the absence of genetically similar individuals from past studies, as "*Thouarella* sp.".

P. laevis and *Thouarella* specimens are deposited at Auburn University Museum of Natural History (AUMNH; Supporting Information Tables 1-2).

2.3.3 *Molecular Data*

One millimeter tissue clips were taken from individual *Polyeunoa* specimens for DNA extraction. Similarly, 3 mm clips were taken from *Thouarella* stalks. For both taxa, whole genomic DNA was obtained using the Qiagen DNAeasy® Blood and Tissue Kit. The standard barcoding region of COI was amplified and sequenced for all *Polyeunoa* using primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994). Polymerase chain reaction (PCR) mix for 25 µl reaction consisted of 1 µl of DNA template, 2.5 µl $Mg(OAc)_2$ (25 mM), 2.5 µl Taq buffer (10X), 2.5 µl dNTPs (10 mM), 1 µl of each primer (10 µM), 0.3 µl Taq DNA polymerase (25mM), and 14.2 µl water (ddH_2O). PCR cycling protocol consisted of denaturation at 96 °C for 4 min followed by 35 cycles of 94 °C for 30 s, 54 °C for 30 s and 72 °C for 1 min, followed by a final elongation at 72 °C for 8 min.

Octocorals have a lower rate of mitochondrial substitution compared to other metazoans (France & Hoover, 2002; McFadden et al., 2011; Shearer, Oppen, Romano, & Worheide, 2002) and thus we used a portion of mtMutS gene for *Thouarella* samples. The mtMutS gene fragment has been proposed to have an appropriate rate of substitution for octocoral species delimitation (McFadden, Ofwegen, Beckman, Benayahu, & Alderslade, 2009; McFadden et al., 2011). PCR used the primers ND42599F (5'-GCCATTATGGTTAACTATTAC-3') (France & Hoover, 2002) and MUT-2458R (5'-TSGAGCAAAAGCCACTCC-3') (McFadden et al., 2006) and a cycling protocol consisted of initial denaturation at 94°C, 35 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min followed by a final extension step of 72° for 5 min. For all PCR samples, products were purified using the Qiagen QIAquick PCR Purification Kit. Purified PCR

products were bidirectionally Sanger sequenced by GENEWIZ Inc. Sequences were assembled and proofread by eye with Geneious R6.

Partial COI sequences for *Polyeunoa* were aligned using default parameters of MUSCLE (Edgar, 2004) and visually inspected (GenBank accession numbers MK593024-MK593134; Supporting Information Table S1). Partial mtMutS sequences for *Thouarella* were aligned with MACSE (Ranwez, Harispe, Delsuc, & Douzery, 2011) using the alignSequences command and the coelenterate mitochondrial code (NCBI code 4). Aligned COI and mtMutS sequences were translated using the invertebrate mitochondrial code and coelenterate mitochondrial code respectively to assure that stop codons or frameshift mutations were not present in alignments (GenBank accession numbers MN121855-MN121936; Supporting Information Table S2). Available sequences from GenBank for *P. laevis* and *Thouarella* were retrieved and included in the analyses (Supporting Information Table S1; Supporting Information Table S2).

2.3.4 *Mitochondrial genome of* Polyeunoa laevis

To further assess genetic variation within *Polyeunoa* individuals, the whole mitochondrial genome was sequenced for 3 specimens (AW14_733.7E, RS19_2375.1E2, NBI20_3227.1E3; Supporting Information Table S1). DNA extractions were performed as described above. Sequencing of total genomic DNA was performed by Novogene Inc. on an Illumina HiSeq 2500 platform, using 2 x 150 pair-end v4 chemistry. Paired-end reads were assembled *de novo* with Ray 2.2.0 (k-mer= 31; Boisvert, Raymond, Godzaridis, Laviolette, & Corbeil, 2012). Potential mitochondrial genomes were identified using BLASTn (Altschul et al., 1997), and previously published COI sequences of *P. laevis* (GenBank accession numbers KU738210-KU738214, KF713377) were used as bait against the assembled data (GenBank

accession numbers MN057924- MN057926). Annotation was conducted using MITOS2 web server (Bernt et al., 2013) and gene boundaries were manually verified. Uncorrected pairwise genetic distances (p) were estimated with MEGA7 (Kumar, Stecher & Tamura, 2016).

2.3.5 *Molecular analyses*

Bayesian inference of nucleotide data was used to reconstructed trees for *Polyeunoa* and *Thouarella* using MrBayes 3.2.6 (Huelsenbeck & Ronquist, 2001) using best-fit partitions as indicated by PartitionFinder (Lanfear, Calcott, Ho, & Guindon, 2012) (i.e., *Polyeunoa* COI: split by codon position; *Thouarella* mtMutS: single partition). Additionally, sequences of *Antarctinoe ferox, Neopolynoe paradoxa* (GenBank accession numbers KJ676611, KT592262; Polynoidae, Annelida), and *Calyptrophora* (GenBank accession numbers DQ297427, DQ234756, JX561183; Primnoidae, Cnidaria) were chosen to serve as outgroups for *P. laevis* and *Thouarella* respectively, based on current understanding of phylogeny (Cairns & Wirshing, 2018; Y. Zhang et al., 2018).

For each partition, a reversible jump Markov chain Monte Carlo (MCMC) was used for model averaging across the GTR models and rate heterogeneity for each partition was modeled using four discrete categories of a gamma distribution as indicated by PartitionFinder. Four independent runs with four Metropolis-coupled chains each were run for 10,000,000 MCMC generations and sampled every 1,000 generations. For each dataset, commands used in MrBayes analyses can be found in Supporting information Data S1. Stationarity of each run was checked with Tracer 1.7.1 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018), and 25% burn-in was determined to be appropriate for each analysis. Both analyses appeared to reach convergences as

all parameters had a potential scale reduction factor of 1.0. A 50% majority rule consensus tree was calculated from each analysis using the "sumt" command in MrBayes.

2.3.6 *Genetic distance, species delimitation and haplotypes*

Uncorrected pairwise genetic distances (*p*) were obtained with MEGA7. Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) was used to conduct a Tajima's D (Tajima, 1989) test to determine whether sequence variation fit the neutral mutation model. Neutrality tests such as Tajima's D can provide some insights about demographic forces affecting a population. Analyses of molecular (AMOVA) variance were also conducted with Arlequin 3.5.2.2 to test for genetic differentiation; regions were defined by sampling locality and then grouped by geographic regions (e.g., Amundsen, Antarctic Peninsula, Bellingshausen, Ross; Supporting Information Table S1). Haplotype diversity and nucleotide diversity were also estimated for each clade.

We used two methods to assess species boundaries for *Polyeunoa* and *Thouarella*. First, barcode gap discovery (ABGD; Puillandre, Lambert, Brouillet, & Achaz, 2012) was used to determine if a barcode gap in percent nucleotide difference existed with the datasets that could be used for species delimitation. ABGD default values were used for *P. laevis*, while 0.3% intraspecific cutoff value was used for *Thouarella* as suggested by Quattrini et al. (2019). However, the lack of a barcode gap in mtMutS (Wirshing & Baker, 2015) might limit resolution of ABGD for *Thouarella* species delimitation. We also used Poisson Tree Processes (PTP; Zhang, Kapli, Pavlidis, & Stamatakis, 2013) which uses non-ultrametric trees with a scale of expected substitutions/site. For this, 10,000 trees from the posterior distribution of MrBayes analyses (i.e., 2,500 from each of the four independent runs) were used as input.

A TCS network was constructed for the largest of three clades containing *Polyeunoa laevis* individuals (see below) was generated using PopART (http://popart.otago.ac.nz), which utilizes statistical parsimony (Clement, Posada, & Crandall, 2000).

2.4 Results

Given the focus of this study, results of both the morphological and molecular analyses for *Polyeunoa* will be described first, followed by results for *Thouarella*.

2.4.1 *Morphological comparisons*

Morphological characters showed different patterns between the *Polyeunoa laevis* from South Argentina and Antarctica. Individuals from South Argentina present between 12-18 alternate pairs of elytra after setiger 32 (determined by the presence of elytra insertions), whereas specimens from the Southern Ocean only have 0, 1 or 2 pairs of elytra after setiger 32, except for one specimen that showed 4 (Supporting Information Table S3). These comparisons were made for individuals with 60-80 segments. Specimens from Antarctica showed morphological variation, but consistent differences were not observed between individuals from different localities. Dorsal coloration for example varies from red horizontal lines, to one red longitudinal line (anterior to posterior region), or most of the dorsal region covered with a red-purple pigment (Figure 1 A-C); cephalic peaks were present in some specimen and absent in others (Supporting Information Table S3). Neurochaetae were also examined for each specimen and were found to be consistent with previous descriptions, with most of the worms having unidentate neurochaetae and some with bidentate tips (Figure 1 D-F). The chaetae, parapodia from the anterior, middle and posterior region of 9 specimens (3 from each clade, see below) were more closely examined

with scanning microscopy (SEM), although no clear differences were observed (Supporting Information Figure S1).

Syntypes deposited at the Natural History Museum of London (NHM) were also examined for this study as a holotype is not available. Given the proximity with samples used in the COI analysis from Serpetti et al. (2016; Figure 2a), we expected the distribution of the elytra after setiger 32 to follow a similar pattern to specimens from southern Argentina and the Indian Ocean. However, the syntypes are morphologically more similar to the specimens from the Southern Ocean and only have a few (0-3) elytra scars after setiger 32.

2.4.2 *Molecular analyses*

The molecular dataset for *Polyeunoa laevis* consisted of a 657 bp fragment of the mitochondrial COI for 115 *Polyeunoa* individuals (including four sequences from GenBank that included 1 Southern Ocean sequence and 3 Indian Ocean sequences: (Gallego, Lavery, & Sewell, 2014; Serpetti et al., 2016; Supporting information Table S1). The dataset included 116 (18%) variable sites, and 82 (12.5%) parsimony informative sites. Out of the 115 sequences, 78 haplotypes were recovered, and only 17 haplotypes were sampled more than once. Bayesian inference recovered 3 main clades (Figure 2b). One clade included most of the samples from the East Antarctic Peninsula, Weddell and Ross Seas (Weddell-Ross), a second clade corresponds to the samples from Argentina and the Indian Ocean, and the third clade includes specimens from the West Antarctic Peninsula, Bellingshausen Sea Amundsen Sea, and Ross Sea (AP-Ross). The same analysis was conducted using midpoint rooting and the recovered topology was the same. Individuals from the two Antarctic clades lacked morphological differentiation. In contrast, the

clade from Argentina showed genetic and morphological distinction (see above morphological comparisons) when compared to the Antarctic clades.

2.4.3 *Mitochondrial genome*

A single contig representing the mitochondrial genome was recovered for all three samples of *Polyeunoa* sequenced in this study. These contigs ranged in length from 15118 to 15123 bp, and had an average GC content of 33.4% (Supporting information Tables S4). MITOS2 annotation recovered 13 protein coding, 22 tRNA and 2 rRNA genes. Composition and gene order (Figure 2c) is consistent with previously published mitochondrial genomes of annelids and specifically polynoids (Boore, Boore, & Brown, 2000; Jennings & Halanych, 2005; Vallès & Boore, 2006; Weigert et al., 2016; Y. Zhang et al., 2018). Uncorrected pairwise genetic distance (p) for complete mitochondrial genomes ranged around 4% among the genomes of representatives from AP-Ross and Argentina-Indian Ocean clades, and 7%, when comparing Weddell-Ross against the other clades (Supporting information Tables S4).

2.4.4 *Genetic distances, species delimitation and haplotype network*

For *Polyeunoa,* the average of uncorrected pairwise genetic distances (p) within groups were less than 1% (Supporting information Tables S4). Between clades, genetic distances ranged from 4-8%. The greater distances were found when comparing the Weddell-Ross clade with the two other clades. Tajima's D test was negative and significant for the AP-Ross clade (Supporting information Tables S4). Tajima's D values were not significant for the other clades. Significant negative values of Tajima's D are the result of rare polymorphism in a population, which could be an indication of purifying selection or a recent population size expansion (Tajima, 1989).

Both of the species delimitation methods, ABGD and PTP, suggest 3 distinct species that are congruent with the clades recovered with Bayesian analysis (Supporting Information Data S2).

Haplotype diversity (h) values were similar for each clade (Table 2). Nucleotide diversity (π) was higher in the AP-Ross clade (0.009), while Argentina-Indian Ocean clade showed the lowest value (0.001; Table 2). A haplotype network was constructed for the AP-Ross clade, and 59 haplotypes out of 90 individuals were identified, this group included most of the specimens from the Western Antarctic Peninsula, Bellingshausen and Amundsen Sea (Figure 3a), although no obvious population structure is observed. This is supported by the results from the AMOVA showing that 98.35% of the variation was found within regions (Supporting information Tables S4). Due to limited numbers of samples, haplotypes network and AMOVA analysis were not constructed for the other clades.

2.4.5 *Biogeography of* Polyeunoa laevis *and* Thouarella

Morphology of 26 *Thouarella* individuals was documented with SEM images (Supporting information Figure S2), although reliable morphological identification of species was not feasible (McFadden, France, Sánchez, & Alderslade, 2006; McFadden et al., 2006) as individuals with identical morphologies seen through SEM were sometimes recovered in different clades and individuals with disparate morphologies were recovered in the same clades. Thus, individuals were identified using a barcode approach. The molecular dataset for *Thouarella* consisted of 756 bp fragment of the mtMutS for 109 specimens, including 27 from GenBank (McFadden et al., 2006; McFadden et al., 2011; Taylor & Rogers, 2015). The dataset included 116 (15%) sites that were variable and 103 (13.6%) that were parsimony informative. Out of the 109 sequences, 35 haplotypes were recovered.

Species delimitation methods for *Thouarella* resulted in at least 12 lineages (Figure 3b), although some of the lineages consist of a single individual. ABGD partitioned the data into 12 different groups, while 27 were found with PTP (Figure 3b), although some partitions from the PTP have low support (Supporting Information Data S3). Additionally, various lineages that included previously identified *Thouarella* species clustered different species together (e.g., *T. variabilis* and *T. chilensis*), while other lineages remain as *Thouarella* sp. since they did not correspond to any previous *Thouarella* barcodes for the Southern Ocean. Given the small sample number of individual groups, we limit our interpretation of these results as resolving phylogenetic relationships among *Thouarella* species is beyond the scope of this study and will likely require several genomic markers (e.g., Quattrini et al., 2019).

Taken as a whole, results show different phylogeographic patterns for *P. laevis* and *Thouarella*. With respect to the worms retrieved directly from *Thouarella,* they were all found to be in the AP-Ross clade (Figure 2b); however, they inhabit different *Thouarella* lineages (Figure 3b). That is, *P. laevis* does not appear to be a species-level specialist in regards to *Thouarella* host/habitat preference although whether their preferences are limited remain unknown.

2.5 Discussion

Polyeunoa laevis represents at least 3 genetic lineages with two lineages from the Southern Ocean and a third from South America and the Indian Ocean. Distinct morphologies were found when comparing specimens from Argentina and the Southern Ocean, but the two genetic lineages from the Southern Ocean show no morphological differences. These results confirm the possibility that *P. laevis* represents a species complex (Barnich et al., 2012) and that current taxonomic delineation does not reflect the actual diversity within the group.

Although most polynoids are considered to have planktotrophic larvae with potential to disperse long distances (Giangrande, 1997; Wilson, 1991), there are no studies on the development of *P. laevis* and its larvae type is unknown. However, independent of the larval type, we found genetic connectivity over 7000 km range (Weddell-Ross lineage) suggesting the capacity of *P.laevis* to disperse long distances, even though barriers have limited dispersion of these organisms in and out of the Southern Ocean. Genetic breaks between South America and Southern Ocean fauna have been reported in similar studies. Thornhill et al. (2008) showed genetic differentiation of populations of the nemertean *Parborlasia corrugatus* in South America, Antarctic, and sub-Antarctic waters, by analyzing mitochondrial 16s rRNA and COI sequence data, and found two lineages to be geographically separated possibly by the Antarctic Polar Front (APF). One lineage included all organisms from Antarctic and sub-Antarctic region, and a second lineage including individuals from South Argentina only. Similarly, Shaw, Arkhipkin, & Al-Khairulla (2004) examined genetic structure of toothfish around the Southern Ocean and South Argentina and found that, despite the potential for high dispersal, there is genetic differentiation between Patagonian and Southern Ocean toothfish. These studies and others (e.g., Hunter & Halanych, 2008) suggested the APF as the main barrier restricting larval dispersion, and consequently limiting gene flow between the adults. Although, the APF is not completely impermeable, and exchange of organisms from Antarctic to South American waters have occurred promoting radiation of the group. Sands et al. (2015) for example, found molecular evidence of migration of the ophiuroid *Ophiura lymanii* moving from the Southern Ocean to South America (also see Galaska, Sands, Santos, Mahon, & Halanych, 2017b). As previously mentioned, *P. laevis* individuals from Argentinean waters do show morphological differences compared to Southern Ocean individuals. Unfortunately, specimens from the Indian

Ocean (Serpetti et al., 2016) could not be examined morphologically, and elytra pattern of these organisms remains unclear. Given this situation, we were not able to determine if the differences found in the distribution of the elytra can be used as a diagnostic character since organisms from Argentina and the Indian Ocean are clustered in the same genetic lineage. Additionally, syntypes from Prince Edward Island are morphologically consistent with Antarctic specimens, but whether these organisms belong to the Argentina-Indian Ocean clade or one of the Southern Ocean clades is unknown. Prince Edward Island is a sub-Antarctic Island located in the Indian region of the Southern Ocean (Ansorge, Froneman, & Durgadoo, 2012). We restrain from describing a new species until a more extensive morphological examination can be completed, especially from organisms from sub-Antarctic regions.

Interestingly, we found genetic connectivity between samples from Ross and Weddell Seas, two regions not currently connected. Previous studies have shown an affinity between the fauna of these regions, including gastropods, bivalves (Linse, Griffiths, Barnes, & Clarke, 2006) bryozoans (Barnes, & Claus-Dieter, 2010), and more recently the ophiuroid *Ophionotus victoriae* (Galaska et al., 2017a). Barnes & Claus-Dieter (2010) suggested that similarity between these regions is to some extent the result of a past connectivity between Ross and Weddell Seas, as a consequence of the collapse of the West Antarctic Ice Sheet (Scherer, Aldahan, Tulaczyk, Engelhardt, & Kamb, 2009), opening a direct trans-Antarctic passage that allowed the exchange of organisms between the West Antarctic waters. Evidence from sediment cores suggests a near complete collapse occurred ~1.1 MYA and modeling suggests a collapse as recent at 125 KYA (Feldmann & Levermann, 2015; Naish et al., 2009; Pollard & DeConto, 2009). This past connectivity would help to explain the occurrence of a geographically discontinuous Antarctic lineage for the *Polyeunoa* worms.

Current morphological examination of *P. laevis* cannot resolve whether the range of variation observed in morphological characters across Antarctic and sub-Antarctic specimens represent one or more species. For several taxa in the Southern Ocean, molecular results yield lineages often unrecognized by morphological characters (e.g., annelid *Glycera kerguelensis*, Schüller, 2011; notothenoid fish Matschiner, Hanel, & Salzburger, 2009; octopuses, Allcock et al., 2011; crinoids, Hemery et al., 2012; pycnogonids, Harder, Halanych, & Mahon, 2016). Several hypotheses have been proposed to account for this discrepancy between morphology and molecular data (Allcock & Strugnell, 2012; Halanych & Mahon, 2018; Thatje et al., 2005). Among the most commonly accepted is the idea that during glaciation events benthic taxa were forced into refugia to survive. Such refugia may have been in deeper water along the continental slope or in polynyas. The resulting isolation (Thatje et al., 2005) would have facilitated establishment of genetically isolated lineages by genetic drift, even though morphologies could remain similar (Janosik & Halanych, 2010). In the case of *Polyeunoa* lineages from the Southern Ocean, the morphology is similar despite genetic differences resulting in cryptic lineages.

Given the association between *P. laevis* and *Thouarella*, we questioned if these taxa would have similar phylogeographic patterns, but that is not the case based on the data collected for this investigation. *P. laevis* is considered a polyxenous polynoid as it has been found on other corals including *Dasytenella* and *Primnoisis* (Barnich et al., 2012; Serpetti et al., 2016; Stiller, 1996), or can be free-living even though the association with *Thouarella* might be more common (Hartmann-Schröeder, 1989). However, the specificity of *P. laevis* for *Thouarella* remains unclear. Importantly, examined specimens from South Argentina showed differences in the number of elytra by having more elytra covering the posterior region. Pettibone “1991” reported that polynoids living in association with soft corals tend to have elytra present only in

their anterior region, while worms showing a more free-living lifestyle have elytra in the anterior and posterior region of the body. *Thouarella* is better represented around the Southern Ocean than South Argentina in terms of number of species (Taylor & Rogers, 2015; Zapata-Guardiola & López-González, 2010) and abundance (K.M. Halanych pers. observation), thus it is possible that the lack of elytra in the posterior region of *Polyeunoa* lineages in the Southern Ocean represents a morphological adaptation to availability of *Thouarella* hosts.

On the other hand, *Thouarella* species delimitation has been considered problematic and many species are in need of revision (Zapata-Guardiola & López-González, 2010). Furthermore, a recent study recovered *Thouarella* as a polyphyletic group (Taylor & Rogers, 2015). Molecular data generated here, for example, recovered specimens previously identified as *T. variabilis*, *T. crenelata* and *T. antarctica* in multiple regions on the tree (Figure 3b), suggesting that current understanding of morphology does not align with genetics. Results from species delineation analyses (i.e., ABGD and PTP) disagree in the species boundaries for some clades (e.g., *T. crenelata*, *T. antarctica*, and *T. chilensis*). These species are placed within the "Antarctica gruppe" which as described by Taylor et al. (2013) represents the *Thouarella* group with the smallest morphological variations between the species. Whether this group represents many or one single variable species has been questioned.

Scale-worms are often are the most common organisms found in association with soft corals, and when present can change coral growth, and alter their morphology. Molodtsova & Budaeva (2007) examined over 300 specimens of black corals and found that taxonomic characters important for species identification were modified by the scale-worms, in some cases identification of the corals was not possible due to the changes caused by the worms.

Morphological modification by *Polyeunoa* in *Thouarella* should be evaluated, as it could represent a source of morphological variation in an already challenging group.

2.6 Acknowledgements

We thank members of the Molette Lab at Auburn University for feedback, and particularly to Damien Waits for proofreading of the manuscript. An anonymous reviewer helped improve the manuscript as well. Emma Sherlock from the Natural History Museum in London (NHM) facilitated the loan of samples of *Polyeunoa laevis*. We also thank Kevin Kocot at the University of Alabama for providing access to the SEM for images of *P. laevis*. Cathy McFadden provided advice for molecular work with *Thouarella*. This research was funded by the National Science Foundation (NSF ANT-1043670 to ARM, NSF ANT-1043745, and OPP-0132032 to KMH). The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. This is Molette Biology Laboratory contribution 94 and Auburn University Marine Biology Program contribution 193.

2.7 References

Allcock, A. L., Barratt, I., Eléaume, M., Linse, K., Norman, M. D., Smith, P. J., … Strugnell, J. M. (2011). Cryptic speciation and the circumpolarity debate : A case study on endemic Southern Ocean octopuses using the COI barcode of life. *Deep-Sea Research II*, *58*, 242–249. https://doi.org/10.1016/j.dsr2.2010.05.016

Allcock, A. L., & Strugnell, J. M. (2012). Southern Ocean diversity: New paradigms from molecular ecology. *Trends in Ecology and Evolution*, *27*(9), 520–528. https://doi.org/10.1016/j.tree.2012.05.009

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Ansorge, I. J., Froneman, P. W., & Durgadoo, J. V. (2012). The Marine Ecosystem of the Sub-Antarctic, Prince Edward Islands. In A. Cruzado (Ed.), *Marine Ecosystems* (pp. 61–76). Rijeka: IntechOpen. https://doi.org/10.5772/36676

Baird, H. P., Miller, K. J., & Stark, J. S. (2011). Evidence of hidden biodiversity, ongoing speciation and diverse patterns of genetic structure in giant Antarctic amphipods. *Molecular Ecology*, *20*(16), 3439–3454. https://doi.org/10.1111/j.1365-294X.2011.05173.x

Barnes, D. K. A., & Claus-Dieter, H. (2010). Faunal evidence for a late quaternary trans-Antarctic seaway. *Global and Planetary Change*, *16*, 3297–3303. https://doi.org/10.1111/j.1365-2486.2010.02198.x

Barnes, D. K. A., & Peck, L. S. (2008). Vulnerability of Antarctic shelf biodiversity to predicted regional warming. *Climate Research*, *37*, 149–163. https://doi.org/10.3354/cr00760

Barnich, R., Gambi, M. C., & Fiege, D. (2012). Revision of the genus *Polyeunoa* McIntosh, 1885 (Polychaeta, Polynoidae). *Zootaxa*, *3523*(2012), 25–38.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., … Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Blake, J. A. (2015). New species of Scalibregmatidae (Annelida, Polychaeta) from the east Antarctic Peninsula including a description of the ecology and post-larval development of species of *Scalibregma* and *Oligobregma*. *Zootaxa*, *4033*(1), 57–93.

https://doi.org/10.11646/zootaxa.4033.1.3

Boisvert, S., Raymond, F., Godzaridis, É., Laviolette, F., & Corbeil, J. (2012). Ray Meta: Scalable de novo metagenome assembly and profiling. *Genome Biology*, *13*(12). https://doi.org/10.1186/gb-2012-13-12-r122

Boore, J., Boore, J. L., & Brown, W. M. (2000). Mitochondrial genomes of *Galathealinum*, *Helobdella*, and *Platynereis*: sequence and gene arrangement comparisons indicate that Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. *Molecular Biology and Evolution*, *17*(1), 87–106. https://doi.org/10.1093/oxfordjournals.molbev.a026241

Brandt, A., Gooday, A. J., Brandão, S. N., Brix, S., Brökeland, W., Cedhagen, T., … Vanreusel, A. (2007). First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature*, *447*(7142), 307–311. https://doi.org/10.1038/nature05827

Brasier, M. J., Wiklund, H., Neal, L., Jeffreys, R., Linse, K., Ruhl, H., & Glover, A. G. (2016). DNA barcoding uncovers cryptic diversity in 50 % of deep-sea Antarctic polychaetes. *Royal Society Open Science*, *3*, 160432. https://doi.org/10.1098/rsos.160432

Brokeland, W., & Rapauch, M. J. (2008). A species complex within the isopod genus *Haploniscus* (Crustacea: Malacostraca: Peracarida) from the Southern Ocean deep sea : a morphological and molecular approach. *Zoological Journal of the Linnaen Society*, *152*, 655–706.

Cairns, S. D., & Wirshing, H. H. (2018). A phylogenetic analysis of the Primnoidae (Anthozoa: Octocorallia: Calcaxonia) with analyses of character evolution and a key to the genera and subgenera. *BMC Evolutionary Biology*, *18*(1), 1–20. https://doi.org/10.1186/s12862-018-1182-5

Clarke, A., Barnes, D. K. A., & Hodgson, D. A. (2005). How isolated is Antarctica? *Trends in Ecology and Evolution*, *20*(1), 1–3. https://doi.org/10.1016/j.tree.2004.10.004

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1659. https://doi.org/10.1016/s0020-7292(02)00021-8

Edgar, R. C. (2004). MUSCLE : multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Feldmann, J., & Levermann, A. (2015). Collapse of the West Antarctic Ice Sheet after local destabilization of the Amundsen Basin. *Proceedings of the National Academy of Sciences U.S.A.*, *112*(46), 14191–14196. https://doi.org/10.1073/pnas.1512482112

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299.

France, S. C., & Hoover, L. L. (2002). DNA sequences of the mitochondrial COI gene have low levels of divergence among deep-sea octocorals (Cnidaria: Anthozoa). *Hydrobiologia*, *471*, 149–155.

Galaska, M. P., Sands, C. J., Santos, S. R., Mahon, A. R., & Halanych, K. M. (2017a). Geographic structure in the Southern Ocean circumpolar brittle star *Ophionotus victoriae* (Ophiuridae) revealed from mtDNA and single-nucleotide polymorphism data. *Ecology and Evolution*, *7*(2), 475–485. https://doi.org/10.1002/ece3.2617

Galaska, M. P., Sands, C. J., Santos, S. R., Mahon, A. R., & Halanych, K. M. (2017b). Crossing the divide: Admixture across the antarctic polar front revealed by the brittle star *Astrotoma agassizii*. *Biological Bulletin*, *232*(3), 198–211. https://doi.org/10.1086/693460

Gallego, R., Lavery, S., & Sewell, M. A. (2014). The meroplankton community of the oceanic Ross Sea during late summer. *Antarctic Science*, *26*(4), 345–360. https://doi.org/10.1017/S0954102013000795

Giangrande, A. (1997). Polychaete reproductive patterns, life cycle and life histories: an overview. *Oceanography and Marine Biology: An Annual Review*, *35*(1975), 323–386.

Halanych, K. M., & Mahon, A. R. (2018). Challenging Dogma Concerning Biogeographic Patterns of Antarctica and the Southern Ocean. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 355–378. https://doi.org/10.1146/annurev-ecolsys-121415-032139

Harder, A. M., Halanych, K. M., & Mahon, A. R. (2016). Diversity and distribution within the sea spider genus *Pallenopsis* (Chelicerata: Pycnogonida) in the Western Antarctic as revealed by mitochondrial DNA. *Polar Biology*, *39*(4), 677–688. https://doi.org/10.1007/s00300-015-1823-8

Hartmann-Schröder, G. (1989) *Polynoe thouarellicola* n. sp. aus der Antarktis, assoziiert mit Hornkorallen, und Wiederbeschreibung von *Polynoe antarctica* Kinberg, 1858 (Polychaeta, Polynoidae). *Zoologischer Anzeiger*, 222, 205–221.

Held, C., Leese, F. (2007) The utility of fast evolving molecular markers for studying speciation in the Antarctic benthos. *Polar Biology*, *30*, 513–521.

Hemery, L. G., Eléaume, M., Roussel, V., Améziane, N., Gallut, C., Steinke, D., … Wilson, N. G. (2012). Comprehensive sampling reveals circumpolarity and sympatry in seven mitochondrial lineages of the Southern Ocean crinoid species *Promachocrinus kerguelensis*

(Echinodermata). *Molecular Ecology*, *21*(10), 2502–2518. https://doi.org/10.1111/j.1365-294X.2012.05512.x

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES : Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*(8), 754–755.

Hunter, R. L., & Halanych, K. M. (2008). Evaluating connectivity in the brooding brittle star *Astrotoma agassizii* across the drake passage in the Southern Ocean. *Journal of Heredity*, *99*(2), 137–148. https://doi.org/10.1093/jhered/esm119

Janosik, A. M., & Halanych, K. M. (2010). Unrecognized Antarctic Biodiversity : A Case Study of the Genus *Odontaster* (Odontasteridae; Asteroidea). *Integrative and Comparative Biology*, *50*(6), 981–992. https://doi.org/10.1093/icb/icq119

Jennings, R. M., & Halanych, K. M. (2005). Mitochondrial genomes of *Clymenella torquata* (Maldanidae) and *Riftia pachyptila* (Siboglinidae): Evidence for conserved gene order in Annelida. *Molecular Biology and Evolution*, *22*(2), 210–222. https://doi.org/10.1093/molbev/msi008

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder : Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses Research article. *Molecular Biology and Evolution*, *29*(October), 1695–1701. https://doi.org/10.1093/molbev/mss020

Lawver, L. A., & Gahagan, L. M. (2003). Evolution of Cenozoic seaways in the circum-Antarctic region. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *198*, 11–37.

https://doi.org/10.1016/S0031-0182(03)00392-4

Linse, K., Cope, T., Lörz, A. N., & Sands, C. (2007). Is the Scotia Sea a centre of Antarctic marine diversification? Some evidence of cryptic speciation in the circum-Antarctic bivalve *Lissarca notorcadensis* (Arcoidea: Philobryidae). *Polar Biology*, *30*(8), 1059–1068. https://doi.org/10.1007/s00300-007-0265-3

Linse, K., Griffiths, H. J., Barnes, D. K. A., & Clarke, A. (2006). Biodiversity and biogeography of Antarctic and sub-Antarctic mollusca. *Deep-Sea Research II*, *53*, 985–1008. https://doi.org/10.1016/j.dsr2.2006.05.003

Matschiner, M., Hanel, R., & Salzburger, W. (2009). Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology*, *18*(12), 2574–2587. https://doi.org/10.1111/j.1365-294X.2009.04220.x

Matsuoka, K., Skoglund, A., & Roth, G. (2018). Quantarctica. Norwegian Polar Institute. https://doi.org/10.21334/npolar.2018.8516e961

McFadden, C. S., Alderslade, P., Ofwegen, L. P. Van, Johnsen, H., & Rusmevichientong, A. (2006). Phylogenetic relationships within the tropical soft coral genera *Sarcophyton* and *Lobophytum* (Anthozoa, Octocorallia). *Invertebrate Biology*, *125*(4), 288–305. https://doi.org/10.1111/j.1744-7410.2006.00070.x

McFadden, C. S., Benayahu, Y., Pante, E., Thoma, J. N., Andrew, P., & France, S. C. (2011). Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources*, *11*, 19–31. https://doi.org/10.1111/j.1755-0998.2010.02875.x

McFadden, C. S., France, S. C., Sánchez, J. A., & Alderslade, P. (2006). A molecular phylogenetic analysis of the Octocorallia (Cnidaria: Anthozoa) based on mitochondrial protein-coding sequences. *Molecular Phylogenetics and Evolution*, *41*(3), 513–527.

https://doi.org/10.1016/j.ympev.2006.06.010

McFadden, C. S., Ofwegen, L. P. Van, Beckman, E. J., Benayahu, Y., & Alderslade, P. (2009). Molecular systematics of the speciose Indo-Pacific soft coral genus , *Sinularia* (Anthozoa: Octocorallia). *Invertebrate Biology*, *128*(4), 303–323. https://doi.org/10.1111/j.1744-7410.2009.00179.x

McIntosh, W.C. (1885) Report on the Annelida Polychaeta collected by H.M.S. Challenger during the years 1873–76. *Report on the Scientific Results of the Voyage of H.S.M. "Challenger", Zoology*, 12, 1–554.

Molodtsova, T., & Budaeva, N. (2007). Modifications of corallum morphology in black corals as an effect of associated fauna. *Bulletin of Marine Science*, *81*(3), 469–479.

Munilla, T., & Soler Membrives, A. (2009). Check-list of the pycnogonids from antarctic and sub-antarctic waters: zoogeographic implications. *Antarctic Science*, *21*(2), 99–111. https://doi.org/10.1017/S095410200800151X

Naish, T., Powell, R., Levy, R., Wilson, G., Scherer, R., Talarico, F., … Williams, T. (2009). Obliquity-paced Pliocene West Antarctic ice sheet oscillations. *Nature*, *458*(7236), 322–328. https://doi.org/10.1038/nature07867

Parapar, J., & Moreira, J. (2008). Redescription of *Terebellides kerguelensis* stat. nov. (Polychaeta: Trichobranchidae) from Antarctic and subantarctic waters. *Helgoland Marine Research*, *62*(2), 143–152. https://doi.org/10.1007/s10152-007-0085-4

Pettibone, M. H. (1969). The genera Polyeunoa McIntosh, Hololepidella Willey, and three new genera (Polychaeta, Polynoidae). *Proceedings of the Biological Society of Washington*, *82*, 43–62.

Pollard, D., & DeConto, R. M. (2009). Modelling West Antarctic ice sheet growth and collapse

through the past five million years. *Nature*, *458*(7236), 329–332. https://doi.org/10.1038/nature07809

Poulin, E., González-wevar, C., Díaz, A., Gérard, K., & Hüne, M. (2014). Divergence between Antarctic and South American marine invertebrates: What molecular biology tells us about Scotia Arc geodynamics and the intensification of the Antarctic Circumpolar Current. *Global and Planetary Change*, *123*, 392–399. https://doi.org/10.1016/j.gloplacha.2014.07.017

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD , Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Biology and Evolution*, *21*, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Quattrini, A. M., Wu, T., Soong, K., Jeng, M. S., Benayahu, Y., & McFadden, C. S. (2019). A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evolutionary Biology*, *19*(1), 1–19. https://doi.org/10.1186/s12862-019-1427-y

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE : Multiple Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. *PLoS ONE*, *6*(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Sands, C. J., O'Hara, T. D., Barnes, D. K. A., & Martín-Ledo, R. (2015). Against the flow: evidence of multiple recent invasions of warmer continental shelf waters by a Southern Ocean brittle star. *Frontiers in Ecology and Evolution*, *3*, 63.

https://doi.org/10.3389/fevo.2015.00063

Scherer, R. P., Aldahan, A., Tulaczyk, S., Engelhardt, H., & Kamb, B. (2009). Pleistocene Collapse of the West Antarctic Ice Sheet. *Science*, *82*(1998), 82–86. https://doi.org/10.1126/science.281.5373.82

Schüller, M. (2011). Evidence for a role of bathymetry and emergence in speciation in the genus *Glycera* (Glyceridae, Polychaeta) from the deep Eastern Weddell Sea. *Polar Biology*, *34*(4), 549–564. https://doi.org/10.1007/s00300-010-0913-x

Serpetti, N., Taylor, M. L., Brennan, D., Green, D. H., Rogers, A. D., Paterson, G. L. J., & Narayanaswamy, B. E. (2016). Deep-Sea Research II Ecological adaptations and commensal evolution of the Polynoidae (Polychaeta) in the Southwest Indian Ocean Ridge: A phylogenetic approach. *Deep-Sea Research II*. https://doi.org/http://dx.doi.org/10.1016/j.dsr2.2016.06.004

Shaw, P. W., Arkhipkin, A. I., & Al-Khairulla, H. (2004). Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: The effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Molecular Ecology*, *13*(11), 3293–3303. https://doi.org/10.1111/j.1365-294X.2004.02327.x

Shearer, T. L., Oppen, M. J. H. Van, Romano, S. L., & Worheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, *11*, 2475–2487.

Stiller, M. (1996). Verbreitung und Lebensweise der Aphroditiden und Polynoiden (Polychaeta) im östlichen Weddellmeer und im Lazarevmeer (Antarktis). *Ber Polarforsch*, *185*, 1–200.

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, *123*, 585–595.

Taylor, M. L., Cairns, S. D., Agnew, D., & Rogers, A. D. (2013). A revision of the genus

Thouarella Gray, 1870 (Octocorallia: Primnoidae), including an illustrated dichotomous key, a new species description, and comments on *Plumarella* Gray, 1870 and *Dasystenella*, Versluys, 1906. *Zootaxa*, *3602*(1), 1–105.

Taylor, M. L., & Rogers, A. D. (2015). Molecular Phylogenetics and Evolution Evolutionary dynamics of a common sub-Antarctic octocoral family. *Molecular Phylogenetics and Evolution*, *84*, 185–204. https://doi.org/10.1016/j.ympev.2014.11.008

Thatje, S., Hillenbrand, C. D., & Larter, R. (2005). On the origin of Antarctic marine benthic community structure. *Trends in Ecology and Evolution*, *20*(10), 534–540. https://doi.org/10.1016/j.tree.2005.07.010

Thornhill, D. J., Mahon, A. R., Norenburg, J. L., & Halanych, K. M. (2008a). Open-ocean barriers to dispersal: a test case with the Antarctic Polar Front and the ribbon worm *Parborlasia corrugatus* (Nemertea: Lineidae). *Molecular Biology and Evolution*, *17*, 5104–5117. https://doi.org/10.1111/j.1365-294X.2008.03970.x

Vallès, Y., & Boore, J. L. (2006). Lophotrochozoan mitochondrial genomes. *Integrative and Comparative Biology*, *46*(4), 544–557. https://doi.org/10.1093/icb/icj056

Weigert, A., Golombek, A., Gerth, M., Schwarz, F., Struck, T. H., & Bleidorn, C. (2016). Evolution of mitochondrial gene order in Annelida. *Molecular Phylogenetics and Evolution*, *94*, 196–206. https://doi.org/10.1016/j.ympev.2015.08.008

Wilson, W. H. (1991). Sexual reproductive modes in polychaetes: classification and diversity. *Bulletin of Marine Science*, *48*(2), 500–516.

Wirshing, H. H., & Baker, A. C. (2015). Molecular and morphological species boundaries in the Gorgonian octocoral genus *Pterogorgia* (Octocorallia: Gorgoniidae). *PLoS ONE*, *10*(7), 1–16. https://doi.org/10.1371/journal.pone.0133517

Zapata-Guardiola, R., & López-González, P. J. (2010). Four new species of *Thouarella* (Anthozoa: Octocorallia: Primnoidae) from Antarctic waters. *Scientia Marina*, *74*(1), 131–146. https://doi.org/10.3989/scimar.2010.74n1131

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, *29*(22), 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Zhang, Y., Sun, J., Rouse, G. W., Wiklund, H., Pleijel, F., Watanabe, H. K., … Qiu, J. W. (2018). Phylogeny, evolution and mitochondrial gene order rearrangement in scale worms (Aphroditiformia, Annelida). *Molecular Phylogenetics and Evolution*, *125*(October 2017), 220–231. https://doi.org/10.1016/j.ympev.2018.04.002

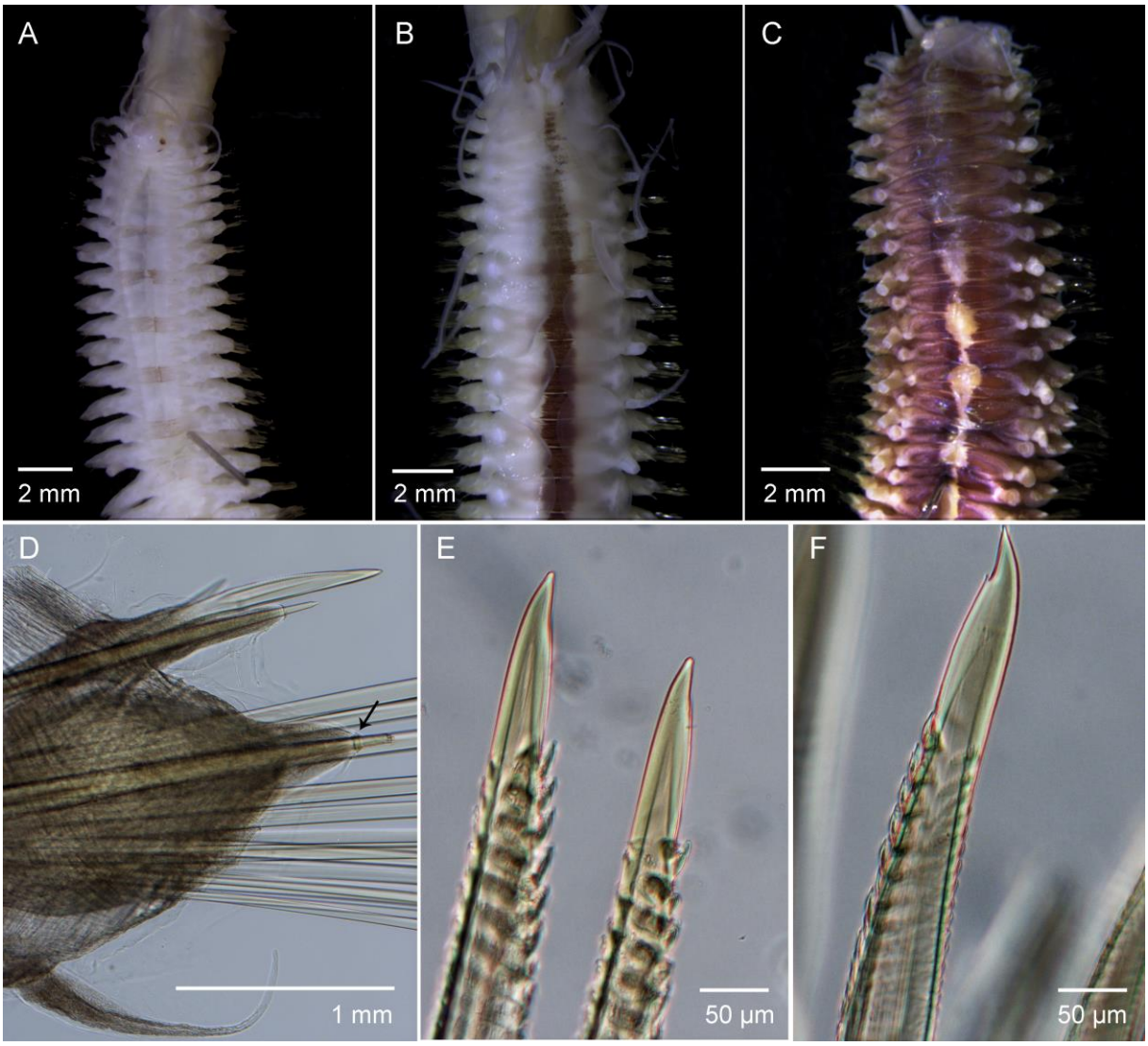

Figure 1. *Polyeunoa laevis*, coloration patterns observed in the AP-Ross clade (A-C); anterior parapodia (D), arrow points to the end of the neuropodial acicular lobe; unidentate chaeta (E); slightly bidentate chaeta (F).

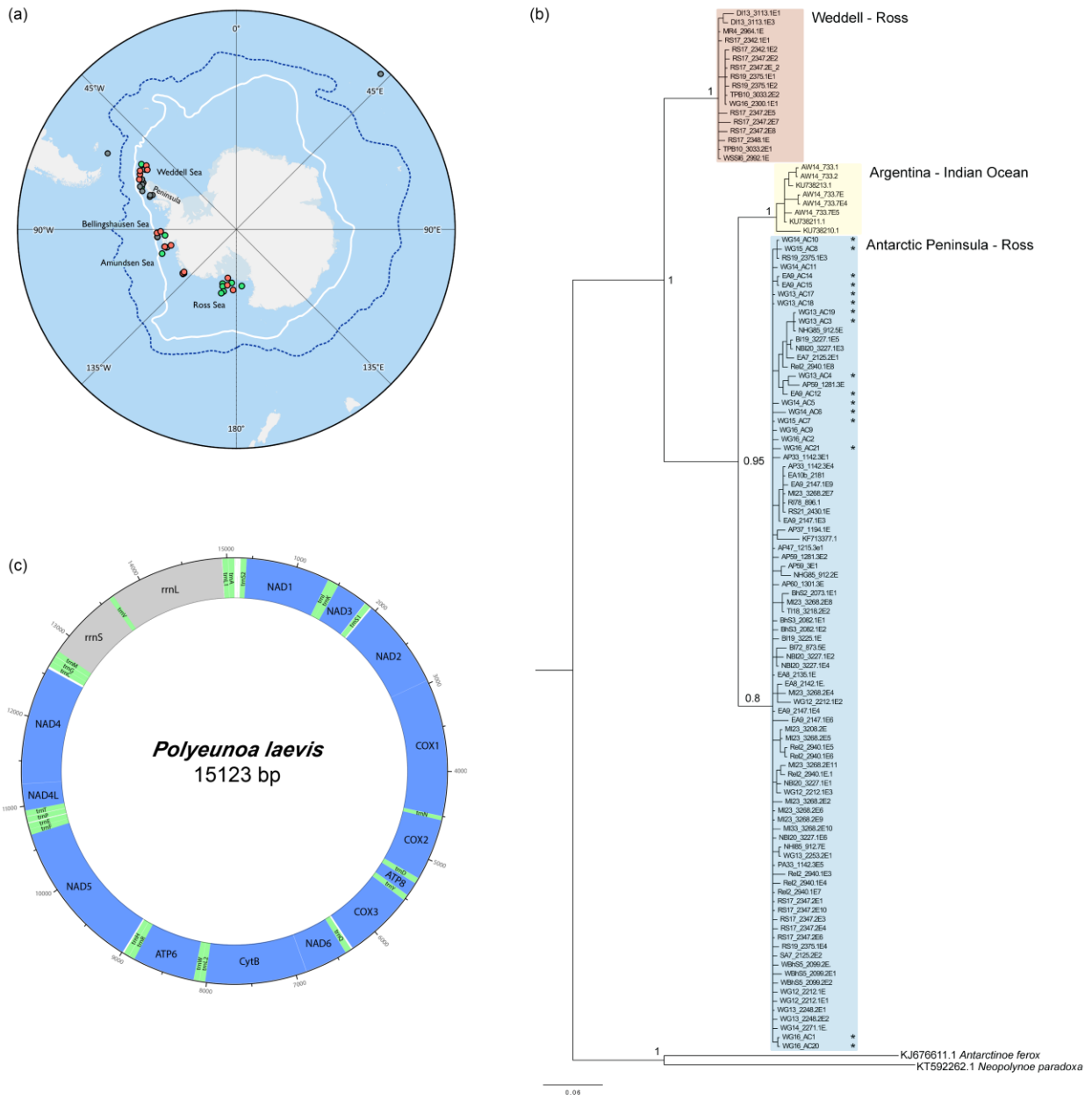

Figure 2. (a) Sampling localities for *Polyeunoa laevis* and *Thouarella*. Gray circles represent sampling localities for *P. laevis* only, green circles are sampling localities for *Thouarella*, and orange circles represent localities were both organisms were collected. White line represents the

Southern boundary of the Antarctic Circumpolar Current, while dotted blue line represents the Antarctic Polar Front. Both of these features vary in position with season and climate. Map created with Quantarctica package (Matsuoka, Skoglund, & Roth, 2018). (b) Bayesian phylogenetic analysis for *Polyeunoa* laevis based on COI data. Asterisks indicate the organisms retrieved directly from *Thouarella*. Topology is consistent with ABGD and PTP species delimitation results (see Supporting information Data S2). (c) Gene order of the mitochondrial genome of *Polyeunoa laevis*. White spaces correspond to unannotated bp.

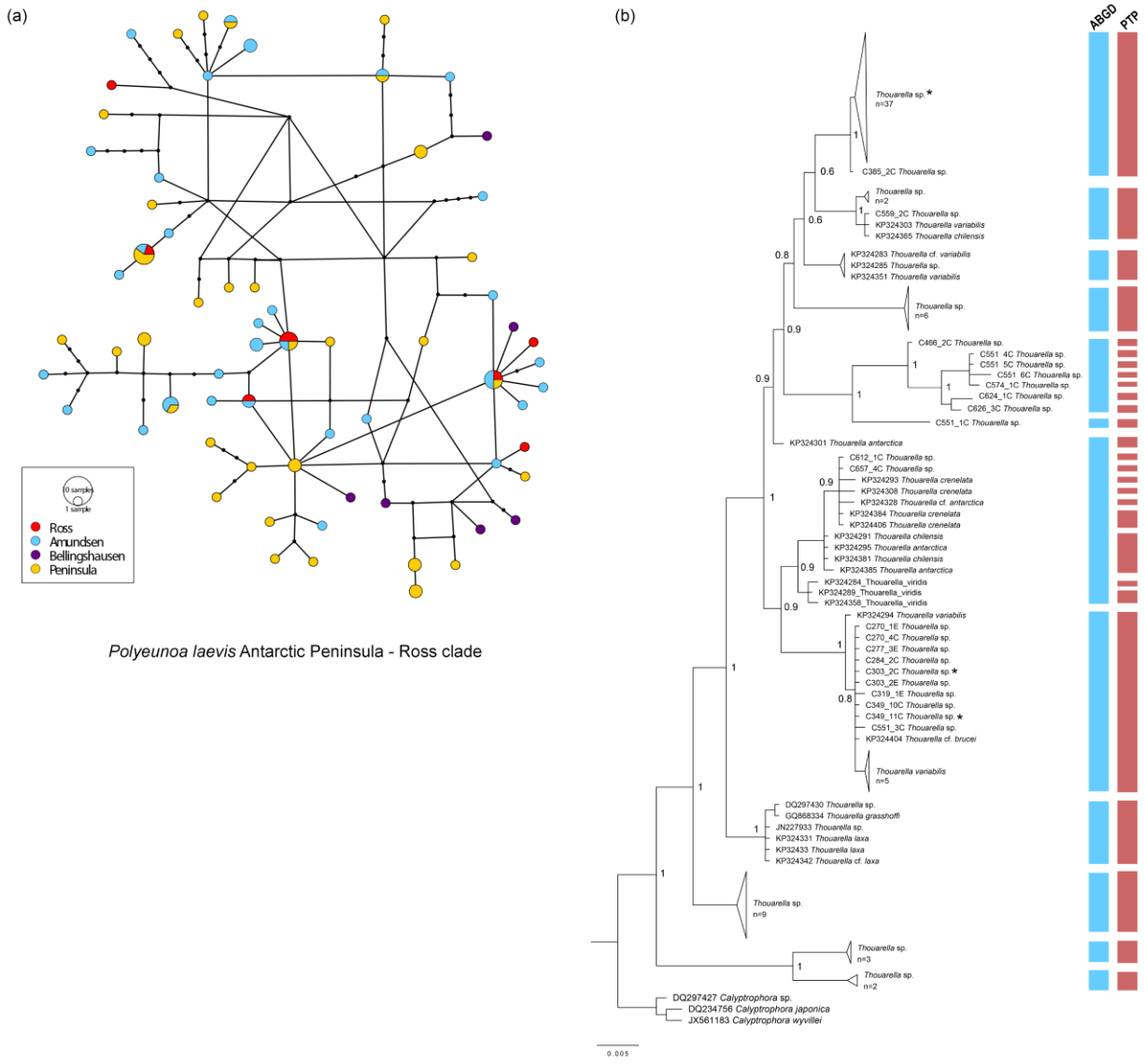

Figure 3. (a) COI haplotype network for AP-Ross clade of *Polyeunoa laevis*. The size of the circles represents the frequency of the haplotype, connecting lines represent one mutation step between haplotypes. Black dots indicate missing haplotypes. (b) Bayesian phylogenetic analysis for *Thouarella* based on mtMutS. Asterisks indicate the corals from which *Polyeunoa* worms were directly retrieved. Vertical bars show species delimitation results from ABDG (red line) and PTP (blue line). Genbank Accession numbers are shown with the taxon labels.

Table 1. Summary of sampling stations and organisms collected in each location. Asterisks indicate samples from Serpetti et al., 2016.

Station / Locality	*Polyeunoa*	*Thouarella*	Latitude	Longitude	Depth
Station 33 Antarctic Peninsula	✓		-67.740	-69.290	122
Station 37 Antarctic Peninsula	✓		-68.186	-67.595	232
Station 47 Antarctic Peninsula	✓		-67.663	-68.245	170
Station 59 Antarctic Peninsula	✓		-64.925	-63.534	360
Station 60 Antarctic Peninsula	✓		-65.026	-63.304	400
Station 14 Argentinean Waters	✓		-54.690	-59.392	207
Station 7 Eastern Amundsen	✓		-72.483	-104.563	591
Station 8 Eastern Amundsen	✓		-72.781	-104.554	496
Station 23 Myriad Islands	✓		-65.021	-64.425	312
Station 85 Near Hugo Island	✓		-64.688	-65.927	368
Station 2 Racovitza Island	✓		-64.411	-61.963	664
Station 78 Renaud Island	✓		-65.624	-67.785	217
Station 10 Tabarin Peninsula	✓		-63.686	-56.859	400
Station 5 Western Bellingshausen	✓		-70.842	-95.411	472
Station 12 Wrights Gulf	✓		-73.159	-129.895	440
Station 16 Wrights Gulf	✓	✓	-73.247	-129.503	478
Station 14 Wrights Gulf	✓	✓	-73.499	-129.919	516
Station 9 Eastern Amundsen	✓	✓	-73.722	-103.617	699
Station 13 Wrights Gulf	✓	✓	-73.297	-129.192	506
Station 15 Wrights Gulf	✓	✓	-73.710	-129.056	655
Station 2 Bellingshausen	✓	✓	-70.812	-92.522	430
Station 3 Bellingshausen	✓	✓	-71.702	-91.504	430
Station 19 Brabant Island	✓	✓	-63.845	-62.624	248
Station 20 North Brabant Island	✓	✓	-63.834	-62.664	256
Station 13 Danvers Island	✓	✓	-63.576	-54.629	227
Station 10b Eastern Amundsen	✓	✓	-72.204	-103.596	612
Station 18 Tower Island	✓	✓	-63.389	-60.120	310
Station 4 Montravel Rock	✓	✓	-62.996	-58.599	320
Station 17 Ross Shelf	✓	✓	-75.330	-176.985	570
Station 19 Ross Shelf	✓	✓	-76.341	-170.850	531
Station 21 Near Ross Shelf	✓	✓	-78.063	-169.991	549
Station 6 Weddell Sea	✓	✓	-64.302	-56.136	290
Station 4 Bellingshausen		✓	-72.699	-94.694	670
Station 5 Western Bellingshausen		✓	-70.842	-95.411	472
Station 10 Eastern Amundsen		✓	-72.177	-103.514	341
Station 11 Eastern Amundsen		✓	-71.147	-108.005	627
Station 20 Ross Shelf		✓	-76.479	-165.738	457
Station 22 Near Ross Shelf		✓	-76.998	-175.093	541
Station 23 Mid-Ross Sea		✓	-76.245	174.504	604
Station 25 North of Ross Island		✓	-75.833	-166.505	552

Table 1 (Continued)

Station / Locality	*Polyeunoa*	*Thouarella*	Latitude	Longitude	Depth
Station 26 Northwest Ross Island		✓	-74.708	-168.408	489
Station 27 Northwest Ross		✓	-74.182	-166.661	390
Station 5 Anderson Island		✓	-63.701	-56.079	293
Station 11 Erebus Terror Bay		✓	-63.935	-56.571	394
Station 12 Dundee Island		✓	-63.754	-55.684	334
Station 14 NE D'Urville Island		✓	-62.442	-55.459	245
Station 25 Hugo Island		✓	-65.087	-65.809	202
Indian Ocean*	✓		-41.346	42.922	1361
Indian Ocean*	✓		-41.346	42.922	1358
Indian Ocean*	✓		-41.380	42.854	1017
Indian Ocean*	✓		-41.357	42.918	917

Table 2. Nucleotide and haplotype diversity for the 3 lineages of *Polyeunoa laevis*, and Tajima's D. Asterisk indicate significant values ($p < 0.005$). ns, number of samples; vs, number of variable sites; nh, number of haplotypes; π, nucleotide diversity; h, haplotype diversity.

Clade	ns	vs	nh	π	h	Tajima's D
AP-Ross	90	56	59	0.00995 ± 0.0005	0.980 ± 0.004	- 1.89399 *
Argentina-Indian Ocean	8	27	6	0.0016 ± 0.004	0.944 ± 0.070	-0.83614
Weddell-Ross	17	22	13	0.0056 ± 0.0008	0.956 ± 0.037	-1.72597

Chapter 3: Diversity and evolution of North Atlantic *Laonice* Malmgren 1867 (Spionidae, Annelida) including description of a novel species

3.1 Abstract

Spionid polychaetes are often dominant members of marine soft-bottom communities. As such, understanding their diversity and evolutionary history is of general interest. One spionid group in particular, *Laonice*, is known from North Atlantic with several species occurring in deeper waters. We explored, as part of the IceAGE project, the biodiversity and evolution of *Laonice* using both morphology and mitochondrial COI gene data. Our data confirm the existence of at least 7 lineages of *Laonice* in the waters surrounding Iceland. Additionally, our sampling suggest species distributions of *Laonice* is similar to previous reports for other annelids, in that the warmer waters south of Iceland appear to harbor more species, but further work is needed to clarify distributional patterns. Although our analysis was hampered by the poor preservation of animals from deep water, we recovered several of species that were previously known to science (e.g. *L.* blakei, *L.* sarsi, *L.* cf. *norgensis*, and *L. cirrata*) and one new species. *L. plumisetosa* sp. nov. is characterized by having u-shaped nuchal organs not exceeding chaetiger 1, and the presence of stout capillaries with plush-like texture in parapodia of anterior chaetigers. Uncorrected genetic distances and phylogenetic analyses of COI data confirm these *Laonice* morphotypes are distinct. However, one morphotype, *L.* cirrata is composed of 3 subclades suggesting unrecognized diversity within this species. Here we aim to provide a preliminary phylogeny for *Laonice* and compare our results with recently proposed subgenera for *Laonice*.

3.2 Introduction

Spionid annelids are among the most common invertebrates occurring in soft sediments. They are usually small (1 mm–5 cm length), sedentary tubiculous (tube-dwelling) annelids that can be recognized in part by a pair of long palps (Blake 1994). Spionidae are usually deposit or suspension feeders, and importantly, they exhibit the highest diversity of reproductive modes among annelids (Wilson 1991), and possess a robust ability to regenerate. These and other traits are thought to contribute to their ecological dominance within benthic assemblages or, in case of the polydorids, as pest species boring into calcareous substrates (Blake 1994). Although also present in the deep sea, most species have been described from shallow waters (Rouse 2001). Unfortunately, spionids are very fragile, and morphological characters used in identification are often lost at the time of collection, making identification challenging and adversely impacting taxonomic treatments in a group that already suffers from limited morphological disparity as adults (Radashevsky et al. 2014). Despite this situation, meaningful studies of spionid fauna in the North Atlantic and the Nordic regions have been undertaken, especially for the deep water *Laonice* (e.g. Wesenberg-Lund 1950; Maciolek 2000; Meißner et al. 2014a; Sigvaldadóttir 2002; Sikorski 2003a; Sikorski 2011; Sikorski and Pavlova 2016; Sikorski et al. 2017).

Laonice was first described in 1867 by Malmgren, and currently this taxon is known to have a worldwide distribution. This group consists of 38 described species with more known to exist (Radashevsky and Lana 2009; Sikorski and Pavlova 2016; Sikorski et al. 2017), especially from deep sea environments. Currently 13 species have been described from North Atlantic and Nordic regions, with most of species being from deep waters (>400 m) and only 3 found within shallow environments (Söderström 1920; Maciolek 2000; Sikorski 2003a; Sikorski et al. 2017). However, the taxonomy of this group is problematic. Even though efforts (Sikorski 2003a;

Sikorski 2011) have improved former species descriptions, the distribution and arrangement of characters between species can overlap or be hard to score, limiting the ability to clearly delineate species (Foster 1971; Meißner et al. 2014a). Phylogenetic studies have not been conducted for this group. Recently, four sub-genera were designated based on hypothesized morphologically distinct groups (Sikorski et al. 2017), that centred on the prostomium being fused or free from the peristomium, arrangement of capillary setae and genital pouches in the anterior segments and length of nuchal organs. Unfortunately, molecular data for the group are very limited (mainly Carr et al. 2011; Brasier et al. 2017) prohibiting independent assessment of such taxonomic hypotheses.

Here, we further explore the biodiversity and evolution of *Laonice* spionids using samples obtained from waters surrounding Iceland as part of the IceAGE (**Ice**landic marine **A**nimals: **G**enetics and **E**cology) project which aims to integrate classical taxonomy, molecular tools and ecological modeling studying marine benthic invertebrates from Icelandic waters (Brix et al. 2014). IceAGE material is valuable not only because it extends previous sampling around Iceland, but also because it provides material fixed in ethanol for molecular work, not available from previous regional surveys (e.g. BIOFAR: Benthic Invertebrate Fauna of the Faore Islands, BIOICE: Benthic Invertebrates of Icelandic waters). To this end, we employ both morphological taxonomy and mitochondrial (mt) COI gene data to assess multiple *Laonice* morphotypes found in the waters surrounding Iceland. Previous analyses (Halanych and Janosik 2006; Carr et al. 2011; Brasier et al. 2016; Meißner et al. 2016) have shown COI to be a useful marker for species identification within annelids. We also aim to produce an initial phylogenetic hypothesis for *Laonice*.

3.3 Materials and Methods

3.3.1 *Study site and sample collection*

Samples were collected during two expeditions around Iceland in 2011 (IceAGE I) and 2013 (IceAGE II), aboard the *RV Meteor (M85/3)* and *RV Poseidon (POS 456),* respectively (Fig. 1). Specimens (Table 1) were collected at depths between 117 and 2780 m using multicorer, boxcorer, epibenthic sledges and trawls, and were fixed in cold 96% ethanol and 4% formalin. For Information on abiotic (including sediment) parameters and habitats of the collection sites see Meißner et al. (2014b; c) and Ostmann et al. (2014) and references therein. Studies on morphology were mainly based on material fixed in formalin; ethanol-fixed material was used for molecular studies and was also used in a reverse taxonomic approach. Samples were managed at the German Centre for Marine Biodiversity Research (DZMB), Hamburg, Germany, while being processed and examined. Final deposition of the material will be at the Senckenberg Museum Frankfurt (SMF), Germany.

3.3.2 *Morphological examination*

Morphology was investigated using both light (LM) and scanning electron microscopy (SEM). Methyl green staining was applied for observation of most characters by means of light microscopy. For SEM studies, specimens were dehydrated in a graded ethanol series, critical-point dried, sputter coated with carbon and examined with a Leo 1525 scanning electron microscope. Drawings were made using a *camera lucida*. Light micrographs were taken with an Olympus SC50 digital camera. Measurements of width refer to the distance between the distal-most structures on the widest chaetiger seen on the anterior end in dorsal view (chaetae not

included in the measurement). All images were post-processed and plates compiled using Adobe TM Adobe Creative Suite 6 (including Photoshop and Illustrator).

The term ‘nuchal organ’ refers to ciliary bands on the dorsum, posterior and posterolateral to the prostomium. Under ‘Material examined’, the following abbreviations are used in case specimens were incomplete: af = anterior fragment, mf = middle fragment. Morphology was assessed independent of knowledge of the molecular phylogeny. However, after both sets of data were brought together, selected specimens were reexamined to reconcile findings. The relevant taxonomic literature (Greaves et al. 2011; Maciolek 2000; Meißner et al. 2014a; Sigvaldadottir and Desbruyeres 2003; Sikorski et al. 1988, 2002, 2003a, b; Sikorski and Pavlova 2016; Sikorski et al. 2017) on *Laonice* was used for identification. In particular, Sikorski’s (2011) Table II catalogues several morphological characters for 31 species of *Laonice*.

3.3.3 *Molecular methods*

Specimens used for molecular analysis were first identified based on morphological characters (Suppl. 1); tissue was later obtained from each specimen for molecular studies. Barcoding with the mitochondrial Cytochrome c Oxidase subunit I (COI) gene was performed at Auburn University and the Smithsonian Institution. For 66 specimens (Table 1), whole genomic DNA was extracted using Qiagen DNeasy® Blood and Tissue Kit (Qiagen Inc., Valencia, CA), following the manufacturer’s protocol. COI was amplified and sequenced using newly designed primers (Table 2) after primers by Folmer et al. (1994) and Carr et al. (2011) failed. The combination of primers 2F-spionid-LCO and 1R-spionid-HCO yielded better results for *Laonice* specimens studied here. These new primers essentially span the same region as the Folmer et al. primers but cover 660bp.

The PCR mix consisted of 14.2 µl water, 2.5 µl $MgCl_2$, 2.5 µl 10x Taq buffer, 2.5 µl dNTP (10 µM), 1 µl of each primer (10mM), 0.3 µl 10x Taq polymerase (VWR) and 1 µl of DNA template. PCR cycling protocol consisted of denaturation at 95 °C for 5 min followed by 35 cycles of 94 °C for 1 min, 50 °C for 1 m and 72 °C for 1:30 min. A final elongation at 72 °C for 8 min was employed. After visualization on a 1% agarose gel, products were purified using the QIAquick PCR Purification Kit (QIAGENE), and bidirectionally Sanger sequenced by GENEWIZ. Chromatograms of forward and reverse sequences were aligned and proofread in Geneious R6.

We also included 42 COI sequences available in GenBank (Suppl. 2) which comprised additional *Laonice* indiviudals (*L. cirrata*, *L. antarcticae*, *L. norgensis* and *L. wedellia*) and outgroups (*Spiophanes*, *Marenzelleria* and *Streblospio*).

Alignment of all *Laonice* sequences was performed using default parameters in MUSCLE (Edgar 2004) and proof read by eye. Aligned COI sequences were translated using the invertebrate mitochondrial code to ensure stop codons or frameshift mutations were not present.

3.3.4 *Phylogenetic analyses*

Laonice relationships were reconstructed from COI data using Bayesian Inference (BI). Best-fit partitions and models were inferred with PartitionFinder 1.1.1 (Lanfear et al. 2012). Trees were reconstructed with MrBayes 3.2.6 (Huelsenbeck and Ronquist 2001) implementing the PartitionFinder results which assigned codon position 1, 2 and 3 to different partition. For each partition, a reversible jump Markov chain Monte Carlo (MCMC) was used to integrate across the selected GTR+I and GTR+G model as implemented in MrBayes. Four runs of four independent chains were run for 10,000,000 MCMC generations and sampled ever 1,000 generations.

Stationarity of each chain was checked with TRACER v.1.6 (Rambaut et al. 2014), and the first 25% discarded as burn-in after visualizing a plot of likelihood scores. Additionally, we report uncorrected and Kimura-2-parameter (K2P) model corrected values for genetic distance between operational taxonomic units for comparison to the literature.

3.4 Results

Below, accounts of morphological and molecular results are followed by a systematic account that integrated findings from both types of data.

3.4.1 *Morphological results*

Specimens were all anterior fragments and often they were not in good condition due to the general fragility of spionids even though careful processing of samples was employed during collection. This impeded the observation of several characters, but still allowed species identification in most cases. Altogether 91 specimens fixed in formalin were available for morphological studies of which 84 were identified to species level. Morphological identification of solely IceAGE material recovered at least 5 different species of *Laonice*. Most common was *L. blakei* Sikorski & Jirkov in Sikorski, Jirkov & Tsetlin, 1988 (N=64). This species was mainly found south of Iceland in both the Iceland and Irminger Basins, but two single records also were from the Denmark Strait and the Norwegian Channel (Fig. 1). Water depths ranged from 683–2750 m. Another twelve specimens were identified as *L. cirrata* (M. Sars, 1851). Accordingly, this species occurred all around Iceland at water depths between 321–1622 m. Less common was *L. sarsi* Söderström, 1920 (N=6). This species was encountered in samples from south of Iceland at the Reykjanes Ridge and at the slopes to the Irminger Basin, at water depths between 214–305

m, another single record also came from the Norwegian Channel at water depths of 303 m. There were two single specimens, one from the eastern part of the study area, and another one from the slopes of the Irminger Basin. They were identified as *L.* cf. *norgensis* Sikorski, 2008 and *L*. cf. *whittardensis* Sikorski et al., 2017. The water depth at these stations was 303 m and 1622 m, respectively. Seven specimens could not be identified to species level, mainly due to their poor condition.

L. plumisetosa sp. nov., a yet unknown species newly described herein, was not found in formalin fixed material and described using ethanol fixed material in a reverse taxonomic approach (see 'systematic account' for details).

3.4.2 *Phylogenetic results*

The final dataset consisted of 108 sequences with an aligned length of 657 nucleotides. Of these, we produced 66 sequences that correspond to 7 morphotypes which included 4 known species and one new species (*L. blakei*, *L. cirrata*, *L.* cf. *norgensis*, *L. sarsi*, *L. plumisetosa* sp. nov.). Some individuals may correspond to known taxa (e.g. *Laonice* sp. a, *Laonice* sp. b, Table 1; Fig. 2) but we did not have adequate morphological information to make reliable species designations. Within *Laonice*, 314 (47%) sites were variable and 281 (42%) were parsimony informative.

Laonice was recovered as monophyletic in BI analyses (Fig. 2; Suppl. 3). One sequence from GenBank (KU697719.1) labeled as *L. cirrata* fell outside all other *Laonice*, but unfortunately there is limited information for this sequence (Miralles et al. 2016) and is likely a misidentified animal. Nodes defining recognized morphotypes were all highly supported (posterior probability= >0.99). Below, we focus our discussion on well supported nodes

(pp>0.95). Uncorrected genetic distances (p) of COI within *Laonice* morphotypes ranged from 0-1.7% except for specimens recognized as having the *L. cirrata* morphotype. Distances between *Laonice* species ranged from 15-24%.

Except for sequence KU697719.1 (see above), *L. cirrata* formed a group composed of three subclades suggesting that it could represent more than one species consistent with genetic distances recovered (10% divergence within group). *L. blakei* and *L. sarsi* were recovered as monophyletic (pp = 98%). The novel morphotype from Iceland (see above) was recovered as monophyletic group based on the COI data. Interestingly, a *Laonice* from Antarctica (*L.* cf. *antarcticae*) was recovered as the sister taxa to *L.* sp. a. Similarly, *L. weddellia,* also found within Antarctic waters, was recovered as the sister taxa of *L. norgensis*. This latter group was composed of sequences from GenBank identified as *L. norgensis* (Meißner et al. 2014a) as well as five specimens from our Icelandic collections. Unfortunately, the morphology of the Icelandic samples was not well preserved, and consisted mainly of anterior ends. Thus, the extension of the nuchal organs, the first appearance of neuropodial hooks, start and extension of interparapodial pouches, the start of sabre chaetae, and the number of apical teeth in the hooks could not be sufficiently examined for all specimens which would have allowed a more reliable determination of these samples as either *L. norgensis* or *L. appelloefi*. Based on the limited morphological assessment together with molecular barcode data, we provisionally called these individuals *L.* cf. *norgensis*. There are also four lineages that only include one or two individuals that fall outside of the 7 morphotypes (starred lineages on Fig. 2). Given their genetic distinctness, they may well represent additional species, but we decline to comment further on them herein as it would be prudent to obtain additional samples.

Our morphological assessment recovered a previously unrecognized and distinct combination of characters within *Laonice* individuals examined. Moreover, these individuals were recovered as a single clade based on COI barcode data which was 15% different (uncorrected distances) from other *Laonice*. Based on these findings, we describe a new species.

3.4.3 Systematic account

GENUS *Laonice* Malmgren, 1867

Laonice Malmgren, 1867: 91; Maciolek 2001: 533–534; Sikorski 2003a: 317–318; Sikorski et al. 2017: 362.

Type species: *Nerine cirrata* M. Sars, 1851 by monotypy, accepted as *Laonice cirrata* (M. Sars, 1851).

Mandane Kinberg, 1866. Type species: Mandane brevicornis Kinberg, 1866. Fide Sikorski 2011.

Spionides Webster & Benedict, 1887. Type species: Spionides cirratus Webster & Benedict, 1887.

Description (amended). Prostomium subtriangular to bell-shaped, anteriorly rounded, straight or medially incised; nuchal organ(s) usually straight, extending posteriorly for a variable number of chaetigers along the middorsum, only exceptionally very short and u-shaped or looped; occipital antenna present. Peristomium free or fused to prostomium. Branchiae from chaetiger 2, apinnate or with digitiform pinnules, separated from or partially fused to notopodial postchaetal lamellae, continuing posteriorly for at least one-half of body length; noto- and neuropodial

postchaetal lamellae large, expanded in anterior chaetigers, reduced posteriorly. Interparapodial lateral pouches present. Notopodia with capillaries; notopodial hooded hooks present or absent; neurochaetae include capillaries, hooded hooks, and sabre chaetae; hooks with main fang and one to several apical teeth. Pygidium with anal cirri.

Laonice plumisetosa sp. nov.

(Figs. 3–5)

Holotype: North Atlantic Ocean, Norwegian Sea, Norwegian Basin: 67°38.63'N 12°09.72'W, Stn. 1184-1 (Area 26), EBS, water depth 1820 m, 20-Sep-2011, fixation 96% ethanol, 1 af (SMF 24376). **Paratypes:** North Atlantic Ocean, Norwegian Sea, Norwegian Basin: 67°38.63'N 12°09.72'W, Stn. 1184-1 (Area 26), EBS, water depth 1820 m, 20-Sep-2011, fixation 96% ethanol, 3 af (SMF 24377), 1 af (SMF 24378); 67°35.29'N 6°57.48'W, Stn. 1165-1 (Area 24), box corer, uppermost 0-7 cm, water depth 2402 m, sandy mud, 18-Sep-2011, fixation 96% ethanol, 1 af (SMF 24379).

Non-type material. North Atlantic Ocean, Iceland Basin, Slope: 62°33.05'N 23°23.33'W, Stn. 1006-1 (Area 6), EBS, water depth 1369 m, 2-Sep-2011, fixation 96% ethanol, 2 af (SMF 24380).

Description. Holotype anterior fragment of 19 chaetigers, 0.9 mm wide and 4.5 mm long. All other specimens anterior fragments of 22 chaetigers at maximum, between 0.4–0.9 mm wide (measured at about chaetiger 4–5), and 2–5 mm long.

Prostomium bell-shaped, anterior margin straight, sometimes with tiny median incision, posteriorly extended into short, slightly elevated caruncle not exceeding chaetiger 1 (Figs. 3a, 4a,

c); papilliform occipital antenna present at the anterior end of the caruncle (Figs. 3a, 4a); eyes absent. Peristomium moderately developed and not clearly separated from the prostomium, anteriolaterally prostomium and peristomium possibly connected via a narrow tissue connection (due to the limited number of specimens and their poor condition this observation remains uncertain at present) (Figs. 3a, 4a). Palps lost in all specimens (only one specimen with very short stump at the position of palps on one side, SMF 24380). Nuchal organ as pair of short u-shaped bands not exceeding chaetiger 1 (Figs. 3a, 4a), of yellowish colour in specimens fixed in 96% ethanol (Fig. 4c). Dorsal branchiae lost in all specimens.

Parapodial postchaetal lamellae well developed, broad, foliaceous (Figs. 3c, 5e) (mostly lost in notopodia of available specimens); notopodial postchaetal lamellae with wide base, rather narrow and tapered in first chaetigers (Fig. 3e), in subsequent chaetigers broad foliaceous with blunt tip (Fig. 3b); neuropodial postchaetal lamellae tapered with wide base in first chaetigers, subsequently rounded, obtuse, with very wide base extending to ventral side of the body, inferior extension most pronounced at chaetigers 2–7 (Figs. 3c, 4b, 5e, f). Prechaetal lamellae absent. Interparapodial lateral pouches first present from between chaetigers 3–4, usually present throughout until the end of the fragment, partially covered by neuropodial postchaetal lamellae. Transversal dorsal ciliated crests not observed.

Capillary chaetae in anterior chaetigers arranged in 2 distinct rows in both rami (Figs. 4a, b, 5e); anterior row with stout chaetae with plush-like texture seen in SEM (Fig. 5d), if seen in LM at high magnification covered by fine hairs (Fig. 5b, c); second row with smooth capillaries without sheaths; capillaries in second row significantly longer and appearing considerably thinner than chaetae in anterior row (Fig. 5f); from about chaetiger 9 rows less distinct and arrangement in irregular bundles predominating, then with simple capillaries without

granulations and sheaths. Neuropodial hooded hooks first present from chaetiger 18, 9–10 hooks per neuropodium; hooks bidentate with two small apical teeth side by side above main fang (Figs. 3d and 5a); hooks accompanied by simple capillaries. Sabre chaetae present in hook-bearing chaetigers; as one or two stout, granulated chaetae in inferior-most position (Fig. 5f). Pygidium not observed (all specimens anterior fragments).

Pigmentation. Pigmentation not observed; yellow colour of nuchal organ in specimens freshly fixed in 96% ethanol conspicuous (Fig. 4c).

Methyl green staining pattern. Inconspicuous. Methyl green bonds to the chaetal surface and thus emphasizes the differing nature of chaetae (Fig. 5e). Hence the stout chaetae with plush-like texture are easily detected. Prostomium and parapodial lamellae are intensively stained (Figs. 4a, b).

Ecology. *Laonice plumisetosa* sp. nov. was collected in deep waters between 1370–2400 m water depth. Sediments were described as very poorly sorted sandy mud with high total organic carbon contents (Meißner et al. 2014b).

Geographical distribution. The species occurs north of Iceland in the Norwegian Sea and also south of Iceland at the slope to the Iceland Basin. The northern areas belong to the coldest (temperatures below zero) and deepest (up to 2400 m) in the study area of the IceAGE project, whereas the southern area is characterized by highly productive surface waters (Meißner et al. 2014b).

Remarks. The most conspicuous characters of *L. plumisetosa* sp. nov. are the u-shaped nuchal organs not exceeding chaetiger 1 (Fig. 3a, 4a, c), and the presence of stout capillaries with plush-like texture in parapodia of anterior chaetigers (Fig. 5b-e). Among *Laonice* spp. looped nuchal

organs are only also described for *L. magnacristata* Maciolek, 2000. However, other characters are not shared between the two species: in *L. magnacristata* the tridentate neuropodial hooks exhibit apical teeth in tandem position whereas apical teeth are positioned side by side in *L. plumisetosa* sp. nov.; moreover, lateral interparapodial pouches start between chaetigers 3 and 4 in the latter species opposed to pouches first appearing between chaetigers 7 and 8 in *L. magnacristata*. The presence of stout chaetae with plush-like texture as described here for *L. plumisetosa* sp. nov. can be regarded a unique character for species of *Laonice*.

Unfortunately, the original description of *L. plumisetosa* sp. nov. has to be based on specimens which are not in very good condition. This hampers the description of characters of taxonomic importance among species of *Laonice*. For example, the fusion between prostomium and peristomium could not be observed sufficiently. In some specimens, an anterior fusion between prostomium and peristomium could not be ruled out entirely, in others this fusion appears to be absent. Also, the absence of branchiae in all specimens has to be mentioned and scars providing evidence that they are lost could not unambiguously determined in any case. However, the fact that almost all notopodial parapodial lamellae are also missing allows the suggestion, that branchiae are rather lost than not having been present. Notably, specimens were collected with epibenthic sledges and fixed onboard with 96% ethanol. This procedure is known to be unfavourable for the collection of fragile polychaetes living in deep-sea habitats.

Given that phylogenetic analyses of COI data (Fig. 2) clearly resolved this taxon as distinct from other *Laonice* species, the formal recognition of the species was appropriate. *L. plumisetosa* sp. nov. from the type locality are deposited under GenBank Accession numbers MG234463-MG234465 and MG234486 (Suppl. 2). In conclusion, despite of the limitations of the material,

we described specimens as a new species of *Laonice* because both morphological and molecular characters allow easy identification.

3.5 Discussion

Using an integrative approach, we found 7 distinct lineages of *Laonice* in the North Atlantic near Iceland (Fig. 1 and 2) and explore the phylogenetic relationships of the group. Although the waters surrounding Iceland are among the best studied in the world, the discovery of more *Laonice* species in the region, and especially in the deep sea, is likely. The novel species described here, *Laonice plumisetosa*, was the only *Laonice* species we found in the deep waters of the Norwegian Sea. Our findings on species distribution (Fig. 1) are similar to previous reports for other annelids in that the warmer waters south of Iceland appear to harbor more species. For example, Oweniidae and Opheliidae were more limited to the southern region (Parapar 2003; Parapar et al. 2011), and Sigvaldadottir (2002) found members of *Prionospio* occur north and south of Iceland. Furthermore, additional work is needed to clarify distributional patterns of these lineages around Iceland and in the North Atlantic (see Stransky and Svavarsson 2006; Dijkstra et al. 2009; Brix and Svavarsson 2010).

Definitive unambiguous morphological characters that distinguish individual species are generally rare in *Laonice*, although, our newly described species has plumose like chaetae representing a unique character. Many *Laonice* species are distinguished based on a combination of characters (e.g., Sikorski 2003a; Sikorski et al. 2011; Meißner et al. 2014a, herein). In combination with the easy fragmentation of *Laonice* specimens and often poor condition of our material, identification based on morphological characters has been challenging. This is the reason *Laonice* sp. a and *Laonice* sp. b could not be identified and more material is needed to

solve their taxonomy. Nonetheless, identification and delineation of North Atlantic *Laonice* species by morphological and molecular criteria produced consistent results in the present study. Identified morphotypes corresponded with genetic clades that showed interspecies uncorrected differences of 15-24% (Table 3) and were defined by well supported nodes (pp =≥0.99). These values are typical of, or greater than, distance values that have been documented by previous studies (Dahlgren et al. 2001; Glover et al. 2005; Meißner and Blank 2009; Carr et al. 2011; Thornhill et al. 2012; Janssen et al. 2015) as distinguishing between different, well-recognized species. For example, spionids typically show a 6-20% difference between congeneric species, but >40% difference when compared with other spionid genera (Mahon et al. 2009; Meißner and Götting 2016; Ye et al. 2017). Thus, based on both morphological and molecular criteria, there are at least 7 species of *Laonice* around Iceland.

However, delineation of *L. cirrata* was more difficult and it may represent several species. CO1 data recovered at least 3 subclades (Fig. 2) with 10-16% difference between them, suggesting an unrecognized diversity within this species. In contrast, other intraspecific distances ranged from 0-2%, which was similar to the subclades within *L. cirrata* (<2%). Additionally, taxonomic characters used for *L. cirrata* show a broad morphological variation with extended ranges for characters (e.g., last segment with branchiae 4-40, last segment with nuchal organs 16-59; Table II in Sikorski 2011) that overlap with other recognized *Laonice* species (Sikorski 2011). Moreover, this species is reported to have a broad distributional range that includes the Arctic, Pacific and North and South Atlantic Ocean (Hartman 1965; Blake 1994; Sikorski 2002; Sikorski 2003a; Radashevsky and Lana 2009; Sikorski 2011). After studying a number of specimens identified as *L. cirrata* from several localities, Sikorski (2002) concluded that this species is probably limited to Norway and adjacent regions, and specimens from other locations

might represent different species. Interestingly, each of the 3 subclades identified by CO1 data represent specimens from distinct geographic regions (Fig. 2), providing additional support for Sikorski's conclusions. Further, studies with specimens from other locations including the type locality are required in order to get a better understanding of the taxonomic status for this species.

Among Spionidae, *Laonice* has been considered a well-defined group (Radashevsky and Lana 2009) but understanding of relationships within *Laonice* have been lacking.

Recently, 28 *Laonice* species were assigned into four new sub-genera (*Laonice*, *Sarsiana*, *Appelloefia* and *Norgensia*) based on morphological evidence (Sikorski et al. 2017). *Sarsiana* comprises 7 species which have the following characters: the prostomium fused to the peristomium, short nuchal organs and notopodial hooded hooks present in posterior segments. In our sampling, this group is represented by *L. sarsi* and *L. antarcticae*, which are recovered within a highly supported clade (pp=0.97) together with *Laonice* sp. a, providing preliminary support for the *Sarsiana* group. The *Laonice* subgenus is differentiated based on the anterior margin of the prostomium being fused to the peristomium in combination with capillary chaetae in anterior region arranged in two vertical rows. This group is composed of 10 described species. However, because we only have molecular data for *L. cirrata* (which likely includes multiple species as currently recognized), we cannot comment on the validity of this subgenus. That said, we do note that all the species with the prostomium fused to the peristomium form a clade (pp=0.92), and include species from the *Laonice* and *Sarsiana* subgenera. Thus, the unfused condition is plesiomorphic for *Laonice* and will not be informative as a diagnostic character (i.e., synapomorphy).

In contrast, two other subgenera, *Appelloefia* and *Norgensia*, were not recovered in our phylogenetic analyses. Our study placed *L. weddellia* and *L. norgensis* as sister taxa with *L. blakei* in a different region of the tree separated by multiple nodes (Fig. 2). However, Sikorski's et al. (2017) subgeneric classification placed both *L. weddellia* and *L. blakei* within the *Appelloefia* subgenus. These subgenera were distinguished by the fact that *Appelloefia* species "often [have] a widening of the body" and *Norgensia* species "often [have] transversal dorsal membranous crests in the post dorsal nuchal organ area" with the description of other characters being largely identical. Given the lack of diagnostic morphological characters supporting these subgeneric designations and results of the molecular phylogenetic analysis, the validity of these subgenera is suspect.

For example, the identification of *L. norgensis* proved difficult. The species is morphologically close to *L. appelloefi* Söderström, 1920 and according to Sikorski (2003a) the main differences are the absence of the visible widening of anterior chaetigers as well as the presence of complete dorsal transverse membranes in *L. norgensis*. In addition, the nuchal organ is supposedly longer in *L. norgensis* (chaetigers 15-28 versus chaetigers 8-14 in *L. appelloefi*), pouches start later (chaetigers 5-15 opposed to 8-17), and hooks are tridentate in *L. appelloefi* but mostly bidentate in *L. norgensis.* Based on this there is a numerical overlap in several characters and these characters may vary with ontogeny (see Radashevsky and Lana 2009; Greaves et al. 2011). *Laonice* worms are usually recovered incomplete as rather short anterior fragments and without characters of taxonomic value like dorsal crests, hooks, pouches. Since COI sequences of our specimens formed a clade with *Laonice* specimens from the Meteor seamounts (Meißner et al. 2014a) identified as *L. norgensis*, here we also provisionally refer our specimens to *L. norgensis*. We would like to stress that this assignment can only regarded as dependable when

barcode data from the type locality of *L. norgensis* become available and provide support for this assignment.

Clearly, additional data and attention are need to more fully understand the evolution with in *Laonice,* but the COI topology presented here represent a primarily view that can be tested in subsequent analyses. For example, the tree indicated that Southern Ocean *Laonice* do not share a recent common ancestor implying that these animals have been able to move into Antarctic waters on separate occasions. With addition of taxa and data, this concept, as well as Sikorski's proposed subgenera, can be revisited to further our understanding of this group.

3.6 Acknowledgements

We would like to thank the technicians of the German Centre of Marine Biodiversity Research (DZMB), Karen Jeskulke and Antje Fischer, for sample management. Dieter Fiege helped with photography and provided the photo of the holotype. Renate Walter (University of Hamburg) kindly assisted with SEM studies. Andrey Sikorski for providing his poster from the IPC2016 meeting as a handout pdf. Saskia Brix and Torben Riehl for facilitating and helping with visits of V.E.B to Hamburg. This work was made possible with the kind support of the crew of the *R/V Meteor* (M85/3) and *R/V Poseidon* (POS456). Visits to Hamburg were partially supported by German Research Foundation (DFG) grant BR3843/6-1. This work was funded by National Science Foundation grant DEB-1036537 to K. M. H., and DFG grant BR3843/4-1. This is Molette Biology Laboratory contribution 75 and Auburn University Marine Biology Program contribution 172.

3.7 References

Blake JA (1994) Family Spionidae. In: Blake JA, Hilbig B, Scott PH (eds) Taxonomic Atlas of the Santa Maria Basin and Western Santa Barbara Channel. Volume 6. The Annelida Part 3 - Polychaeta: Orbiniidae to Cossuridae. Santa Barbara Museum of Natural History, Santa Barbara, CA, pp 81–224

Brasier MJ, Wiklund H, Neal L, Jeffreys R, Linse K, Ruhl H, Glover AG (2016) DNA barcoding uncovers cryptic diversity in 50 % of deep-sea Antarctic polychaetes. R Soc Open Sci 3:160432. doi: 10.1098/rsos.160432

Brix S, Meißner K, Bente S, Halanych KM (2014) The IceAGE project – a follow up of BIOICE. Polish Polar Res 35:141–150

Brix S, Svavarsson J (2010) Distribution and diversity of desmosomatid and nannoniscid isopods (Crustacea) on the Greenland – Iceland – Faroe Ridge. Polar Biol 33:515–530

Carr CM, Hardy SM, Brown TM, Macdonald TA, Hebert PDN (2011) A Tri-Oceanic Perspective: DNA Barcoding Reveals Geographic Structure and Cryptic Diversity in Canadian Polychaetes. PLoS One. doi: 10.1371/journal.pone.0022232

Dahlgren TG, Akesson B, Schander C, Halanych KM, Sundberg P (2001) Molecular phylogeny of the model annelid *Ophryotrocha*. Biological Bulletin, 201(2): 193-203. Biol Bull 201:193–203

Dijkstra HH, Warén A, Gudmundsson G (2009) Pectinoidea (Mollusca: Bivalvia) from Iceland. Mar Biol 5:207–243

Edgar RC (2004) MUSCLE : multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of

mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Foster NM (1971) Spionidae (Polychaeta) of the Gulf of Mexico and the Caribbean Sea. Studies on the fauna of Curaçao and other Caribbean Islands 36:1–183.

Glover AG, Goetze E, Dahlgren TG, Smith CR (2005) Morphology, reproductive biology and genetic structure of the whale-fall and hydrothermal vent specialist, *Bathykurila guaymasensis* Pettibone, 1989 (Annelida: Polynoidae). Mar Ecol 26:223–234

Greaves E, Meissner K, Wilson R (2011) New *Laonice* species (Polychaeta: Spionidae) from western and northern Australia. Zootaxa 2903:1-20

Halanych KM, Janosik AM (2006) A review of molecular markers used for Annelid phylogenetics. Integr Comp Biol 46:533–543

Hartman O (1965) Deep-water benthic polychaetous annelids off New England to Bermuda and other North Atlantic areas. Occas Pap Allan Hancock Found 1–384

Huelsenbeck JP, Ronquist F (2001) MRBAYES : Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Janssen A, Kaiser S, Meißner K, Brenke N, Menot L, Martínez Arbizu P (2015) A reverse taxonomic approach to assess macrofaunal distribution patterns in abyssal pacific polymetallic nodule fields. *PLoS ONE* 10:e0117790. doi: 10.1371/journal.pone.0117790

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

Maciolek NJ (2000) New species and records of *Aonidella*, *Laonice*, and *Spiophanes* (Polychaeta: Spionidae) from shelf and slope depths of the western north Atlantic. Bull Mar

Sci 67:529–547

Mahon AR, Mahon HK, Dauer DM, Halanych KM (2009) Discrete genetic boundaries of three *Streblospio* (Spionidae, Annelida) species and the status of *S. shrubsolii*. Mar Biol Res 5:172–178

Malmgren AJ (1867) Annulata Polychaeta Spetsbergiæ, Grœnlandiæ, Islandiæ et Scandinaviæ. Hactenus Cognita. Ex Officina Frenckelliana, Helsingforslæ. 127 pp.

Meißner K, Bick A, Götting M (2016) Arctic *Pholoe* (Polychaeta: Pholoidae): When integrative taxonomy helps to sort out barcodes. Zool J Linn Soc. doi: 10.1111/zoj.12468

Meißner K, Bick A, Guggolz T, Götting M (2014a) Spionidae (Polychaeta: Canalipalpata: Spionida) from seamounts in the NE Atlantic. Zootaxa 3786:201–245

Meißner K, Blank M (2009) *Spiophanes norrisi* sp. nov. (Polychaeta:Spionidae) a new species from the NE Pacific coast, separated from the *Spiophanes bombyx* comples on both morphological and genetic studies. Zootaxa 2278: 25

Meißner K, Brenke N, Svavarsson J (2014b) Benthic habitats around Iceland investigated during the IceAGE expeditions. Polish Polar Res 35:177–202

Meißner K, Fiorentino D, Schnurr S, Martinez Arbizu P, Huettmann F, Holst S, Brix S, Svavarsson J (2014c) Distribution of benthic marine invertebrates at northern latitudes - An evaluation applying multi-algorithm species distribution models. Journal of Sea Research 85:241-254

Miralles L, Ardura A, Arias A, Borrell YJ, Clusa L, Dopico E, Hernandez de Rojas A, Lopez B, Muñoz-Colmenero M, Roca A, Valiente AG, Zaiko A, Garcia-Vazquez E (2016) Barcodes of marine invertebrates from north Iberian ports : Native diversity and resistance to biological invasions. Mar Pollut Bull 112:183–188

Ostmann A, Schnurr S, Arbizu PM (2014) Marine Environment Around Iceland : Hydrography, Sediments and First Predictive Models of Icelandic Deep-sea sediment characteristics. Polish Polar Res 35:151–176

Parapar J, Moreira J, Helgason GV. (2011) Distribution and diversity of the Opheliidae (Annelida, Polychaeta) on the continental shelf and slope of Iceland, with a review of the genus *Ophelina* in northeast Atlantic waters and description of two new species. Org Divers Evol 11:83–105

Parapar J, (2003) Oweniidae (Annelida, Polychaeta) from Icelandic waters, collected by the BIOICE project , with a description of *Myrioglobula islandica* n. sp. Sarsia. doi: 10.1080/00364820310002506

Radashevsky VI, Lana PC (2009) *Laonice* (Annelida: Spionidae) from South and Central America. Zoosymposia 2:265–295

Radashevsky VI, Neretina TV, Pankova VV, Nuzhdin SV (2014) Molecular identity, morphology and taxonomy of the *Rhynchospio glutaea* complex with a key to *Rhynchospio* species (Annelida, Spionidae). Syst Biodivers 12:37–41

Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6

Rouse GW (2001) Spionidae. In: Rouse GW, Pleijel F (eds) Polychaetes. Oxford University Press, New York, pp 269–272

Sars M (1851) Beretning om en i Sommeren 1849 foretagen zoologisk Reise i Lofoten og Fimarken. Nyt Magazin for Naturvidenskaberne 6:121-211

Sigvaldadóttir E (2002) Polychaetes of the genera *Prionospio* and *Aurospio* (Spionidae, Polychaeta) from Icelandic waters. Sarsia 87:207–215

Sigvaldadóttir E, Desbruyéres D (2003) Two new species of Spionidae (Annelida: Polychaeta)

from Mid-Atlantic Ridge hydrothermal vents. Cah Biol Mar 44:219-225

Sikorski AV (2002) On distinguishing the morphologically close species, *Laonice cirrata* and *L. bahusiensis* (Polychaeta, Spionidae). Zoologichesky Zhurnal 81:406-419

Sikorski AV (2003) *Laonice* (Polychaeta, Spionidae) in the Arctic and the North Atlantic. Sarsia 88:316–345

Sikorski AV (2003b) On the fauna of the genus *Laonice* (Polychaeta, Spionidae) in the northern Pacific. Zoologichesky Zhurnal 82:1179-1190

Sikorski AV (2011) Review of *Laonice* (Spionidae, Annelida) with remarks on several species and a description of a new species from South Africa. Ital J Zool 78:37–41

Sikorski AV, Gunton LM, Pavlova L (2017) *Laonice* species (Polychaeta, Spionidae) from the Whittard Canyon (NE Atlantic) with descriptions of two new species. J Mar Biol Assoc UK 97:961-973

Sikorski AV, Jirkov IA, Tzetlin AB (1988) The genus *Laonice* (Polychaeta, Spionidae) in the Arctic Ocean: weighing the taxonomic characters and species composition. Zoologichesky Zhurnal 67:826-838

Sikorski AV, Pavlova L (2016) Three new species of *Laonice* (Polychaete: Spionidae) from West and Southwest Africa. Zootaxa 4097:353–368

Söderström A (1920) Studien über die Polychätenfamilie Spionidae. Dissertation Uppsala Almquist & Wicksells:1-286

Stransky B, Svavarsson J (2006) *Astacilla boreaphilis* sp. nov. (Crustacea: Isopoda: Valvifera) from shallow and deep North Atlantic waters. Zootaxa 1259:1–23

Thornhill DJ, Struck TH, Ebbe B, Lee RW, Mendoza GF, Levin LA, Halanych KM (2012) Adaptive radiation in extremophilic Dorvilleidae (Annelida): diversification of a single

colonizer or multiple independent lineages? Ecol Evol 2:1958–1970

Wesenberg-Lund E (1950) Polychaeta. The Danish Ingolf-Expedition 4:1-81

Wilson WH (1991) Sexual reproductive modes in polychaetes: classification and diversity. Bull Mar Sci 48:500–516.

Ye L, Cao C, Tang B, Yao T, Wang R, Jiangyong W (2017) Morphological and molecular characterization of *Polydora websteri* (Annelida: Spionidae), with remarks on relationship of dult worms and larvae using mitochondrial COI gene as a molecular marker. Pak J Zool 49:699–710

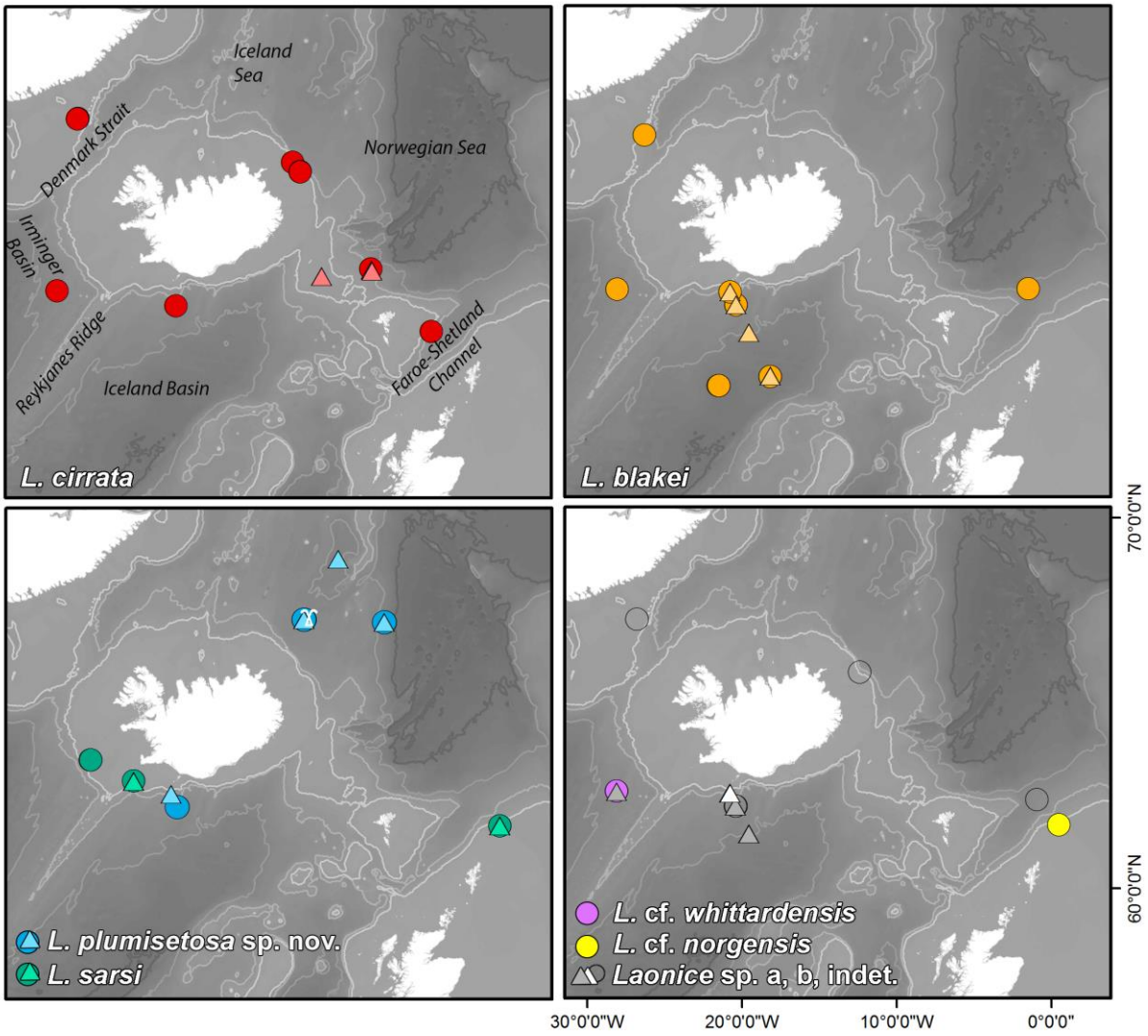

Figure 1. Sample sites for *Laonice* spp. in the study area: circle for records from morphological studies, triangle for records based on molecular studies; Ϫ indicates type locality for *L. plumisetosa* sp. nov., for color coding see lower left corner of the maps.

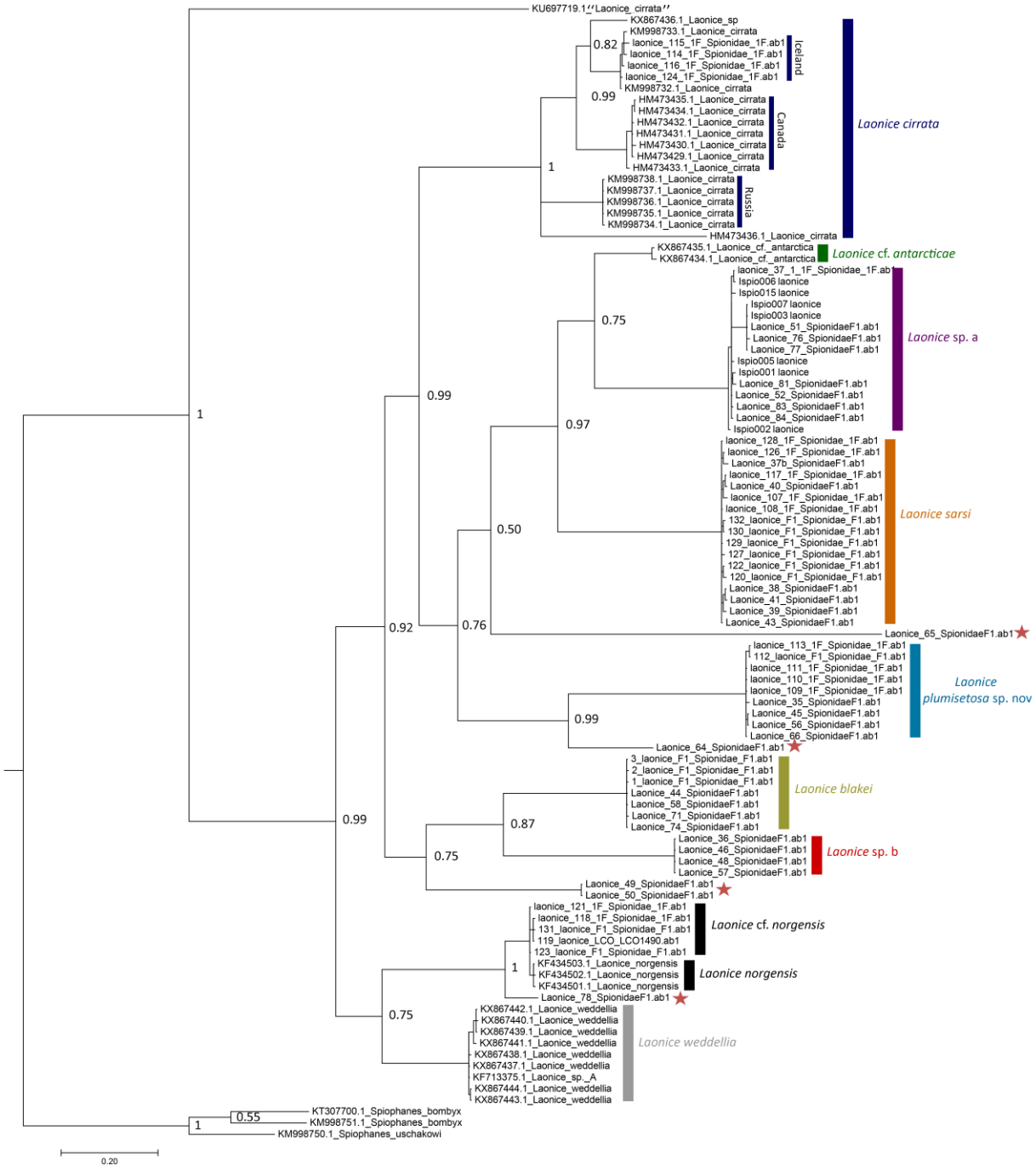

Figure 2. Bayesian inference analysis based on mitochondrial COI data. Posterior probabilities (pp) shown next to the relevant node. Starred lineages represent singletons lineages. Within *Laonice cirrata*, locality of the specimens is indicated.

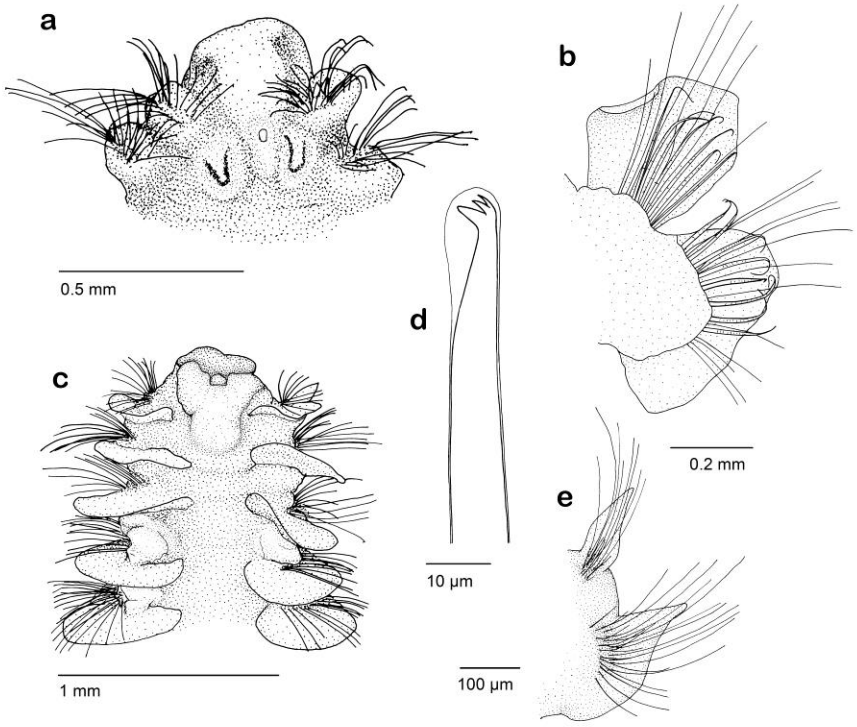

Figure 3. *Laonice plumisetosa* sp. nov.: a. Anterior end, close-up dorsal view (note u-shaped nuchal organs which appear yellow in specimens fixed in 96% ethanol). b. Parapodium from chaetiger 10, anterior view. c. Anterior end, ventral view. d. Neuropodial hooded hook from chaetiger 19. e. Parapodium from 1st chaetigers, anterior view. — a, c SMF 24378, b, e SMF 24380, d SMF 24379; a, c, d paratypes.

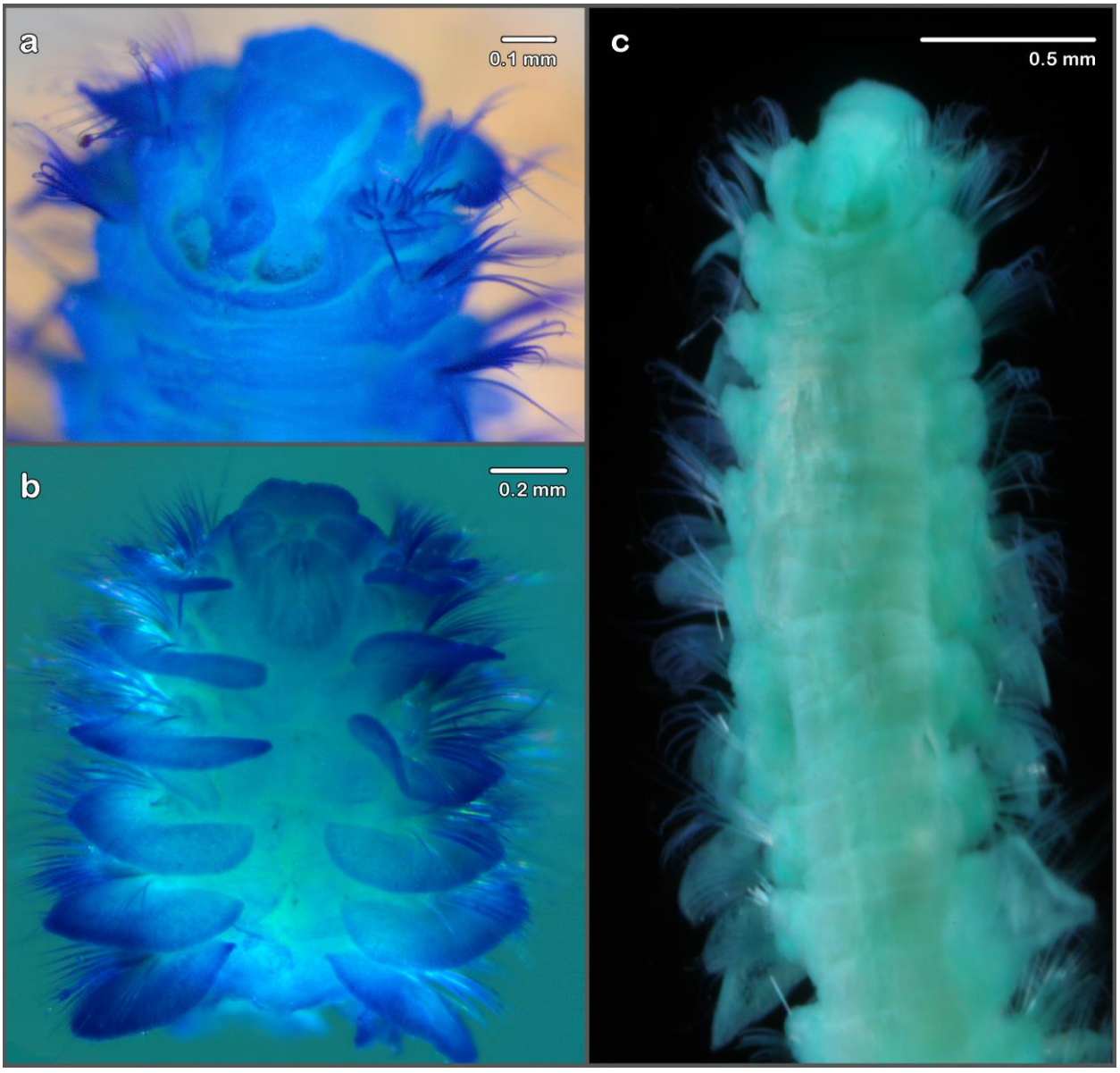

Figure 4. *Laonice plumisetosa* sp. nov.: a. Anterior end, dorsal-oblique view; note u-shaped nuchal organs, specimen stained with methyl green. b. Anterior end, ventral view; stained with methyl green. c. Anterior end, dorsal view. — a, c holotype SMF 24376, b paratype SMF 24378.

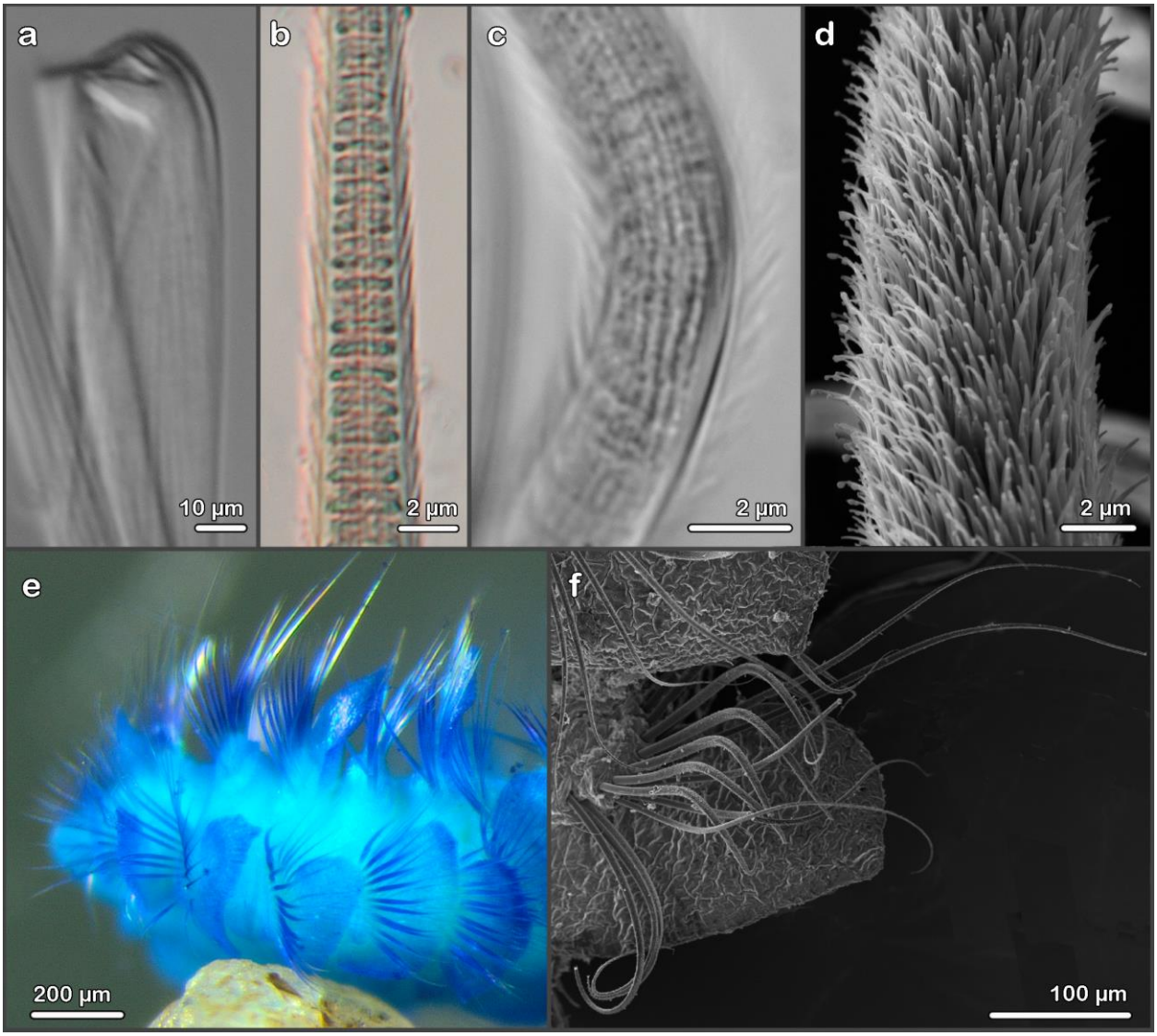

Figure 5. *Laonice plumisetosa* sp. nov.: a. Neuropodial hooded hook from chaetiger 19. b. Micrograph of neurochaeta with plush-like texture from anterior row of 8th chaetiger. c. Same of notochaeta from 6th chaetiger. d. Same type of chaeta studied with SEM. e. Anterior end, lateral view; specimen stained with methyl green, majority of notopodial lamellae lost. f. SEM study of middle chaetiger. — a–c. SMF 24379, d, f SMF 24380, e SMF 24378; a–c, e all paratypes.

Table 1. Taxa found in Icelandic waters, collection site and number of specimens used for genetic and morphological studies.

Species_ID	Latitude	Longitude	Depth (m)	Specimens for genetics	Specimens for morphology
Laonice blakei	60.358	-18.137	2567.6	4	0
	62.55	-20.38	1386.8	1	0
	61.71	-19.55	1912.3	1	0
	62.94	-20.74	913.6	1	0
	60.046	-21.488	2750.4	0	37
	60.045	-21.482	2749.4	0	14
	60.342	-18.144	2571.7	0	3
	62.558	-20.353	1392.4	0	2
	62.940	-20.736	913.6	0	1
	63.013	-28.060	1593.8	0	2
	67.213	-26.272	683.1	0	4
	63.030	-1.451	1842.0	0	1
Laonice cirrata	63.577	-7.711	1043.6	1	0
	63.609	-7.753	1056.2	1	0
	62.99	-28.10	1588.2	2	0
	62.552	-20.388	1384.8	0	1
	62.999	-28.079	1621.8	0	1
	67.646	-26.746	320.6	0	3
	66.539	-12.867	315.9	0	1
	66.301	-12.373	730.8	0	1
	66.290	-12.351	579.1	0	3
	61.777	-3.873	835.1	0	1
	63.641	-7.783	1073.4	0	1
Laonice cf. *norgensis*	61.997	0.506	302.5	5	0
Laonice norgensis	61.996	0.507	302.5	0	1
Laonice plumisetosa	67.643	-12.162	1819.3	4	0
	69.093	-9.933	2173.4	1	0
	62.94	-20.74	913.6	2	0
	67.59	-6.96	2402	2	0
Laonice sarsi	63.314	-23.160	288.5	8	1
	61.997	0.506	302.5	9	2
	63.922	-25.964	215	0	3
Laonice sp. a	61.710	-19.549	1912.3	12	0
	62.551	-20.395	1384.8	2	0
	62.99	-28.10	1588.2	1	0
Laonice sp. b	62.93	-20.77	891.7	4	0
Laonice whittardensis	62.999	-28.079	1622	0	1
Laonice 49	63.417	-10.97	440.5	1	0
Laonice 50	63.417	-10.97	440.5	1	0

Table 1. continued

Species_ID	Latitude	Longitude	Depth (m)	Specimens for genetics	Specimens for morphology
Laonice 64	62.55	-20.38	1386.8	1	0
Laonice 65	67.59	-6.96	2402	1	0
Laonice 78	61.71	-19.55	1912.3	1	0

Table 2. Primers designed in this study for the mitochondrial COI region.

Primer name	Sequence
1F–spionid–LCO	5'– TACTCAACYAAYCAYAAAGACATTGG–3'
2F–spionid–LCO	5'–TACWCMWCYAAYCAYAAASRMATTGG–3'
1R–spionid–HCO	5' –TAYACTTCDGGRTGTCCRAARAATCA–3'
2R–spionid–HCO	5'– CCAATWCCDARYATDGCRTAAATTAT–3'

Table 3. Uncorrected pairwise genetic distances (p) above diagonal and K2P genetic distances in bold below diagonal within *Laonice*.

	1	2	3	4	5	6	7	8	9	10
1. *L. weddellia*	0.57 **0.58**	19.9	20.1	20.4	17.5	14.7	21.5	19.7	14.7	20.2
2. *L. cirrata*	**23.28**	9.48 **10.59**	21.4	23.5	22.5	19.8	21.7	21.5	20.1	22.1
3. *L.* cf. *antarcticae*	**23.76**	**25.51**	0.62 **0.62**	15.0	17.5	19.0	19.9	20.7	19.1	21.8
4. *L.* sp. a	**23.88**	**28.62**	**16.97**	1.66 **1.70**	18.6	21.9	20.9	22.6	22.0	23.3
5. *L. sarsi*	**20.03**	**26.88**	**20.01**	**21.54**	0.60 **0.61**	19.1	18.8	20.3	18.4	20.2
6. *L.* cf. *norgensis*	**16.39**	**23.14**	**22.09**	**25.98**	**22.07**	0.41 **0.41**	19.7	20.1	1.1	21.2
7. *L. plumisetosa*	**25.47**	**25.77**	**23.18**	**24.61**	**21.69**	**22.99**	0.37 **0.37**	20.8	19.6	22.0
8. *L. blakei*	**22.92**	**25.43**	**24.34**	**27.13**	**23.81**	**23.61**	**24.52**	0.12 **0.12**	20.0	17.4
9. *L. norgensis*	**16.46**	**23.56**	**22.20**	**26.23**	**21.18**	**1.10**	**22.90**	**23.46**	0.28 **0.28**	21.3
10. *L.* sp. b	**23.66**	**26.32**	**25.90**	**28.27**	**23.71**	**25.15**	**26.30**	**19.93**	**25.22**	0.00 **0.00**

Chapter 4: Larval development of ciliated, muscular and nervous organ systems in *Pseudopolydora paucibranchiata* (Spionidae, Annelida)

4.1 Abstract

Given the diversity of annelid forms, relatively little is known about the developmental processes that transform a larva into an adult. Here, we describe the early development of ciliation, musculature and serotonergic elements of the nervous system during larval formation in *Pseudopolydora paucibranchiata* using compound light (DIC) and confocal laser scanning microscopy (CLSM). Ciliation is extensive and includes multiple ciliary cells around the head, stomodeum and gut regions, and on the pygidium. Completely circularized trochal bands were not observed. Interestingly, no apical tuft is distinguishable. There is a simultaneous development of longitudinal muscles associated with the body wall and sections of digestive system, and distinctive musculature associated with chaetal sacs is the most prominent during early stages. Similar to the spionid *Malacoceros fulginosus*, early larval stages showed no serotonergic activity, and consequently it was not possible to identify the origin of the first serotonergic cells in *P. paucibranchiata*. Comparative studies using CLSM with understudied groups will broaden our understanding of evolutionary developmental patterns across Annelida.

4.2 Introduction

Annelida is a highly diverse phylum with over 21,000 species described, with a broad variety of contrasting life history patterns and ecological relationships (Purschke et al. 2020). Within annelids, Spionidae is one of the most speciose families, and they are commonly found within shallow benthic sediments where they are one of the dominant groups of marine invertebrates (Blake 1996). They are small (1 mm – 5 cm), sedentary, tubiculous (tube-dwelling) worms. Spionids are also present in the deep sea, and a few species have been reported in fresh water. Because of their diversity and abundance, there have been many studies on spionids, including comprehensive reviews of their reproduction, development and systematics, although most of these topics have never been fully reviewed (Blake and Arnofsky 1999, Blake 2006). Literature on spionid development, for example, is one of the most extensive among annelids (Blake and Arnofsky 1999); however, previous studies were conducted with compound light microscopy (Hannerz 1956; Simon 1967; Blake 1969; Blake and Woodwick 1975). and only a few studies have investigated spionid development using confocal laser scanning microscopy (Forest & Lindsay 2008; Vortsepneva et al. 2009).

The spionid *Pseudopolydora paucibranchiata* (Okuda, 1937) is a widely distributed species that can be found in intertidal sandy beaches. It was originally described from Japan (Okuda 1937), but there are records from the Northeast and Southwest Pacific, and Northeast Atlantic, and more recently it has been reported also in the Northwest Atlantic (Blake and Woodwick 1975; Dagli and Çinar 2008; Radashevsky 1993; Simboura et al. 2010; Çinar 2013; Bogantes et al. *In review*). Reproduction and development of this species has been previously studied with light microscopy by Blake and Woodwick (1975), Myohara (1980), and Radashevsky (1983), providing a comprehensive background on fertilization, egg cleavage,

timing of development and the formation of external morphological characters throughout different larval stages.

Studies on annelid development using modern microscopy techniques have increased in the last decade (Boyle and Rice 2014; Boyle and Seaver 2010; Carrillo-Baltodano et al. 2019; Helm et al. 2013; Meyer et al. 2010; Meyer et al. 2015; Starunov et al. 2017), although, spionid development and life history diversity has not been investigated with confocal laser scanning microscopy. Moreover, all spionid studies have been focused on development of the nervous system (Forest and Lindsay 2008; Kumar et al. 2020; Vortsepneva et al. 2009), while other organ systems remain understudied. Comparative studies of morphological, embryological, and developmental data are required for a proper understanding of the broader evolutionary patterns across annelids (Halanych 2016). Here we describe the development of musculature, ciliation patterns, and serotonergic elements of the nervous system during larval formation of *Pseudopolydora paucibranchiata*.

4.3 Methods

4.3.1 *Collection of specimens*

This study occurred between June and August of 2017, at the Smithsonian Marine Station (SMS) at Fort Pierce, Florida. Intertidal specimens were collected by sieving (300 µm) shovel loads of sediment, removing worms by hand, and transporting the samples in seawater for further processing in the laboratory. Tubes comprised of sand and sediments containing adult worms of *Pseudopolydora paucibranchiata* were identified under a stereoscope and isolated for imaging and dissection. The tubes were opened and examined for egg capsules or larvae, and if present, those specimens were fixed for compound light and confocal laser scanning microscopy.

4.3.2 *Fixation and immunostaining*

Embryos and larvae were removed from their egg or brood capsule using an eyelash mounted to the tip of a glass pipette. Specimens were relaxed in $MgCl_2$ for about 10 minutes (or until movement was not observed), followed by fixation with 4% paraformaldehyde (pfa) in filtered sea water (FSW) for 1 hour at room temperature. After fixation, specimens were washed 3 times FSW to remove fixative, then washed 3 times in phosphate buffered saline (PBS) and stored in PBS at 4˚C. Specimens were then transferred into phosphate-buffered saline with 0.1% Triton X-100 (PBT), followed by incubation in 5 µg/mL RNase A solution in PBT for 1 hour at 37°C. RNaseA treatments were terminated with PBT washes. Larvae and embryos were then blocked in PBT with 5% heat inactivated goat serum (HIGS) for 2 hours at room temperature while rocking. Following blocking treatment, the specimens were incubated overnight at 4°C with primary antibodies in blocking solution (anti-tyrosinated tubulin, anti-serotonin). Specimens were washed 6 times in PBT and incubated overnight at 4°C with secondary antibodies in blocking solution. The following day the specimens were incubated overnight at 4°C with Propidium Iodide and BD Phallacidin. Larvae and embryos were washed with PBS, dehydrated and mounted on a non-coated glass slides.

4.3.3 *Confocal microscopy*

Mounted specimens were examined and scanned with a Zeiss LSM510 confocal laser scanning microscope (CLSM). LSM files were analyzed and processed with ImageJ (Rueden et al. 2017). Channel-specific digital Z-stack projections and color-merge projections were prepared in ImageJ, followed by editing and formatting in Photoshop CS5 and Illustrator CS5.

4.4 Results

4.4.1 *Development of* Pseudopolydora paucibranchiata

Based on morphological features described by Blake and Woodwick (1975), 5 different stages (I, II, III, IV, V; Fig. 1) were identified. Stage I represents postembryonic early stage pre-larval form (Fig. 1A). During this stage, the specimens ciliated and have a bilaterally symmetric oval shape with a mouth opening and identifiable axes; no prominent internal features are detectable. Stage II represents the pre-setiger larva (Fig. 1B). During stage II, the posterior region becomes narrower and chaetal musculature is visible. Stage III represents larvae with 3 segments, 2 eyes and short chaetae extending from the first chaetiger (Fig. 1C). Stage IV larvae have 3 segments, 4 eyes are present, and chaetae can now be observed within the first three chaetigers (Fig. 1D). Lastly, stage V represents specimens with more than 4 segments, 6 eyes and a noticeable elongation of the body, and with anterior and middle chaetae extending beyond the most posterior region of the larva (Fig. 1E). Stages I-IV occur in the egg capsule, while stage V larvae have typically emerged from the egg capsule.

The following sections describe observable patterns in the development of cilia, muscles and serotonergic elements of the nervous system in the different stages of *P. paucibranchiata*.

4.4.2 *Cilia development*

Stage I shows three ventral patches of cilia in the region around the mouth opening, with no other ciliation arrangements observed during this stage (Fig. 2A). At stage II, there are now two ventral patches of cilia near the mouth opening (Fig. 2B). Additionally, an akrotroch is observed as a discrete ciliary band encircling ventral and lateral sides of the anterior end or head region. Dorsal images show the presence of two ciliated pits (Fig. 2C), given their location they

might represent the beginning of nuchal organs. Early ciliation is also visible at the location of the future prototroch, and on the posterior end where the telotroch will develop. During stage III, a second anterior ciliary band is present, which constitutes the prototroch, and other ciliary bands including the neurotroch, and telotroch start to develop (Fig. 2D). In stage IV the ciliated pits are still present and can be observed in dorsal images (Fig. 2E). During this stage the mouth is fully develop and heavily ciliated, long tactile cilia can be observed extending from the anterior region of the head (Fig. 2F). The last stage (V) maintains the same ciliary bands observed during the previous stage but there is a noticeable extension in the length of the cilia (Fig. 2G). Interestingly, no apical tuft was observed during any of the stages.

4.4.3 *Muscle development*

During stage I, F-actin staining is stronger around the mouth opening, along cell margins (e.g. cytoskeleton) and muscle joints of the body wall (Fig. 3A). Specimens assigned to stage II show no signs of segmentation when viewed under a compound light microscope, although, the muscles associated with the chaetal sacs are already starting to develop (Fig. 3B). During stage III, four longitudinal muscle fibers have begun to form, with some fibers reaching almost to the posterior end of the larva. In addition, chaetal retractor muscles have developed (Fig. 3C-D). Even though there are no signs of an apical tuft, a set of muscle cells near the location of the apical organ are present (Fig. 3C). In stage IV, as the larvae elongates, more muscles associated with the posterior region of the digestive system start to form, several longitudinal fibers are visible connecting the anterior and posterior region, and there are additional muscle fibers around the anus (Fig. 3E). Staining is strong around the muscles associated with the first chaetal sac. In stage V there is an increase of longitudinal muscle fibers around the digestive system and body

wall, and now transverse bands of muscles have developed (Fig. 3F). At this point, the most prominent muscles are associated with the chaetal sacs, especially in the first segment. Although, segments four and five are delimited, no visible muscles associated with chaetal sacs in either of those segments were observed.

4.4.4 *Serotonergic elements of nervous system development*

Stages I and II show no signs of serotonergic activity. However, at stage III, positive serotonergic-like immunoreactive elements (5HT-lir) were observed along the larval head and trunk body (Fig. 4A). More specifically, during stage III, two serotonergic neurons were present within the brain region and serotonergic circumesophageal connectives extended from the brain to the anterior trunk, where they connect across the ventral-anterior trunk in combination with two pairs of cell bodies on left and right sides. During stage IV, serotonergic activity increases in the anterior region, with additional pairs of 5HT-lir neurons in the region of the brain neuropil, including an incomplete 5HT-lir ring around the stomodeum (Fig. 4B). Moreover, 3 5HT-lir neurons were visible on each side of the anterior trunk, where a pair of ventral nerve cord (vnc) connectives extend from anterior to posterior along left and right sides of the trunk. At stage V there is more serotonergic activity in the brain region, with several connectives converging toward the apical organ to the anterior of the brain neuropil (Fig. 4C). As the larvae elongates, vnc connectives are closer to the ventral midline, although still distinct as left and right longitudinal nerves each containing 2-3 individual neurites. At stage V, there were approximately 8 5HT-lir cell bodies in the head, and 5 cell bodies along each side of the trunk. Additionally, a complete serotonergic ring around the pygidium was developed and left and right ventral cords connected posteriorly in a terminal looping pattern, anterior to the pygidium. 5HT-

lir growth cones were still visible at the end of each vnc connective. No serotonergic neurons were observed in the posterior-most region, within the pygidium.

4.5 Discussion

4.5.1 *Cilia development*

Cilia development of *Pseudopolydora paucibranchiata* from Florida is consistent with previous descriptions from Blake and Woodwick (1975). Ventral ciliated patches and an anterior circumferential band of cilia constitute the first ciliated structures to develop. The patches have been observed in 3 other spionid genera, and their function has been postulated as providing a means of movement by the embryos in the egg capsules before ciliary bands are formed (Blake and Woodwick 1975), although is not clear if the ventral patches are also present in other spionid genera. Blake and Woodwick (1975) mentioned that the most anterior band of cilia to develop is the first prototrochal ring; however, based on the location of the first circumferential ciliated band in *P. paucibranchiata*, it more specifically represents an akrotroch. This term was introduced by Hacker (1896), and it is defined as a distinct ciliated band encircling the anterior end between the prototroch and the apical plate. Akrotrochs have only been observed in trochophore larvae of polychaetous annelids, and they have been reported in 9 families Phillips and Pernet 1996; Rouse 1999; Fischer et al. 2010;), although not yet for spionids.

This investigation and previous studies indicate that larvae of *P. paucibranchiata* do not develop an apical tuft (Blake and Woodwick 1975). This structure is commonly found in annelids with trochophore larvae, and generally is considered a sensory organ, with mechanosensory or chemosensory functions (Lacalli 1981), but the role of the apical tuft is not well understood. Moreover, it is known that for some annelids the apical tuft is replaced by

secondary sensory structures. Lacalli (1981), studied larval behavior of *Spirobranchus polycerus* (Serpulidae) and found no obvious function for the apical tuft; although, after the loss of the apical tuft other sensory structures were developed, and he attributed these new sensory structures as responsible for the more complex behaviors observed. *P. paucibranchiata* develops long tactile cilia in the anterior region, thus it is possible that these cilia may have important sensory functions.

The development of other ciliary bands such as the prototroch, neurotroch and telotroch reflect a common pattern of trochal band formation during annelid trochophore development, but developmental timing of ciliated bands differs when compared to other taxa. *Scoloplos armiger* (Orbiniidae) for example, develops the prototroch, followed by metatroch, neurotroch, telotroch, and later the akrotroch is observed. Additionally, the akrotroch remains present during late larval stages (Anderson, 1959), while in *P. paucibranchiata* can only be observed during early larval development.

4.5.2 *Muscle development*

This study *P. paucibranchiata* revealed simultaneous development of longitudinal muscles associated with the body wall and internal musculature of the digestive system, as both sets of muscles can be observed before larval elongation starts. The larvae of *P. paucibranchiata* are planktotrophic, meaning that larvae will be swimming and feeding at the time of release from the egg capsule. Previous studies have found that during the development of planktotrophic larvae in annelids, such as observed with *Sabellaria alveolata* (Sabellaridae; Brinkmann and Wanninger 2008) and *Pomatoceros lamarkii* (Serpulidae; McDougall et al. 2006), the muscles associated with development of the digestive system are formed first, while muscles associated

with the body wall start to developed during larval elongation. In contrast, the lecithotrophic larvae of *Capitella* showed that body wall muscles developed before the muscles associated with the digestive system. McDougall et al. (2006) hypothesized that the difference in timing of development of the digestive system and body wall muscles is correlated with contrasting feeding strategies of planktotrophy vs lecithotrophy. However, Helm et al. (2013) studied development in the planktotrophic larva of *Phyllodoce groenlandica* and found that earlier stages first developed the muscles associated with body wall, contradicting the hypothesis of McDougall et al. (2006). Moreover, this study shows an alternate pattern in which digestive system and body wall muscles develop synchronously.

While muscles associated with the prototrochal ring tend to be the most prominent muscles during the development of planktotrophic trochophores (McDougall et al. 2006; Helm et al. 2013), the most prominent muscles observed during larval development in *P. paucibranchiata* are associated with the first chaetal sac. The overall shape of these muscles is similar to those described for *S. alveolata* (Brinkmann and Wanninger 2010), in that they form a conspicuously large basket-like structure. Larvae of *S. alveolata* are also similar to *P. paucibranchiata* in that the length of the first chaetae extend beyond the posterior region; moreover, they also show similar swimming behavior where the first chaetae are extended laterally in fast movements.

4.5.3 *Serotonergic elements of the nervous system development*

Based on some of the annelid literature on early neural development in larvae, there are at least four recognizable patterns for where the first serotonergic elements appear during development: (1) in Polygordiidae, the first serotonergic cells were found to be associated with

the developing ventral nerve cords (Hay-Schmidt 1995); (2) in Phyllodocidae and Serpulidae, the first serotonergic cells appeared at the posterior end of the larva (Voronezhskaya et al. 2003; McDougall et al. 2006); (3) in Sabellariidae, the first serotonergic elements were detected at later larval stages at the tip of each chaetal sac (Brinkmann and Wanninger 2008); and (4) in *Capitella*, the first serotonergic elements were found associated with the brain region once the cilia of the prototroch became visible (Meyer et al. 2015). In contrast, at an earlier stage of *P. paucibranchiata* that showed some initial evidence of prototroch development (stage II), serotonergic cells or neurites were not yet visible in the episphere. And, within pre-larval stages of *P. paucibranchiata*, there were also no signs of serotonergic cells, which may be just prior to detection of serotonergic-like immunoreactivity. Then, in early larval stages serotonergic cells and neurites associated with the first commissure of the development ventral nerve cord were detected, along with serotonergic growth cones in the region of the brain, and the longitudinal nerve cords extending toward the pygidial region. Thus early serotonergic development appears to align with pattern 1 for Polygordiidae, although only with a general similarity. Clearly, at least with respect to serotonergic neural development, there is notable variation within Annelida.

Subsequently, once serotonergic cells were consistently detectable, serotonergic activity continued to expand within anterior, middle, and posterior regions of the larva. Thus, it was not possible to identify where the first serotonergic elements appear in *P. paucibranchiata*. One possible explanation is that more stages between II and III need to be identified in order to find the origins of serotonergic activity in this species. Interestingly, a recent study of the spionid *Malacoceros fulginosus* found a similar pattern and suggested that early larval stages do not express serotonin or FMRFamide (Kumar et al. 2020). Thus, it appears that to study the development of the nervous system in spionid larvae, other markers will certainly be required.

4.6 Acknowledgements

We thank the staff from SMS for their valuable assistance with logistics and sampling. We also thank Greg Rouse for his guidance in the identification of major ciliary bands. This study was funded by the Link Fellowship.

4.7 References

Anderson DT (1959) The Embryology of the Polychaete *Scoloplos armiger*. *Quarterly Journal of Microscopical Science*, 100(1): 89–166.

Blake JA & Arnofsky PL (1999) Reproduction and larval development of the spioniform Polychaeta with application to systematics and phylogeny. *Hydrobiologia* 402: 57–106.

Blake JA & Woodwick KH (1975) Reproduction and larval development of *Pseudopolydora paucibranchiata* (Okuda) and *Pseudopolydora kempi* (Southern) (Polychaeta: Spionidae). Biol. Bull. 149: 109–127.

Blake JA (1969) Reproduction and larval development of *Polydora* from northern New England (Polychaeta: Spionidae). *Ophelia* 7: 1–63.

Blake JA (1996) Family Spionidae. p: 81-223. In JA Blake, B Hilbig & PH Scott (eds), Taxonomic Atlas of the Santa Maria Basin and Western Santa Barbara Channel. Vol. 6. Annelida Part 3. Polychaeta: Orbiniidae to Cossuridae. Santa Barbara Museum of Natural History.

Blake JA (2006) Spionida. p: 565-638. In Jamieson, B.G., G. Rouse & F. Pleijel (eds.), Reproductive Biology and Phylogeny of Annelida. Science Publishers, USA.

Boyle MJ & Rice ME (2014) Sipuncula: an emerging model of spiralian development and evolution. *International Journal of Developmental Biology*, 98: 485–499.

Boyle, MJ & Seaver EC (2010) Expression of FoxA and GATA transcription factors correlates with regionalized gut development in two lophotrochozoan marine worms: *Chaetopterus* (Annelida) and *Themiste lageniformis* (Sipuncula). *EvoDevo*, 1, 2. https://doi.org/10.1186/2041-9139-1-2

Brinkmann N & Wanninger A (2008) Larval neurogenesis in *Sabellaria alveolata* reveals plasticity in polychaete neural patterning. *Evolution and Development*, 10:606–18

Carrillo-Baltodano AM, Boyle MJ, Rice ME & Meyer NP (2019) Developmental architecture of the nervous system in *Themiste lageniformis* (Sipuncula): New evidence from confocal laser scanning microscopy and gene expression. *Journal of Morphology*, 280(11), 1628–1650.

Çinar, M. E. (2013). Alien polychaete species worldwide: Current status and their impacts. *Journal of the Marine Biological Association of the United Kingdom*, 93(5), 1257–1278.

Dagli E & Çinar ME (2008) Invasion of polluted soft substratum of Izmir Bay (Aegean Sea, eastern Mediterranean) by the spionid polychaete worm, *Pseudopolydora paucibranchiata* (Polychaeta: Spionidae). *Cahiers de Biologie Marine*, 49: 87–96.

Delgado-Blas VH (2009) Spionidae Grube, 1850. p: 589-613. In de León-González JA, Bastida-Zavala JR, Carrera-Parra LF, García-Garza ME, Peña-Rivera A, Salazar-Vallejo SI & Solís-Weiss V (eds.), Poliquetos (Annelida: Polychaeta) de México y América Tropical. Universidad Autónoma de Nuevo León, Monterrey, Mexico.

Fischer AHL, Henrich T & Arendt D (2010) The normal development of *Platynereis dumerilii* (Nereididae, Annelida). *Frontiers in Zoology*, 7(1), 31. https://doi.org/10.1186/1742-9994-7-31

Forest DL & Lindsay SM (2008) Observations of serotonin and FMRFamide-like immunoreactivity in palp sensory structures and the anterior nervous system of spionid polychaetes. *Journal of Morphology*, 269(5), 544–551. https://doi.org/10.1002/jmor.10605

Häcker V (1896) Pelagische Polychäten-Larven. Zur Kenntnis der Neapler Frühjahrs-Auftriebs. *Zeitschriftfür wissenschaftliche Zoologie* 62: 74–168.

Halanych KM (2016) How our view of animal phylogeny was reshaped by molecular approaches: Lessons learned. *Organisms Diversity and Evolution*, 16: 319-328.

Hannerz L (1956) Larval development of the polychaete families Spionidae Sars, Disomidae Mesnil, and Poecilochaetidae N. Fam. in the Gullmar Fjord (Sweden). *Zoologiska bidrag från Uppsala*, 31:1-204.

Hay-Schmidt A (1995) The larval nervous system of *Polygordius lacte*us Scheinder, 1868 (Polygordiidae, Polychaeta): Immunocytochemical *Data. Acta Zoologica*, 76(2), 121–140. https://doi.org/10.1111/j.1463-6395.1995.tb00987.x

Helm C, Schemel S & Bleidorn C (2013) Temporal plasticity in annelid development-ontogeny of *Phyllodoce groenlandica* (Phyllodocidae, Annelida) reveals heterochronous patterns. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 320(3), 166–178. https://doi.org/10.1002/jez.b.22492

Kumar S, Tume S, Helm C & Hausen H (2020) The neuron pioneering the ventral nerve cord does not follow the common pathway of neurogenesis in the polychaete M*alacoceros fuliginosus*. *In review* 1–35. https://doi.org/10.21203/rs.2.19981/v1

Lacalli TC (1981) Structure and Development of the Apical Organ in Trochophores of *Spirobranchus polycerus*, *Phyllodoce maculata* and *Phyllodoce mucosa* (Polychaeta).

Proceedings of the Royal Society of London Series B: Biological Sciences, 212(212), 381–402. https://doi.org/10.1098/rspb.1981.0045

McDougall C, Chen WC, Shimeld SM & Ferrier DEK (2006) The development of the larval nervous system, musculature and ciliary bands of *Pomatoceros lamarckii* (Annelida): Heterochrony in polychaetes. *Frontiers in Zoology*, 3, 1–14. https://doi.org/10.1186/1742-9994-3-16

Meyer NP, Boyle MJ, Martindale MQ & Seaver EC (2010) Lineage analysis of ectoderm and nervous system in the polychaete annelid *Capitella teleta*. *Developmental Biology*, 344(1), 530. https://doi.org/10.1016/j.ydbio.2010.05.391

Meyer NP, Carrillo-Baltodano A, Moore RE & Seaver EC (2015) Nervous system development in lecithotrophic larval and juvenile stages of the annelid *Capitella teleta*. *Frontiers in Zoology*, 12(1). https://doi.org/10.1186/s12983-015-0108-y

Myohara M (1980) Reproduction and development of *Pseudopolydora paucibranchiata* (Polycheata: Spionidae) under laboratory conditions, with special regard to the polar lobe formation. *Journal of the Faculty of Science Hokkaido University*, 22(2), 145–155.

Okuda S (1937) Spioniform polychaetes from Japan. *Journal of the Faculty of Science Hokkaido University*, 5:217–54.

Phillips NE & Pernet B (1996) Capture of large particles by suspension-feeding scaleworm larvae (Polychaeta: Polynoidae). *Biological Bulletin*, 191(2), 199–208. https://doi.org/10.2307/1542923

Purschke G, Böggemann M & Westheide W (2019) Handbook of Zoology: Annelida, 1: Annelida Basal Groups and Pleistoannelida, Sedentaria I. De Gruyter, Berlin, pp. 313–324.

Radashevsky VI (1993) Revision of the genus *Polydora* and related genera from the north west Pacific (Polychaeta: Spionidae). *Publications of the Seto Marine Biological Laboratory*, 36(1–2): 1–60.

Rouse GW (1999) Trochophore concepts: ciliary bands and the evolution of larvae in spiralian Metazoa. *Biological Journal of the Linnean Society*, 66: 411–464. https://doi.org/10.1111/j.1095-8312.1999.tb01920.x

Rueden CT, Schindelin J & Hiner MC, DeZonia BE, Walter AE, Arena ET & Eliceiri KW (2017) ImageJ2: ImageJ for the next generation of scientific image data, *BMC Bioinformatics* 18:529, PMID 29187165, doi:10.1186/s12859-017-1934-z

Simboura N, Kurt-Sahin G, Panagoulia A & Katsiaras N (2010) Four new alien species on the coasts of Greece (Eastern Mediterranean). *Mediterranean Marine Science*, 11(2), 341–352. https://doi.org/10.12681/mms.81

Simon JL (1967) Reproduction and larval development of *Spio setosa* (Spionidae: Polychaeta). *Bulletin of Marine Science* 17: 398–431.

Starunov VV, Voronezhskaya EE & Nezlin LP (2017) Development of the nervous system in *Platynereis dumerilii* (Nereididae, Annelida). *Frontiers in Zoology*, 14(1), 1–20. https://doi.org/10.1186/s12983-017-0211-3

Voronezhskaya EE, Tsitrin EB, Nezlin LP (2003) Neuronal development in larval polychaete *Phyllodoce maculata* (Phyllodocidae). *Journal of Comparative Neurology*. 455:299-309.

Vortsepneva E, Tzetlin A & Tsitrin E (2009) Nervous system of the dwarf ectoparasitic male of *Scolelepis laonicola* (Polychaeta, Spionidae). *Zoosymposia*, 2: 437–445.

Wilson WH (1991) Sexual reproductive modes in polychaetes: classification and diversity. *Bulletin of Marine Science*, 48: 500–516.

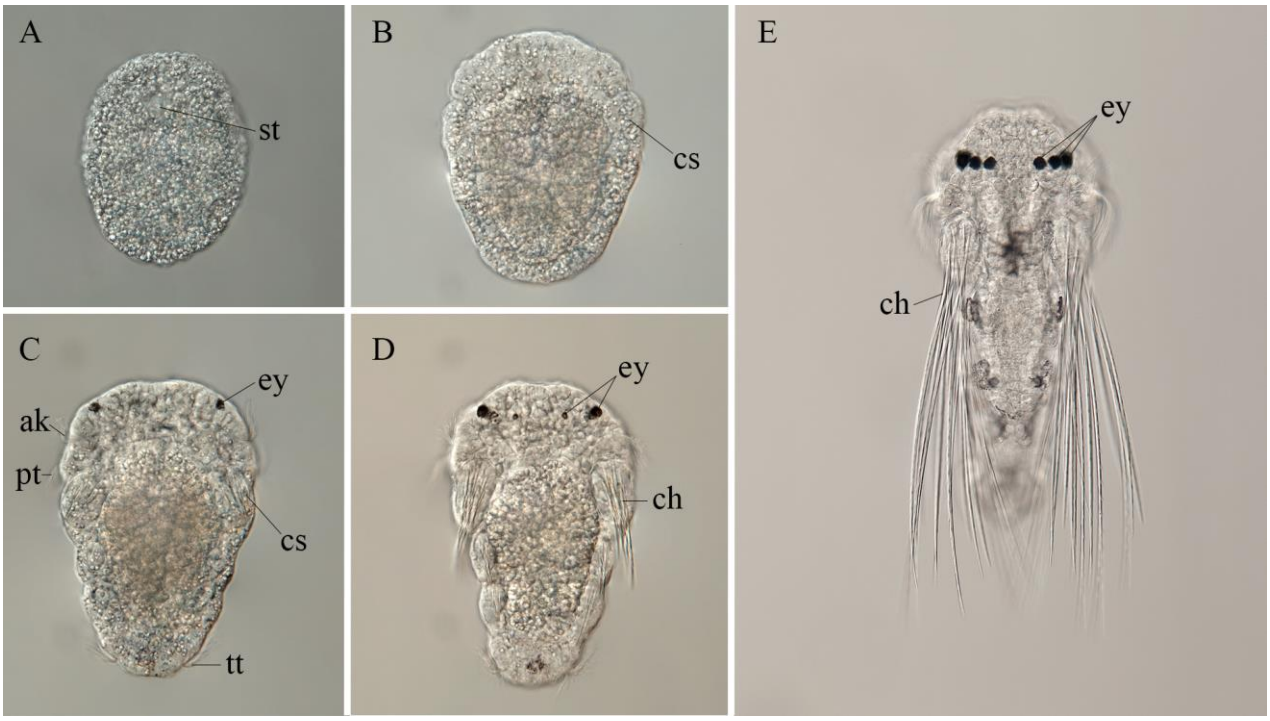

Figure 1. Five stages of growth during larval development of *Pseudopolydora paucibranchiata*. Differential interference contrast (DIC) micrographs. A, ventral view; B-E, dorsal views; all stages with anterior to the top. (A) Stage I, early pre-larva with mouth (stomodeum) visible. (B) Stage II, pre-setiger larva with first chaetal sac, yolk-filled coelom and distinct head region. (C) Stage III, larva with 3 developing segments, one pair of eyes, akrotroch, prototroch, telotroch and first set of chaetae (1st setiger). (D) Stage IV, 3-setiger larva with four eyes (small inner pair forming), elongation of chaetae from 1st setiger, telotroch and pygidial pigmentation. (E) Stage V, swimming larva with six eyes (3 pairs), more than 4 segments, chaetae extending beyond the posterior end and distinct patches of pigmentation along the body. ak: akrotroch; cs: chaetal sac; ch: chaetae; ey: eyes; pt: prototroch; st: stomodeum.

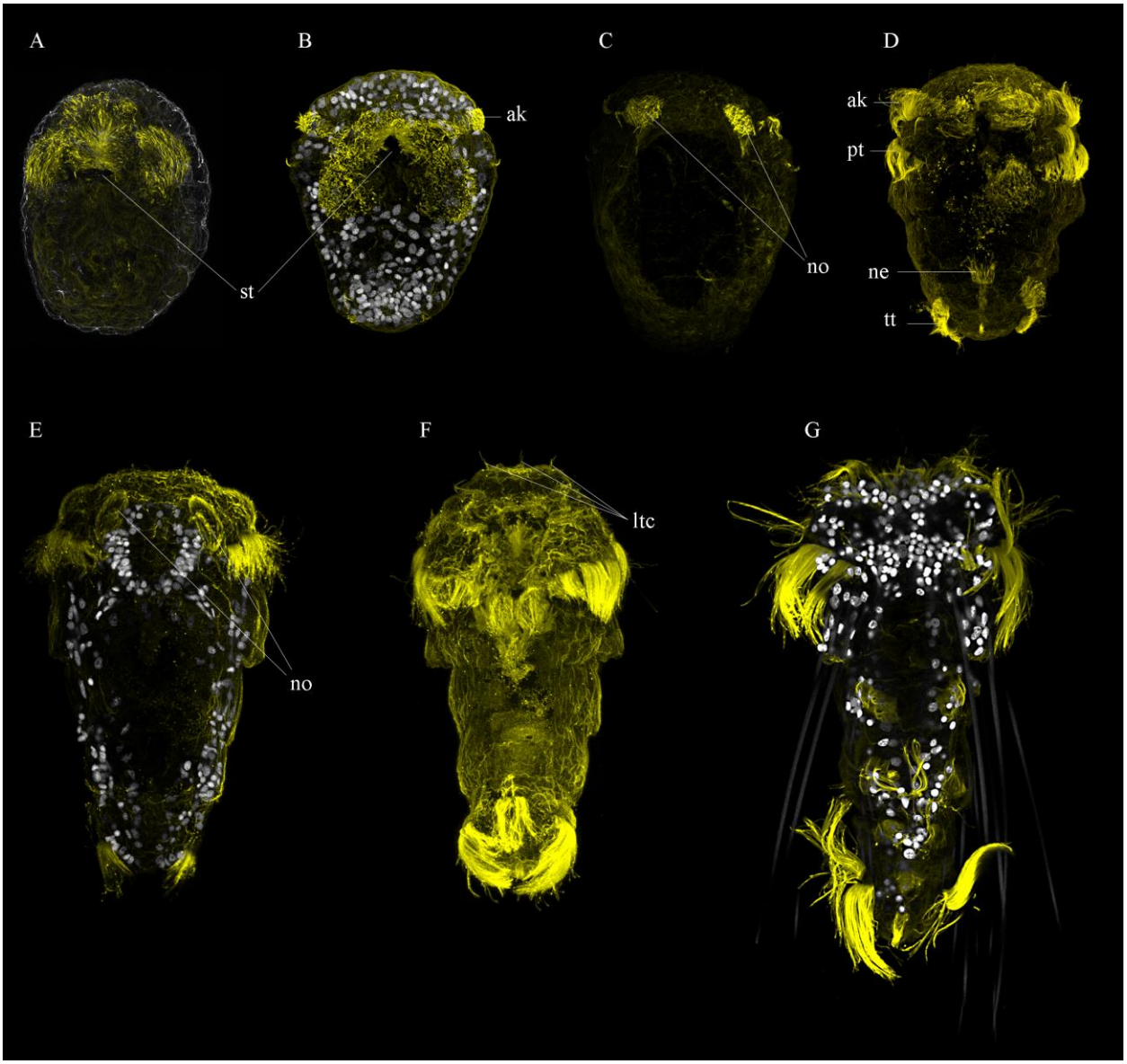

Figure 2. Patterns of ciliation during larval development in *Pseudopolydora paucibranchiata*. Confocal laser scanning microscopy (CLSM); z-stack projection micrographs of larvae labeled with anti-tyrosinated Tubulin for cilia (yellow), and propidium iodide for DNA (grayscale); all specimens with anterior to the top. (A) Stage I in ventral view with patches of cilia flanking the stomodeum. Stage II in ventral view (B) showing an akrotroch and first cilia of the prototroch, and in dorsal view (C) with pair of ciliated pits. (D) Stage III in ventral view showing major ciliary bands, including akrotroch, prototroch, telotroch, and neurotroch. (E-F) Stage IV in dorsal

(E) and ventral (F) views, now with long tactile cilia that are visible extending anteriorly from the head. (G) Stage V in ventral view showing notable extension in the length of cilia in the prototroch and telotroch. ak: akrotroch; ltc: long tactile cilia; ne: neurotroch; no: pre-nuchal organs; pt: prototroch; st: stomodeum; tt: telotroch.

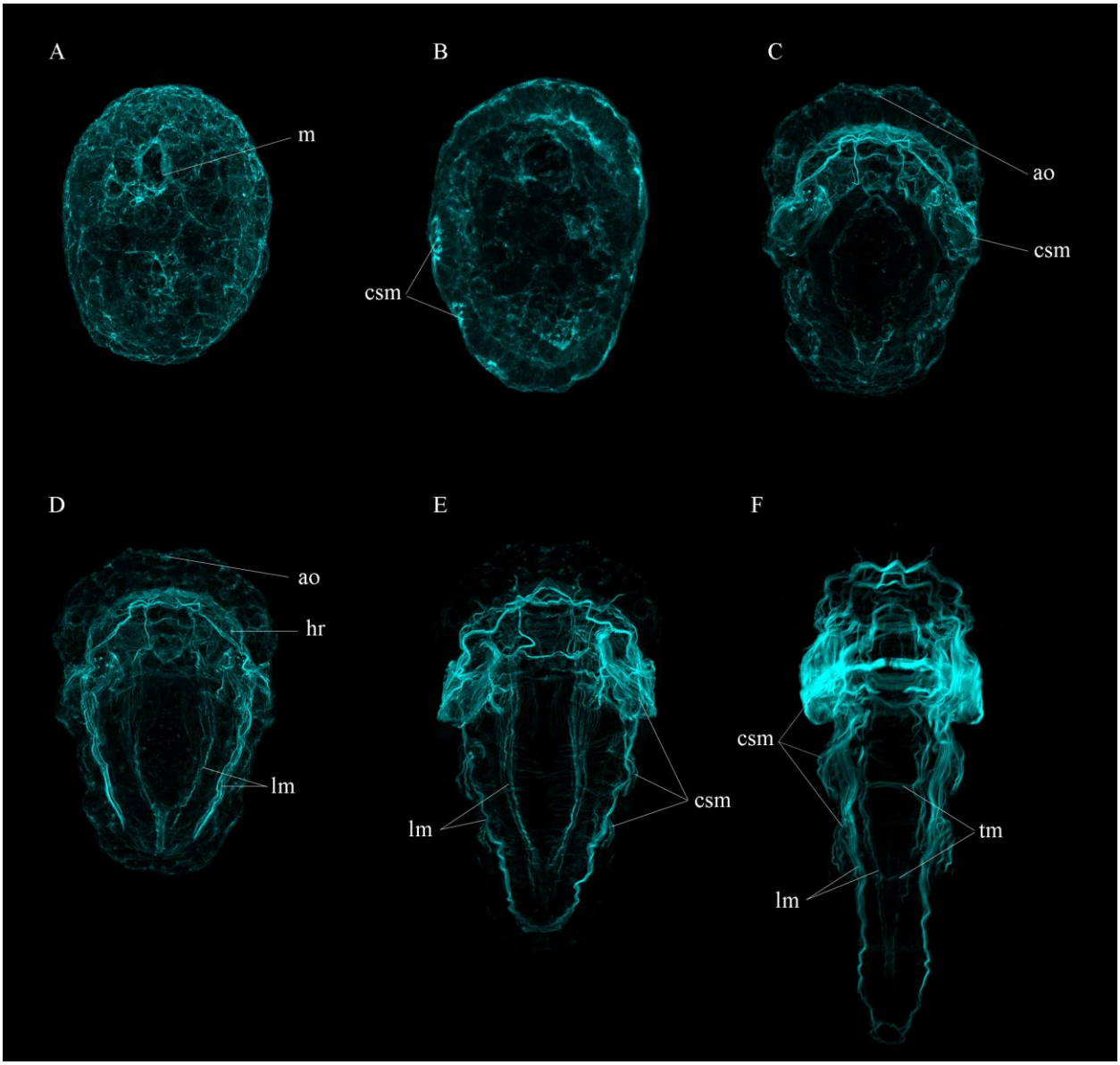

Figure 3. Muscle formation during larval development in *Pseudopolydora paucibranchiata*. Confocal laser scanning microscopy (CLSM); z-stack projection micrographs of larvae labeled with phallacidin for F-actin (cyan); all specimens with anterior to the top. (A) Stage I in ventral view with labeling around the mouth opening and non-distinct labeling of cellular margins, externally and internally. (B) Stage II in ventral view showing muscles associated with the chaetal sacs, and the anterior boundary of the body coelom. (C-D) Stage III in dorsal (C) and ventral (D) views. Muscles of the first chaetal sac, foregut and head retractors are visible.

Longitudinal fibers of internal and trunk-body retractors are developing. There is also actin labeling at the location of a putative apical organ. (E) Stage IV in dorsal view with chaetal sac muscles for setigers 1-3, anterior fibers extending into the head, and longitudinal muscles flanking the gut (visceral) and body (somatic) regions. The anterior chaetal sac musculature is extensive. (F) Stage V in dorsal view showing longitudinal muscles, transverse muscles and chaetal sac muscles. Larval head musculature is now fully developed. ao: apical organ; csm: chaetal sac muscles; hr: head retractor muscles; lm: longitudinal muscles; m: mouth; tm: transverse muscle.

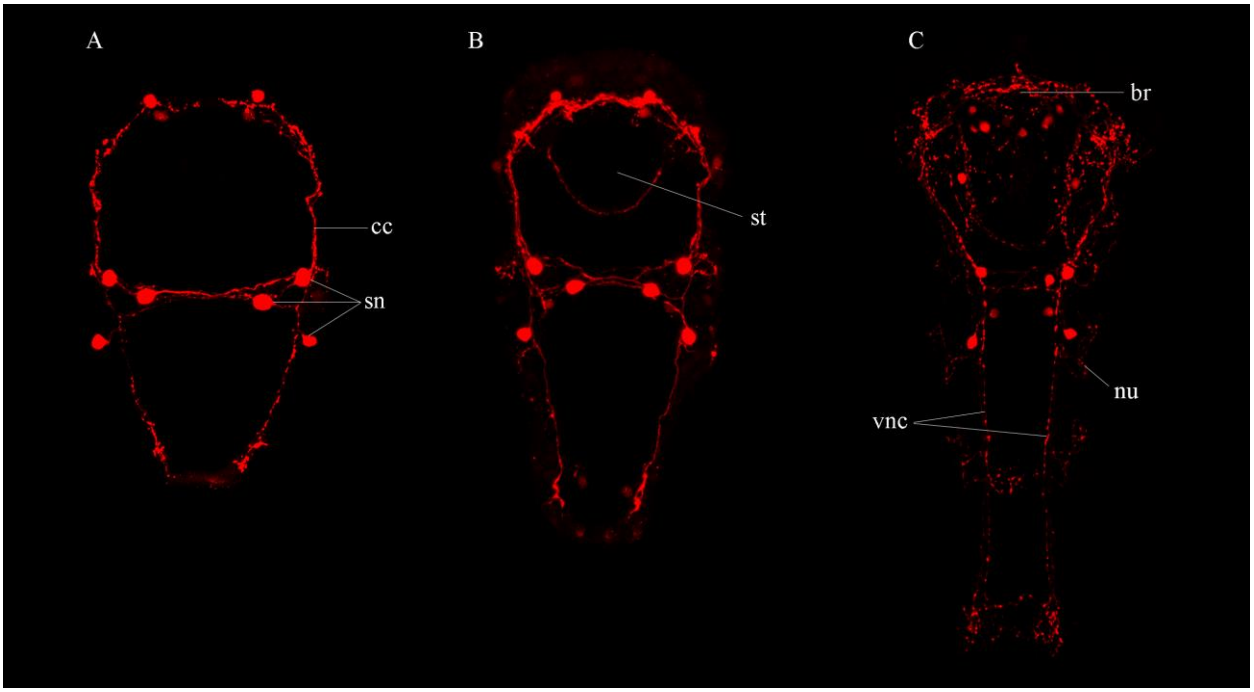

Figure 4. Serotonergic elements of the nervous system during larval development in *Pseudopolydora paucibranchiata*. Confocal laser scanning microscopy (CLSM); z-stack projection micrographs of larvae labeled with anti-serotonin antibodies for 5HT-lir (serotonin-like immunoreactivity, red); all specimens with anterior to the top. (A) Stage III showing a bilaterally symmetric pattern of 5HT-lir in cell bodies within the brain region, and at the ventral commissure where circumesophageal connectives meet two longitudinal connectives of the ventral nerve cord (vnc). Both termini of the vnc exhibit growth cones. (B) Stage IV with increased serotonergic activity in all previous areas, including additional cell bodies, and a U-shaped 5HT-lir ring around the stomodeum. Small neurites also project from the vnc toward chaetal-sac regions on left and right sides. (C) Stage V showing more serotonergic activity in the brain region, around the foregut, along and ventral nerve cord, within chaetal sacs, and now with posterior cross-connections on the anterior side of the pygidium. br: brain region; cc: circumesophageal connectives; nu: neurites; sn: serotonergic cells; st: stomodeum; vnc: ventral nerve cord connectives.

Chapter 5: Phylogenetic relationships within Spionidae Grube, 1850

5.1 Abstract

Spionidae is one of the largest and most diverse groups within annelids with 590 species and 38 recognized genera. They are commonly found in marine environments and are dominant members of the invertebrate community. Spionids are ranked among the most recognized annelid invasive species with some having economic impacts in the millions of dollars. Aside from their diversity and ecological impact, a robust phylogenetic understanding of the group remains unresolved. The most recent phylogenetic analysis is based on a parsimony framework using 38 characters derived from reproduction, larval development, and adult morphology. However, this analysis recovered spionids as a paraphyletic group. To elucidate the evolution of one of the most diverse groups of annelids, and test previous hypotheses, we studied phylogenetic relationships of Spionidae using a combination of whole genome sequencing (WGS) data and previously collected transcriptome data. Taxon sampling includes 28 Operational Taxonomic Units (OTUs) representing 17 spionid genera, as well as 1 trochochaetid, and 2 sabellarid annelids as the outgroups. We perform *de novo* assemblies of transcriptomic and genomic reads using Trinity and Ray respectively, and putative genes were extracted from fragmented genome assemblies with AUGUSTUS using *Schistosoma* (dataset S) or human (dataset H) as the model. Additionally, resulting predictions from both modeling runs were concatenated and use as a third dataset (dataset SH). Homology inference was conducted for each of the datasets using Orthofinder, and PhyloPyPruner was used for paralog removal. Moreover, three supermatrices for each of the datasets (S, H, SH) were constructed by varying the minimal gene occupancy threshold to 30% (S30, H30, SH30), 40% (S40, H40, SH40), and 50% (S50, H50, SH50). Phylogenetic reconstruction was conducted with IQ-tree. Maximum likelihood phylogenetic

recovered two major groups with high support. Clade 1 includes the *Polydora*-complex, *Pygospio, Spio, Scolecolepides, Scolelepis, Dispio, Lindaspio* and *Rhynchospio*. Clade 2 includes the *Prionospio*-complex, *Aonides, Spiophanes*, and *Laonice*. Importantly, *Trochochaeta* was recovered as the sister taxon to all spionid taxa, directly contrasting previous phylogenetic studies which suggested Spionidae represented a paraphyletic group. Topologies reconstructed from datasets SH30 and SH40, had higher nodal support values and included more taxa in the resulting species trees when compared with all other datasets, suggesting that combining gene predictions from multiple models provided a more robust dataset for resolving evolutionary relationships within spionids. In addition to supporting Spionidae as monophyletic using WGS datasets, we provide a well-resolved backbone within the family. This approach is highly reproducible and given its feasibility future studies can easily be combined with our dataset.

5.2 Introduction

Spionids are small (1 mm – 5 cm) sedentary tubicolous annelid worms recognized in part by the pair of long palps that can be used to collect particles from water sediment (Blake 1996; Blake & Arfnosky 1999). They are commonly found in shallow benthic sediments but are also present in the deep ocean. These worms can be dominant members of the invertebrate community (Blake 1996) and can act as ecosystem engineers altering environments (Elías et al. 2015). Spionids are ranked among the most recognized annelid invasive species (Çinar 2013) with some having economic impacts in the millions of dollars (e.g., to mollusk aquaculture; Blake, 1994; Simon et al. 2009; Sato-Okoshi et al., 2017). With approximately 590 species and 38 genera currently recognized (Blake et al. 2020), spionids exhibit the highest diversity of reproductive modes among annelids (Wilson 1991), have a robust ability to regenerate, have two types of asexual reproduction, and can possesses multiple patterns of larval development within the same species, or poecilogony (Chia *et al*., 1996). Their diversity is evident in spermiogenesis and oogenesis (Blake & Arnofsky 1999). Despite their ecological dominance and importance, their biological diversity, and their economic impact, a robust phylogenetic understanding of the group is wanting.

Spionidae was established by Grube (1850) and originally included the genera *Spio*, *Polydora*, *Malacoceros* and *Scolelepis*. Among the first works on spionid systematics were Mesnil (1896) and later Söderström (1920) who proposed three major groups within spionids: Nerininae, Laonicinae and Spioninae. The next substantial works on the group were by Hannerz (1956) and Orrhage (1964) who proposed alternative classifications using information of reproductive and larval traits or anatomy and morphology of adults, respectively. However, the first phylogenetic analysis for the family, not conducted until 1997 (Sigvaldadóttir *et al.* 1997),

included 25 morphological characters across 28 genera, and recovered 4 major clades of spionids, albeit with weak bootstrap support (0.43-0.65). This study further suggested that Uncispionidae was nested within Spionidae. In 1999, Blake and Arnofsky expanded this work using 38 characters focused on reproduction, larval development, and adult morphology from 35 taxa. They recovered three clades with high support (Pygospiopsis, Spioninae, Nerininae), but also hypothesized that Spionidae may be paraphyletic, with other recognized families (e.g. Uncispionidae, Poecilochaetidae, and Trochochaetidae) nested within Spionidae. With no other significant efforts since their work, current knowledge of Spionidae phylogeny is based on morphological parsimony analyses conducted over two decades ago.

Among spionids two main groups have been historically recognized, these are *Prionospio*-complex and *Polydora*-complex (or polydorids). Importantly, they comprised around 50% of the species recognized. *Polydora*-complex is unique among spionids by having the fifth chaetiger modified, it is the largest group with 115 species and 9 genera that includes: *Amphipolydora*, *Boccardia*, *Boccardiella*, *Caraziella*, *Dipolydora*, *Polydora*, *Polydorella*, *Pseudopolydora* and *Tripolydora*. This group includes most of shell borer species such as *Polydora websteri*, which has been extensively studied because of its damages to the aquaculture industries (Martinelli et al. 2020). The *Prionospio*-complex on the other hand has been the subject of different studies and genera included under this group varies by different authors (Foster 1971, Macioleck 1985, Blake and Kudenov 1978, Yokoyama 2007). In the last review by Blake et al. (2020), 9 genera and 100 species are recognized as part of *Prionosprio*-complex, including *Prionospio*, *Paraprionospio*, *Aquilaspio*, *Minuspio*, *Aurospio*, *Apoprionospio*, *Laubierellus*, *Orthoprionospio*, and *Streblospio*. They can be identified based on a composition of a rather general characters including: branchiae first present from chaetiger 1-3 limited to

anterior part of the body and of different forms, prostomium simple without frontal horns, bi- or multidentate hooded hooks in noto and neutopodia, and notopodial lamellae in anterior segments well develop. Currently, there is no consensus on the support for the *Prionospio*-complex or whether this group represents a monophyletic assemblage.

Phylogenetic assessment of across Spionidae using molecular data is lacking (Blake 2006), even though molecular phylogenetic analyses are common for other annelid groups (e.g., Aguado et al. 2011; Weigert et al. 2014; Li et al. 2017; Eilertsen et al. 2017; Stiller et al. 2020). Moreover, molecular data, in general, for this group is limited. Previous studies including molecular data have focused on problematic taxa such as invasive species (Wiiliams et al. 2017; Radashevsky et al. 2019, Martinelli et al. 2020), and species complexes with genera (Blank et al. 2007; Rice et al. 2008; Simon et al. 2009; Sato-Okoshi and Abe 2013, Simon et al. 2017), which are usually tackled using a limited set of markers (e.g. COI, 16S, 18S, 28S), and only small number of studies have aimed to resolve phylogenetic relationships of specific genera (Blank and Bastrop 2009; Abe et al. 2016; Radashevsky et al. 2016; Bogantes et al. 2018; Guggolz et al. 2019). The first mitochondrial genome was only recently published (Gastineau et al. 2019). Spionids may be particularly challenging because their small size limits the ability to identify and preserve them for genetic analyses before nucleic acid residues are damaged. This is exacerbated by subtle and fragile morphological features that are important to taxonomic identification (Radashevsky et al. 2014; Bogantes et al. 2018). Unraveling spionid phylogeny will likely require tools that overcome these challenges and allow data from new samples to be easily integrated with existing data.

Numerous phylogenetic studies on a variety of organisms have used techniques to partition or reduce the complexity of genomes (e.g., transcriptome sequencing, anchored

enrichment, RAD-tags and target capture) in part because of their advantages over whole genome sequencing (WGS) in terms of sequencing and computational cost (Cruaud et al. 2014; Weigert et al. 2014; Fitz-Gibbon et al. 2017). However, newer platforms (e.g. Hiseq X Ten, NovaSeq 6000, Nanopore) have reduced sequencing costs, making them comparable to approaches that only sequence a portion of the genome (Zhang et al. 2019). In addition, WGS requires no prior bioinformatic work before sequencing allowing access to study taxa with limited prior genomic resources. Importantly there is a larger utility of data generated as it can be used to study non targeted genes, assembly of mitochondrial genomes and identification of microbial associates among others (Allen et al. 2017; Zhang et al. 2019). Continued decline of sequencing costs (Wetterstrand 2020) should facilitate implementation of WGS for phylogenomic studies in a broad variety of taxa.

Potential disadvantages of working with WGS data involve higher computational costs and processing time of downstream tasks such as genome assembly, and identification of ortholog sequences required for phylogenetic analyses. However, full genome assemblies are not required, and thousands of orthologs can be extracted from fragmented assemblies (Allio et al. 2019). Moreover, a variety of bioinformatics tools for orthology inference have been developed (e.g. OrthoMCL - Li et al. 2003, Orthofinder - Emms and Kelly 2019). Orthology inference however can be influence by many steps involve during curation (e.g. cleaning, assembly) and annotation of the data of interested. For example, annotation tools like gene prediction software (e.g. Augustus Stanke et al. 2004, MAKER Cantarel et al. 2008) can predict very different amounts of putative genes in the same dataset (Veeckman et al. 2016). Additionally, many of these tools have been trained on a small number of taxa or model organisms only; this will likely

influence the accuracy and breadth of predicted genes for taxa for which genomic annotations are lacking, and further affecting orthology inference.

Here, we studied phylogenetic relationships of Spionidae using a combination of WGS data augmented with previously collected transcriptome data, to better understand the evolution of one of the most diverse groups of annelids. Specific hypothesis to be addressed include: Is Spionidae a paraphyletic group? Does the *Prionospio*-complex represent a monophyletic clade? A better understanding of Spionidae phylogeny, will facilitate interpretation of plesiomorphic characters. Moreover, phylogenetic inferences from different gene model predictions are compared, and an approach for improving such predictions is discussed.

5.3 Methods

5.3.1 *Taxon sampling and preservation*

Twenty-nine Operational Taxonomic Units (OTUS) representing 17 spionid genera, 1 trochochaetid, and 2 outgroups were included in this study (Table 1). Specimens were preserved in 95% ethanol, RNA later, or frozen at -80 °C for molecular work and preserved in ~4% formalin for vouchers. Voucher specimens are deposited at Auburn University Museum of Natural History (AUMNH) and Senckenberg Museum Frankfurt (SMF). Our Spionidae dataset includes 24 genomic sequences generated by this study and 5 transcriptomes previously sequenced, 4 of these transcriptomes were obtained as part of the Wormnet II project, and 1 transcriptome (*Rhynchospio* sp.) was provided by Dr. Kevin Kocot. A transcriptome of *Trochochaeta* sp. was included, based on a previous morphological phylogenetic hypothesis that indicated Trochochaetidae is nested within Spionidae (Blake and Arnofsky 1999). Available

transcriptomes of two sabellarids (*Neosabellaria cementarium* NCBI SRR2017810; *Idanthyrsus* sp.) were used as outgroups based on phylogenetic results of Weigert et al. (2014).

5.3.2 *DNA isolation and sequencing*

For spionid samples, one-millimeter tissue clips, or whole organisms in the case of small specimens, were used for DNA extraction. Total genomic DNA was extracted using the Qiagen DNeasy® Blood and Tissue Kit or the MicroElute genomic DNA kit (for the smaller worms) following manufacturer's protocols. Library preparation and sequencing of total genomic DNA was performed by HudsonAlpha Discovery, Huntsville, Alabama. Sequencing was conducted on Illumina NovaSeq 6000 platform, using S4 300 chemistry run in a 2x150bp configuration.

5.3.3 *RNA extraction and sequencing*

RNA extraction, cDNA preparation, and high-throughput sequencing was conducted following Whelan et al. (2015) modified from Kocot et al. (2011). Briefly, total RNA was extracted from tissue clips of the body wall or from whole specimens in the case of smaller worms, using the Invitrogen TRIzol® protocol, and purified with Qiagen RNeasy® kit with on-column DNase digestion. Clontech SMART cDNA library construction kit was used to reverse single stranded RNA template, and double stranded cDNA synthesis was conducted using Clontech Advantage Clonetech SMART cDNA library construction kit 2 PCR system. Library preparation and sequencing was performed by The Genomic Services Lab at the HudsonAlpha Discovery (formerly HudsonAlpha Institute of Biotechnology), Huntsville, Alabama. Sequencing was conducted on Illumina Hiseq 2000 or 2500 platform, using 2 x 100 or 2 x 125 paired-end runs on either v3 or v4 chemistry.

5.3.4 *Assembly, gene prediction and ortholog assignment of genomic and transcriptomic data*

The bioinformatic pipeline used herein is provided in Fig. 1. Genomic paired-end reads were trimmed with Trimmomatic 0.38 (Bolger et al. 2014) to remove low quality bases from the start and end of reads (LEADING:3, TRAILING:3), remove reads below 50 bp (MINLENGTH: 50), and evaluate read quality using the sliding window trimming approach (SLIDINGWINDOW: 4:15). Trimmed reads were assembled *de novo* using Ray 2.20 (k-mer= 31; Boisvert et al. 2010). Use of other assemblers, including SOAPdenovo2 (Luo et al. 2012) and DISCOVAR de novo, was explored for a subset of samples. However, in comparison to Ray, SOAPdenovo2 produced smaller contigs, and DISCOVAR assembly was not finalized for any of the samples. Reads from transcriptomes were trimmed, normalized, and assembled with Trinity 2.8.5 (Grabherr et al. 2011).

To estimate sequence coverage reads were mapped to assembled contigs with Bowtie2 (Langmead and Salzberg 2012). SAMtools (Li et al. 2009) was used to convert SAM files to BAM files followed by bam sorting. Estimated coverage per contig was obtained with BEDtools (Quinlan and Hall 2010), and averaged depth per contig was calculated to get total sequence coverage.

Gene prediction for genomic assemblies was conducted with Augustus 3.3 using default settings (Stanke et a. 2008). Three datasets were generated, the first dataset contains gene predictions when using *Schistosoma* as the gene model (dataset: S), for the second dataset human was used as the gene model (dataset: H). A 3rd dataset was generated by concatenating predictions from *Schistosoma* and human (dataset: SH), and predicted duplicates were removed

with cd-hit (-c 1.0; Fu et al. 2012). Predictions from Augustus were extracted as gene sequences with the gffread utility.

Candidate coding regions from the 24 genomic and 7 transcriptomic data were predicted using TransDecoder (Haas et al. 2014) and the longest Open Reading Frames (ORF) were used for further analyses. Duplicates of predicted ORF were removed using cd-hit (-c 1.0), keeping only unique sequences for each taxon. Putative homologue detection was conducted with OrthoFinder (Emms and Kelly 2015; 2019).

5.3.5 *Phylogenomic analyses*

Gene trees inferred for each orthogroup (OG) and corresponding multiple sequence alignments (MSA) generated from OrthoFinder were used as the input for PhyloPyPruner for paralogue removal and further filtering of OGs. PhyloPyPruner (Thalén et al. *in prep.*) is a tree-based orthology inference tool that uses the output of graph-based homology inference approaches including OrthoFinder to select orthologs. PhyloPyPruner is a new tool based on PhyloTreePruner (Kocot et al. 2013) with additional features based on Yang and Smith 2014 and Roure et al. 2007; and novel features that enhance the specificity of paralog and contaminant detection and removal. Using PhyloPyPruner, three supermatrices for each of the datasets (S, H, SH) were constructed by varying the minimal gene occupancy threshold by 30% (dataset: S30, H30, SH30), 40% (dataset: S40, H40, SH40) and 50% (dataset: S50, H50, SH50). Other thresholds (60-70%) were also tested but without good results (see results section), thus they were not included in further analyses. Other filter settings used included removal of sequences shorter than 95 amino acids (min-len=95), removal of branches longer than 5 times the standard deviation of all branches (trim-lb=5), collapse of nodes with less than 75% bootstrap support

into polytomies (min-support=0.75). Maximum inclusion (MI) algorithm was selected for paralogy pruning (prune=MI) and minimal number of taxa for each ortholog was set to 16 (mintaxa=16). Taxa less than 10% occupancy in each of the supermatrices were omitted from the species tree reconstruction. Additionally, removal of missing data from supermatrices generated by PhyloPyPruner was explored with MARE for the nine datasets (Misof et al. 2013).

Maximum-likelihood analyses for the 9 datasets were conducted using IQ-TREE 1.6.12 (Nguyen et al. 2015). Best partition schemes were selected by PartitionFinder (Lanfear et al. 2012), and model selection was determined by ModelFinder (Kalyaanamoorthy et al. 2007) included in IQ-tree (-m MPF+MERGE). Nodal support was assessed with 1000 replicates using the ultrafast bootstrapping setting (-bb 1000).

Results

5.4.1 *Genome assembly and gene prediction*

A median of 71.4 million reads per taxon were generated from Illumina sequencing, and 64.6 million reads remained after cleaning (Table 2). Statistics for the genomes generated in this study are provided in Table 2.

Numbers of gene predictions for the genomic data using different gene models (*Schistosoma*, human) varied. Overall, the Augustus output produced using only *Schistosoma* data as gene model yielded fewer (2,415 - 44,466) gene predictions compared to human (2,507 – 95,018) (Fig. 2). For 17 of the genomic data, number of gene predictions increased over 2 or 3 times when using human as a gene model (e.g. *Boccardia proboscidea*, *Dipolydora commensalis*, *Polydora cornuta*). However, seven genomic assemblies showed a reduced number of gene predictions under this scenario. Two taxa in particular, *Paraprionospio* sp. and

Scolelepis chilensis, had gene predictions reduced by half (Fig. 2) with human as a model. Importantly, the number of gene predictions when combining predictions from *Schistosoma* and human were higher for all taxa (Fig. 2).

5.4.2 *Orthology assignment and data matrix assembly*

Number of resulting homologs from OrthoFinder of genomic and transcriptomic data were lower for dataset S (28,095) and H (51,214), and higher for dataset HS (56,744), indicating, that the combined dataset allowed for more retrieval of OGs. Similarly, orthology filtering and paralogue removal for the different minimal gene occupancies using PhyloPyPruner resulted in higher number of ortholog extraction for the HS datasets, followed by H dataset (table 3). Taxa with less than 10% occupancy in each of the supermatrices (*Laonice* sp.b, *L. sarsi*, and *Pseudopolydora floridensis)* were omitted from the species tree reconstruction. These three taxa also had the lowest coverage, N50 values, and number of gene predictions. Number of raw reads was similar across *Laonice* sp. b, *Spio decoratus*, and *Dipolydora commensalis*, but coverage was higher for the two later taxa. Minimal gene occupancy of 60 - 70% were also explored, and the resulting supermatrices were the same composed of 18 orthologs, and a concatenated alignment length of 10502 for the dataset HS, number of resulting orthologs was even lower for datasets S (6 orthologs, 4289 alignment length) and H (8 orthologs, 5479 alignment length). Despite the relatively low setting for minimal gene occupancy, average matrices completeness for the nine datasets range from 64 -70% (Table 3), but number of orthologs was highly reduced when specifying 50% minimal gene occupancy, resulting in supermatrices composed of 8 (S50), 17 (H50), and 34 (SH50) orthologs.

To reduce amount of missing data, supermatrices were filtered with MARE, but there was a tradeoff between matrix completeness and number of taxa. For all the datasets, average completeness increased to 85 to 90%, but the number of included taxa was reduced (from 31 to 20-16). Given that a goal of this study is to resolve phylogenetic relationships of spionids, taxanomic inclusiveness was prioritized over matrix completeness. Thus, datasets filtered with MARE were excluded from phylogenetic reconstruction.

5.4.3 *Phylogenomic analyses*

Resulting topologies for the 9 datasets are shown in Fig. 3. Topologies reconstructed from datasets SH30, SH40, had higher nodal support values than other data sets (Fig. 3A-F). These datasets also included more taxa in resulting species trees than the other datasets (S30, S40, S50, H30, H40, H50; Fig 3. A-F) suggesting that combining predictions from different models provided a more robust dataset for resolving evolutionary relationships within spionids. Topologies reconstructed with datasets S and H recovered similar branching order for most of the groups, but deep nodes were weakly supported (Figs. 3A-F). In general, resulting trees from datasets produced with *Schistosoma* as gene model prediction had the lowest nodal support values (Figs. 3B-C).

The SH30 and SH40 topologies have highest node support value and are shown in figure 3A-B. Two major groups were recovered, clade 1 includes the *Polydora*-complex, *Pygospio*, *Spio*, *Scolecolepides*, *Scolelepis*, *Dispio*, *Lindaspio* and *Rhynchospio*. While clade 2 includes *Prionospio*-complex, *Aonides*, *Spiophanes*, and *Laonice*. *Trochochaeta* was recovered as the sister taxon to all spionids, suggesting that Spionidae does not represent a paraphyletic group. Notably, tree reconstruction using SH50 dataset results in a topology with overall weaker node

support (in comparison to SH30 and SH40) and recovered *Trochochaeta* as nested within spionids (Fig. 3I).

Relationships among clade 1 were consistent across all trees except for the position of *Scolecolepides*. *Lindaspio* and *Rhynchospio* were always recovered together (BS=100) and are the sister group to all other genera in clade 1. Other well supported clades include *Scolelepis* and *Dispio* which clustered together in all the analyses, the *Polydora*-complex (highly supported in all topologies, BS=100), and placement of *Pygospio* as the sister taxon to this clade. In contrast, the position of *Scolecolepides* is unclear, as it is recovered with *Spio* albeit with week support (BS=62; Fig. 3H), or as sister taxon to *Spio* + *Pygospio* + *Polydora*-complex (Fig. 3G, Fig. 4). Interestingly, *Scolecolepides* was excluded from several reconstructions (S30, S40, S50, S40, S50) due to high amounts of missing data. Thus, the recovered positions for this taxon needs further investigation.

Clade 2 supports monophyly of *Prionospio*-complex with *Streblospio* found as the most basally branching taxon for this group. *Aonides*, *Spiophanes* and *Laonice* were recovered as the sister taxa to *Prionospio*-complex including *Streblospio* (H30, H40, SH30, SH40, SH50). The position of *Spiophanes* or *Aonides* was unstable, with Spiophanes as the sister taxon to *Aonides* + *Laonice* (Fig. 3G, Fig. 4), or *Aonides* as the sister group to *Spiophanes* + *Laonice* (Fig. 3H).

5.5 Discussion

Our study, which provides a framework for spionid phylogeny (Fig 4), shows that tree reconstruction was influenced by three main factors: sequence data coverage, gene model used for gene prediction, and number of resulting orthologs in supermatrices. The three taxa excluded from reconstructions (*Laonice sarsi*, *Laonice* sp. b, *Pseudopolydora floridensis*) had lowest N50

values and data coverage (3x-5x), followed by *Scolecolepides viridis* which was recovered as a problematic taxon. Lower coverage resulted in more fragmented genome assemblies, consequently reducing number of gene predictions for these taxa and ultimately lower ortholog extraction. These findings are consistent with those of Zhang et al. (2019) who explored impacts of varying sequence coverage of 1-30x from whole genome sequencing for phylogenomic analyses. They found that a minimal of 10x was recommended because lower coverages resulted in datasets with reduced number of ortholog extraction. Similarly, Allen et al. (2017) compared recovery of 100 targeted genes from whole genome assemblies at different sequence coverages (1x, 5x, 10x, 20x, 40x) from 15 species of lice, and found that coverage affected the number of genes assembled. As expected, lower coverages recovered less genes, however coverages of 10x and above allowed for the recovery of almost all genes for close and distantly related taxa.

Furthermore, orthology inference was affected by the gene prediction, and specifically the model organism used to aid annotation. Topologies reconstructed with datasets S and H recovered similar branching order for most of the groups, but deep nodes were weakly supported (Figs. 3A-F). In general, resulting trees (Figs. 3B-C) produced with *Schistosoma* as gene model prediction had the lowest nodal support values. Even though *Schistosoma* is more closely related to annelids than human, using human as the gene model provided a more robust annotation in 70% of the genomic assemblies. This might be explained by the fact that the human genome is arguably the best annotated genome (Salzberg et al. 2019) within eukaryotes, increasing the resolution for gene prediction even for evolutionarily distant groups. However, that was not the case for all taxa. *Paraprionospio* had 42% less gene predictions with human as the gene model than with *Schistosoma*, which caused the removal of *Paraprionospio* from tree reconstruction in

datasets H40, and H50. Moreover, topology from H30 was the only case that *Paraprionospio* was recovered as the sister taxon to *Prionospio*-complex.

An alternative approach used in this study to improve gene prediction in all taxa was to combine gene predictions from the different models. Dunne and Kelly (2017) used a similar approach and found that the completeness of genome annotations in plants and fungi improved significantly when using gene predictions from multiple taxa, allowing genes that escape detection when using one gene model to be annotated with another. Better annotation results in identification of more orthologs. This result is important because a greater number of loci positively affects phylogenetic "accuracy" (Rokas et al. 2003; Phillips et al. 2004; Rokas and Carroll 2005; Yang and Smith 2014). For example, trees recovered with datasets SH30 (280 orthologs, Fig3G) and SH40 (114 orthologs, Fig 3H) have topologies and higher nodal support than SH50 (Fig 3I) which includes only 38 orthologs and suggests Spionidae is paraphyletic due to the placement of *Trochochaetidae*. Combining gene predictions from multiple species can be particularly useful when working with genomes from different sequencing coverage and fragmented assemblies.

Spionidae phylogeny

This study recovered 2 clades with high support (Fig. 4). Clade 1 includes Spioninae, this group has been supported in all systematic studies of spionids except for Sigvaldadottir et al. (1997), and it has been considered the "best defined group" (Soderstrom 1920; Hannerz 1956, Orrhage 1964; Blake and Arnofsky 1999). Spioninae was established based on similarities of the shape of the nephridia (Söderström 1920), all species having thin membrane eggs, long headed sperm, and egg capsules incubated by females within their tubes. Within Spioninae there is

Polydora-complex, this data supports the modification in the 5th segment as a synapomorphy for this group.

Interestingly, clade 2 from our study supports clade 2 from Sigvaldadottir et al. 1997. This group is composed by *Prionospio*-complex, *Aonides, Laonice* and *Spiophanes*, and resembles the subfamily Laonicinae as delimitated by Orraghe (1964). Additionally, *Prionospio*-complex was also recovered by Blake and Arnofsky (1999). *Prionospio*-complex has been considered morphologically a diverse but distinct group among spionids (Foster 1971; Macioleck 1985). Based on Sigvaldadottir et al. (1997), the *Prionospio*-complex was supported by two synapomorphies, the presence of multiple sharp secondary teeth in neuropodial hooded hooks and the number of anal cirri from 1 to 3. Although the presence of sabre setae was considered the unifying character of clade 2, some species of *Dispio, Spio*, and *Scolecolepides* also have sabre setae and this character likely represents a homoplasy. No morphological characters are uniquely shared among this group.

This study provides a robust backbone for Spionidae phylogeny that includes hypotheses on the evolutionary relationships for 17 of its genera. The genomic approach using whole genome sequencing represented the most flexible and cost-effective option, and it facilitated the inclusion of taxa from museum collections. Previous bioinformatic work for identification of targeted regions was not required, reducing processing time while still allowing for the extraction of hundreds of orthologs required for phylogenetic reconstruction. Moreover, the data generated can be used to explore other genomic aspects of spionids like mitochondrial composition and gene order, among others. This approach is highly reproducible and given its feasibility future studies can easily be combined with our dataset.

5.6 Acknowledgements

I would like to thank all my friends and colleagues who send me samples: Karin Meiβner, Jason Williams, Andrew Davis, Mauricio Shimabukuru, Cinthya Santos, Allan Carrillo, Liron Goren, William Walton and Scott Rikard. I also thank Christina Zakas for sharing *Streblospio* data. Thanks to Kevin Kocot for sharing the transcriptome of *Rhynchospio* and deep-sea samples with spionids, in addition to his valuable feedback for methods. I also thank Nicole Harrison for discussing bionformatic methods and help with pipeline designs. I would like to thank James Blake for sharing his knowledge and answering all my questions about spionids and polychaetes in general.

5.7 References

Abe H, Kondoh T, Sato-Okoshi W. First report of the morphology and rDNA sequences of two *Pseudopolydora* species (Annelida: Spionidae) from Japan. Zoolog Sci.

Aguado M.T., Siddall M.E. 2011. Cladistics. 27:1–17.

Allen J.M., Boyd B., Nguyen N.P., Vachaspati P., Warnow T., Huang D.I., Grady P.G.S., Bell K.C., Cronk Q.C.B., Mugisha L., Pittendrigh B.R., Leonardi M.S., Reed D.L., Johnson K.P. 2017. Phylogenomics from whole genome sequences using aTRAM. Syst. Biol. 66:786–798.

Allio R., Scornavacca C., Nabholz B., Clamens A.L., Sperling F.A.H., Condamine F.L. 2020. Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. Syst. Biol. 69:38–60.

Blake J.A. & Arnofsky P.L. 1999. Reproduction and larval development of the spioniform Polychaeta with application to systematics and phylogeny. *Hydrobiologia,* 402: 57–106.

Blake J.A. 1994. Family Spionidae. In: Blake J.A., Hilbig B., Scott P.H., editors. Taxonomic Atlas of the Santa Maria Basin and Western Santa Barbara Channel. Volume 6. The Annelida Part 3 - Polychaeta: Orbiniidae to Cossuridae. Santa Barbara, CA: Santa Barbara Museum of Natural History. p. 81–224.

Blake J.A. 1996. Family Spionidae. p: 81-223. In Blake, J. A., B. Hilbig & P. H. Scott (eds), Taxonomic Atlas of the Santa Maria Basin and Western Santa Barbara Channel. Vol. 6. Annelida Part 3. Polychaeta: Orbiniidae to Cossuridae. Santa Barbara Museum of Natural History.

Blake J.A. 2006. Spionida. p: 565-638. In Jamieson B.G., Rouse G. & Pleijel F. (eds.), Reproductive Biology and Phylogeny of Annelida. Science Publishers, USA.

Blake J.A., Macioleck N.J. & Meiβner K. 2020. Spionidae Grube, 1850. In: Handbook of Zoology. Annelida. Volume 2: Pleistoannelida, Sedentaria II. Purschke G., Böggemann, M., Westheide, W. (eds.) DeGruyter, Berlin.

Blank M., Bastrop R. 2009. Phylogeny of the mud worm genus *Marenzelleria* (Polychaeta, Spionidae) inferred from mitochondrial DNA sequences. :313–321.

Bogantes V.E., Halanych K.M., Meißner K. 2018. Diversity and phylogenetic relationships of North Atlantic *Laonice* Malmgren, 1867 (Spionidae, Annelida) including the description of a novel species. Mar. Biodivers. 1867:737–749.

Cantarel B.L., Korf I., Robb S.M.C., Parra G., Ross E., Moore B., Holt C., Alvarado A.S., Yandell M. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18:188–196.

Chia F.S., Gibson G. & Quian P.Y. 1996. Poecilogony as a reproductive strategy of marine invertebrates. Oceanol. Acta. 19: 285-294.

Çinar M.E. 2013. Alien polychaete species worldwide: Current status and their impacts. J. Mar. Biol. Assoc. United Kingdom. 93:1257–1278.

Cruaud A., Gautier M., Galan M., Foucaud J., Sauné L., Genson G., Dubois E., Nidelet S., Deuve T., Rasplus J.Y. 2014. Empirical assessment of rad sequencing for interspecific phylogeny. Mol. Biol. Evol. 31:1272–1274.

Eilertsen M.H., Kongsrud J.A., Alvestad T., Stiller J., Rouse G.W., Rapp H.T. 2017. Do ampharetids take sedimented steps between vents and seeps? Phylogeny and habitat-use of Ampharetidae (Annelida, Terebelliformia) in chemosynthesis-based ecosystems. BMC Evol. Biol. 17:1–15.

Elías R., Jaubet M.L., Llanos E.N., Sanchez M.A., Rivero M.S., Garaffo G. V., Sandrini-Neto L. 2015. Effect of the invader *Boccardia proboscidea* (Polychaeta: Spionidae) on richness, diversity and structure of SW Atlantic epilithic intertidal community. Mar. Pollut. Bull. 91:530–536.

Emms D.M., Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biol. 20:1–14.

Fitz-Gibbon S., Hipp A.L., Pham K.K., Manos P.S., Sork V.L. 2017. Phylogenomic inferences from reference-mapped and de novo assembled short-read sequence data using RADseq sequencing of California white oaks (*Quercus* section *Quercus*). Genome. 60:743–755.

Foster N.M. 1971. Spionidae Gulf Mexico Caribbean. Stud. Fauna Curaçao other Caribb. Islands. 36:1–183.

Gastineau R., Wawrzyniak-Wydrowska B., Lemieux C., Turmel M., Witkowski A. 2019. Complete mitogenome of a Baltic Sea specimen of the non-indigenous polychaete *Marenzelleria neglecta*. Mitochondrial DNA Part B Resour. 4:581–582.

Grabherr M.G.., Brian J. Haas, Moran Yassour Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, Lin Fan, Raktima Raychowdhury, Qiandong Zeng, Zehua Chen, Evan Mauceli, Nir Hacohen, Andreas Gnirke, Nicholas Rhind, Federica di Palma, Bruce W. N., Friedman and A.R. 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29:644–652.

Grube A.E. 1850. Die Familien der Anneliden. *Archiv für Naturgeschichte, Berlin.* 16(1): 249-364

Guggolz T., Meißner K., Schwentner M., Brandt A. 2019. Diversity and distribution of *Laonice* species (Annelida: Spionidae) in the tropical North Atlantic and Puerto Rico Trench. Sci. Rep. 9:1–12.

Haas B.J., Papanicolaou A., Yassour M., Grabherr M., Philip D., Bowden J., Couger M.B., Eccles D., Li B., Macmanes M.D., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R., William T., Dewey C.N., Henschel R., Leduc R.D., Friedman N., Regev A. 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity.

Hannerz, L. 1956. Larval development of the polychaete families Spionidae Sars, Disomidae Mesnil and Poecilochaetidae n. fam. in the Gullmar Fjord (Sweden). *Zoologiska Bidrag från Uppsala* 31: 1–204.

Kalyaanamoorthy, S., Minh, B., Wong, T. von Haeseler, A., Jermiin, L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14**:** 587–589.

Kocot K.M., Citarella M.R., Moroz L.L., Halanych K.M. 2013. PhyloTreePruner: A phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. Evol. Bioinforma. 2013:429–435.

Kudenov J.D., Blake J.A., 1978. The Spionidae (Polychaeta) from southeastern Australia and adjacent areas with a revision of the genera. Mem. Mus. Vic. 39: 171-280.

Lanfear R., Calcott B., Ho S.Y.W., Guindon S. 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses Research article. Mol. Biol. Evol. 29:1695–1701.

Langmead B., Salzberg S. 2013. Bowtie2. Nat. Methods. 9:357–359.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.

Li L., Stoeckert C.J.J., Roos D.S. 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res. 13 (9): 2178–2189.

Luo R., Liu B., Xie Y., Li Z., Huang W., Yuan J., He G., Chen Y., Pan Q., Liu Y., Tang J., Wu G., Zhang H., Shi Y., Liu Y., Yu C., Wang B., Lu Y., Han C., Cheung D.W., Yiu S.M., Peng S., Xiaoqian Z., Liu G., Liao X., Li Y., Yang H., Wang J., Lam T.W., Wang J. 2015. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. GigaScience. 4:1.

Maciolek N. J. (1985). A revision of the genus *Prionospio* Malmgren, with special emphasis on species from the Atlantic Ocean, and new records of species belonging to the genera *Apoprionospio* Foster and *Paraprionospio* Caullery (Polychaeta, Annelida, Spionidae). Zool. J. Linn. Soc. 84: 325-383.

Martinelli J.C., Lopes H.M., Hauser L., Jimenez-Hidalgo I., King T.L., Padilla-Gamiño J.L., Rawson P., Spencer L.H., Williams J.D., Wood C.L. 2020. Confirmation of the shell-boring oyster parasite *Polydora websteri* (Polychaeta: Spionidae) in Washington State, USA. Sci. Rep. 10:1–14.

Mesnil, F. 1896. Études de morphologie externe chez les annélides. I. Les spionidens des côtes de la marche. Bull. biol. Fr. Bel. 29: 110–287.

Misof B., Meyer B., von Reumont B.M., Kück P., Misof K., Meusemann K. 2013. Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. BMC Bioinformatics. 14:29–36.

Nguyen L.T., Schmidt H.A., Von Haeseler A., Minh B.Q. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268–274.

Orrhage, L. 1964. Anatomische und morphologische Studien über die Polychätenfamilien Spionidae, Disomidae, und Poecilochaetidae. Zoologiska Bidrag från Uppsala. 36: 335–405.

Philippe H., Snell E.A., Bapteste E., Lopez P., Holland P.W.H., Casane D. 2004. Phylogenomics of eukaryotes: Impact of missing data on large alignments. Mol. Biol. Evol. 21:1740–1752

Quinlan A.R., Hall I.M. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842.

Radashevsky V.I., Neretina T. V, Pankova V. V, Tzetlin A.B., Choi J. 2014. Molecular identity, morphology and taxonomy of the *Rhynchospio glutaea* complex with a key to *Rhynchospio* species (Annelida, Spionidae). Syst. Biodivers. 12:37–41.

Radashevsky V.I., Pankova V. V., Neretina T. V., Stupnikova A.N., Tzetlin A.B. 2016. Molecular analysis of the *Pygospio elegans* group of species (Annelida: Spionidae). Zootaxa. 4083:239–250.

Radashevsky V.I., Pankova V.V., Malyar V.V., Neretina T.V., Wilson R.S., Worsfold T.M., Diez M.E., Harris L.H., Hourdez S., Labrune C., Houbin C., Kind B., Kuhlenkamp R., Nygren A., Bonifácio P., Bachelet G. 2019. Molecular analysis and new records of the invasive polychaete *Boccardia proboscidea* (Annelida: Spionidae). Mediterr. Mar. Sci. 20.

Rice S.A., Karl S., Rice K.A. 2008. The *Polydora cornuta* complex (Annelida: Polychaeta) contains populations that are reproductively isolated and genetically distinct. Invertebr. Biol. 127:45–64.

Rokas A., Williams B.L., King N., Carroll S.B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798–804.

Rokas A**.,** and Carroll S.B. 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22: 1337-1344.

Roure B., Rodriguez-Ezpeleta N., Philippe H. 2007. SCaFoS: A tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol. Biol. 7:1–12.

Salzberg S.L. 2019. Next-generation genome annotation: We still struggle to get it right. Genome Biol. 20:19–21.

Sato-Okoshi W., Abe H. 2013. Morphology and molecular analysis of the 18S rRNA gene of oyster shell borers, *Polydora species* (Polychaeta: Spionidae), from Japan and Australia. J. Mar. Biol. Assoc. United Kingdom. 93:1279–1286.

Sato-Okoshi W., Abe H., Nishitani G., Simon C.A. 2017. And then there was one: *Polydora uncinata* and *Polydora hoplura* (Annelida: Spionidae), the problematic polydorid pest species represent a single species. J. Mar. Biol. Assoc. United Kingdom. 97:1675–1684.

Sigvaldadóttir, E., Mackie, A.S. & Pleijel, F. 1997. Generic interrelationships within the Spionidae (Annelida: Polychaeta). Zool J Linnean Soc. 119: 473–500.

Simon C.A. 2009. *Pseudopolydora* species associated with mollusc shells on the south coast of South Africa, with the description of *Ps. dayii* sp. nov. J. Mar. Biol. Assoc. United Kingdom. 89:681–687.

Simon C.A., Thornhill D.J., Oyarzun F., Halanych K.M. 2009. Genetic similarity between *Boccardia proboscidea* from Western North America and cultured abalone, *Haliotis midae*, in South Africa. Aquaculture. 294:18–24.

Söderström, A. 1920. Studien über die Polychätenfamilie Spionidae. Inaugural Dissertation, Uppsala, Almquist and Wicksells. 288 pp.

Stanke M., Steinkamp R., Waack S., Morgenstern B. 2004. AUGUSTUS: A web server for gene finding in eukaryotes. Nucleic Acids Res. 32:309–312.

Stiller J., Tilic E., Rousset V., Pleijel F., Rouse G.W. 2020. Spaghetti to a Tree: A Robust Phylogeny for. Biology. 9:1–28.

Veeckman E., Ruttink T., Vandepoele K. 2016. Are we there yet? Reliably estimating the completeness of plant genome sequences. Plant Cell. 28:1759–1768.

Weigert A., Helm C., Meyer M., Nickel B., Arendt D., Hausdorf B., Santos S.R., Halanych K.M., Purschke G., Bleidorn C., Struck T.H. 2014. Illuminating the base of the Annelid tree using transcriptomics. Mol. Biol. Evol. 31:1391–1401.

Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: **www.genome.gov/sequencingcostsdata**. Accessed May-10-2020.

Williams L.G., Karl S.A., Rice S., Simon C. 2017. Molecular Identification of Polydorid Polychaetes (Annelida: Spionidae): Is there a Quick Way to Identify Pest and Alien Species? African Zool. 52:105–118.

Wilson W.H. 1991. Sexual reproductive modes in polychaetes: classification and diversity. Bull. Mar. Sci. 48:500–516.

Yang Y., Smith S.A. 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: Improving accuracy and matrix occupancy for phylogenomics. Mol. Biol. Evol. 31:3081–3092.

Yokoyama H. 2007. A revision of the genus *Paraprionospio* Caullery (Polychaeta: Spionidae).

Zhang F., Ding Y., Zhu C.D., Zhou X., Orr M.C., Scheu S., Luan Y.X. 2019. Phylogenomics from low-coverage whole-genome sequencing. Methods Ecol. Evol. 10:507–517.

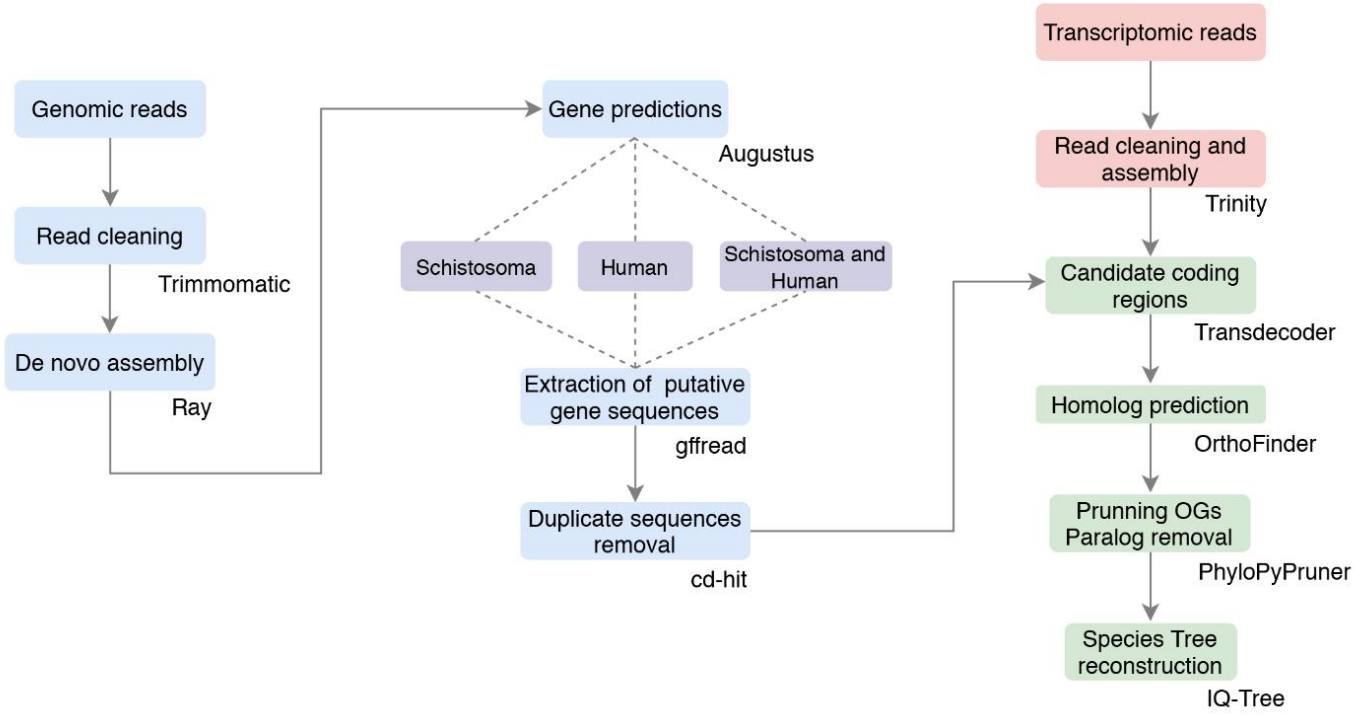

Figure 1. Summary of bioinformatic pipeline used in this study.

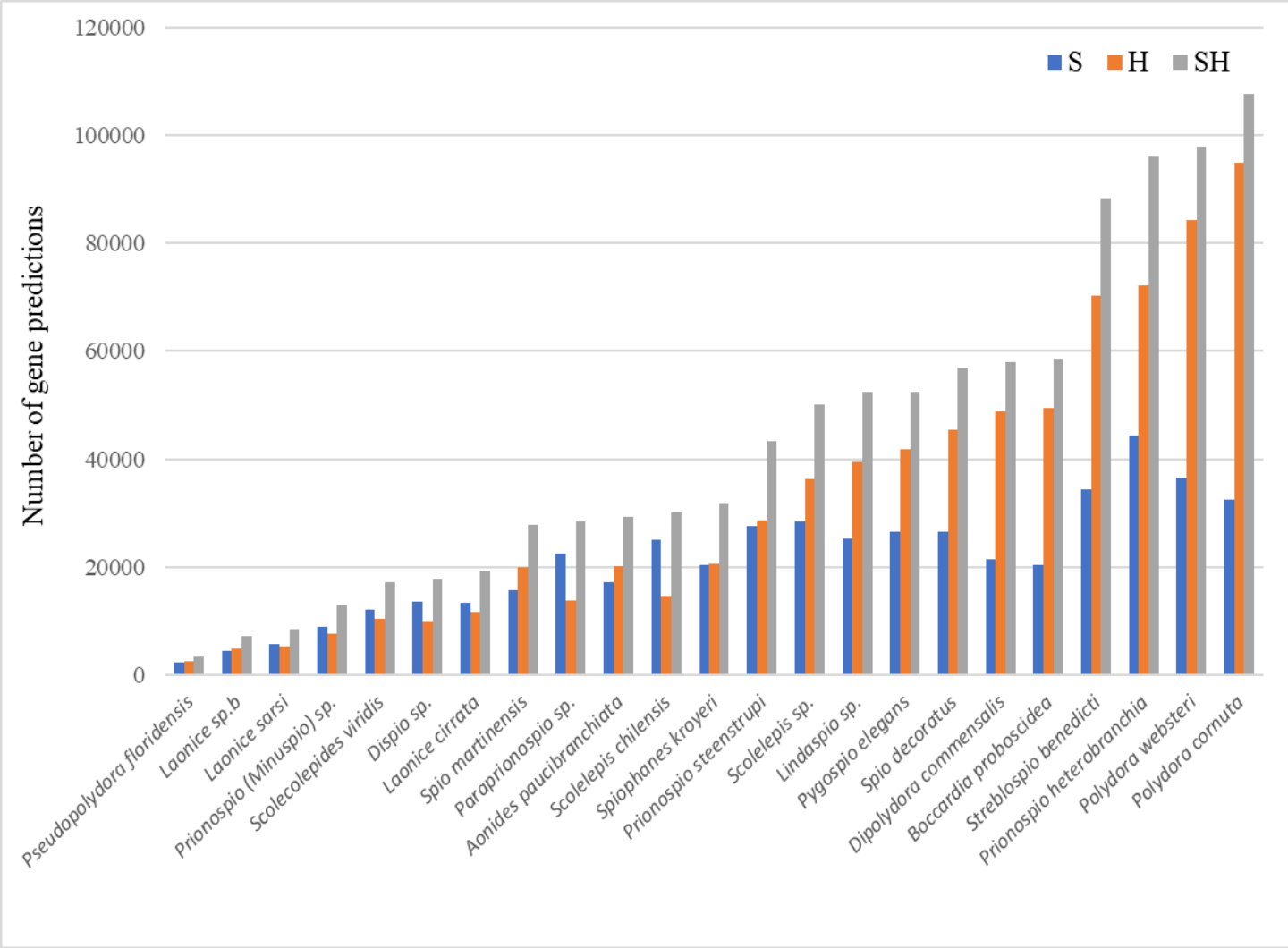

Figure 2. Comparison of number of gene predictions when using different gene models. S=*Schistosoma*, H=Human, SH=*Schistosoma* and Human.

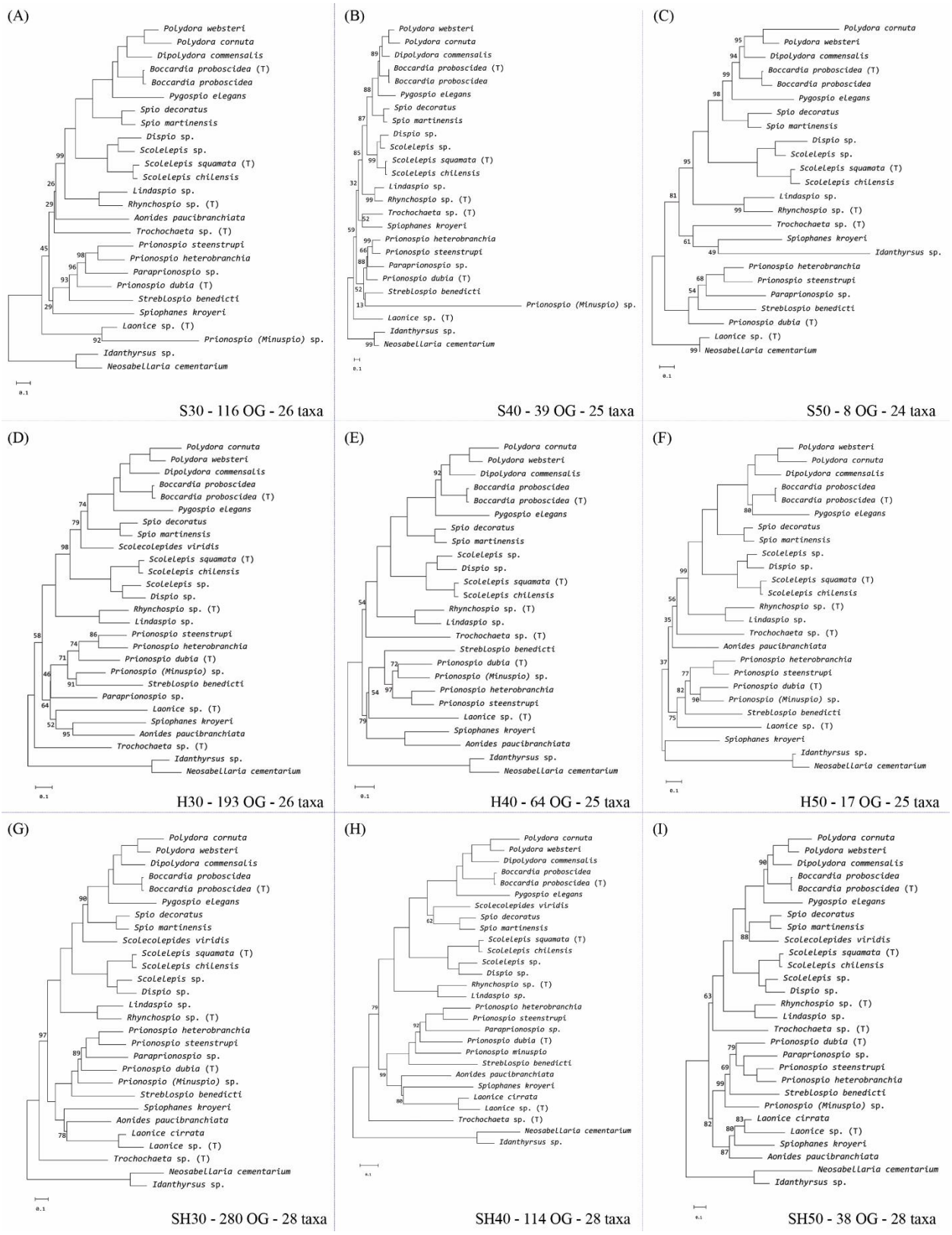
(A)
Polydora websteri
Polydora cornuta
Dipolydora commensalis
Boccardia proboscidea (T)
Boccardia proboscidea
Pygospio elegans
Spio decoratus
Spio martinensis
Dispio sp.
Scolelepis sp.
Scolelepis squamata (T)
Scolelepis chilensis
Lindaspio sp.
Rhynchospio sp. (T)
Aonides paucibranchiata
Trochochaeta sp. (T)
Prionospio steenstrupi
Prionospio heterobranchia
Paraprionospio sp.
Prionospio dubia (T)
Streblospio benedicti
Spiophanes kroyeri
Laonice sp. (T)
Prionospio (Minuspio) sp.
Idanthyrsus sp.
Neosabellaria cementarium
S30 - 116 OG - 26 taxa
(B)
Polydora websteri
Polydora cornuta
Dipolydora commensalis
Boccardia proboscidea (T)
Boccardia proboscidea
Pygospio elegans
Spio decoratus
Spio martinensis
Dispio sp.
Scolelepis sp.
Scolelepis squamata (T)
Scolelepis chilensis
Lindaspio sp.
Rhynchospio sp. (T)
Trochochaeta sp. (T)
Spiophanes kroyeri
Prionospio heterobranchia
Prionospio steenstrupi
Paraprionospio sp.
Prionospio dubia (T)
Streblospio benedicti
Prionospio (Minuspio) sp.
Laonice sp. (T)
Idanthyrsus sp.
Neosabellaria cementarium
S40 - 39 OG - 25 taxa
(C)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea (T)
Boccardia proboscidea
Pygospio elegans
Spio decoratus
Spio martinensis
Dispio sp.
Scolelepis sp.
Scolelepis squamata (T)
Scolelepis chilensis
Lindaspio sp.
Rhynchospio sp. (T)
Trochochaeta sp. (T)
Spiophanes kroyeri
Idanthyrsus sp.
Prionospio heterobranchia
Prionospio steenstrupi
Paraprionospio sp.
Streblospio benedicti
Prionospio dubia (T)
Laonice sp. (T)
Neosabellaria cementarium
S50 - 8 OG - 24 taxa
(D)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Scolecolepides viridis
Scolelepis squamata (T)
Scolelepis chilensis
Scolelepis sp.
Dispio sp.
Rhynchospio sp. (T)
Lindaspio sp.
Prionospio steenstrupi
Prionospio heterobranchia
Prionospio dubia (T)
Prionospio (Minuspio) sp.
Streblospio benedicti
Paraprionospio sp.
Laonice sp. (T)
Spiophanes kroyeri
Aonides paucibranchiata
Trochochaeta sp. (T)
Idanthyrsus sp.
Neosabellaria cementarium
H30 - 193 OG - 26 taxa
(E)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Scolelepis sp.
Dispio sp.
Scolelepis squamata (T)
Scolelepis chilensis
Rhynchospio sp. (T)
Lindaspio sp.
Trochochaeta sp. (T)
Streblospio benedicti
Prionospio dubia (T)
Prionospio (Minuspio) sp.
Prionospio heterobranchia
Prionospio steenstrupi
Laonice sp. (T)
Spiophanes kroyeri
Aonides paucibranchiata
Idanthyrsus sp.
Neosabellaria cementarium
H40 - 64 OG - 25 taxa
(F)
Polydora websteri
Polydora cornuta
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Scolelepis sp.
Dispio sp.
Scolelepis squamata (T)
Scolelepis chilensis
Rhynchospio sp. (T)
Lindaspio sp.
Trochochaeta sp. (T)
Aonides paucibranchiata
Prionospio heterobranchia
Prionospio steenstrupi
Prionospio dubia (T)
Prionospio (Minuspio) sp.
Streblospio benedicti
Laonice sp. (T)
Spiophanes kroyeri
Idanthyrsus sp.
Neosabellaria cementarium
H50 - 17 OG - 25 taxa
(G)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Scolecolepides viridis
Scolelepis squamata (T)
Scolelepis chilensis
Scolelepis sp.
Dispio sp.
Lindaspio sp.
Rhynchospio sp. (T)
Prionospio heterobranchia
Prionospio steenstrupi
Paraprionospio sp.
Prionospio dubia (T)
Prionospio (Minuspio) sp.
Streblospio benedicti
Spiophanes kroyeri
Aonides paucibranchiata
Laonice cirrata
Laonice sp. (T)
Trochochaeta sp. (T)
Neosabellaria cementarium
Idanthyrsus sp.
SH30 - 280 OG - 28 taxa
(H)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Scolecolepides viridis
Spio decoratus
Spio martinensis
Scolelepis squamata (T)
Scolelepis chilensis
Scolelepis sp.
Dispio sp.
Rhynchospio sp. (T)
Lindaspio sp.
Prionospio heterobranchia
Prionospio steenstrupi
Paraprionospio sp.
Prionospio dubia (T)
Prionospio minuspio
Streblospio benedicti
Aonides paucibranchiata
Spiophanes kroyeri
Laonice cirrata
Laonice sp. (T)
Trochochaeta sp. (T)
Neosabellaria cementarium
Idanthyrsus sp.
SH40 - 114 OG - 28 taxa
(I)
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Scolecolepides viridis
Scolelepis squamata (T)
Scolelepis chilensis
Scolelepis sp.
Dispio sp.
Rhynchospio sp. (T)
Lindaspio sp.
Trochochaeta sp. (T)
Prionospio dubia (T)
Paraprionospio sp.
Prionospio steenstrupi
Prionospio heterobranchia
Streblospio benedicti
Prionospio (Minuspio) sp.
Laonice cirrata
Laonice sp. (T)
Spiophanes kroyeri
Aonides paucibranchiata
Neosabellaria cementarium
Idanthyrsus sp.
SH50 - 38 OG - 28 taxa

Figure 3. Maximum likelihood trees for the nine datasets. BS values not shown indicate BS=100. Taxa with the letter “T” represent transcriptomic data. Dataset, number of orthologs included in each supermatrix, and number of taxa are indicated in the figure.

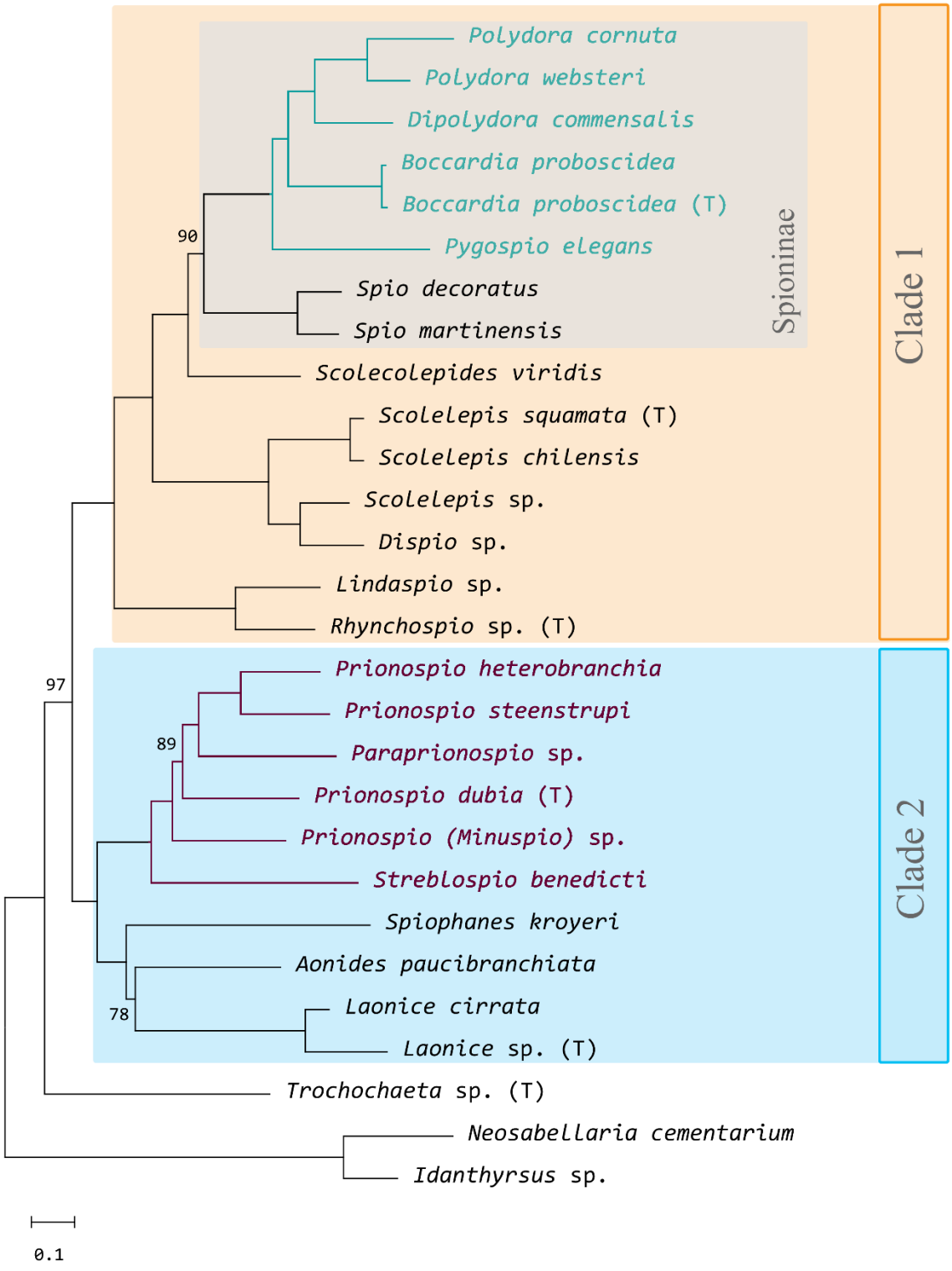
Polydora cornuta
Polydora websteri
Dipolydora commensalis
Boccardia proboscidea
Boccardia proboscidea (T)
Pygospio elegans
Spio decoratus
Spio martinensis
Spioninae
90
Scolecolepides viridis
Scolelepis squamata (T)
Scolelepis chilensis
Scolelepis sp.
Dispio sp.
Lindaspio sp.
Rhynchospio sp. (T)
Clade 1
Prionospio heterobranchia
Prionospio steenstrupi
Paraprionospio sp.
89
Prionospio dubia (T)
Prionospio (Minuspio) sp.
Streblospio benedicti
Spiophanes kroyeri
Aonides paucibranchiata
Laonice cirrata
78
Laonice sp. (T)
Clade 2
97
Trochochaeta sp. (T)
Neosabellaria cementarium
Idanthyrsus sp.
0.1

Figure 4. Maximum likelihood tree from dataset SH30 which includes 280 orthologs with 28 taxa. BS values equal to 100 not shown. Taxa with the letter “T” represent transcriptomic data. Taxa in green in clade 1 represent *Polydora*-complex, and taxa in purple in clade 2 represent *Prionospio*-complex. *Neosabellaria cementarium* and *Idanthyrsus* sp. (Sabellaridae) were used as the outgroups.

Table 1. Sampling information for the taxa included in this study. Species with a “T” at the end of the name indicates the data originated from transcriptomes.

Taxon	ID number	Latitude	Longitude	Locality	Depth (m)
Aonides paucibranchiata	DZMB-HH 62465	43° 27.550' N	008° 20.200' W	Spain	10.5
Boccardia proboscidea	A1833.3E	32°59.862' S	17°52.672' E	South Africa	---
Dipolydora commensalis		40°35'38.08" N	73°36'56.04" W	New York	1-5
Dispio sp.	A1819.4R	32°59.908' S	17°57.524' E	South Africa	?
Laonice cirrata	DZMB-HH-42927	63°25.039' N	010°58.200' W	Iceland	440.5
Laonice sarsi	DZMB-HH-57474	63° 18.88' N	023°09.61' W	Iceland	288.5
Laonice sp. b	DZMB-HH-57465	62°56'24.0" N	20°44'24.0" W	Iceland	288.5
Lindaspio sp.	AD4834-PC1-1A	39°06' N	069°24' W	New England Seep 2	1130
Paraprionospio sp.	A416.6C	37°58.612' N	122°53.558' W	California	45
Polydora cornuta	---	40°35'38.08" N	73°36'56.04" W	New York	1-5
Polydora websteri	---	30°14'58.78" N	88°04'35.83" W	Alabama	0
Prionospio (Minuspio) sp.	A1383.6E	48°27.327' N	128°42.5394' W	Canada	2408
Prionospio cf. *steenstrupi*	DZMB-HH 62467	43°28.020' N	008° 16.700' W	Spain	8
Prionospio heterobranchia	---	29°05'44.55" N	83°03'58.64" W	Florida	1
Pseudopolydora floridensis	---	27 27'30.29" N	80 18'38.73" W	Florida	1
Pygospio elegans	DZMB-HH 62474	53° 57.233' N	010° 54.164' E	Germany	
Scolecolepides viridis	A22.9C	41º41.484' N	070°37.612' W	Massachusetts	?
Scolelepis chilensis	---	22°45'S	41°55'W	Brazil	1-5
Scolelepis sp.	---	40°35'6.87"N	73°35'5.48" W	New York	1-5
Spio decoratus	DZMB-HH 62463	54° 19.730' N	006° 59.680' E	North Sea	39
Spio martinensis	DZMB-HH 62468	43° 27.890' N	008° 17.250' W	Spain	0
Spiophanes kroyeri	DZMB-HH 62464	66° 18.060' N	012° 22.400' W	Iceland	732.1

Streblospio benedicti	---	33°20'47.69" N	79°11'31.86" W	South Carolina	1
Laonice sp.T	A388.3C	36°49.847' N	122°01.729' W	California	92
Boccardia proboscidea T	A1793	48°28'11.97" N	123°03'24.56" W	Washington	1
Rhynchospio sp. T	AD4834-PC8-8	39°06' N	069°24' W	New England Seep 2	1130
Trochochaeta sp. T	A422.6C	37°58.612'N	122°53.558'W	California	45
Prionospio dubia T	A619.4C	39º56.172'N	69º34.573'W	Massachusetts	253
Idanthyrsus sp.	A354.5C	36º23.381'N	121º57.974'W	California	107
Neosabellaria cementarium	A133.14C	48º 34.231'N	123º 02.247'W	Washington	1

Table 2. Genome assembly statistics on the taxa included in this study.

OTU	Number raw reads	coverage	N50	Number contigs	Total contig length in bp	Mean %GC
Pseudopolydora floridensis	5,180,073	3	148	681,407	113,630,300	37.46
Laonice sp. *b*	51,847,133	6	155	3,569,419	614,343,368	37.91
Laonice sarsi	44,456,328	5	158	3,467,739	604,798,508	36.94
Scolecolepides viridis	88,343,970	8	229	4,613,802	1,028,880,780	34.08
Laonice cirrata	105,120,637	9	231	4,734,116	1,074,674,129	36.24
Prionospio minuspio	67,695,010	8	243	2,922,467	678,334,899	37.41
Paraprionospio sp.	112,125,774	9	265	4,702,738	1,173,775,491	35.31
Dispio sp.	75,079,419	8	269	3,328,708	832,670,627	31.5
Aonides paucibranchiata	64,217,359	9	320	3,149,519	891,696,092	37.25
Spiophanes kroyeri	77,829,705	11	322	3,122,064	889,096,562	36.68
Spio decoratus	51,000,756	12	382	1,958,988	636,901,065	38.25
Prionospio steenstrupi	84,577,386	14	389	2,219,536	728,409,189	37.35
Spio martinensis	46,325,116	11	403	1,847,026	620,045,972	36.49
Pygospio elegans	65,754,111	19	593	1,247,838	531,980,836	38.21
Lindaspio sp.	84,395,999	20	639	1,216,160	578,350,385	38.44
Scolelepis chilensis	57,802,291	11	691	1,916,847	930,318,202	28.93
Boccardia proboscidea	60,989,670	19	703	1,016,885	478,353,126	40.84
Scolelepis sp.	134,086,094	16	760	2,533,873	1,423,497,514	32.83
Dipolydora commensalis	51,339,212	17	885	968,225	580,316,744	38.63
Polydora cornuta	62,207,176	18	981	939,705	684,835,060	42.3
Polydora websteri	89,074,630	21	1242	1,041,707	906,610,025	38.88
Streblospio benedicti	96083709	26	1431	777,872	642,137,505	38.12
Prionospio heterobranchia	121,013,017	46	2288	452,941	595,774,552	36.39

Table 3. Summary of matrices composition and best fit model selected by IQ-TREE.

Gene model	**Dataset**	**Number of OGs**		**Concatenated alignment length**	**Average matrix completeness**	**Model**
		Input	Output			
Schistosoma	S30	28095	116	50331	0.67	LG+F+I+G4
	S40		39	15075	0.70	LG+F+R4
	S50		8	4779	0.69	LG+F+I+G4
Human	H30	51241	193	80092	0.64	LG+F+R4
	H40		64	23604	0.70	LG+F+R4
	H50		17	8132	0.70	LG+F+R4
Schistosoma and human	SH30	56744	280	113833	0.65	LG+F+R5
	SH40		114	41634	0.67	LG+F+R5
	SH50		38	16092	0.67	LG+F+R4