School of Doctoral Studies in Biological Sciences
University of South Bohemia in České Budějovice
Faculty of Science

Evolution and Phylogeny of Mesozoa

Ph.D. Thesis

Mgr. Marie Drábková

Supervisor: doc. RNDr. Jan Štefka, Ph.D.

Institute of Parasitology, Biology Centre of the Czech Academy of Sciences
Department of Parasitology, Faculty of Science, University of South Bohemia

České Budějovice 2022

This thesis should be cited as:

Drábková, M., 2022: Evolution and phylogeny of Mesozoa. Ph.D. Thesis Series, No.10 University of South Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, 169 pp.

Annotation

This thesis focuses on the phylogenetic position of Mesozoa (Orthonectida and Dicyemida) based on phylogenomics, and on dicyemid life-history traits revealed by molecular methods used in population genetics. The thesis is introduced by the review of biology of both groups, complemented by up to now development of views on their phylogenetic position and notes concerning the study of the population structure of marine invertebrates. The introduction is followed by a study focusing on the phylogenetic position of Mesozoa, a comparison of population structure between the cephalopod host and its dicyemid parasite, and a case study of dicyemid parasite infrapopulation. The thesis wraps up with a review on cephalopod parasites and a summary.

Declaration

I hereby declare that I am the author of this dissertation and that I have used only those sources and literature detailed in the list of references.

České Budějovice, 5. 5. 2022

……

Marie Drábková

This thesis originated from a partnership between the Faculty of Science, University of South Bohemia, and the Institute of Parasitology, Biology Centre CAS.

Financial support

The studies were financially supported by the Czech Science Foundation (GACR 15-08717S).

Cover image: Marie Drábková

Acknowledgments

First, I would like to thank my supervisor for support, guidance, and patience, and also for showing great courage by venturing into uncharted dicyemid waters. Great thanks also go to our lab members, university staff, and people at the Biology Centre for keeping things running.

Second, I would like to thank my fellow women scientists group `the Vierklang`. Without you, the world would be boring and grey! Thank you for the fun (and science, and coffee, and wine) we shared. A life well enjoyed in good company ☺.

Last but not least I am eternally grateful to my family for being relentlessly and unconditionally supported by love and all the necessary things. You are the best!

List of papers and author's contribution

The thesis is based on the following papers (listed in order of appearance):

I. Drábková, M., Kocot, K.M., Halanych, K.M., Oakley, T.H., Moroz, L.L., Cannon, J.T., Kuris, A., Garcia-Vedrenne, A.E., Pankey, S.M., Ellis, E.A., Varney, R., Štefka, J., Zrzavý, J., Different Phylogenomic Methods Support Monophyly of Enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa). (manuscript)
Drábková Marie with ŠJ designed and carried out sample collection and sequencing for the dicyemid genome included in the study. DM performed dicyemid genome assembly. DM and KMK analyzed the data and wrote the manuscript with ZJ, and ŠJ. DM contributed 75%.

II. Drábková, M., Jachníková, N., Tyml, T., Sehadová, H., Ditrich, O., Myšková, E., Hypša, V., Štefka, J. 2019. Population co-divergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea. Scientific Reports 9, 14300 (IF=4.379).
Drábková Marie designed the study together with ŠJ, obtained the samples together with TT and ME, carried out the laboratory work, analyzed the data with the contribution of JN, and wrote the manuscript with ŠJ and HV. DM contributed 60%.

III. Drábková, M., Flegrová, T., Myšková, E., Hypša, V., Štefka, J. 2021 Genetic analysis of dicyemid infrapopulations suggests sexual reproduction and host colonization by multiple individuals is common. Organisms Diversity & Evolution 21, 437–446 (IF=2.153).
Drábková Marie designed the study together with ŠJ, obtained the samples with ME, carried out the laboratory work with FT, analyzed the data with FT, and wrote the manuscript with ŠJ. DM contributed 60%.

IV. Roumbedakis, K., Drábková, M., Tyml, T., di Cristo, C., 2018 A perspective around cephalopods and their parasites, and suggestions on how to increase knowledge in the field. Frontiers in Physiology 9, 1573 (IF=4.566).

RK and CC designed the study. Drábková Marie wrote the manuscript with RK and TT. DM contributed 25%.

Co-author agreement

Jan Štefka, the supervisor of this thesis and co-author of presented manuscripts, fully acknowledges the contribution of Marie Drábková as the first author and her contributions as stated above.

.....

Doc. RNDr. Jan Štefka, Ph.D.

Jan Zrzavý, the co-author of the Chapter I, fully acknowledges the contribution of Marie Drábková as the first author and her contributions as stated above.

.....

Prof. RNDr. Jan Zrzavý, CSc.

Katina Roumbedakis, the first author of the Chapter IV, fully acknowledges the contribution of Marie Drábková as the co-author and her contributions as stated above.

.....

Katina Roumbedakis, Ph.D.

Contents

Introduction **1**

1.
   1.1 Mesozoa 3
   1.2 Dicyemida 4
   1.3 Orthonectida 13
2. Phylogenetic inference and Mesozoa 16
3. Study of population structure based on genetic markers 19
4. Aims and scope of the thesis 20
5. References 22

Chapter I **31**

Different Phylogenomic Methods Support Monophyly of Enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa).

Chapter II **65**

Population co-divergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea.

Chapter III **101**

Genetic analysis of dicyemid infrapopulations suggests sexual reproduction and host colonization by multiple individuals is common.

Chapter IV **127**

A perspective around cephalopods and their parasites, and suggestions on how to increase knowledge in the field.

Summary **155**

1. Summary of results 157
2. Future perspectives 160
3. References 163

Author's Curriculum vitae **165**

INTRODUCTION

1.1 Mesozoa

Mesozoan parasites intrigued biologists since their discovery more than a hundred years ago but some aspects of their life remain enigmatic (e.g., mode of transmission) and they still present a challenge for current biology. This study aims to build up on the classical natural history research on Mesozoa (works by van Beneden, Nouvel, and Furuya, to name just a few). Through the use of modern molecular methods, it brings new pieces of information helping to elucidate some aspects of mesozoan life and filling in the gaps in knowledge in this often overlooked but interesting part of the Tree of Life.

Mesozoa traditionally contains two groups of small ciliated marine parasites of invertebrates, Orthonectida and Dicyemida. Originally, Dicyemida and Orthonectida were placed together in one phylum named Mesozoa by Van Beneden (1876) and were thought to be a link between the unicellular Protozoa and multicellular Metazoa because of their simple structure and ciliated cells. However, these two clades differ in some aspects of their life and their common origin has since been contested. Shared features of Dicyemids and Orthonectids include simple body organization (lack of proper tissues and organs), parasitic lifestyle (both use marine invertebrates as their hosts), cilia on the surface of the cells, and complicated life cycles. Nonetheless, there are doubts that this assembly reflects their true evolutionary history and therefore the name mesozoa (without capital M) is sometimes used just to refer to their level of body organization (Hochberg, 1982, Pawlowski et al., 1996).

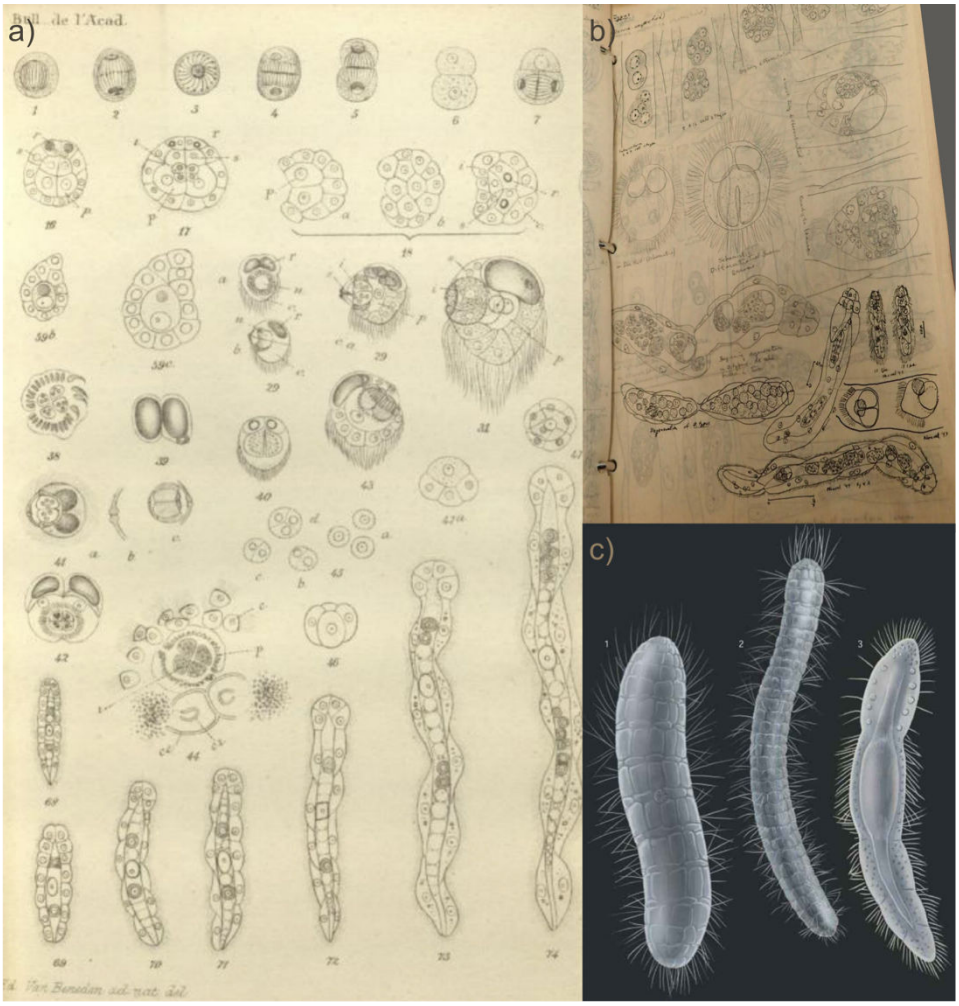


Figure 1. Schematic drawings of dicyemids that used to serve as a basis for mesozoan research. a) drawings from the collections of Santa Barbara Museum of Natural History b) drawings of dicyemids by Van Beneden (1876) c) schematic drawing of orthonectids by John Megahan (accessed on 09 2019 at https://www.guwsmedical.info/reproductive-biology/orthonectidans.html)

1.2 Dicyemida

Dicyemids are tiny wormlike animals with simple body structure. They live in the renal sacs of cephalopod hosts that prefer a benthic lifestyle. The body length reaches from 0.1 to 8.0 mm and the majority of species measures up to 3 mm (Furuya & Tsuneki, 2003). Their body consists usually of only 10-40 cells (Fururya & Tsuneki, 2003). However, what they

lack in the complexity of body structure, they make up for in the elaborated life cycle. The name Dicyemida (from greek dis kyema=two embryos) was suggested by von Kölliker (1849 in Stunkard, 1954) to express the fact that dicyemids produce two kinds of embryos. The vermiform embryo looks like an adult dicyemid and should serve for multiplication to increase the density of infection in the host. The second kind of embryo is termed infusoriform embryo and is thought to be a means of dispersion infecting new hosts. Infusoriform embryo can survive for a few hours swimming freely in seawater (pers. obs.), but no other life stages were ever found outside of their cephalopod host.

Body structure and reproduction

The apical end of a dicyemid is termed calotte and consists of two layers of cells covered in seemingly thick cilia (Ridley, 1968). The first tier of cells is called propolars and the second tier is called metapolars (Nouvel, 1947). The position and number of propolar and metapolar cells forming calotte and the overall shape of the calotte are the main morphological characters used for genus and species identification. The calotte is used to attach to the surface of the renal tissue of a host. At the interface of a calotte and the host's border tissue a slight erosion can be seen (Ridley, 1968). The body consists of one elongated cell, termed axial cell, which is covered by approximately 20 ciliated cells (named either coat or jacket cells or diapolars). Usually, more than one developing embryo can be seen inside the axial cell. Two special coat cells containing refractive material might be present at the posterior end (uropolars).

Dicyemids either reproduce asexually by producing vermiform embryos or they switch to sexual reproduction (possibly by self-fertilization) and produce infusoriform embryos. The switch is believed to be triggered by reaching a high density of dicyemid population in the renal organ (Lapan & Morowitz, 1972) but it also might be correlated with the maturation of the host or other environmental factors (Finn et al., 2005).

All embryos develop from a reproductive cell called axoblast, which is contained inside the axial cell of the dicyemid from the early stages of development. Axoblast is only partially enveloped by a membrane (Ridley, 1968; Matsubara & Dudley, 1976). Resulting embryos, therefore, develop

inside the parent axial cell until they are released into the host urine by eclosion, i.e., an embryo slips through an opening between the coat cells.

Vermiform embryos resemble miniature adult dicyemids in their appearance (one elongated axial cell covered by ciliated coat cells). During their development, they only increase in size but do not undergo significant structural changes. Compared to vermiform embryos, infusoriform embryos are organized very differently and are thought to be the most complex stage of a dicyemid's life cycle (Matsubara & Dudley, 1976). The infusoriform embryo is bilaterally symmetrical and consists of two apical cells containing refringent bodies, inner cells with urn cells containing germ cells, and approximately twelve ciliated coat cells (Ridley, 1969). Refringent bodies are made of dense material and aid the embryo in maintaining buoyancy. Vermiform embryos were observed to swim near the bottom of the container where their host was kept during laboratory experiments (Lapan & Morowitz, 1975; Stunkard, 1954). Swimming round-shaped, but tapered at the end, ciliated infusoriform embryos are reminiscent of little drakes/kites (rhombos) and gave this group their alternative name Rhombozoa.

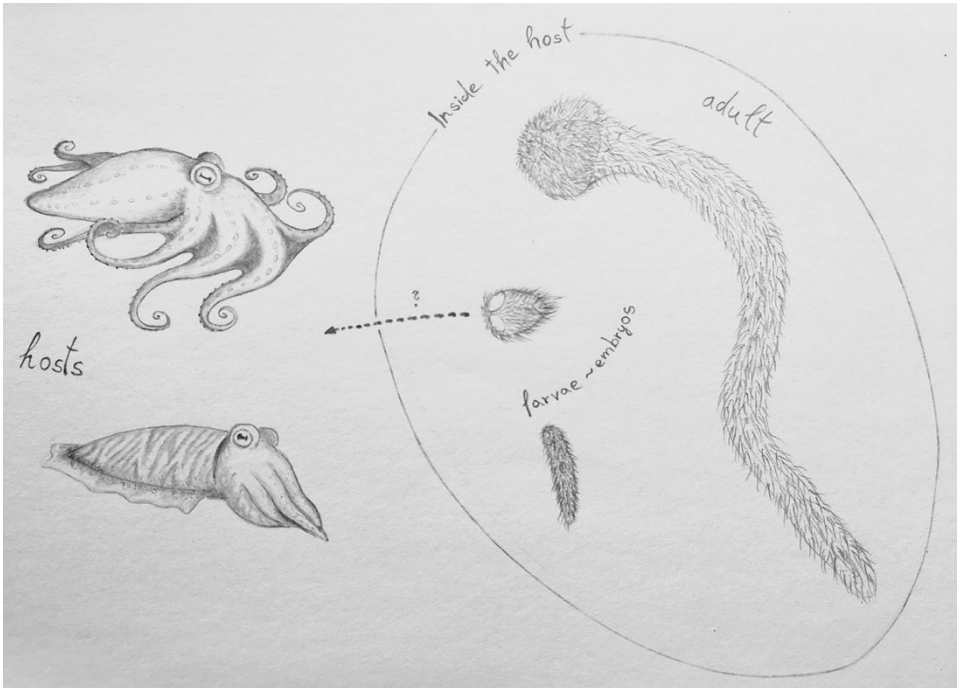


Figure 2. A schematic drawing of Dicyemid's life cycle. Both types of dicyemid embryos and adults can be found in the renal organs of their cephalopod host. A way of infection of a new host is unknown. However, infusoriform embryos were suggested as a possible transmission route. The drawing of hosts and dicyemids is not in scale. Adult dicyemids can reach about 1-2 mm in length. Original drawing by Marie Drábková.

Transmission

The mode of transmission of dicyemids to a new host is still mostly unknown. An infusoriform embryo is supposed to be a means of dispersion. Infusoriforms can be isolated from seawater in which a mature cephalopod host was held for a few hours (Stunkard, 1954). Usually, infusorifoms remain near the bottom, thanks to the buoyancy caused by refringent bodies and move by the locomotion of cilia (Stunkard, 1954). After a while (a couple of minutes to hours) in laboratory conditions, infusoriforms start to disintegrate. This process might be a natural step in the development in which germ cells are released and infect a new host. Experimental infection under laboratory conditions, where young cephalopods reared in aquaria were exposed to infusoriform embryos, has been tried at least two times, however, without convincing results (Nouvel, 1947; Lapan & Morowitz, 1972). When a batch of cuttlefish's eggs was raised separated from the adults, none of the progeny was infected (Nouvel, 1947; Lapan &

Morowitz, 1972). When the eggs were raised with an adult already infected by dicyemids, the progeny became infected as well (Lapan & Morowitz, 1972). However, a direct infection with infusoriform embryos was not successful (Lapan & Morowitz, 1972, 1975; Nouvel, 1947). In early works, an intermediate host was suggested to exist but nowadays it is widely accepted that it is not required for the development of dicyemids. Transmission of dicyemids from adults to the next generation of hosts is probably not realized via infection of the host's eggs, even though it is almost impossible to be absolutely certain that a yet unknown stage did not escape detection. An experiment by Catalano et al. (2013) aimed at the detection of dicyemid DNA in cuttlefish eggs and seawater from the breeding site of *Sepia apama* failed to recover any traces of dicyemid DNA either from eggs or seawater.

Nouvel (1947) claimed that the first stage infecting a new host is a stage called stem nematogen (or in original "larva fondatrice") that he observed in very young specimens of cuttlefish and octopus. At that time, there was a debate about whether this stage existed or not, later the observation was confirmed by McConnaughey (in Stunkard, 1954). However, Lapan & Morowitz (1975) stated that its existence remains uncertain. Stem nematogen differs from the usual adult in having three axial cells instead of one. Nevertheless, how an infusoriform embryo infects a new host and develops into a stem nematogen remains unknown and the stem nematogen stage still requires a better description of its organization, function, development, and mode of reproduction.

The way of life

According to Hochberg (1982), the cephalopod's renal organs are the best environment for parasites. He argued so because renal organs provide for all parasite's needs, that is: they provide substrate for attachment, constant fluid bath, source of nutrients, and easy exit for dispersal stages. Hochberg based his observation on his study of the adaptation of dicyemids and protozoan opalinids that inhabit the renal organs of pelagic cephalopods. However, in terms of dicyemid's exact needs and requirements, not much is known. In an experiment focused on maintaining dicyemids under laboratory conditions, Lapan & Morowitz (1975) used a very rich medium

based on the chemical properties of cephalopod's urine and successfully maintained them, slowly reproducing for at least three months. In our experiment, where we followed the previously mentioned method as closely as possible, we were able to maintain dicyemids for a maximum of two weeks (Drabkova, pers. obs.). However, in the original study, it was not determined which of the components of the medium were essential, therefore we might have missed the key component of the medium. Furthermore, the way how dicyemids feed, which might be crucial for maintaining them in laboratory conditions, still remains to be revealed. The surface of an adult dicyemid is furrowed by ruffles that can merge into smaller or larger vesicles to import material into coat cells. In an experiment done by Ridley (1968), dicyemids were shown to take up ferritin from solution. Therefore, a possible way of sustenance might be endocytosis of particulate matter from the surrounding fluid.

To what extent do dicyemids harm or help their host remains to be discovered. Where calotte is inserted into the folds of the renal tissue a little abrasion can be seen (Ridley, 1968). However, there is a need for a thorough study of the biochemical pathways used by dicyemids to show whether they are parasites, commensals, or mutualists.

To date, two studies reported dicyemids to be parasitized by microsporidian hyperparasite (Czaker, 2000; Ogino et al., 2007). Ogino et al. (2007) observed considerable damage to dicyemid's tissue caused by a microsporidian parasite in *Callistoctopus minor* host. However, whether the infection is anecdotal (i.e., the microsporidian parasite is shared with the host) or widespread remains to be seen.

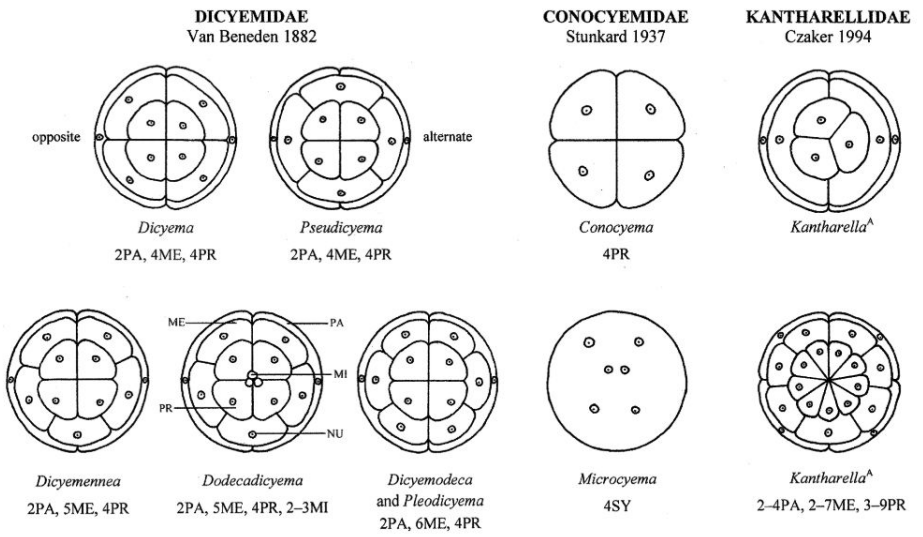
Diversity and taxonomy

Up to now, 136 species of dicyemids have been described (Catalano, 2012, 2013a, 2013b; Catalano & Furuya, 2013; Castellanos-Martinez et al., 2016; Furuya, 2018). Catalano (2012) reviewed available species descriptions, discussed the validity of described species, and advocated for the separation of dicyemid species into three families according to the position of propolar and metapolar cells composing a calotte (fig. 3), originally suggested by Whitman (1882), with the addition of Kantharellidae (Czaker, 1994). The most commonly found and the richest in species number is the

family Dicyemidae (116 species in 5 genera). In the family Conocyemidae, coat cells are usually merged into syncytium and this family is represented by two genera. The last family, Kantharellidae, is represented by only one species, *Kantharella antarctica*, described by Czaker (1994) from two specimens of antarctic octopus *Pareledone turqueti*.

Similarly, to the family level, individual dicyemid species are discerned by the position of cells in the calotte (propolars to metapolars) and by some additional morphological structures (shape of the calotte, length of an adult or embryo, presence of uropolar cells, number of coat cells; Furuya et al., 2001). Catalano (2012) expressed doubts about the validity of 20% of the described species because of invalid or incomplete descriptions. For European waters, 16 dicyemid species were reported in a recent review by Furuya & Souidenne (2019).

Because dicyemids lack “hard” morphological features and those that are used can be highly variable (i.e., number of coat cells, body length) or subjectively viewed (shape of a calotte) molecular methods are often mentioned as the way to better understand dicyemid's diversity. A study by Eshragh & Leander (2014) showed that when molecular markers are taken into account (18S marker), some dicyemid species (although morphologically different) may be genetically similar (e.g., *Dicyemennea rossiae* and *Dicyemennea brevicephaloides*, both infecting *Rossia pacifica*). In general, dicyemid taxonomy still waits for wider implementation of molecular markers, and the fusion of traditional taxonomy with modern methods would be highly beneficial.



^ = Alternative views of calotte arrangement for *Kantharella* with the minimum (top) and maximum (bottom) number of PA, ME and PR drawn.

Figure 3. Position of cells in calotte region in dicyemid's families, adopted from Catalano (2012). PA parapolar cells, ME metapolar cells, PR propolar cells, SY syncytium, MI micropolar cells.

Geographical distribution

The geographical range of dicyemids extends from the Southern Ocean to the Arctic Ocean. The main center of occurrence was thought to be in the temperate seas mainly in the Mediterranean, North Atlantic, and North-West Pacific (chapter 5 by Furuya in Rhode, 2005; Finn et al., 2005). This remarkably correlates with the position of established biological marine stations (e.g., Roscoff, Naples, Woods Hole, Japanese institute). In recent years, increasingly more records come from other parts of the world ocean, e.g., from Chile (Muñoz et al. 2013), Australia (Finn et al., 2005), Antarctic Ocean (Czaker, 1994), and the Arctic Ocean (Furuya et al., 2002, Furuya, 2010). In early studies, it was suggested that the prevalence is highest in the temperate seas (in Mediterranean area up to 100%, pers. obs.) and declines towards the equator (in subtropical seas about 20% to no infection in tropical seas; Hochberg 1990). However, Finn et al. (2005) show that the prevalence in tropical hosts is almost similar to temperate areas, even though the infection is more difficult to detect (less dense infection, individuals more difficult to dislodge). Records from polar seas are too

scarce to provide an informed estimate of the prevalence of infection in cold seas.

Host specificity

To what extent are dicyemids host-specific is not very clear. Van Beneden thought that each species of dicyemid coexists with one species of a host and suggested a taxon structure corresponding to a strict host-specific relationship (Van Beneden, 1876). Later some species were shown to be shared, usually between closely related host taxa (Furuya, 1999), whereas McConnaghuey (in Stunkard, 1954) reported differences in dicyemid fauna between two cryptic species of an octopus.

If more than one species of a dicyemid infects a given host, there is a possibility of mixed infection. Furuya (1999) examined cephalopods from seas surrounding Japan and found two to three species of dicyemids in one species of host, or even in one host individual. However, Furuya also states that the shape of a calotte can differ according to the part of the surface of the renal organ where the parasite is attached. Given that the main morphological feature determining species is the shape of a calotte and that two species with different calotte shapes can share the same genetic information (Eshragh & Leander, 2014), it remains to be seen if these species represent morphologically adapted individuals of a single species or true, separate species. In other works (e.g., Finn et al., 2005; Catalano, 2013a) mixed infection is reported much less often. This could possibly be resolved by examination of single individuals from one host species or by metagenomic studies.

In cephalopods, the renal organs are found in pairs. A difference in dicyemid populations between the two renal organs in one host was reported by Finn et al. (2005) and Furuya (2006). Each renal organ could differ in the contained life stages of dicyemids, their abundance, or even the presence or absence of the species (Finn et al., 2005) or in the species assembly (Furuya, 2006). This suggests that infection of each renal organ happens individually.

Available molecular data

The first molecular sequence published for dicyemids was the 5S fragment of ribosomal RNA as part of an early phylogenetic work attempting to

classify multicellular animals based on molecular sequences (Ohama et al., 1984, Hori & Osawa 1987). However, this marker showed not to be suitable for such deep phylogenetic questions and was replaced by the then-popular 18S ribosomal sequence (18S dicyemid sequences were included in Katayama et al., 1995, Hanelt et al., 1996 and Pawlowski et al., 1996). Additionally, the 18S marker was used in the comparison of morphospecies by Eshragh & Leander (2014) concerning the dicyemids from British Colombia and was also used as a part of the redescription of Mediterranean species *Dicyemennea eledones* (Souidenne et al., 2016). The sequence of COI gene of dicyemid species originating in the seas surrounding Australia was published in a study focused on the structure of fragmented dicyemid mitochondrial genome (Catalano et al., 2015). Few other sequences of other dicyemid genes were published as a part of developmental studies (i.e., Pax6, Zic, hox genes, innexin, tektins, alpha-tubulin, and beta-tubulin; Aruga et al., 2007, Ogino et al., 2007, Kobayashi et al., 2009, Suzuki et al., 2010). These single gene sequences were recently supplemented by publicly available transcriptomic and genomic data for *Dicyema japonicum* in studies focused on the structure of the dicyemid genome and dicyemid phylogenetic position (Lu et al., 2017 and 2019).

1.3 Orthonectida

Orthonectids are tiny marine parasites infecting diverse groups of marine invertebrates. So far, orthonectids were reported from Echinodermata, Platyhelminthes, Annelida, Mollusca, Tunicata, and Nemertea. Recently *Rhopalura xenoturbellae* was reported infecting *Xenoturbella* (Nakano & Miyazawa, 2019). The group was first described and named "Orthonectida" (straight + swimming) by Giard (1877) to emphasize their characteristic movement in a straight line (Giard, 1877). Orthonectids are categorized into 4 genera (*Rhoplaura*, *Intoshia*, *Stoecharthrum,* and *Ciliocincta*) and 27 species (mostly in the genus *Rhopalura*; WoRMS, 2019) with uncertain addition of the genus *Pelmatosphaera* (Kozloff, 1992). Orthonectid species description is usually based on the pattern of bands of surface cilia, but also on the position of genital opening and disposition of eggs and sperms (Kozloff, 1992). Compared to dicyemids, their prevalence is low (about a few %) and they usually occur only in some areas of the distribution of their host species (e.g., *Ciliocincta sabellariae*

is typical for polychaete worm *Sabellaria cementarium* syn. *Neosabellaria cemntarium* from San Juan archipelago, Washington; Kozloff, 1992). Their rarity and unpredictability make the acquisition of samples very difficult.

Life cycle and body structure

Adult orthonectids (sexually mature males and females) leave their host and mate outside in the seawater. Eggs are fertilized by internal fertilization and develop into ciliated larvae that infect a new host. When an orthonectid larva infects a new host through either a genital opening or a gut, a plasmodium stage is created where multiple males or females are produced in one plasmodium. Sexually mature orthonectids leave their host and complete their life cycle (Kozloff, 1994; fig. 4). Orthonectids cause damage of the tissues of their hosts, even leading to castration if the infection happens in a reproductive tissue (e.g., observed in sea star *Amphiophiuris squamata*). Most orthonectids (except for the genus *Stoecharthrum* which is hermaphroditic) display sexual dimorphism with females being about two to three times larger than the males. Females of two sizes in one infection were reported for *Rhopalura ophiocomae* (Kozloff, 1992). The body of orthonectids is simply organized, but muscles and traces of nerve tissue are present in the free-living stage that can be visualized with microscopy based on fluorescent antibody staining (Slyusarev & Starunov, 2016). Information on host specificity, population structure, species distribution, or mixed infections has not been reported in orthonectids.

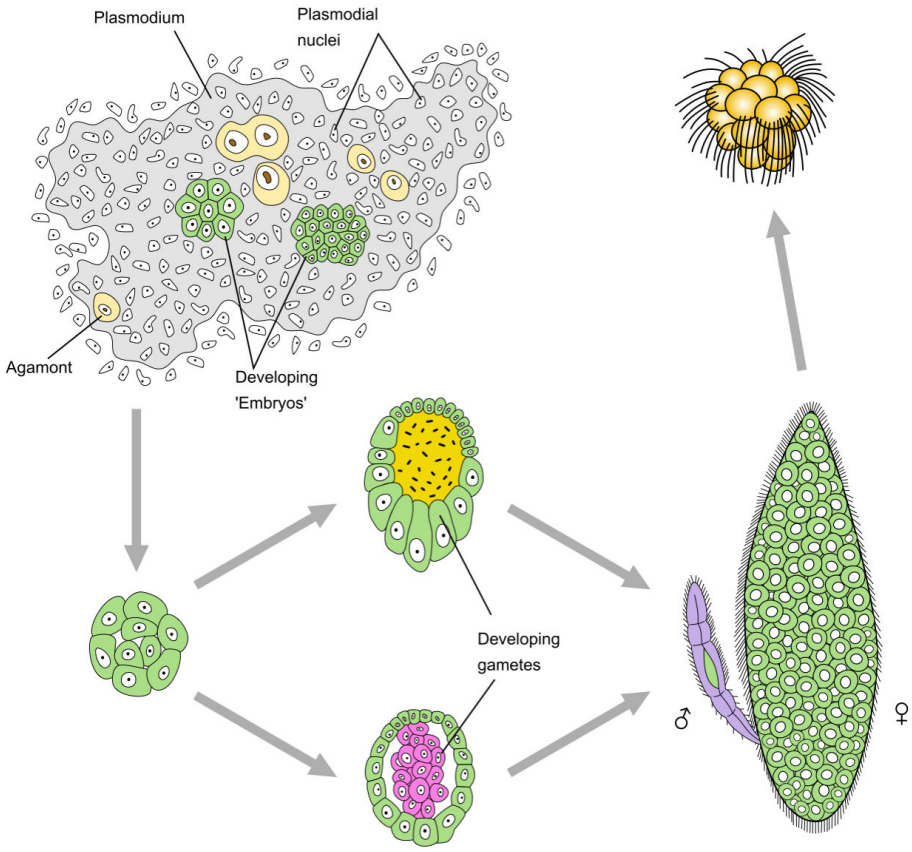


Figure 4. Schematic representation of orthonectid life cycle (Kaidor, 2016, Wikimedia Commons).

Available molecular data

For a long time, molecular data have been rare for orthonectids, with only an 18S sequence published (for *Rhopalura ophiocomae* both in Hanelt et al., 1996 and Pawlowski et al., 1996). However, recently, due to phylogenetic interest in this group, available data were expanded by the genome of *Intoshia linei* (Mikhailov et al., 2016) and 3 mitochondrial genomes (*Intoshia linei*, Schiffer et al., 2018; *I. variabilis*, *Rhopalura litoralis*, Bondarenko et al., 2019).

2 Phylogenetic inference and Mesozoa

The evolutionary tree of life has been of interest to biologists (or natural historians) since the Darwinian times, starting with the famous doodle of a tree with the inscription "I think...". Phylogenetic placement of organisms in the tree of life enables us to interpret organismal traits in an evolutionary view. It can help to explain the evolution of structures and to understand the process in which traits evolve from simple to complex or, the other way around, from complex to simple, as we can see for example in the case of either entire loss or simplification of some features in parasitic lineages. Phylogenetic trees can also bring more light into the cases when traits just simply change (without any alteration to complexity level).

First phylogenetic studies were based on morphology and structural (or behavioral) similarity. With the development of modern molecular methods and most importantly those that include sequences of DNA/RNA, the molecular methods almost completely overtook morphological studies. The most important finding at the beginning of the usage of molecular methods was the discovery that the tree of life splits into three main kingdoms: Eubacteria, Eukaryota, and Archea (Pace et al., 2012; original study by Woese & Fox 1977). Since then, molecular phylogenetics was employed to study the evolution of genes, identification of pathogens, the role of evolutionary relationships in conservation biology, and of course, in the systematics itself (Soltis & Soltis 2003, Stoakes 2019, Vezquez & Gittleman 1998). Even though the methods for inferring phylogenetic relationships changed, the evolutionary questions they try to answer remained the same.

One of the areas in the tree of life that is evolutionarily important, but understudied, is Lophotrochozoa, a clade that should also contain both Mesozoan groups according to recent studies (Schiffer et al., 2018; Zverkov et al., 2019). This clade of invertebrates was grouped together based on molecular sequences (Halanych et al., 1995). It encompasses animals both simple in structure (e.g., gnathiferans, parasitic platyhelminths) as well as complex animals like annelids and cephalopods with highly developed nervous tissues and locomotory skills. This clade presents an invaluable resource for comparative biology, especially in the

area of the development of nervous tissue structures, the evolution of complex body plans and life strategies along with their simplification due to miniature body size and/or parasitic lifestyle. However, the study of Lophotrochozoa is complicated by the rarity of samples from some major clades, coupled with their small body size (e.g., phoronids, cycliophorans, gastrotrichas), and by insufficient scientific focus. Thus, Lohpotrochozoan evolutionary relationships still await a clearer understanding.

Employment of big data, such as transcriptomes and genomes, showed great promise for resolving difficult phylogenetic questions with information from hundreds of genes. Indeed, in some cases, the phylogenomic approach was successful in resolving phylogenetic relationships in previously uncertain groupings (e.g., monophyly of Excavata, deep relationships in seed plants, the phylogenetic backbone of sea urchins, systematic relationships in arthropods; Hampl et al., 2009, Ran et al., 2018, Koch et al., 2018, Meusemann et al., 2010). However, some cases proved to be hard to solve even with a high abundance of data (e.g., in birds, relationships on the base of Neoaves, in mammals interrelationships in Laurasiatheria, relationships in early metazoans; Suh 2016, Chen et al., 2017, Philippe et al., 2011). Reasons behind this failure to resolve some phylogenetic questions may be multiple and are often generally alleged to be caused by "systematic errors". Specifically, the main players that can affect the inference of phylogenetic trees from hundreds of genes are the selection of appropriate orthologous genes (and removal of paralogous genes, which become manually impossible due to the scale of studies), the choice of fitting model of sequence evolution and the effect of missing data (for an overview see Philippe et al., 2017). The selected method of inference can also have a considerable effect on the resulting tree (e.g., Bayesian-based methods, Maximum likelihood, or coalescent approach). Further sources of incongruence in results can stem from undetected contamination in the datasets, biased taxon sampling, incomplete lineage sorting, uneven rate of sequence evolution, and, more simply, from the ancient age of radiation. One such exemplary case of the difficult phylogenetic question with incongruent results from different datasets and data treatments is the resolving of deep metazoan relationships, known as the "Ctenopohrora versus Porifera first" question

(for comparison of studies and methods see Philippe et al., 2011). Even though this question is important in the evolutionary sense (it can, for example, show us if complex nervous tissue evolved independently in Ctenophora), it still awaits conclusive resolution. Similarly, pinpointing the phylogenetic position of both groups of Mesozoa proved to be a difficult task.

Mesozoa was originally considered a link between protozoa and multicellular metazoa because of their simplicity (Van Beneden, 1876). However, ideas that their life history may be similar to other parasites (i.e., like Platyhelminthes, being not truly simple but descended from more complex metazoans), appeared in the literature early on (Metschnikoff, 1881, for early views on mesozoan position see Nouvel, 1948). Due to the lack of hard morphological features in Mesozoa (shells, scales, etc.), great anticipation was present at the dawn of the phylogenetic molecular era. First studies based on 18S that included members of mesozoa showed them as early-branching metazoans, which corresponded to the Protozoa-Metazoa link hypothesis (Pawlowski et al., 1996). However, Mesozoa always showed exceptionally long branches, making the results less trustworthy due to the suspicion of long branch attraction artifact confounding the results. The advancement of molecular methods and the usage of genome-wide data again showed that the question of mesozoan placement in the tree of life is a hard problem to tackle (Lu et al., 2017, Schiffer et al., 2018, Zverkov et al., 2019). Phylogenomic methods agree on placing Mesozoa in Lophotrochoza but further position and their monophyly are still contested. Possible sister groups of Mesozoa could be Gnathifera or Platyhelminthes. Orthonectida alone is also sometimes linked with Annelida. The main suspected issues hindering the recovery of mesozoan phylogenetic position in the current phylogenomic studies are data quality, taxonomic sampling, and systematic errors or other shortcomings of modern methods.

The question is, how to deal with such uncertainties in the tree of life? We can try to embrace the uncertainty and work with more scenarios. Crucial point is to report results correctly and fully and carefully examine them, rather than pitch for either of the hypotheses (King & Rokas, 2017).

Research into these difficult phylogenetic cases may sometimes seem futile but it helps us to identify weak points of the currently used methods and to spot areas that require improvement and methods that need further testing. New methods, more data coupled with conscientious reporting of the results should, hopefully, be a way forward to the reliable phylogenetic inference that can help us answer interesting evolutionary questions like the inner relationships in Lophotrochozoa and the exact position of mesozoan clades in the tree of life.

3 Study of population structure based on genetic markers

The study of genetic population structure and its dynamics across species distribution range has the potential to reveal interesting phylogeographical patterns, current or past demography, and ecological relationships. Population genetic structure is defined by life-history traits of a given organism, such as the mode and ability of dispersal or its mating system. Additionally, in an organism where life history is not completely known, the tools of population genetics can help us study the unknown aspects of the species' history and bring more information about its life cycle. For example, analyses of mitochondrial and nuclear markers in sea turtles showed population structure corresponding to the presence of a strong homing impulse in females, but not in males (Bowen & Karl, 2007). This approach to studying life-history traits and populations might prove to be especially useful in small invertebrates that lack strong morphological features, such as orthonectids and dicyemids.

In the marine environment. populations tend to be homogeneous on a wide geographic scale, due to greater uniformity of the environment (Palumbi, 2003). However, some barriers (e.g., salinity or temperature clines, currents, oceanographic fronts) may present unexpected boundaries for a given species (Galarza et al., 2009). Therefore, we cannot always reliably predict population structure without careful study (Wright et al., 2015). For example, the giant squid, *Architeuthis dux*, shows a remarkably homogeneous population across the world seas (Winkelmann et al., 2013), but the Atlantic cod has an unexpectedly structured population on a fine-scale along the short span of the Skagerrak coast without obvious dispersal barriers (Knutsen et al., 2003).

In an organism closely bound to its host (parasite, symbiont, or commensal) population genetic structure is also affected by the life-history traits of its host. One of the drivers of population structuring could be the type of host included in the parasite's life cycle, host's habits (e.g., way of feeding, prey selection, mating habits), and mode of dispersal. Such a relationship can be seen in parasitic trematodes, where genetic population structure is determined by the type of host dispersal. Trematodes with an allogenic cycle (including a bird host) tend to have homogeneous populations but those with an autogenic cycle (without a bird host) display structured populations, due to the lower dispersal capabilities of the hosts (Blasco-Costa & Poulin, 2013). However, in this case, data for marine trematodes are scarce and it is not clear if in the marine environment the difference between the allogenic and autogenic cycle would be clear cut. In general, the marine environment supports long distance dispersal, hence marine parasites might have homogeneous populations regardless of the type of their cycle.

Studies of parasite population structure and their use as tags to discern host origin or specific feeding habits in cohorts or subpopulations can also reveal the true connectivity of the host populations. This approach is crucial for the evaluation of fish stock in a commercial setting as well as for conservation planning or epidemiology and management of pathogen outbreaks in fish (or future cephalopod) aquaculture (Palumbi, 2003; Tully & Nolan, 2002). In contrast to fish, such studies remain rare in cephalopods and marine invertebrates in general (Catalano et al., 2014). Particularly for dicyemids and orthonectids, no such study was previously performed.

4 Aims and scope of the thesis

This thesis aims to elucidate several aspects of mesozoan life traits and their evolutionary history through the use of traditional and novel molecular methods. First, it tackles the question of dicyemid and orthonectid position in the evolutionary tree of life. Through the use of up to date phylogenomic methods, using genomic and transcriptomic data for both mesozoan clades and a wide array of lophotrochozoan representatives, it tries to resolve

mesozoan relationship to other animals and investigate the question of Mesozoan monophyly, providing further insights into evolutionary implications such as the evolution of complex parasitic lifecycles and their simplification. Next, the thesis aims to elucidate aspects of the dicyemid life cycle through the study of their populations. By comparison of the genetic population structure of the host and its respective dicyemids, it aims to explore the patterns of reinfection and differentiation in the selected geographical area. Through the exploration of genetic profiles in individual dicyemids inside one host, this thesis aims to provide information on dicyemid modes of reproduction and reinfections inside its host. Selected chapters are also accompanied by confocal microscopy pictures of dicyemids providing a visual embellishment for mesozoan research.

Overview of chapters:

The first chapter focuses on the phylogenetic position of both mesozoan groups, dicyemids and orthonectids, and the question of their monophyly. It contains an extensive comparison of results based on different phylogenomic approaches.

Chapters two and three examine the population structure of dicyemids, first in comparison with their host, which is followed by an examination of the local population inside one host. Specifically, chapter two shows genetic structure based on the study of the COI gene in the case of *Sepia officinalis* in the Mediterranean area and its respective dicyemids. Chapter three focuses on the examination of individual dicyemids based on microsatellite markers comparing several local infrapopulations from the host *Eledone moschata* from two localities (Naples, Italy, and Pula, Croatia).

The final fourth chapter frames dicyemid position among other cephalopod parasites and contains an overview of future challenges in cephalopod parasitology.

5 References

Aruga, J., Odaka, Y. S., Kamiya, A., & Furuya, H. (2007). Dicyema Pax6 and Zic: tool-kit genes in a highly simplified bilaterian. BMC Evolutionary Biology, 7(1), 201. https://doi.org/10.1186/1471-2148-7-201

Blasco-Costa, I., & Poulin, R. (2013). Host traits explain the genetic structure of parasites: A meta-analysis. Parasitology, 140(10), 1316–1322. https://doi.org/10.1017/S0031182013000784

Bondarenko, N., Bondarenko, A., Starunov, V., & Slyusarev, G. (2019). Comparative analysis of the mitochondrial genomes of Orthonectida: insights into the evolution of an invertebrate parasite species. Molecular Genetics and Genomics, 294(3), 715–727. https://doi.org/10.1007/s00438-019-01543-1

Bowen, B. W., & Karl, S. A. (2007). Population genetics and phylogeography of sea turtles. Molecular Ecology, 16(23), 4886–4907. https://doi.org/10.1111/j.1365-294X.2007.03542.x

Castellanos-Martinez, S., & , Aguirre-Macedo L., & Furuya, H. (2016). Two new species of dicyemid mesozoans (Dicyemida : Dicyemidae) from *Octopus maya* Voss & Solis-Ramirez (Octopodidae) off Yucatan , Mexico. Systematic Parasitology, 93(6), 551–564. https://doi.org/10.1007/s11230-016-9644-x

Catalano, S. R. (2012). A review of the families, genera and species of Dicyemida Van Beneden, 1876. Zootaxa, 32(3479), 1–32. https://doi.org/10.11646/zootaxa.3646.1.11

Catalano, S. R. (2013a). First descriptions of dicyemid mesozoans (Dicyemida : Dicyemidae) from Australian octopus (Octopodidae) and cuttlefish (Sepiidae), including a new record of Dicyemennea in Australian waters. 5683, 306–320.

Catalano, S. R. (2013b). Five new species of dicyemid mesozoans (Dicyemida: Dicyemidae) from two Australian cuttlefish species, with comments on dicyemid fauna composition. Systematic Parasitology, 86(2), 125–151. https://doi.org/10.1007/s11230-013-9443-6

Catalano, S. R., & Furuya, H. (2013). Two new species of dicyemid (Dicyemida: Dicyemidae) from two Australian cephalopod species: *Sepioteuthis australis* (Mollusca: Cephalopoda: Loliginidae) and *Sepioloidea lineolata* (Mollusca: Cephalopoda: Sepiadariidae). The Journal of Parasitology, 99(2), 203–211. https://doi.org/10.1645/GE-3252.1

Catalano, S. R., Whittington, I. D., Donnellan, S. C., & Gillanders, B. M. (2013). Using the giant Australian cuttlefish (*Sepia apama*) mass breeding aggregation to explore the life cycle of dicyemid parasites. Acta Parasitologica, 58(4), 599–602. https://doi.org/10.2478/s11686-013-0186-y

Catalano, S. R., Whittington, I. D., Donnellan, S. C., & Gillanders, B. M. (2014). Parasites as biological tags to assess host population structure: Guidelines, recent genetic advances and comments on a holistic approach. International Journal for Parasitology: Parasites and Wildlife, 3(2), 220–226. https://doi.org/10.1016/j.ijppaw.2013.11.001

Catalano, S. R., Whittington, I. D., Donnellan, S. C., Bertozzi, T., & Gillanders, B. M. (2015). First comparative insight into the architecture of COI mitochondrial minicircle molecules of dicyemids reveals marked inter-species variation. Parasitology, 142(8), 1066–1079. https://doi.org/10.1017/S0031182015000384

Chen, M. Y., Liang, D., & Zhang, P. (2017). Phylogenomic resolution of the phylogeny of laurasiatherian mammals: Exploring phylogenetic signals within coding and noncoding sequences. Genome Biology and Evolution, 9(8), 1998–2012. https://doi.org/10.1093/gbe/evx147

Czaker, R. (1994). *Kantharella antarctica*, a new and unusual dicyemid mesozoan from the Antarctic. Zoologischer Anzeiger 232, 151-158.

Czaker, R. (2000). Extracellular matrix (ECM) components in a very primitive multicellular animal, the dicyemid mesozoan *Kantharella antarctica*. Anatomical Record, 259(1), 52–59. https://doi.org/10.1002/(sici)1097-0185(20000501)259:1<52::aid-ar6>3.0.co;2-j

Eshragh, R., & Leander, B. S. (2014). Molecular contributions to species boundaries in dicyemid parasites from eastern Pacific cephalopods. Marine Biology Research, 11(November), 414–422. https://doi.org/10.1080/17451000.2014.943241

Finn, J. K., Hochberg, F. G., & Norman, M. D. (2005). Phylum Dicyemida in Australian Waters: First Record and Distribution Across Diverse Cephalopod Hosts. Phuket Marine Biology Center Research Bulletin, 66, 83–96.

Furuya, H. (1999). Fourteen New Species of Dicyemid Mesozoans from Six Japanese Cephalopods, with Comments on Host Specificity. Species Diversity, 4, 257–319.

Furuya, H., (2005). Chapter: Minor groups and fossils, Mesozoa Dicyemida in Marine parasitology (editor Rhode K.). CSIRO, ISBN: 9780643090255: 185-189.

Furuya, H. (2006). Three new species of dicyemid mesozoans (phylum Dicyemida) from *Amphioctopus fangsiao* (Mollusca: Cephalopoda), with comments on the occurrence patterns of dicyemids. Zoological Science, 23(1), 105–119. https://doi.org/10.2108/zsj.23.105

Furuya, H. (2010). A new dicyemid from *Benthoctopus sibiricus* (Mollusca: Cephalopoda: Octopoda). The Journal of Parasitology, 96(6), 1123–1127. https://doi.org/10.1645/GE-2411.1

Furuya, H. (2018). Eleven new species of dicyemids (phylum Dicyemida) from *Octopus longispadiceus* and *O. tenuicirrus* (Mollusca: Cephalopoda: Octopoda) in Japanese waters. Species Diversity, 23(2), 143–179. https://doi.org/10.12782/SPECDIV.23.143

Furuya, H., Hochberg, F. G., & Short, R. B. (2002). *Dicyemennea* (Dicyemida) from *Bathypolypus arcticus* (Mollusca : Cephalopoda : Octopoda). Journal of Parasitology, 88(1), 119–123.

Furuya, H., & Souidenne, D. (2019). Chapter 11: Dicyemids in *Handbook of Pathogens and Diseases in Cephalopods*. Handbook of Pathogens and Diseases in Cephalopods, Springer, ISBN 978-3-030-11330-8, 159–168. https://doi.org/10.1007/978-3-030-11330-8

Furuya, H., Hochberg, F. G., & Tsuneki, K. (2001). Developmental patterns and cell lineages of vermiform embryos in dicyemid mesozoans. Biological Bulletin, 201(3), 405–416.

Furuya, H., & Tsuneki, K. (2003). Biology of dicyemid mesozoans. Zoological Science, 20(5), 519–532. https://doi.org/10.2108/zsj.20.519

Galarza, J. A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G. F., & Rico, C. (2009). The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. Proceedings of the National Academy of Sciences of the United States of America, 106(5), 1473–1478. https://doi.org/10.1073/pnas.0806804106

Giard, A., (1877). Sur les Orthonectida, classe nouvelle d'animaux parasites des Échinodermes et des Turbellariés. Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, 85(18). 812-814.

Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M. A., Liva, S. M., Hillis, D. M., & Lake, J. A. (1995). Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. Science, 267(5204), 1641–1643. https://doi.org/10.1126/science.7886451

Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G. B., & Roger, A. J. (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups." Proceedings of the National Academy of Sciences of the United States of America, 106(10), 3859–3864. https://doi.org/10.1073/pnas.0807880106

Hanelt, B., Van Schyndel, D., Adema, C. M., Lewis, L. A., & Loker, E. S. (1996). The phylogenetic position of *Rhopalura ophiocomae* (Orthonectida) based on 18S ribosomal DNA sequence analysis. Molecular Biology and Evolution, 13(9), 1187–1191. https://doi.org/10.1093/oxfordjournals.molbev.a025683

Hochberg, F.G. (1982). The "kidneys" of cephalopods: a unique habitat for parasites. Malacologia 23: 121-134.

Hochberg, F. G. (1990). "Diseases of Mollusca: Cephalopoda: diseases caused by protistans and metazoans," in Diseases of Marine Animals, Vol. III, Cephalopoda to Urochordata, ed. O. Kinne (Hamburg: Biologisches Anstalt Helgoland), 47–227.

Hori, H., & Osawa, S. (1987). Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Molecular Biology and Evolution, 4(5), 445–472. https://doi.org/10.1093/oxfordjournals.molbev.a040455

Katayama, T., Wada, H., Furuya, H., Satoh, N., & Yamamoto, M. (1995). Phylogenetic position of the dicyemid mesozoa inferred from 18S rDNA sequences. Biological Bulletin, 189(2), 81–90. https://doi.org/10.2307/1542458

King, N., & Rokas, A. (2017). Embracing Uncertainty in Reconstructing Early Animal Evolution. Current Biology, 27(19), R1081–R1088. https://doi.org/10.1016/j.cub.2017.08.054

Knutsen, H., Jorde, P. E., André, C., & Stenseth, N. C. (2003). Fine-scaled geographical population structuring in a highly mobile marine species: The Atlantic cod. Molecular Ecology, 12(2), 385–394. https://doi.org/10.1046/j.1365-294X.2003.01750.x

Kobayashi, M., Furuya, H., & Wada, H. (2009). Molecular markers comparing the extremely simple body plan of dicyemids to that of lophotrochozoans: Insight from the expression patterns of Hox, Otx, and brachyury. Evolution and Development, 11(5), 582–589. https://doi.org/10.1111/j.1525-142X.2009.00364.x

Koch, M. N., Coppard, S. E., Lessios, H. A., Briggs, D. E. G., Mooi, R., & Rouse, G. W. (2018). A phylogenomic resolution of the sea urchin tree of life. BMC Evolutionary Biology, 18(1), 1–18. https://doi.org/10.1186/s12862-018-1300-4

Kozloff, E. N. (1992). The genera of the phylum Orthonectida. Cahiers de Biologie Marine, 33, 377–406.

Kozloff, E. N. (1994). The structure and origin of the plasmodium of *Rhopalura ophiocomae* (Phylum Orthonectida). Acta Zoologica, 75(3), 131–199. https://doi.org/10.1111/j.1463-6395.1994.tb01206.x

Lapan, E. A., & Morowitz, H. J. (1972). The Mesozoa. Scientific American, 227, 94–101.

Lapan, E. A., & Morowitz, H. J. (1975). The dicyemid mesozoa as an integrated system for morphogenetic studies. I. Description isolation and maintenance. Journal of Experimental Zoology, 193(2), 147–159. https://doi.org/10.1002/jez.1401930204

Lu, T. M., Kanda, M., Satoh, N., & Furuya, H. (2017). The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. Zoological Letters, 3(1), 1–9. https://doi.org/10.1186/s40851-017-0068-5

Lu, T. M., Kanda, M., Furuya, H., & Satoh, N. (2019). Dicyemid Mesozoans: A Unique Parasitic Lifestyle and a Reduced Genome. Genome Biology and Evolution, 11(8), 2232–2243. https://doi.org/10.1093/gbe/evz157

Matsubara, J. A., & Dudley, P. L. (1976). Fine Structural Studies of the Dicyemid Mesozoan, *Dicyemmenea californica* McConnaughey. I. Adult Stages. The Journal of Parasitology, 62(3), 377–289. https://doi.org/10.2307/3279146

Metschnikoff, E. (1881). Untersuchunger uber Orthonectiden. Ztschr. Wiss. Zool., 32, 349–387.

Meusemann, K., Von Reumont, B. M., Simon, S., Roeding, F., Strauss, S., Kück, P., Ebersberger, I., Walzl, M., Pass, G., Breuers, S., Achter, V., Haesler, A., Burmester, T., Hadrys, H., Wägele, J. W., & Misof, B. (2010). A phylogenomic approach to resolve the arthropod tree of life. Molecular Biology and Evolution, 27(11), 2451–2464. https://doi.org/10.1093/molbev/msq130

Mikhailov, K. V., Slyusarev, G. S., Nikitin, M. A., Logacheva, M. D., Penin, A. A., Aleoshin, V. V., & Panchin, Y. V. (2016). The Genome of *Intoshia linei* Affirms Orthonectids as Highly Simplified Spiralians. Current Biology, 26(13), 1768–1774. https://doi.org/10.1016/j.cub.2016.05.007

Muñoz, G., Rodríguez, S., Guerra, R., & Guisado, C. (2013). Primer registro en Chile de diciémidos simbiontes en el pulpo pigmeo *Robsonella fontaniana* (Cephalopoda : Octopodidae). Revista de Biologia Marina y Oceanografia, 48(3), 647–651. https://doi.org/10.4067/S0718-19572013000300022

Nakano, H., & Miyazawa, H., (2019). A new species of Orthonectida that parasitizes *Xenoturbella bocki*: Implications for studies on Xenoturbella. Biological Bulletin, 236, 66–73. Doi:10.1086/700834

Nouvel, H., (1947). Les dicyémides 1er partie: systématique, générations vermiformes, infusorigène et sexualité. Archives de Biologie, 58, 59–214.

Nouvel, H., (1948). Les dicyémides 2em partie: Infusoriforme, Teratologie, Specificite du parasitisime, Affinites. Archives de Biologie. 59:147–223.

Ogino, K., Tsuneki, K., & Furuya, H. (2007). The expression of tubulin and tektin genes in dicyemid mesozoans (Phylum : Dicyemida). Journal of Parasitology, 93(3), 608–618.

Ohama, T., Kumazaki, T., Hori, H., & Osawa, S. (1984). Evolution of multicellular animals as deduced from SS rRNA sequences: A possible early emergence of the mesozoa. Nucleic Acids Research, 12(12), 5101–5108. https://doi.org/10.1093/nar/12.12.5101

Pace, N. R., Sapp, J., & Goldenfeld, N. (2012). Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. Proceedings of the National Academy of Sciences of the United States of America, 109(4), 1011–1018. https://doi.org/10.1073/pnas.1109716109

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. Ecological Applications, 13(1 SUPPL.), S146–S158. https://doi.org/10.1890/1051-0761(2003)013[0146:pgdcat]2.0.co;2

Pawlowski, J., Montoya-Burgos, J. I., Fahrni, J. F., Wüest, J., & Zaninetti, L. (1996). Origin of the Mesozoa inferred from 18S rRNA gene sequences. Molecular Biology and Evolution, 13(8), 1128–1132. https://doi.org/10.1093/oxfordjournals.molbev.a025675

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving difficult phylogenetic questions: Why more sequences are not enough. PLoS Biology, 9(3), e1000602. https://doi.org/10.1371/journal.pbio.1000602

Philippe, H., Vienne, D. M. de, Ranwez, V., Roure, B., Baurain, D., & Delsuc, F. (2017). Pitfalls in supermatrix phylogenomics. European Journal of Taxonomy, (283), 1–25. https://doi.org/10.5852/ejt.2017.283

Ran, J. H., Shen, T. T., Wang, M. M., & Wang, X. Q. (2018). Phylogenomics resolves the deep phylogeny of seed plants and indicates partial convergent or homoplastic evolution between Gnetales and angiosperms. Proceedings of the Royal Society B: Biological Sciences, 285(1881). https://doi.org/10.1098/rspb.2018.1012

Ridley, R. K. (1968). Electron microscopic studies on Dicyemid Mesozoa. I. Vermiform Stages. The Journal of Parasitology, 54(5), 975–998.

Ridley, R. K. (1969). Electron Microscopic Studies on Dicyemid Mesozoa. II. Infusorigen and Infusoriform Stages. The Journal of Parasitology, 55(4), 779–793.

Schiffer, P. H., Robertson, H. E., & Telford, M. J. (2018). Orthonectids Are Highly Degenerate Annelid Worms. Current Biology, 28(12), 1970-1974.e3. https://doi.org/10.1016/j.cub.2018.04.088

Slyusarev, G. S., & Starunov, V. V. (2016). The structure of the muscular and nervous systems of the female *Intoshia linei* (Orthonectida). Organisms Diversity and Evolution, 16(1), 65–71. https://doi.org/10.1007/s13127-015-0246-2

Soltis, D. E., & Soltis, P. S. (2003). The Role of Phylogenetics in Comparative Genetics. Plant Physiology, 132(4), 1790–1800. https://doi.org/10.1104/pp.103.022509

Souidenne, D., Florent, I., Dellinger, M., Romdhane, M. S., Grellier, P., & Furuya, H. (2016). Redescription of *Dicyemennea eledones* (Wagener, 1857) (Phylum Dicyemida) from *Eledone cirrhosa* (Lamarck, 1798) (Mollusca: Cephalopoda: Octopoda). Systematic Parasitology, 93(9), 905–915. https://doi.org/10.1007/s11230-016-9659-3

Stoakes, S. F., (2019). Importance of Phylogeny in Microbiology. accessed at news-medical.net, https://www.news-medical.net/life-sciences/Importance-of-Phylogeny-in-Microbiology.aspx

Stunkard, H. W. (1954). The life history and systematic relations of the mesozoa. The Quarterly Review of Biology, 29(3), 230–244.

Suh, A. (2016). The phylogenomic forest of bird trees contains a hard polytomy at the root of Neoaves. Zoologica Scripta, 45(S1), 50–62. https://doi.org/10.1111/zsc.12213

Suzuki, T. G., Ogino, K., Tsuneki, K., & Furuya, H. (2010). Phylogenetic analysis of dicyemid mesozoans (phylum Dicyemida) from innexin amino acid sequences: dicyemids are not related to Platyhelminthes. The Journal of Parasitology, 96(3), 614–625. https://doi.org/10.1645/GE-2305.1

Tully, O., & Nolan, D. T. (2002). A review of the population biology and host–parasite interactions of the sea louse *Lepeophtheirus salmonis* (Copepoda: Caligidae) . Parasitology, 124(7), 165–182. https://doi.org/10.1017/s0031182002001889

Van Beneden, M. E. (1876). Recherches sur le Dicyemides, survivants actuels d'un embranchement des Mésozoaires. Bull. Acad. R. Sci. Lett. Beaux Arts Belgique, 45 année, 2me ser., XLI, 1160–1205, and XLII, 35–97
(suite) https://www.biodiversitylibrary.org/item/28491#page/9/mode/1up https://www.biodiversitylibrary.org/item/107938#page/2/mode/1up

Vázquez, D. P., & Gittleman, J. L. (1998). Biodiversity conservation: Does phylogeny matter? Current Biology, 8(11), 379–381. https://doi.org/10.1016/s0960-9822(98)70242-8

Whitman, C. O. (1882). A contribution to the embryology, life history and classification of the Dicyemids. Publisher W. Englemann, Ohio State university, reprint by Breitkopf & Hartel, Leipzig.

Winkelmann, I., Campos, P. F., Strugnell, J., Cherel, Y., Smith, P. J., Kubodera, T., Allcock, L., Kampmann, M., Schroeder, H., Guerra, A., Norman, M., Finn, J., Ingrao, D., Clarke, M., & Gilbert, M. T. P. (2013). Mitochondrial genome diversity and population structure of the giant squid Architeuthis: Genetics sheds new light on one of the most enigmatic marine species. Proceedings of the Royal Society B: Biological Sciences, 280(1759). https://doi.org/10.1098/rspb.2013.0273

Wright, D., Bishop, J. M., Matthee, C. A., & von der Heyden, S. (2015). Genetic isolation by distance reveals restricted dispersal across a range of life histories: Implications for biodiversity conservation planning across highly variable marine environments. Diversity and Distributions, 21(6), 698–710. https://doi.org/10.1111/ddi.12302

Zverkov O.A., Mikhailov K.V., Isaev S.V., Rusin L.Y., Popova O.V., Logacheva M.D., Penin A.A., Moroz L.L., Panchin Y.V., Lyubetsky V.A., & Aleoshin V.V. (2019). Dicyemida and Orthonectida: Two stories of body plan simplification. Frontiers in Genetics, 10(MAY),443. https://doi.org/10.3389/fgene.2019.00443

CHAPTER I.

Different Phylogenomic Methods Support Monophyly of Enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa)

Manuscript under review in
Proceedings of the Royal Society B (2022)

Different Phylogenomic Methods Support Monophyly of Enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa)

Marie Drábková*[1, 2], Kevin M. Kocot[3], Kenneth M. Halanych[4], Todd H. Oakley[5], Leonid L. Moroz[6], Johanna T. Cannon[5], Armand Kuris[5], Ana Elisa Garcia-Vedrenne[5], M. Sabrina Pankey[7], Emily A. Ellis[5], Rebecca Varney[5], Jan Štefka[1, 2] and Jan Zrzavý[8]

[1] Department of Parasitology, Faculty of Science, University of South Bohemia, České Budějovice 37005, Czech Republic
[2] Laboratory of Molecular Ecology and Evolution, Institute of Parasitology, Biology Centre CAS, České Budějovice 37005, Czech Republic
[3] Department of Biological Sciences, The University of Alabama, Campus Box 870344, Tuscaloosa, AL 35487, USA
[4] The Center for Marine Science, University of North Carolina Wilmington, 5700 Marvin K. Moss Lane, Wilmington, NC 28409, USA
[5] Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA
[6] Department of Neuroscience, and the Whitney Laboratory for Marine Bioscience, University of Florida, 9505 Ocean Shore Blvd, St Augustine, FL 32080, USA
[7] Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham NH 03824
[8] Department of Zoology, Faculty of Science, University of South Bohemia, České Budějovice 37005, Czech Republic

* corresponding author

Abstract

Dicyemids and orthonectids were traditionally classified in a group called Mesozoa, but their placement in a single clade has been contested and their position(s) within Metazoa is uncertain. Here, we assembled a comprehensive lophotrochozoan matrix and investigated the position of Dicyemida (= Rhombozoa) and Orthonectida employing multiple phylogenomic approaches. We sequenced seven new transcriptomes and one draft genome from dicyemids (*Dicyema*, *Dicyemennea*), two transcriptomes from orthonectids (*Rhopalura*), and selected lophotrochozoan phyla. Using these and published data, we assembled and analyzed contamination-filtered datasets with up to 987 genes. Our results recover Mesozoa monophyletic and as a close relative of Platyhelminthes or Gnathifera. Because of the tendency of the long-branch mesozoans to group with other long-branch taxa in our analyses, we explored the impact of approaches purported to help alleviate long branch attraction (e.g., taxon removal, coalescent inference, gene targeting). None of these were able to break the association of Orthonectida with Dicyemida in the maximum likelihood trees. Contrastingly, the Bayesian analysis and site-specific frequency model in maximum likelihood did not recover a monophyletic Mesozoa, but only when using a specific 50 gene matrix. The classic hypothesis on monophyletic Mesozoa is possibly reborn and should be further tested.

Introduction

Phylogenomic studies have dramatically improved our understanding of deep metazoan phylogeny, however, some key branches remain controversial (see e.g., Dunn et al., 2014; Halanych, 2016; Kocot, 2016; Giribet & Edgecombe, 2017; Bleidorn, 2019; Marlétaz et al., 2019). Incongruence in phylogenomic analyses could be due to insufficient phylogenetic signal (e.g., due to closely spaced branching events), limited taxon and/or gene sampling, or systematic error, particularly the long-branch attraction (LBA) artifact (Felsenstein, 1978; reviewed by Bergsten, 2005; Philippe et al., 2017).

The early radiation of Lophotrochozoa (549-535 Mya old; Qun et al., 2007), a major clade of protostome metazoans including annelids, molluscs, brachiopods, bryozoans, flatworms and rotifers among others, has been especially challenging to uncover (Kocot, 2016, Bleidorn, 2019, Laumer et al., 2019, Marlétaz et al., 2019, Zverkov et al., 2019). This is probably due to relatively rapid diversification (Rokas et al., 2005) and fast evolutionary rates in some lineages, leading to long terminal branches in phylogenetic analyses possibly prone to LBA. Differences in branch lengths can be caused by differences in the rate of molecular evolution and generation time, as both are known to be variable for invertebrate phyla (Thomas et al., 2010), especially in parasites (Haraguchi & Sasaki, 1996). This may aggravate the inherent difficulty of resolving ancient radiations because of short internal branches among deep nodes (Budd & Jackson, 2016).

The placement of Orthonectida and Dicyemida, two clades of enigmatic morphologically simple parasites, could be seen as examples of such a controversial case (Fig. 1). Orthonectids (ca. 25 spp. (WoRMS Editorial Board 2018: van der Land, Furuya & Decock, 2018)) are parasites of a variety of marine invertebrates; their adults are free-swimming and produce ciliated larvae that enter the host tissues to form amoeboid trophic syncytia that effectively castrate their hosts (Deheyn et al., 1998). Dicyemida (syn. Rhombozoa, especially if including still unsequenced Heterocyemida; ca.

100 spp. (Catalano, 2012)) are symbionts inhabiting the renal organs of benthic cephalopods. Orthonectids and dicyemids were originally united as Mesozoa by Van Beneden (Van Beneden, 1876) who viewed them as a 'link' between Protozoa and Metazoa. In contrast, a hypothesis on their convergent evolution from more complex animals prevails in modern literature, even though their phylogenetic origins remain uncertain.

Phylogenomic analyses by Mikhailov et al. (Mikhailov et al., 2016; 22,909 amino acid positions, 61 taxa), including the first published mesozoan genome (*Intoshia*, Orthonectida) recovered Orthonectida as a clade of uncertain position within Lophotrochozoa. Lu et al. (Lu et al., 2017) conducted phylogenomic analyses (58,124 amino acid positions, 29 taxa) with two genomes from both Orthonectida (*Intoshia*) and Dicyemida (*Dicyema*) and recovered Mesozoa as a monophyletic group close to gastrotrichs and/or platyhelminths. On the contrary, Schiffer et al. (Schiffer et al., 2018) proposed, based on both mitochondrial (2,969 amino acid positions, 69 taxa) and nuclear (190,027 amino acid positions, 45 taxa) data, that Mesozoa is not monophyletic: orthonectid *Intoshia* was recovered as nested within Annelida while the position of Dicyemida (three spp. of *Dicyema*) within Lophotrochozoa was considered unresolved. Zverkov et al. (Zverkov et al., 2019) reported similar results in their analysis of nuclear protein-coding genes (87,610 amino acid positions, 73 taxa including *Intoshia* and two spp. of *Dicyema*). Taken together, all phylogenomic analyses to date (and most molecular analyses in general) have converged to the scenario that both mesozoan groups represent secondarily simplified lophotrochozoans, but their phylogenetic position within Lophotrochozoa and whether or not they form a monophyletic group remain unresolved. Difficulties in the placement of these enigmatic marine groups may be due to their fast rate of evolution, as evidenced by extremely long branches of both 'mesozoan' groups in all analyses.

Multiple approaches exist to tackle systematic errors in phylogenomics. One straightforward way to treat LBA and test leaf stability is to remove other long branches, thereby decreasing the 'pull' of these long branches

and compare the resulting position of the taxon in question (e.g., the position of Orthonectida after exclusion of Dicyemida from the matrix, and vice versa (Schiffer et al., 2018). However, if more than two long-branch taxa are present, interpretation of the results becomes difficult. The site-specific models of sequence evolution, such as the CAT+GTR model (Lartillot & Philippe, 2004) which is implemented in a Bayesian framework in PhyloBayes (Lartillot et al., 2013), and the PMSF model(Wang et al., 2018), implemented in a maximum likelihood (ML) framework in IQ-TREE (Nguyen et al., 2015), have been purported to reduce LBA by modeling the actual complexity of the data resulting from biological processes. Other approaches thought to help mitigate LBA are based on selecting molecular markers based on properties such as branch-length homogeneity or compositional homogeneity. Balanced sequence composition has also been suggested to be a good predictor of phylogenetic signal (e.g., Shen et al., 2016), and taking steps to reduce compositional heterogeneity was advocated by Nesnidal et al. (Nesnidal et al., 2010) in order to help overcome systematic artifacts affecting inference of relationships within Lophotrochozoa. Another approach for phylogeny reconstruction that may be useful for genomic datasets containing heterogeneous signal is the reconstruction of the species tree based on coalescent theory. This method has been argued to perform well in cases where heterogeneous single-gene trees are present in the analyzed set (Liu et al., 2009), (Kapli et al., 2020).

Here, we sought to explore the phylogenetic position(s) of Orthonectida and Dicyemida in the context of lophotrochozoan phylogeny, focusing on extensive comparisons of results from different approaches purported to reduce LBA. To this end, we assembled a dataset with significantly increased taxon sampling for Mesozoa (genome/transcriptome data from eight dicyemids and three orthonectids) and diverse representatives of other lophotrochozoan phyla (32 spp.) that we carefully screened to exclude exogenous contamination. We constructed several data matrices from different sets of genes, selecting genes according to different criteria purported to help reduce LBA and other artifacts thought to have impacted

earlier studies of lophotrochozoan phylogeny. We performed phylogenetic analyses on these datasets using the following approaches: (i) maximum likelihood (ML) using the best-fitting site-homogeneous model for each gene as implemented in RaxML (Stamatakis, 2014); (ii) Bayesian inference (BI) using the CAT+GTR empirical profile mixture model to account for site-specific evolutionary rate heterogeneity (Lartillot et al., 2007) as implemented in PhyloBayes (Lartillot et al., 2013); (iii) ML inference using site-specific PMSF profile mixture model (Wang et al., 2018) as implemented in IQ-TREE (Nguyen et al., 2015), enabling us to compare the use of site-specific models in both a ML and Bayesian framework; and (iv) a coalescent approach for phylogenomics as implemented in ASTRAL II (Mirarab & Warnow, 2015).

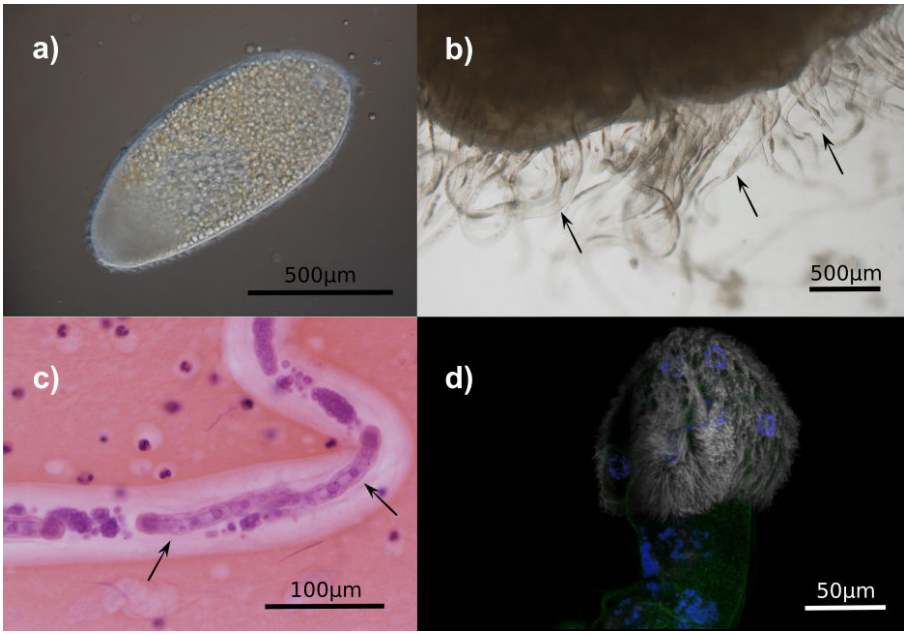


Figure 1. Photos of mesozoans (a and b *in vivo*, c and d fixed and stained). a) Orthonectida sp. indet. from *Ophionotus victoriae* (Echinodermata: Ophiuroidea) collected near Hugo Island, Antarctica. Scale bar is approximate. b) Many specimens of Dicyemida sp. indet. from *Eledone moschata* (Mollusca: Cephalopoda) attached to the host's renal organ. c) Vermiform embryos (stained dark purple) developing within the axial cell of *Dicyema moschatum* (smears on coverslips, fixed by Bouin's fluid, stained with haematoxylin-eosin). d) Cilia on calotte of *Dicyema moschatum* specimen stained with the tubulin-specific dye phalloidin and the nuclear stain DAPI (visualized by fluorescent confocal microscopy).

Material and methods

Taxon Sampling

We sought to broadly sample the diversity of Lophotrochozoa with an emphasis on Dicyemida and Orthonectida, while avoiding including too many terminals in order to facilitate the computationally intensive analyses (e.g., PhyloBayes analyses with the site-heterogeneous CAT model). Thus, we opted to broadly sample each relevant phylum with the minimum number of taxa necessary to capture the deepest node in that phylum while using only high-quality (e.g., deeply sequenced and contamination-free) data. We constructed a dataset with seven newly sequenced dicyemid transcriptomes, one newly sequenced dicyemid genome, two newly sequenced orthonectid transcriptomes, and representatives of all other lophotrochozoan phyla, including three newly sequenced bryozoan transcriptomes (Table S1 in ESM1).

To test the historical hypothesis that the mesozoans are closely related to flukes (Trematoda), *Schistosoma* was included as a representative (Platyhelminthes: Neodermata). Because orthonectids have been recovered as closely related to or even placed within Annelida, representatives of Palaeoannelida (*Magelona*, *Owenia*), Chaetopteriformia (*Phyllochaetopterus*), Sipunculida (*Phascolosoma*), and Pleistoannelida (*Capitella*) were included. Representatives of Cnidaria (*Nematostella*), Xenacoelomorpha (*Xenoturbella* and *Meara*), Ambulacraria (*Ptychodera* and *Strongylocentrotus*), and Ecdysozoa (*Priapulus*) were used as outgroups.

Molecular Laboratory Techniques

To produce new transcriptome data for this study, total RNA was extracted from *Alcyonidium* sp. (Bryozoa), *Cristatella mucedo* (Bryozoa), *Pectinatella magnifica* (Bryozoa), two different collections of *Rhopalura* cf. *ophiocomae* (Orthonectida; both from *Amphipholis squamata*), and dicyemids *Dicyema* sp. 1 (from Enteroctopus *dofleini*), *Dicyema* sp. 2 (from *Octopus bimaculoides*), Dicyemida sp. 3 (from *Megaleledone*

setebos), Dicyemida sp. 4 (from *Octopus vulgaris*), and *Dicyemennea brevicephaloides* and *D. rossiae* (both from *Rossia pacifica*). For assembly of dicyemid genome individuals of *Dicyema moschatum* from *Eledone moschata* were used. Specimen collection data, details of sample processing and sequence assembly (sample handling, extraction and cDNA library kits used, sequencing strategy, sequence assembly and annotation) for all newly sequenced taxa are provided in ESM1.

Contamination Filtering

Sequences were compared to the NCBI Protein and Nucleotide databases (accessed March 2017) using the blastp and blastn algorithms (Altschul et al., 1990) to check for possible contamination. First, a custom script (available in ESM2) was used to evaluate taxonomic distribution of hits in blastp output (Altschul et al., 1990) and the results were visualized in MEGAN v.3 (Huson et al., 2016). Taxa with a relatively high number of suspicious hits were exchanged for available closely related species. Datasets retained after this preliminary screening were compared to the NCBI nucleotide database (accessed March 2017) with blastn (Altschul et al., 1990). Using the custom script described above, the output of blastn was semi-manually evaluated, checking the similarity scores and taxonomy ID of hits. Briefly, blastn output was screened with stricter settings than the original blastp search: it was further filtered to consider only hits with percent identity above 95 and e-value less than 1e-50. The remaining hits were sorted by bitscore and only the top hit for a given gene was extracted. If the extracted hits were similar to closely related species, they were retained, but when it was clear they originated from contamination (e.g., sequences >95% similar to bacterial sequences, parasitic protists and other non-metazoans, or cephalopods in the case of the dicyemids), they were removed (for details see ESM2). Contamination-filtered nucleotide sequences were then translated into amino acids with Transdecoder (http://transdecoder.sf.net; script available in ESM2).

Dataset Construction

Translated transcripts for all taxa were searched against 2,259 lophotrochozoan-specific core orthologous groups (OGs) as profile hidden Markov models (Lophotrochozoa-Kocot pHMMs (Kocot et al., 2017) in HaMStR 13.2.6 (Ebersberger et al., 2009). Briefly, this dataset is based on genes identified to be single-copy in genomes or deeply-sequenced transcriptomes from representatives of the phylum Annelida, Brachiopoda, Entoprocta, Mollusca, Nemertea, Platyhelminthes, Phoronida, and Rotifera. Sequences matching an OG's pHMM were compared to the proteome of *Lottia* using blastp. If the *Lottia* amino acid sequence contributing to the pHMM was the best blastp hit in each of these back-BLASTs, the sequence was then assigned to that OG.

In order to reduce missing data, OGs sampled for fewer than 20 taxa (631 OGs) were discarded. Redundant sequences that were identical where they overlapped were then removed, leaving only unique sequences for each taxon. Each OG was then aligned with MAFFT (Katoh et al., 2005) using the automatic alignment strategy with a 'maxiterate' value of 1,000. Alignments were then trimmed with Zorro (Wu et al., 2012) with the default options to remove ambiguously aligned regions (score below 0.5). At this point, OGs with alignments shorter than 50 amino acids in length were discarded (248 OGs discarded). Only sequences overlapping with all other sequences in the alignment by at least 20 amino acids were kept. To select the best sequence for each taxon and to help exclude overlooked paralogs or exogenous contamination, we built approximate maximum likelihood trees in FastTree 2 (Price et al., 2010) and used PhyloTreePruner (Kocot et al., 2013) to select the best sequence for each taxon. Further screening for paralogs and exogenous contamination was implemented using TreSpEx 1.0 (Struck, 2014). Potential paralogs were removed and excluded from further analysis and only OGs with at least 20 taxa remaining were retained (for details on OG selection see ESM3). Remaining OGs were concatenated to create the complete phylogenomic matrix (abbreviated COMP; for overview of matrices and proportion of missing data see Table S2 in ESM4).

To reduce homoplasy that may be introduced by including distant outgroups, we excluded, for most analyses, all outgroups except the very slowly evolving ecdysozoan *Priapulus* (matrices with names ending in R have a Reduced outgroup, e.g., COMPR). We assembled several different matrices targeting genes with specific qualities such as missing data, phylogenetic signal, branch-length heterogeneity, and compositional heterogeneity (Table S2 in ESM4). We used Matrix Reduction (MARE; Meyer et al., 2011) to exclude OGs with low phylogenetic signal and to reduce missing data (a dataset called 'MARER' hereafter). OGs were sorted by branch length heterogeneity ('CxLBR' hereafter) and by compositional heterogeneity ('CxHR' hereinafter) to select the best 50 and 100 OGs according to each criterion (C50LBR, C100LBR, C50HR, C100HR). We used TreSpEx 1.0 (Struck, 2014) and the single-gene ML trees described above to assess branch length heterogeneity of each OG and BaCoCA (Kück & Struck, 2014) to assess the compositional heterogeneity of each OG according to the relative composition frequency variability (RCFV) metric. We also assembled a matrix of the best 50 OGs of the MARE-reduced dataset (MARER) according to compositional heterogeneity (M50HR). Both C50LBR and MARER datasets were used for our numerous taxon removal experiments (see below) and Bayesian inference analyses using the computationally demanding CAT+GTR model.

For an overview of matrices produced in this study and a flowchart explaining data matrix production see ESM4. All data matrices analyzed in this study and corresponding partition data are available in ESM4.

Phylogenomic Analyses

Maximum Likelihood Analyses

Phylogenetic analyses of all datasets were conducted using maximum likelihood (ML) with RAxML 8.2.4 (Stamatakis, 2014). Matrices were partitioned by gene (=orthologous group or OG) and the AUTO option was used to select the best-fitting model for each partition. For each ML analysis, the tree with the best likelihood score after ten random addition

sequence replicates was retained and topological robustness (i.e., nodal support) was assessed by the optimal number of rapid bootstraps. Furthermore, some matrices were analyzed with posterior mean site frequency (PMSF) model in ML (LG+C60+G+F; Wang et al., 2018) as implemented in IQ-TREE 1.5.3 (Nguyen et al., 2015), as an approximation of the CAT site-specific model (Lartillot & Philippe, 2004),(Si Quang et al., 2008).

Bayesian Inference

Bayesian analyses (BI) were performed in Phylobayes MPI 1.5 or 1.7 (Lartillot et al., 2013) using the CAT+GTR model (Lartillot & Philippe, 2004; Si Quang et al., 2008). Because of the computationally intensive nature of PhyloBayes analyses, only datasets with a highly reduced number of genes sampled were analyzed (see Results). Each analysis was run simultaneously in four chains. Analyses were periodically checked for progress and convergence among chains was assessed with the bpcomp program from the PhyloBayes package after discarding a burn-in of 1500 cycles. Convergence was assumed if the maxdiff dropped below 0.3, as advocated in the Phylobayes manual. Chains were run until convergence or for at least 12,000 generations, but often for longer if resources were available.

Comparison of ML and BI tree-building methods

In order to inspect differences between tree-building methods without the confounding factor of the long-branch mesozoans, we analyzed the C50LBR matrix (the only BI analysis that successfully converged) by RAxML, IQ-TREE (PMSF model), and BI (CAT model). This data matrix was also modified by exclusion of Dicyemida, Orthonectida or both.

Taxon Removal Experiments

To test the robustness of relationships within Lophotrochozoa, we carried out series of taxon removal experiments based on the MARER dataset. Topological effects of the presence/absence of several clades of interest or their combinations (Platyhelminthes, Gastrotricha, Rouphozoa, Gnathifera,

Gnathifera + Platyhelminthes, Entoprocta–Cycliophora, Bryozoa, and Annelida, i.e., all possible sister groups of Mesozoa, and/or clades containing long-branch taxa; for details see ESM5) were tested in ML (IQ-TREE).

Gene signal analysis

We examined conflicting signal among genes by the difference in log-likelihood score (ΔGLS) between two competing tree topologies (as in (Shen et al., 2017; Ballesteros & Sharma, 2019)). The two topologies correspond to the unconstrained tree resulting from ML analysis (T1; including monophyletic Mesozoa) and the constrained tree (T2) where Orthonectida were forced to group with Annelida. ML trees were estimated from the 202-gene (MARER) dataset under automatically selected model per partition in RAxML. Site likelihood scores for contrasting gene topologies needed for ΔGLS computation were estimated under LG model in RAxML. We also applied the approximately unbiased (AU) test in the software package CONSEL version 0.20. The AU test was conducted using the multi-scale bootstrap technique based on the site-wise log-likelihood scores.

Coalescent approach

Coalescent phylogenetic analysis of single-gene trees generated in RAxML (987 gene trees) was run in ASTRAL-II 4.11.1 (Mirarab & Warnow, 2015) with default settings. Supports were computed by gene resampling as suggested by Simmons et al. (Simmons et al., 2019) in 1000 replicates. Presented majority rule consensus tree was computed by Phyutility (Smith & Dunn, 2008).

Tree plotting

All trees were plotted with ETE Toolkit version 3 (Huerta-Cepas et al., 2016) and FigTree (http://tree.bio.ed.ac.uk/software/figtree/) and adjusted in Inkscape (https://inkscape.org/nl/) or Vectornator (https://www.vectornator.io). All scripts used for phylogenetic analyses and for plotting trees are provided in ESM6.

Results

Matrix Construction

Our bioinformatic pipeline for ortholog clustering with strict paralogy filtering and decontamination resulted in a matrix with 49 taxa and 987 OGs that was 171,791 amino acid positions long, with 31% missing data (COMP). This set of OGs was further reduced to test the effects of analyzing subsets of the best OGs in terms of phylogenetic signal, resulting in a MARE-reduced matrix (MARER) of 202 OGs, as well as branch-length heterogeneity and compositional heterogeneity, resulting in matrices composed of the best 100, and 50 OGs according to these criteria (C100LBR, C50LBR, C100HR and C50HR matrices). Details on all matrices analyzed herein are presented in ESM4.

Comparison of Tree Reconstruction Approaches

Maximum Likelihood (ML)

ML analysis of the complete matrix (COMP) in RAxML, using the best-fitting model for each gene, recovered Mesozoa monophyletic with maximal support and sister to Gnathostomulida with strong bootstrap support (bs=98; Fig. S1 in ESM7). Overall, Lophotrochozoa was split into two subclades: 1) Trochozoa (Annelida, Brachiopoda, Phoronida, Mollusca, and Nemertea) and 2) a clade we refer provisionally to as 'Platyzoa *s.l.*,' which includes the platyzoan as well as polyzoan phyla. Polyzoa (Bryozoa + Entoprocta–Cycliophora) and Gnathifera (Gnathostomulida + Micrognathozoa–Rotifera) were not recovered monophyletic. Removal of distant outgroup taxa in favor of the short-branch ecdysozoan *Priapulus* did not have any effect on the resulting branching order and support values were comparable (Fig. S2 in ESM7). Thus, all outgroup taxa except *Priapulus* were excluded from most subsequent analyses to decrease computational demands.

To examine the effect of using a site-heterogeneous model in a ML framework, we analyzed the COMP matrix with the posterior mean site frequency model (PMSF; Wang et al., 2018) with 60 rate categories as

implemented in IQ-TREE 1.5.3 (Nguyen et al., 2015). The resulting tree topology (Fig. 2) slightly differed from that of the RAxML analysis: when the site-specific PMSF model was used, monophyletic Mesozoa was again maximally supported but recovered sister to Platyhelminthes (bs=90), with that clade sister to monophyletic Gnathifera (bs=99). For comparison of site specific models in ML and BI we also analyzed a smaller matrix (C50LBR that was used also for BI analyses) with PMSF model. The resulting tree recovered Mesozoa as two separate clades in Platyzoa *s.l.*: Orhonectida were sister to Platyhelminthes and Dicyemida sister to Gnathifera (albeit both with a low support: bs=49 and 54, respectively; Fig. S3 in ESM7).

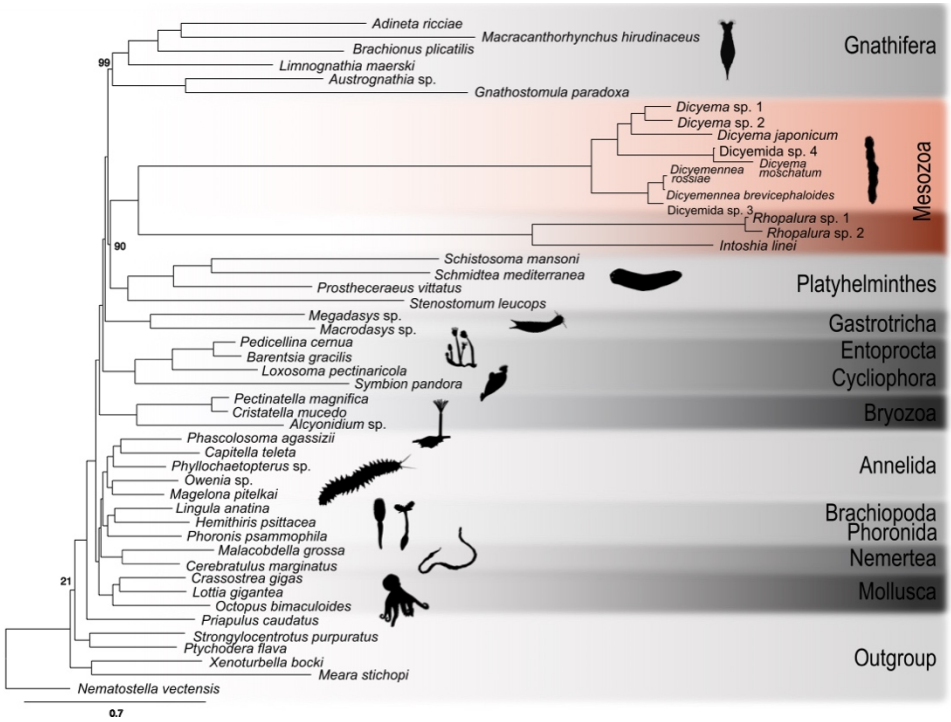


Figure 2. Tree representing mesozoan position in Lophotrochozoa based on 987 genes, computed in Maximum Likelihood framework with site specific model with 60 categories (IQtree c60 PMSF COMP), only bootstrap supports lower than 100 showed.

ML analysis of the MARE-reduced matrix (MARER; Fig. S4 in ESM7) also recovered Mesozoa as a monophyletic group (with maximal support) within paraphyletic Gnathifera, as the sister group of Gnathostomulida (bs=92). The high similarity of the MARER and COMPR topologies suggests that the reduced matrix represents well the whole dataset and could substitute it where the lower number of genes is essential to carry out the analysis.

In an attempt to reduce potential sources of systematic error, we assembled datasets based on reduced subsets of best 100 genes with the lowest branch heterogeneity (C100LBR; Fig. S5 in ESM7) and the lowest compositional heterogeneity (C100HR; Fig. S6 in ESM7), and then compare RAxML trees derived from both matrices. Results of both analyses recovered Mesozoa monophyletic within Platyzoa, either as a sister group of Platyhelminthes, with monophyletic Gnathifera sister to the Mesozoa–Platyhelminthes clade (C100LBR), or as a sister group of Gnathostomulida, and Platyhelminthes and Micrognathozoa–Rotifera as successive sister groups to the Mesozoa–Gnathostomulida clade (C100HR).

Bayesian Inference

As the site-heterogeneous CAT+GTR model has been suggested to be more robust against long-branch attraction artifacts than site-homogeneous models (Lartillot & Philippe, 2004; but see Whelan & Halanych, 2017), we performed Bayesian inference (BI) analyses with this model in Phylobayes. Matrices with reduced numbers of genes (C50LBR, C100LBR, C100HR, and M50HR) and just *Priapulus* as the outgroup were chosen for BI, due to the computational demands of the complex CAT+GTR model. In all of the analyses of C50LBR, the four chains converged on one solution, according to the Phylobayes bpcomp maxdiff statistic(Lartillot et al., 2013), but the other analyses failed to converge after > 30,000 generations (see ESM8 for details and resulting trees). The BI analysis of the C50LBR dataset (Fig.S7 in ESM7) recovered Mesozoa polyphyletic within monophyletic Platyzoa, with Dicyemida in a poorly-supported polytomy

with Gnathostomulida and Micrognathozoa–Rotifera, and Orthonectida sister to Platyhelminthes (with posterior probability, pp=0.98, a moderate support by Bayesian standards).

Comparison of ML and BI tree-building methods

In order to inspect differences between tree-building methods, we analyzed the C50LBR matrix (the only BI analysis that successfully converged) with or without Dicyemida and Orthonectida, respectively, by RAxML (or with similar setting in IQ-TREE with automatically selected gene models), IQ-TREE (PMSF model), and BI. Both ML trees with excluded mesozoans were identical (Fig. S8 and Fig S9 in ESM7): Platyhelminthes and Gnathifera formed a clade, and Entoprocta–Cycliophora, Bryozoa–Gastrotricha, Nemertea, Annelida–Brachiozoa, and Mollusca were successive sister groups to it (i.e., there were monophyletic Platyzoa *s.l.* within paraphyletic Trochozoa). On the contrary, in the BI tree excl. Mesozoa (Fig. S10 in ESM7) there were monophyletic Trochozoa and Platyzoa *s.l.*, both weakly supported, and within Platyzoa *s.l.*, a very weakly supported basalmost position of Gnathostomulida (i.e., polyphyletic Gnathifera). When only Orthonectida were included, they grouped with Platyhelminthes in all trees (Fig. S11, Fig. S12, and Fig. S13 in ESM7). Dicyemida alone grouped with Gnathifera (IQ-TREE, Fig. S14 in ESM7), with Gnathostomulida within monophyletic Gnathifera (RAxML, Fig. S15 in ESM7), or in an unresolved trichotomy with Gnathostomulida and Micrognathozoa–Rotifera (BI; Fig. S16 in ESM7). When both mesozoan groups were added to the analyses, they became diphyletic, both placed at the same positions as individual mesozoan clades alone (Dicyemida next to Gnathifera, or in polytomy with them, Orthonectida sister to Platyhelminthes; Fig. S3 and S7 in ESM7). Only in the RAxML tree the dicyemids were attracted towards orthonectids to form monophyletic Mesozoa sister to Platyhelminthes (Fig. S17 in ESM7).

Coalescent Approach

The coalescent tree (Fig. S18. in ESM7) recovered mesozoans monophyletic with maximal support. They were nested in polytomy with Gnathostomulida, Rotifera–Micrognathozoa and Platyhelminthes with relatively high support (bs=97). Overall, the coalescent tree showed a similar branching pattern as ML and BI analyses (Lophotrochozoa split into Platyzoa *s.l.* and Trochozoa). In this case, however, Bryozoa was not associated with Platyzoa *s.l.* as in ML and BI trees, but was recovered as an early branching clade of Trochozoa.

Taxon removal experiments

In order to examine the effect of taxon sampling on the position of long-branch clades, we performed a series of taxon removal experiments, based on the 202-gene MARER dataset, analyzed by both RAxML and IQ-TREE (Figures S19-S37 in ESM5). The IQ-TREE-based tree with both mesozoan groups excluded was split to Trochozoa and Platyzoa *s.l.*; the latter group included Platyhelminthes and Gnathifera as sister groups, followed by Gastrotricha, Entoprocta–Cycliophora, and Bryozoa. Orthonectida alone groups with Gnathostomulida (within Gnathifera), Dicyemida alone with Platyhelminthes (or as a sister group of Catenulida within Platyhelminthes; RAxML). When both mesozoan clades were present they formed monophyletic groups, sister to Gnathostomulida (i.e., at the same position as Orthonectida alone). However, all deeper nodes were weakly supported. Subsequent exclusion of Platyhelminthes, Gastrotricha, Rouphozoa, Gnathifera, Gnathifera + Platyhelminthes, Entoprocta–Cycliophora, Bryozoa, and Annelida had no effects on the tree topology. Only when Gnathostomulida were excluded, Mesozoa did not stay as a sister group of the rest of Gnathifera but transferred towards Platyhelminthes. On the contrary, removing the Micrognathozoa and Rotifera had no effect and Mesozoa remained close to the Gnatostomulida. Importantly, in all experiments Mesozoa remained a single clade within Platyzoa s.l. Only when all Platyzoa and Dicyemida were excluded (merely 17 spp. of Trochozoa and Orthonectida plus *Priapulus* remained), Orthonectida was

recovered as nested within Annelida, albeit with low support (Figure S34 in ESM5).

Gene-conflict analyses

The likelihood score of the constrained topology T2 (Orthonectida grouping with Annelida), compared by AU test, was significantly worse than of the unconstrained topology with monophyletic Mesozoa (T1) in the case of MARER matrix. Out of 202 genes in the data matrix, 148 support topology T1 and only 54 support the constrained topology T2 based on ΔGLS (Fig. S38 in ESM9). However, the RAxML tree based on genes supporting T1 includes, as usually, monophyletic Mesozoa within Gnathifera (as a sister group of Gnathostomulida), but the tree based on T2-supporting genes includes Dicyemida as a sister group of Catenulida within Platyhelminthes, Orthonectida sister to the Platyhelminthes–Dicyemida clade and monophyletic Gnathifera sister to Rouphozoa (the latter including both mesozoan groups). The T1-supporting genes are insignificantly longer than T2-compatible ones (mean 159 > 139); also taxonomic representativeness is comparable in both groups of genes (mean 36.5 : 34.5) but compositional heterogeneity is slightly higher in T1-supporting genes (T1 0.163, T2 0.149).

Discussion

Phylogenetic Relationships

Consistent with previous studies (Pawlowski et al., 1996; Lu et al., 2017; Schiffer et al., 2018), all our results confirm that both Orthonectida and Dicyemida are secondarily morphologically simplified bilaterians, not a primitive missing link between Protozoa and Metazoa as occasionally suggested earlier (Van Beneden, 1876; Czaker, 2000). Both Orthonectida and Dicyemida were recovered as monophyletic groups belonging to Lophotrochozoa in all analyses performed in this study. In general, our results represent a ‘traditional view’ on lophotrochozoan relationships with

small- and simple-bodied phyla forming a clade (Platyzoa *s.l.*) apart from Trochozoa. Importantly, this pattern was not affected by presence of mesozoans.

Dicyemids fundamentally differ from orthonectids by having no cuticle, nervous system, muscles, or true tissues at all. Furthermore, their life-histories are extremely derived yet quite different (host-parasite relationships, host spectrum, position of parasitic stage in the life cycle, etc.). The only potentially shared ultrastructural character, a unique type of ciliary rooting (Ax, 1996), has been challenged by the discovery of more diverse ciliary rootlets in Orthonectida (Slyusarev & Kristensen, 2003). Consequently, the monophyletic phylum Mesozoa has recently been doubted or rejected by most authors. However, the genomic data so far provided equivocal results (monophyly: Lu et al., 2017; polyphyly: Schiffer et al., 2018; Zverkov et al., 2019). Nevertheless, the old hypothesis about monophyly of Mesozoa seems to be (quite surprisingly) supported by most of our phylogenomic analyses. All ML and coalescent analyses of the complete 987-gene (COMP) and 202-gene (MARER) datasets (BI analyses did not converge) recovered Mesozoa monophyletic, usually with strong support, and close to Gnathifera or Gnathostomulida (COMP: RAxML, MARER: RAxML, IQ-TREE, coalescent), Platyhelminthes (COMP: IQ-TREE), or in the polytomy with both (COMP: coalescent). The same applied also to RAxML analysis of the C50LBR dataset: Mesozoa were found monophyletic and sister to Platyhelminthes.

Contrarily, the IQ-TREE (with PMSF model) and BI analyses of the C50LBR data matrix recovered Mesozoa split into two separate clades. Dicyemida was found either within Gnathifera in a poorly supported polytomy with Gnathostomulida and Micrognathozoa–Rotifera (BI, with only very low support pp=0.62), or as a sister group of Gnathifera (IQ-TREE, again with very low support bs=54). The position of Orthonectida as a sister group of Platyhelminthes was relatively well supported in the BI tree (pp=0.98) but not in the IQ-TREE one (bs=49). However, both gene subsampling strategies (branch length homogeneity and compositional

homogeneity), used in an attempt to alleviate LBA, were not able to disintegrate monophyletic Mesozoa or to reduce its strong support.

The possibility that Mesozoa are two separate groups attracted to each other due to long-branch artifact was tested by including just one mesozoan group in absence of the other. Both Dicyemida and Orthonectida consistently grouped within the Gnathifera–Platyhelminthes clade. The taxon-removal analyses suggest that Gnathostomulida (but not Micrognathozoa and Rotifera) is the only attractor strong enough to change position of mesozoans from Platyhelminthes to Gnathifera. Thus, we can assume that gnathiferan or even within-gnathiferan affinities of Mesozoa are likely artificial, and the alternative, i.e., close relation between Mesozoa and Platyhelminthes, is more probable. The hypothesis considering Dicyemida and Orthonectida as extremely simplified parasitic flatworms (Neodermata, the group including flukes and tapeworms) has actually been considered for a long time (see Whitman, 1883), based on putative similarities between trematode miracidia and mesozoan infective larvae and more recently on the ciliary rootlet system putatively shared by Neodermata and Mesozoa (but see above). On the contrary, Mesozoa have the orthodox mitochondrial genetic code rather than the apomorphic code characteristic for Rhabditophora (that includes Neodermata), which rejects the idea that they might be derived flatworms (Telford et al., 2000). Even if Dicyemida were occasionally recovered as platyhelminths here (T2 tree in ΔGLS analysis, or IQ-TREE analysis of the MARER dataset with orthonectids excluded), they were a sister group of Catenulida, not an in-group of Rhabditophora.

The instability of position of Orthonectida was evident in published studies. Mikhailov et al. (2016) sequenced the genome of Intoshia linei and recovered Orthonectida as a sister group of Platyhelminthes in Bayesian analysis with the CAT+GTR model as well as in ML analysis, yet the use of CAT model with a flat rate in the Bayesian analysis resulted in Orthonectida sister to Annelida. With addition of a dicyemid genome, Lu et al. (2017) recovered Orthonectida (together with Dicyemida in

monophyletic Mesozoa) as close relatives of Gastrotricha and/or Platyhelminthes, while Schiffer et al. (2018) and Zverkov et al. (2019) found them as early-branching annelids, unrelated to Dicyemida, and Bondarenko et al. (2019) even as a sister group of leeches (Annelida: Pleistoannelida: Clitellata).

The annelid affinities of the Orthonectida, based on morphology, have been proposed previously, starting with Metschnikoff (1881). In more recent studies, it has been hypothesized on the basis of annelid-like microvilli-formed cuticle in the free-living stages of Orthonectida (Slyusarev & Kristensen, 2003). Orthonectids (contra dicyemids) also have a true epidermis (but their epidermal basal lamina is reduced or absent). Possible trace of an ancestral segmented body plan is seen as the series of regularly spaced circular muscles along the anterior-posterior axis of Orthonectida (along with repeated bands of cilia and paired serotonin-like immunoreactive nerve cells; Slyusarev et al., 2022). However, orthonectids use cilia for locomotion, and their muscle system is reduced to serve almost exclusively for copulation and hatching of larvae. The annelid muscular pattern consists of two muscle layers, external circular/transverse and internal longitudinal. In Orthonectida, the circular muscles are situated inversely, inside the longitudinal muscles. The only annelid with external longitudinal muscles sunken into the epidermis is non-segmented meiofaunal *Lobatocerebrum* (Kerbl et al., 2015), a member of Dinophiliformia (Martín-Durán et al., 2021). The possible relationships of Orthonectida and basal annelids are compatible with the results published by Schiffer et al. (Schiffer et al., 2018) who found in the Orthonectida a short stretch of mitochondrial genes (nad1, nad6, cob) in the same order as in *Owenia* but not as in the pleistoannelids.

In our study, the scenario of Orthonectida related or belonging to Annelida was recovered only in one unconverged BI analysis (C100HR matrix) and in a taxon removal experiment where no dicyemid and platyzoan taxa were present, both results being extremely problematic to interpret. The AU test showed that topology with orthonectids related to annelids is significantly

worse, based on our data, than topology with monophyletic Mesozoa within Platyzoa. Furthermore, ΔGLS analysis revealed that monophyly of Mesozoa is supported by more genes than its polyphyly with annelid affinities of Orthonectida (148:54). Moreover, even in the genes compatible with the annelid hypothesis, the 'annelid signal' is minor, and the resulting tree included platyzoan Mesozoa, albeit diphyletic.

Different methods, different views

Two main approaches are typically used in analyzing phylogenomic matrices, Bayesian inference and maximum likelihood, which have a wide, but not fully overlapping range of available models of sequence evolution. Site-specific frequency models (or empirical profile mixture models) based on modeling of each alignment site individually are considered to better model sequence evolution and therefore be more reliable in obtaining correct phylogeny in difficult cases (Lartillot et al., 2007; Roure et al., 2013). Until recently, the site-specific frequency model (CAT; Si Quang et al., 2008) was available only in the Bayesian framework, but an approximation of such a model has recently been implemented in IQ-TREE in the ML framework as a posterior mean site frequency model (PMSF; Wang et al., 2018).

In the case of Mesozoa, almost all maximum likelihood-based methods, regardless of the number of genes used, recovered a similar topology; contrarily, Bayesian analyses tended to recover a different topology. Based on the 50-gene matrix (C50LBR, the only matrix that converged in BI), both BI and IQ-TREE (with PMSF model) analyses recovered topology with diphyletic Mesozoa (Dicyemida close to Gnathifera, Orthonectida to Platyhelminthes), while in the RAxML tree Mesozoa is monophyletic and sister to Platyhelminthes. It could suggest that site specific model is a key for recovery of this specific topology in this 50 gene matrix. However, when the same model was used in ML on bigger matrices, it again consistently showed Mesozoa monophyletic.

Similar inconsistency was recovered by Zverkov et al. (2019) who presented BI phylogeny with polyphyletic Mesozoa and paraphyletic Platyzoa but their ML trees are similar to ours: they recovered monophyletic Mesozoa within monophyletic Platyzoa in both RAxML and IQ-TREE topology (Supplementary Figures S5 and S4, respectively). Both Zverkov et al.'s and the present analyses suggest a consistent conflict between ML and BI topologies. Surprisingly, ML methods with both classic and site-specific PMSF models showed similar scenarios.

The inference of relationships among lophotrochozoan clades is thought to be hindered by the effect of LBA due to (sometimes dramatic) differences in molecular evolutionary rate, the application of inaccurate models of sequence evolution, and by the inherent difficulty of resolving ancient radiations because of short stem branches in deep nodes (Budd & Jackson, 2016). Using genes with specific characteristics (branch length homogeneity, compositional homogeneity) has been suggested to be more suited for phylogenetic inference of taxa with uneven branch lengths (Struck, 2013; Shen et al., 2016). However, targeting genes with these specific qualities did not have a clear effect in the case of mesozoan phylogeny, similarly as in an earlier phylogenomic analysis of Lophotrochozoa (Kocot et al., 2017). Also, in our case, the phylogenetic position of the long-branch mesozoan groups was not affected by taxon-removal experiments. This suggests that either LBA is not the main player causing significant errors in the case of lophotrochozoan and mesozoan phylogeny, or that LBA in the case of Mesozoa is too strong to be overcome by any of the currently used methods.

Conclusions

After summarizing all the results of different analyses, the old hypothesis on monophyletic Mesozoa is possibly reborn and should be further tested. The results of our phylogenomic analyses can be summarized as follows:

(i) both Dicyemida and Orthonectida are secondarily modified lophotrochozoans;

(ii) both Dicyemida and Orthonectida are closely related either to Platyhelminthes or to Gnathifera, the latter position could be caused by long-branch attraction towards Gnathostomulida as indicated by taxon-removal analyses;

(iii) monophyly of Mesozoa is corroborated by most analyses.

Data availability

Data are available in Electronic Supplemental Material and on https://drive.google.com/drive/folders/1YolJxNRV_2ioVpG02xgcJcQMDjWIaRwp?usp=sharing

Funding

This work was supported by the Czech Science Foundation (GACR grant nos. 15-08717S to JZ and MD, and 19-28399X to JŠ) and, in part, by the United States National Science Foundation (grants 1146575, 1557923, 1548121 and 1645219 to LLM and 1846174 to KMK).

Acknowledgements

We would like to thank colleagues and institutions who let us use their facilities for sample processing, namely Graziano Fiorito and Giovanna Ponte (Stazione Zoologica Anton Dohrn, Naples, Italy); Andrea De Lucia, Andrea Camedda and Stefano Marra (Institute for Coastal Marine Environment, CNR, Oristano, Sardinia, Italy) and we acknowledge the Laboratory of Microscopy and Histology of the Biology Centre CAS for technical help with the guiding through immunohistochemical visualization and for allowing the use of their facilities. We thank Deb Crocker and Robert Griffin for support of the University of Alabama High-Performance Computing cluster (UAHPC). Computational resources were also supplied by the project ‘e-Infrastruktura CZ’ (e-INFRA LM2018140)

provided within the program Projects of Large Research, Development and Innovations Infrastructures, as well as the National Science Foundation (CNS-1725797) administered by UCSB's Center for Scientific Computing (CSC).

References

Altschul, S.F. *et al.* (1990) 'Basic local alignment search tool.', *Journal of molecular biology*, 215(3), pp. 403–410. doi:10.1016/S0022-2836(05)80360-2.

Ax, P. (1996) *Multicellular animals: a new approach to the phylogenetic order in Nature, vol 1.* Berlin–Heidelberg: Springer Vrlg.

Ballesteros, J.A. and Sharma, P.P. (2019) 'A Critical Appraisal of the Placement of Xiphosura (Chelicerata) with Account of Known Sources of Phylogenetic Error', *Systematic Biology*, 68(6), pp. 896–917. doi:10.1093/sysbio/syz011.

Bergsten Johannes (2005) 'A review of long-branch attraction', *Cladistics*, 21(2), pp. 163–193.

Bleidorn, C. (2019) 'Recent progress in reconstructing lophotrochozoan (spiralian) phylogeny', *Organisms Diversity & Evolution*, 19(4), pp. 557–566. doi:10.1007/s13127-019-00412-4.

Bondarenko, N. *et al.* (2019) 'Comparative analysis of the mitochondrial genomes of Orthonectida: insights into the evolution of an invertebrate parasite species', *Molecular Genetics and Genomics*, 294(3), pp. 715–727. doi:10.1007/s00438-019-01543-1.

Budd, G.E. and Jackson, I.S.C. (2016) 'Ecological innovations in the Cambrian and the origins of the crown group phyla', *Philosophical transactions of royal society B*, 371(1685), p. 20150287.

Catalano, S.R. (2012) 'A review of the families, genera and species of Dicyemida Van Beneden, 1876', *Zootaxa*, 32(3479), pp. 1–32. doi:10.11646/zootaxa.3646.1.11.

Czaker, R. (2000) ‘Extracellular matrix (ECM) components in a very primitive multicellular animal, the dicyemid mesozoan Kantharella antarctica’, *Anatomical Record*, 259(1), pp. 52–59. doi:10.1002/(SICI)1097-0185(20000501)259:1<52::AID-AR6>3.0.CO;2-J.

Deheyn, D., Watson, N.A. and Jangoux, M. (1998) ‘Symbioses in Amphipholis squamata (Echinodermata, Ophiuroidea, Amphiuridae): Geographical variation of infestation and effect of symbionts on the host’s light production’, *International Journal for Parasitology*, 28(9), pp. 1413–1424. doi:10.1016/S0020-7519(98)00119-2.

Dunn, C.W. *et al.* (2014) ‘Animal Phylogeny and Its Evolutionary Implications’, *Annu. Rev. Ecol. Evol. Syst*, 45, pp. 371–395. doi:10.1146/annurev-ecolsys-120213-091627.

Ebersberger, I., Strauss, S. and von Haeseler, A. (2009) ‘HaMStR: Profile hidden markov model based search for orthologs in ESTs’, *BMC Evolutionary Biology*, 9(1), p. 157. doi:10.1186/1471-2148-9-157.

Felsenstein J (1978) ‘Cases in which parsimony or compatibility methods will be positively misleading’, *Systematic zoology*, (27), pp. 401–410.

Giribet, G. and Edgecombe, G.D. (2017) ‘Current Understanding of Ecdysozoa and its Internal Phylogenetic Relationships’, *Integrative and Comparative Biology*, 57(3), pp. 455–466. doi:10.1093/icb/icx072.

Halanych, K.M. (2016) ‘How our view of animal phylogeny was reshaped by molecular approaches: lessons learned’, *Organisms Diversity & Evolution*, 16(2), pp. 319–328. doi:10.1007/s13127-016-0264-8.

Haraguchi, Y. and Sasaki, A. (1996) ‘Host–Parasite Arms Race in Mutation Modifications: Indefinite Escalation Despite a Heavy Load?’, *Journal of Theoretical Biology*, 183(2), pp. 121–137. doi:10.1006/jtbi.1996.9999.

Huerta-Cepas, J., Serra, F. and Bork, P. (2016) ‘ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data’, *Molecular Biology and Evolution*, 33(6), pp. 1635–1638. doi:10.1093/molbev/msw046.

Huson, D.H. *et al.* (2016) ‘MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data’, *PLoS Computational Biology*, 12(6), pp. 1–12. doi:10.1371/journal.pcbi.1004957.

Kapli, P., Yang, Z. and Telford, M.J. (2020) 'Phylogenetic tree building in the genomic age', *Nature Reviews Genetics*, 21, pp. 428–444. doi:10.1038/s41576-020-0233-0.

Katoh, K. *et al.* (2005) 'MAFFT version 5: improvement in accuracy of multiple sequence alignment', *Nucleic Acids Research*, 33(2), pp. 511–518. doi:10.1093/nar/gki198.

Kerbl, A. *et al.* (2015) 'Detailed reconstruction of the nervous and muscular system of Lobatocerebridae with an evaluation of its annelid affinity.', *BMC evolutionary biology*, 15, p. 277. doi:10.1186/s12862-015-0531-x.

Kocot, K.M. *et al.* (2013) 'PhyloTreePruner: A Phylogenetic Tree-Based Approach for selection of Orthologous sequences for phylogenomics', *Evolutionary Bioinformatics*, 9, p. EBO.S12813. doi:10.4137/EBO.S12813.

Kocot, K.M. (2016) 'On 20 years of Lophotrochozoa', *Organisms Diversity & Evolution*, 16(2), pp. 329–343. doi:10.1007/s13127-015-0261-3.

Kocot, K.M. *et al.* (2017) 'Phylogenomics of lophotrochozoa with consideration of systematic error', *Systematic Biology*, 66(2), pp. 256–282. doi:10.1093/sysbio/syw079.

Kück, P. and Struck, T.H. (2014) 'BaCoCa – A heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions', *Molecular Phylogenetics and Evolution*, 70, pp. 94–98. doi:10.1016/j.ympev.2013.09.011.

Lartillot, N. *et al.* (2013) 'PhyloBayes MPI: Phylogenetic Reconstruction with Infinite Mixtures of Profiles in a Parallel Environment', *Systematic Biology*, 62(4), pp. 611–615. doi:10.1093/sysbio/syt022.

Lartillot, N., Brinkmann, H. and Philippe, H. (2007) 'Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model', *BMC Evolutionary Biology*, 7(Suppl 1), p. S4. doi:10.1186/1471-2148-7-S1-S4.

Lartillot, N. and Philippe, H. (2004) 'A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process', *Molecular Biology and Evolution*, 21(6), pp. 1095–1109. doi:10.1093/molbev/msh112.

Laumer, C.E. *et al.* (2019) 'Revisiting metazoan phylogeny with genomic sampling of all phyla', *Proceedings of the Royal Society B: Biological Sciences*, 286(1906). doi:10.1098/rspb.2019.0831.

Liu, L. *et al.* (2009) 'Coalescent methods for estimating phylogenetic trees', *Molecular Phylogenetics and Evolution*, 53(1), pp. 320–328. doi:10.1016/j.ympev.2009.05.033.

Lu, T.M. *et al.* (2017) 'The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution', *Zoological Letters*, 3(1), pp. 1–9. doi:10.1186/s40851-017-0068-5.

Marlétaz, F. *et al.* (2019) 'A New Spiralian Phylogeny Places the Enigmatic Arrow Worms among Gnathiferans', *Current Biology*, 29(2), pp. 312-318.e3. doi:10.1016/j.cub.2018.11.042.

Martín-Durán, J.M. *et al.* (2021) 'Conservative route to genome compaction in a miniature annelid', *Nature Ecology & Evolution*, 5(2), pp. 231–242. doi:10.1038/s41559-020-01327-6.

Metschnikoff, E. (1881) 'Untersuchunger uber Orthonectiden', *Zeitschrift für wissenschaftliche Zoologie*, 32, pp. 349–387.

Meyer, B., Meusemann, K. and Misof, B. (2011) *MARE v0.1.2-rc.xxx*.

Mikhailov, K.V. *et al.* (2016) 'The Genome of Intoshia linei Affirms Orthonectids as Highly Simplified Spiralians', *Current Biology*, 26(13), pp. 1768–1774. doi:10.1016/j.cub.2016.05.007.

Mirarab, S. and Warnow, T. (2015) 'ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes', *Bioinformatics*, 31(12), pp. i44–i52. doi:10.1093/bioinformatics/btv234.

Nesnidal, M.P. *et al.* (2010) 'Compositional Heterogeneity and Phylogenomic Inference of Metazoan Relationships', *Molecular Biology and Evolution*, 27(9), pp. 2095–2104. doi:10.1093/molbev/msq097.

Nguyen, L.T. *et al.* (2015) 'IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies', *Molecular Biology and Evolution*, 32(1), pp. 268–274. doi:10.1093/molbev/msu300.

Pawlowski, J. *et al.* (1996) 'Origin of the Mesozoa inferred from 18S rRNA gene sequences', *Molecular biology and evolution*, 13(8), pp. 1128–32. doi:10.1093/oxfordjournals.molbev.a025675.

Philippe, H. *et al.* (2017) 'Pitfalls in supermatrix phylogenomics', *European Journal of Taxonomy*, (283), pp. 1–25. doi:10.5852/ejt.2017.283.

Price, M.N., Dehal, P.S. and Arkin, A.P. (2010) 'FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments', *PLoS ONE*. Edited by A.F.Y. Poon, 5(3), p. e9490. doi:10.1371/journal.pone.0009490.

Qun, Y. *et al.* (2007) 'Phylochronology of early metazoans: combined evidence from molecular and fossil data', *Geological Journal*, 42, pp. 281–295. doi:10.1002/gj.

Rokas Antonis, Krüger Dirk, and Carroll Sean B. (2005) 'Animal Evolution and the Molecular Signature of Radiations Compressed in Time', *Science*, 310(5756), pp. 1933–1938. doi:10.1126/science.1116759.

Roure, B., Baurain, D. and Philippe, H. (2013) 'Impact of missing data on phylogenies inferred from empirical phylogenomic data sets', *Molecular Biology and Evolution*, 30(1), pp. 197–214. doi:10.1093/molbev/mss208.

Schiffer, P.H., Robertson, H.E. and Telford, M.J. (2018) 'Orthonectids Are Highly Degenerate Annelid Worms', *Current Biology*, 28(12), pp. 1970-1974.e3. doi:10.1016/j.cub.2018.04.088.

Shen, X.-X., Hittinger, C.T. and Rokas, A. (2017) 'Contentious relationships in phylogenomic studies can be driven by a handful of genes', *Nature Ecology & Evolution*, 1(5), p. 0126. doi:10.1038/s41559-017-0126.

Shen, X.-X., Salichos, L. and Rokas, A. (2016) 'A Genome-Scale Investigation of How Sequence, Function, and Tree-Based Gene Properties Influence Phylogenetic Inference', *Genome Biology and Evolution*, 8(8), pp. 2565–2580. doi:10.1093/gbe/evw179.

Si Quang, L., Gascuel, O. and Lartillot, N. (2008) 'Empirical profile mixture models for phylogenetic reconstruction', *Bioinformatics*, 24(20), pp. 2317–2323. doi:10.1093/bioinformatics/btn445.

Simmons, M.P. *et al.* (2019) 'Gene-wise resampling outperforms site-wise resampling in phylogenetic coalescence analyses', *Molecular Phylogenetics and Evolution*, 131(March 2018), pp. 80–92. doi:10.1016/j.ympev.2018.10.001.

Slyusarev, G.S. *et al.* (2022) 'The structure of the muscular and nervous systems of the orthonectid Rhopalura litoralis (Orthonectida) or what parasitism can do to an annelid', *Organisms Diversity & Evolution*, 22(1), pp. 35–45. doi:10.1007/s13127-021-00519-7.

Slyusarev, G.S. and Kristensen, R. (2003) 'Fine structure of the ciliated cells and ciliary rootlets of Intoshia variabili (Orthonectida)', *Zoomorphology*, 122(1), pp. 33–39. doi:10.1007/s00435-002-0065-9.

Smith, S.A. and Dunn, C.W. (2008) 'Phyutility: a phyloinformatics tool for trees, alignments and molecular data', *Bioinformatics*, 24(5), pp. 715–716. doi:10.1093/bioinformatics/btm619.

Stamatakis, A. (2014) 'RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics*, 30(9), pp. 1312–1313. doi:10.1093/bioinformatics/btu033.

Struck, T.H. (2013) 'The Impact of Paralogy on Phylogenomic Studies – A Case Study on Annelid Relationships', *PLoS ONE*, 8(5), p. e62892. doi:10.1371/journal.pone.0062892.

Struck, T.H. (2014) 'Trespex-detection of misleading signal in phylogenetic reconstructions based on tree information', *Evolutionary Bioinformatics*, 10, pp. 51–67. doi:10.4137/EBo.s14239.

Telford, M.J. *et al.* (2000) 'Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms.', *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), pp. 11359–64. doi:10.1073/pnas.97.21.11359.

Thomas, J.A. *et al.* (2010) 'A Generation Time Effect on the Rate of Molecular Evolution in Invertebrates', *Molecular Biology and Evolution*, 27(5), pp. 1173–1180. doi:10.1093/molbev/msq009.

Van Beneden, M.E. (1876) 'Recherches sur le Dicyemides, survivants actuels d'un embranchement des Mésozoaires', *Bulletins de l'Académie royale des sciences, des lettres et beaux-arts de belgique*, Bruxelles, pp. 1160–1250.

Wang, H.-C. *et al.* (2018) 'Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation', *Systematic Biology*, 67(2), pp. 216–235. doi:10.1093/sysbio/syx068.

Whelan, N.V. and Halanych, K.M. (2017) 'Who Let the CAT Out of the Bag? Accurately Dealing with Substitutional Heterogeneity in Phylogenomic Analyses', *Systematic Biology*, 66(2), pp. 232–255. doi:10.1093/sysbio/syw084.

Whitman C.O. 1883. A contribution to the embryology, life-history and classification of the dicyemids. Mittheilungen Zool. Station Napels 4:1–90.

WoRMS Editorial Board 2018: van der Land, Furuya and Decock (2018) 'World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2018-05-18.', p. Accessed 2018-05-18. doi:doi:10.14284/170.

Wu, M., Chatterji, S. and Eisen, J.A. (2012) 'Accounting For Alignment Uncertainty in Phylogenomics', *PLoS ONE*. Edited by M. Salemi, 7(1), p. e30288. doi:10.1371/journal.pone.0030288.

Zverkov, O.A. *et al.* (2019) 'Dicyemida and Orthonectida: Two stories of body plan simplification', *Frontiers in Genetics*, 10(MAY). doi:10.3389/fgene.2019.00443.

CHAPTER II.

Population co-divergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea

Scientific Reports (2019) 9, 14300
https://doi.org/10.1038/s41598-019-50555-9

Population co-divergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea

Drábková Marie*[1,2], Nikola Jachníková[1], Tomáš Tyml[1,3], Hana Sehadová[2], Oleg Ditrich[1], Eva Myšková[1], Václav Hypša[1], Jan Štefka[1,2]

[1] Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic
[2] Biology Centre, CAS, v.v.i., České Budějovice, Czech Republic Budejovice, Czech Republic
[3] Faculty of Science, Masaryk University, Brno, Czech Republic

* corresponding author

Abstract

Population structure and biogeography of marine organisms are formed by different drivers than in terrestrial organisms. Yet, very little information is available even for common marine organisms and even less for their associated parasites. Here we report the first analysis of population structure of both a cephalopod host (*Sepia officinalis*) and its dicyemid parasite, based on a homologous molecular marker (cytochrome oxidase I). We show that the population of common cuttlefish in the Mediterranean area is fragmented into subpopulations, with some areas featuring restricted level of gene flow. Amongst the studied areas, Sardinia was genetically the most diverse and Cyprus the most isolated. At a larger scale, across the Mediterranean, the population structure of the parasite shows co-diversification pattern with its host, but a slower rate of diversification. Differences between the two counterparts are more obvious at a finer scale, where parasite populations show increased level of fragmentation and lower local diversities. This discrepancy can be caused by local extinctions and replacements taking place more frequently in the dicyemid populations, due to their parasitic lifestyle.

Introduction

In marine organisms, genetic structure is usually supposed to be determined by various extrinsic factors unique to this environment. Among the most typical factors are the lack (or rare occurrence) of obvious dispersal barriers or boundaries compared to terrestrial systems (Palumbi, 1994). The lack of dispersal barriers should, in theory, lead to the maintenance of large effective populations sizes, spanning vast areas of suitable habitats and showing low level of inter-population genetic variation (Palumbi, 1992). However, while empirical data confirmed this view for some organisms (i.e., *Architeuthis dux*, *Homarus gammarus*, *Thunnus alalunga*; Winkelmann et al., 2013, Watson et al., 2016, Laconcha et al., 2015, respectively), others showed surprisingly high diversification on a smaller scale than would be expected (i.e., Dinoflagellate *Alexandrium minutum*, *Sepia esculenta*; Casabianca et al., 2012, Zheng et al., 2009, respectively; for more examples see Palumbi 1994). Palumbi (1994) listed several factors possibly responsible for such diversification. They include biological traits as well as physical barriers (mainly ocean/sea currents). Since then, many studies were carried out on a broad taxonomic range of marine organisms in different oceans and seas, revealing a high variety of reconstructed genetic patterns and their relationships to the oceanographic conditions, showing that genetically homogeneous populations are not the only option in marine organisms (Hauser & Carvalho, 2008).

Mediterranean Sea, with its extremely rich biodiversity and a long history of research interest, belongs among the best mapped marine regions. As a consequence, the oceanographic processes (currents and discontinuities) are well known (Robinson et al., 2001, Ayata et al., 2018) and their possible influence on population genetic connectivity has been investigated for many organisms (e.g. reviews focused on fish and benthic invertebrates, Dalongeville et al., 2016, Pascual et al., 2017). Mediterranean Sea has been traditionally divided into Western and Eastern Basins (here shortened as WB and EB), connected by the Strait of Sicily and possibly also by the Strait of Messina. Water currents in the WB are defined by two major oceanographic fronts, the Almeria-Oran front (AO front) near Gibraltar and

the North Balearic front (NB front) near Balearic islands. In the EB, the two major barriers are represented by the Otranto Strait, separating Adriatic Sea, and the Greek islands, forming the Aegean front (Robinson et al., 2001, Ayata et al., 2018). In a recent meta-analysis Pascual et al. (2017) showed that the relationship between population structure and oceanographic features in Mediterranean Sea varies considerably across different species and is largely determined by the life history of the given organism. Particularly, the presence/absence of a pelagic larva, and duration of this phase, is an important factor in the dispersal capacity and therefore determines population structure. For example, the organisms with low dispersal capabilities show significant genetic differentiation, but their population structure is not determined by the oceanographic fronts (Pascual et al., 2017). High variety of the reconstructed genetic patterns for different organisms shows that understanding genetic differentiation and gene flow in marine conditions will require number of genetic studies on a rich variety of biologically different organisms.

Considering strong dependence of parasites on their hosts, and obvious role of the host in parasites' dispersal, the factors driving population diversity of free-living marine organisms should, in turn, affect the genetic diversity and structure of their parasites. However, the degree of such interdependency is yet unclear. While in the terrestrial systems, comparative studies on both host and parasite population structures are more common (e.g. Levin &Parker, 2013, Martin; et al., 2018), marine surveys are limited to a few studies, usually involving complex multi-host systems (Valdivia et al., 2014), or parasites of sessile hosts, lacking opportunities for co-dispersal (Lane et al., 2018). The studies performed so far indicate that the answer to this question is likely to be dependent on the particular model and its biological traits. For example, Blasco-Costa and Poulin (2013), reviewing studies on 16 trematode species, concluded that the host mobility is the main determinant of the parasite genetic diversification. A slightly different view was presented by Maze-Guilmo et al. (2016). Based on their meta-analysis of a broader spectrum of parasites, they demonstrated that the outcome of such parasite-host comparison is

dependent on various biological traits (i.e. reproduction mode, presence/absence of a larval stage). For example, they suggest that, generally, parasites tend to show lower genetic differentiation than the host, particularly in hermaphroditic parasites with asexual reproduction phase, whereas for gonochoristic groups the genetic differentiation is often the same or even higher than in the host. Maze-Guilmo et al. (2016) explain this discrepancy by a lower number of dispersal events required for successful host colonization by hermaphroditic parasites.

Analysing the impact of host population structure in multi-host systems, such as in digenean parasites, is complicated by the presence of intricate networks often including migratory or terrestrial hosts (Keeney et al., 2009). On the contrary, comparative studies of single-host-parasite associations provide a more straightforward approach, with the capacity to address such questions as: Is the overall population structure of a parasite mirroring that of its host due to their shared dispersal? May the structure of parasite's local subpopulation differ from the host due to local extinctions and reinfections? Would such extinctions result in a decrease of genetic diversities within local populations of the parasites? Here, we address these questions using a model of host-parasite pair, Common cuttlefish (*Sepia officinalis*) and its parasites from a rarely studied group Dicyemida. This model allows for addressing the general issues of genetic diversification in marine environment but also specifically the relationship between diversities of the host and parasite. To our best knowledge, this study introduces the first single-host model entirely bound to the marine environment and involving a free-living host.

Common cuttlefish (*S. officinalis*) is an important species for fisheries in the Mediterranean Sea and the neighbouring Atlantic coast, which makes it one of the few marine organisms for which considerable amount of data is available. Considering the span and connectivity of the suitable habitats, together with the mobility of adults, cuttlefish could in theory maintain large continuous population, with isolation by distance as a main pattern of the genetic structure. However, the known biological features of the

cuttlefish suggest that more complex population structure, affected by other factors than mere geographic distance, might be expected. Cuttlefish is a benthic cephalopod, that, unlike most octopuses and squids, does not possess an obvious pelagic early phase of life (e.g. larvae) as a mean of dispersal. Moreover, its lifespan of only one to two years suggests a quick turnover of the residing population. These life history traits may predispose cuttlefish to formation of fragmented populations (e.g. a network of relatively isolated subpopulations with poor genetic exchange). The few studies carried out for this species in general support this view, indicating substantial level of structuring across the whole Mediterranean Sea and the effect of Isolation By Distance (IBD). However, they do not provide any conclusive view, since based on the used molecular marker, they provide slightly different pictures. The first indication of significant subpopulation structuring in this organism came from the allozyme analysis carried out by Perez-Losada et al. (1999), which revealed clinal changes of the analyzed allozymes between Mediterranean and Atlantic localities but did not indicate any clear subdivision associated with any oceanic front in the Mediterranean. A significant increase in the genetic distances across the AO front was indicated by the following microsatellite study, which, however, covered only the coasts of the Iberian peninsula (Perez-Losada et al., 2002). The third study, using mitochondrial gene for COI, brought the most complex analysis covering a large area from Greece to the Atlantic coast of Portugal. It confirmed the general picture of a highly fragmented population with restricted gene flow and possible effect of some oceanographic factors (Perez-Losada et al., 2007).

Even less information is available on intraspecific genetic diversity of the organisms associated with the cuttlefish as a host. Perhaps the closest topic addressed for the Mediterranean cephalopods is the study of symbiotic *Vibrio* in sepiolid squids, showing independence of the host and symbiont genetic structures (Zamborsky & Nishiguchi, 2011). In dicyemids, their life history parameters indicate much higher degree of host-dependence. Dicyemids are endogenous parasites only found in the renal organs of benthic cephalopods, such as cuttlefish or octopus, and they are completely

dependent on their host for growth and reproduction (Lapan & Morowitz, 1975). Both asexual and sexual reproduction take place inside the host, whereas only a short-lived larva is expected to serve as a means of dispersal within the host population (Furuya & Tsuneki, 2003). This dependency of dicyemids on their host indicates that their inter-host transfer can only occur on short distances and their population structure will thus be strongly determined by the host. More specifically, we expect that their dispersion across long distances is driven entirely by the host and the gross genetic structure of the dicyemids will thus correspond to that of the cuttlefish. However, since they are spreading between hosts by environment (possibly by a free swiming larva, that hovers near sea floor) and do not transmit vertically (see Catalano et al., 2013) their population structure could deviate from the host's pattern within local populations. Almost no information is currently available on dicyemid population structure or genetic diversity to test these questions. Up to now, a little more than 120 species of dicyemids have been described (Catalano, 2013), whereas only a few 18S rDNA sequences from dicyemids associated with eastern Pacific cephalopods (Eshragh & Leander, 2015) and from *Dicyemennea eledones* associated with *Eledone cirrhosa* in the Mediterranean (Souidenne et al., 2016) have been published. Despite the fact, that dicyemids were also suggested as suitable tags in phylogeographic parasitological research (Catalano et al., 2014, Catalano et al., 2013), only several sequences usable for population studies are publicly available in GenBank, but has not been published (such as the sequences of dicyemids from Australian cephalopods; Catalano 2013 unpublished).

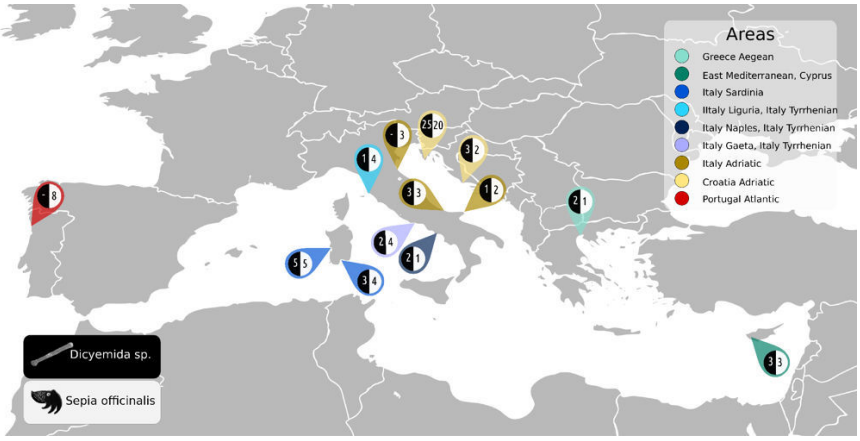


Figure 1. Map of sampling area showing the number of host specimens and parasite samples from each locality. Dicyemids white font, black background, their hosts common cuttlefish (*Sepia officinalis*) black font white background. The colours of the pins correspond to population network in Fig. 2.

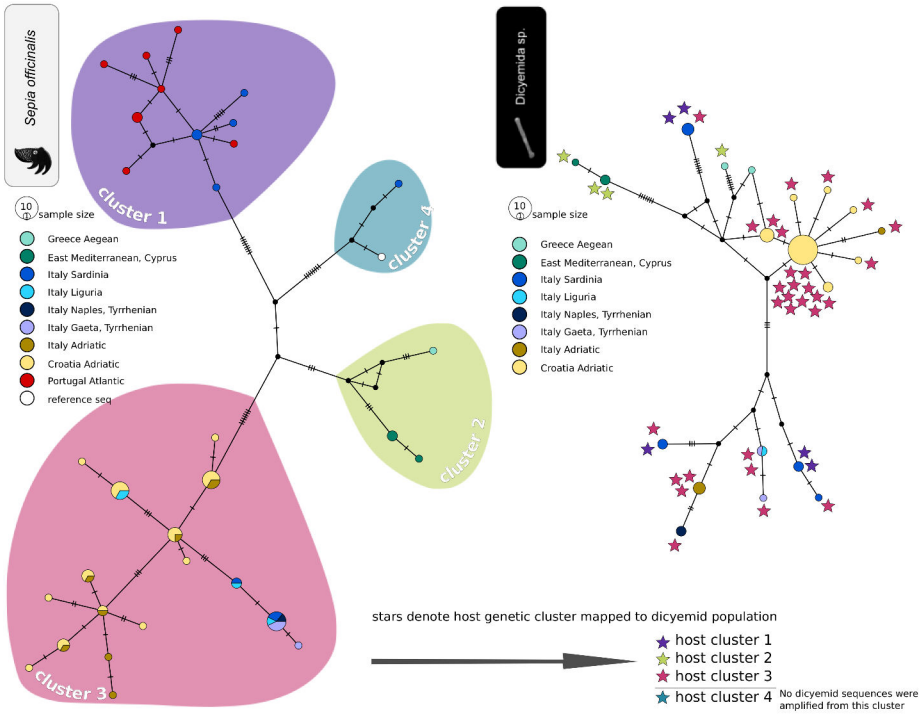


Figure 2. Population net (PopART, median joining network) of the host, *Sepia officinalis* (on the left), and its associated dicyemids (on the right). Stars denote the host cluster that dicyemids were associated with.

Material and methods

Samples were collected at 13 localities in the in the Mediterranean area and the Atlantic coast of Portugal (Fig. 1) during multiple sampling trips from 2014 to 2017. Complete list of the samples and localities is available in Supplementary Table S1. Freshly killed cephalopods were purchased on local fish markets and geographic origin of the samples was confirmed with the retailers to rule out mixing samples from the adjacent localities. The purchased samples were kept on ice until dissection. A piece of arm muscle tissue was excised and stored in pure ethanol as a reference for the host. Mantle was opened from the ventral side and renal tissue was carefully transferred into a dish or a falcon tube containing artificial seawater (prepared according to Lapan & Morowitz, 1975) to dislodge attached dicyemids. Liquid with the dislodged dicyemids was transferred into a micro-tube and centrifuged at low speed to concentrate dicyemids in the solution. Supernatant was removed, additional amount of liquid with dicyemids was added and the centrifugation was repeated. Finally, supernatant was removed and 1 ml of pure ethanol was added for storage.

Slides for morphological examination were prepared according to the protocol of Pascual et al. (1997). In summary, a piece of renal tissue was smeared on a coverslip and fixed in Boiun's fluid for up to 20 minutes, then stored in 70% ethanol until staining with hematoxylin-eosin and mounting in Canada balsam.

One sample with live dicyemids in good condition (SOIN1) was fixed and stained for confocal microscopy to provide illustrative photos of dicyemids from *Sepia officinalis*.

Prior to DNA extraction, samples were removed from ethanol and let to dry. For the hosts, DNA extractions from individual specimens were performed by DNeasy Blood and tissue kit (QIAGEN). For the parasites, pooled samples (each obtained from an individual host, but containing multiple individual dicyemids) were extracted by QIAamp DNA Micro Kit (QIAGEN). In both cases, the protocol provided by the manufacturer was

followed. For cephalopod hosts, universal primers F1490 (Folmer et al., 1994) and H7005 (Hafner et al., 1994) were used for amplification of a 1100 bp fragment of the mitochondrial cytochrome oxidase I (COI) gene. For dicyemids, new specific primers were designed based on COI gene sequence extracted from a preliminary draft assembly of a cuttlefish renal tissue transcriptome (see online Supplementary Methods for details). To supplement COI based species delimitation of the studied dicyemid parasite with a nuclear locus, we amplified and sequenced 1330 bp fragment of 18S rDNA in 23 samples across our population sets with primers F3 and R2 published in Eshragh & Leander (2015). Standard PCR protocol with Taq polymerase was followed (for PCR conditions and primer sequences see Supplementary Table 2). PCR products were visualised on 1% agarose gel, enzymatically cleaned (by 2 µl FastAP and 0.5 µl ExoI enzymes with 2.5 µl H_2O per reaction added and incubated for 15 minutes at 37°C in thermocycler) and sequenced with PCR primers on the ABI analyzer (Thermo Fisher). Each sample was sequenced in both directions. Where necessary (in the case of the cephalopod host), specific sequencing primers were used to obtain the full length of the sequence (specifically designed primers for sequencing cephCOI11F, cephCOI14R, for primer sequences see Supplementary Table 2).

The sequences were assembled and trimmed in Geneious v. 11.1 (https://www.geneious.com/). Mafft v. 7 (Katoh et al., 2002) was used for translation guided alignment of sequences. To evaluate *Sepia* diversity in a broader context, publicly available sequences were added to the sequences obtained in this study (Accession numbers in Supplementary Table S3) and phylogenetic trees were computed with IQ-TREE web server (Trifinopouloset al., 2016) with auto model selection (Kalyaanamoorthy et al., 2017). Ultrafast boostrap and aLRT statistics were computed as node supports (Hoang et al., 2018).

To visualize the geographic diversity of populations, haplotype networks were designed in PopArt software (Leigh & Bryant, 2015) using Median Joining Network algorithm (Bandelt et al., 1999). One reference sequence (AB2401555) of *Sepia officinalis* covering the whole length of the alignment was added to the sequences obtained in this study to verify correct assignment of all samples to *S. officinalis* species in the population network. To show *S. officinalis* population structure across its dispersal area, publicly available sequences were added to the alignment, which was however shorter (473 bp) due to only partial overlap between the sequences (Accession numbers in the Supplementary Table S3). As an alternative to haplotype networks, the population structure information of the host and parasite was complemented using a Principal Coordinates Analysis (PCoA, implemented as cmdscale in R package stats) and AMOVA with significance calculated by a permutation test (calculated in R package poppr; Kamvar et al., 2014). Portuguese populations were omitted from AMOVA analyses due to the lack of parasite data.

To characterize the diversity of populations, summary statistics (Haplotype diversity, Nucleotide diversity, Theta , Dxy; Nei, 1987), neutrality tests (Tajima's D; Tajima, 1989), Fu and Li's D* and F* (Fu & Li, 1993) and Fu's Fs (Fu, 1997) and several population size change statistics (Raggedness, Rogers & Harpeding, 1992), Ramos-Onsins test (Ramos-Onsins & Rozas, 2002) were obtained in DNASP ver. 5 (Rozas et al., 2003). Pairwise Fst values and their significance was obtained by 10000 permutations in Arlequin (Excoffier & Lischer, 2010). Due to the fact that population sample sizes vere highly unequal in some cases (see Table 1), we also calculated these statistics for populations randomly downsized to a maximum of 5 sequences per population.

To test for correlation between population divergence of the host and its parasite, we estimated a linear regression of their pairwise population genetic distances (average number of nucleotides and Nei's Da; Nei, 1987) in R software (R core Team, 2017). As in AMOVA, Portuguese population was not included in the correlations due to the lack of parasite data. Due to

the extremely small sample size Greek population was excluded from the regression analysis. Mantel test implemented in adegenet R package (Jombart, 2008) was used to test for Isolation By Distance (IBD) between individuals, separately for parasite and host (again excluding Portuguese population). This test compared the similarity of matrices of genetic (Gst Nei, 1973) and geographic (marine distance in km, i.e. shortest straight sea route along the coast between localities measured manually on a map) distances for individual samples (R code available in the Supplementary Methods). To control for the effect of distant populations with small sample size on the results of Mantel test, we re-calculated the test on a dataset excluding Greek and Cyprus populations.

Data Accessibility. COI haplotypes newly acquired in this study of both host and parasite and 18S (small ribosomal subunit) sequence of parasite are stored in GenBank (Accession numbers: dicyemid 18S MN066345-MN0666367, dicyemid COI MN069252-MN069301 and MN310702-MN310704, cuttlefish host COI MN069190-MN069251). Overview of all sequences used in this study with corresponding accession numbers (both from public database and produced in this study) are also provided in Supplementary Table S3.

general overview		all sampled populations		Portugal	Sardinia		Italy Tyrrhenian		Italy Adriatic		Croatia Adriatic		East Mediterranean		Greece Aegean	
		host	parasite	host	host	parasite	host	parasite	host	parasite	host	parasite	host	parasite	host	parasite
number of sequences		60	50	8	9	8	9	5	8	4	22	28	3	3	1	2
positions		864	777	864	864	777	864	777	864	777	864	777	864	777	864	777
Polymorphic sites	S	61	41	11	32	19	8	6	8	9	15	6	1	1	-	8
number of haplotypes	h	32	20	7	7	4	4	3	7	2	12	7	2	2	1	2
haplotype diversity	Hd	0,96	0,86	0,964	0,944	0,821	0,694	0,8	0,964	1	0,926	0,577	0,667	0,667	-	1
nucleotide diversity	Pi	0,013	0,008	0,004	0,014	0,013	0,004	0,004	0,004	0	0,005	0,001	0,001	0,001	-	0,01
Theta (from S)	Theta-w	13,1	9,2	4,2	11,8	7,3	2,9	2,9	3,1	5	4,1	1,5	0,7	0,7	-	8
average pairwise differences	k	11,4	6	3,2	12,3	9,8	3,1	3,4	3,2	5	4	0,7	0,7	0,7	-	8
Raggedness	r	0,01	0,03	0,06	0,077	0,208	0,169	0,28	0,086	-	0,024	0,154	-	-	-	-
Ramos-Onsins and Rozas	R2	0,089	0,07	0,115	0,168	0,257	0,181	0,259	0,166	-	0,119	0,067	-	-	-	-
Fu's	Fs	-6,404	-3,568	*-3,05*	0,719	4,117	1,337	1,569	*-3,05*	-	-3,191	*-4,386*	-	-	-	-
Fu and Li	D*	-1,411	-0,289	-1,299	-0,092	*1,582*	1	1,241	0,219	-	-0,804	-1,832	-	-	-	-
Fu and Li	F*	-1,286	-0,743	-1,417	-0,013	*1,795*	0,895	1,286	0,238	-	-0,775	-2,061	-	-	-	-
Tajima's D	D	-0,526	-1,213	-1,21	0,239	1,716	0,173	1,241	0,202	-	-0,336	-1,628	-	-	-	-

values significant at P<0.02 in italics

*r. R2, Fu and Li test, Fu's Fs and Tajima's D statistics were run for populations with at least 5 sequencess

Table 1. Results of population statistic approaches: basic population metrics, population size changes and dynamics. values significant at P < 0.02 in italics. *r. R2, Fu and Li test, Fu's Fs and Tajima's D statistics were run for populations with at least 5 sequencess.

Results

In this study, we successfully sequenced and examined 60 individuals of the cephalopod host and 50 samples of their dicyemid parasites.

Two cuttlefish specimens from Sicily (Accession numbers MN069250 and MN069251) were found to be genetically distinct from the rest of the *S. officinalis* specimens in this study and from available *S. officinalis* sequences in public databases (approximate distance of 10% compared to other *S. officinalis* samples based on a phylogenetic analysis, Supplementary Figure S1). The Sicilian samples possibly represent a different *Sepia* species or subspecies and were therefore not included in the subsequent population analyses.

Haplotypes of common cuttlefish in the studied area form a fragmented structure with many missing haplotypes and most haplotypes are represented by one or a few individuals (Fig. 2), indicating high overall genetic diversity within the sampled area. At the most general level, the population structure is comprised of four main distinct clusters seen both in the haplotype network (Fig. 2) and in the PCoA (Supplemetary Figure S2). Two of the clusters encompass the majority of the obtained samples; one is found mostly in Sardinia and at the Atlantic coast of Portugal, and the other is found in the Adriatic and Tyrrhenian Seas. Two additional isolated groups are represented by only a few samples. One is represented by a single specimen found in Sardinia and is most similar to the selected *Sepia officinalis* reference sequence (see the methods), the other includes a few samples from the Eastern Mediterranean in Cyprus and Aegean sea in Greece (median joining network, Fig. 2 left). Sardinia shows surprisingly diverged set of haplotypes (Table 1) belonging to several COI clusters (Supplementary Figure S3). In contrast, the Adriatic Sea population, is characterized by higher number but lower nucleotide diversity of their haplotypes (Table 1). Correspondingly, all four neutrality tests produced negative values (albeit statistically non-significant) for the Croatian population and significantly negative value of the Fu's Fs in the Italian

Adriatic population. Although the Fst values expressing connectivity to the closest mainland shore, may potentially be biased by low sample numbers, they point into the same direction as the nucleotide diversities (Fst 0.409 for Sardinia vs. Tyrrhenian Sea and 0.031 for Adriatic Sea between Italian and Croatian coasts; Table 2). When the dataset is extended with the *S. officinalis* sequences across the distribution range available in the GenBank, the population network consists of the same four clusters as seen in our data (Fig. 3). In the extended dataset, Adriatic and Atlantic clusters mix in the WB, one cluster is distributed only in the East Mediterranean (Cyprus, Greece) and one cluster, originally represented by only one sample found on Sardinia, is now also found on the North African coast together with Atlantic cluster.

Fst	Croatia Adriatic (22)	Sardinia (9)	Italy Tyrrhenian (9)	Portugal (8)	Italy Adriatic (8)	Cyprus (3)
Croatia Adriatic (28)		0,546	0,352	0,826	0,031	0,785
Sardinia (8)	0,599		0,409	0,238	0,476	0,493
Italy Tyrrhenian (5)	0,853	0,251		0,839	0,45	0,861
Portugal -	-	-	-		0,849	0,86
Italy Adriatic (4)	0,788	0,162	0,121	-		0,844
Cyprus (3)	0,935	0,469	0,822	-	0,777	

Dxy	Croatia Adriatic (22)	Sardinia (9)	Italy Tyrrhenian (9)	Portugal (8)	Italy Adriatic (8)	Cyprus (3)
Croatia Adriatic (28)		0,018	0,006	0,025	0,004	0,019
Sardinia (8)	0,011		0,015	0,012	0,018	0,020
Italy Tyrrhenian (5)	0,010	0,012		0,022	0,007	0,021
Portugal -	-	-	-		0,025	0,021
Italy Adriatic (4)	0,008	0,012	0,006	-		0,019
Cyprus (3)	0,019	0,017	0,018	-	0,017	

legend: parasite | host

Table 2. Population pairwise comparison for host (upper right triangle) and its parasite (lower right triangle) by Fst and Dxy. Italy Tyrrhenian contains localities of Liguria, Gaeta and Naples. Greek samples are not included because of low sample size. Significant values of Fst in italics at level p = 0.05.

Based on morphological examination of the stained preparations, 18S sequence and confocal microscopy (Fig. 4), we tentatively determined the species of COI sequenced dicyemids as *Pseudicyema truncatum* Whitman, 1883. Despite there is a possibility of mixed infections in the host renal organ [25], the specificity of the designed COI primer pair seems to be very high and seems to amplify mostly one prevalent species of dicyemid. Only in 3 cases (samples SOIC4, SOIR11 and SOIR13; not included in the population dataset) the resulting sequences seemed to belong to a different dicyemid species, possibly *Dicyema* sp. or *Dicyemennea* sp. (Accession numbers MN310702-MN310704). In three cases (SOCT1, SOIC3, SOKL1) the resulting sequencing chromatogram showed double peaks in one position (a different position in each sequence) which suggests that more than one COI haplotype may be sometimes present in the sample, possibly because multiple individuals were used for DNA extraction. We arbitrarily chose the higher peak for nucleotide assignment for further analyses. We verified the species identity of our COI sequenced samples with 18S sequencing. All 18S sequences, except one (sample originating in Italia, Adriatic coast, Rimini; SOIR11), were identical to each other and were also identical to one of the *Pseudicyema truncatum* sequences available in genbank (LT669919, 1017 bp, Souidenne et al. 2017 unpublished). However, there is only 97% match to the other available *Pseudicyema truncatum* sequence (LT669870, 1175 bp, Souidenne et al. 2017 unpublished). SOIR11 sample showed only 94% match to other acquired 18S sequences and was not closely related to any available dicyemid 18S sequence, supporting the view that this cuttlefish specimen was infected by a different dicyemid species to the one this study focused on. A list of samples for which 18S sequence is available is provided in Supplementary Table S1. It is worth noting, that the number of reliably described sequences available for dicyemid species in public databases is very limited and most of them originate from Australian, Japanese and North Pacific waters making the comparisons to the Mediterranean of limited use.

Population structure of the parasite showed an interesting difference in comparison to the host. In majority of the comparisons, the basic population characteristics (Hd, Pi statistics, Table 1) suggested lower level of the local diversities in the parasite subpopulations when compared to the host (whole dataset, Sardinia and Croatian Adriatic). In the haplotype network, this reduced diversity is reflected by more compact clustering, particularly well expressed by the central haplotype from the Adriatic Sea, which represents 18 samples.

Dicyemid population structure was also partitioned into several clusters (haplotype network in Fig. 2 right, PCoA plot in Supplementary Figure S2). Although only some of the clusters showed geographical specificity, high level of haplotype sorting between sampling sites was seen (Fig. 2 right). With the exception of two Italian localities (Liguria and Tyrrhenian), no localities shared COI haplotypes. Haplotypes from East Mediterranean (Cyprus) were separated by several mutations, similarly to the host dataset. Largest diversity was seen in the samples from Sardinia with haplotypes scattered in several parts of the network. Unfortunately, no parasite data were available from Portugal to compare with the host. Samples from Croatian Adriatic formed a compact star-like cluster suggesting possible recent expansion. Unlike the host population, Fst value between the two Adriatic coasts (Croatia and Italy) was high (Fst = 0.788, Table 2). This is in agreement with the lack of haplotype sharing between the coasts. Calculation using downsampled population sizes (N=5) produced similar values of pairwise Fst distances (Supplementary Table S4).

Neutrality tests for parasites showed significant values in two cases. On Sardinia two tests (Fu an Li's D and F) revealed lack of low and high frequency polymorphisms, which is usually interpreted either as a sign of past population bottlenecks or population admixture (Biswas & Akey, 2006). Similarly to the host, all neutrality test statistics showed negative values in the Croatian Adriatic, with the Fu's Fs producing statistically significant value (Table 1).

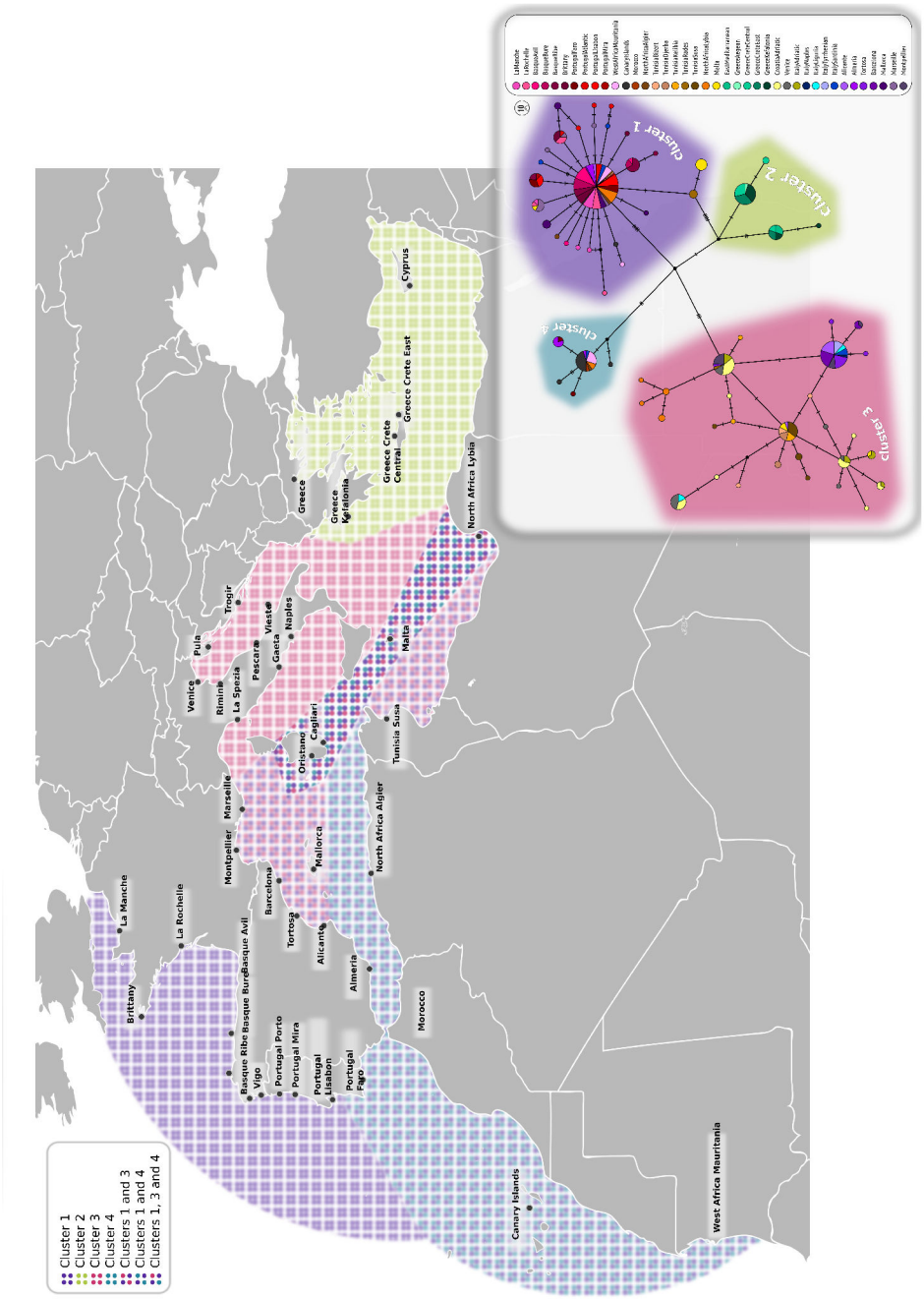


Figure 3. Sepia officinalis structure throughout its distribution. Colored areas correspond to genetic clusters shown on population network.

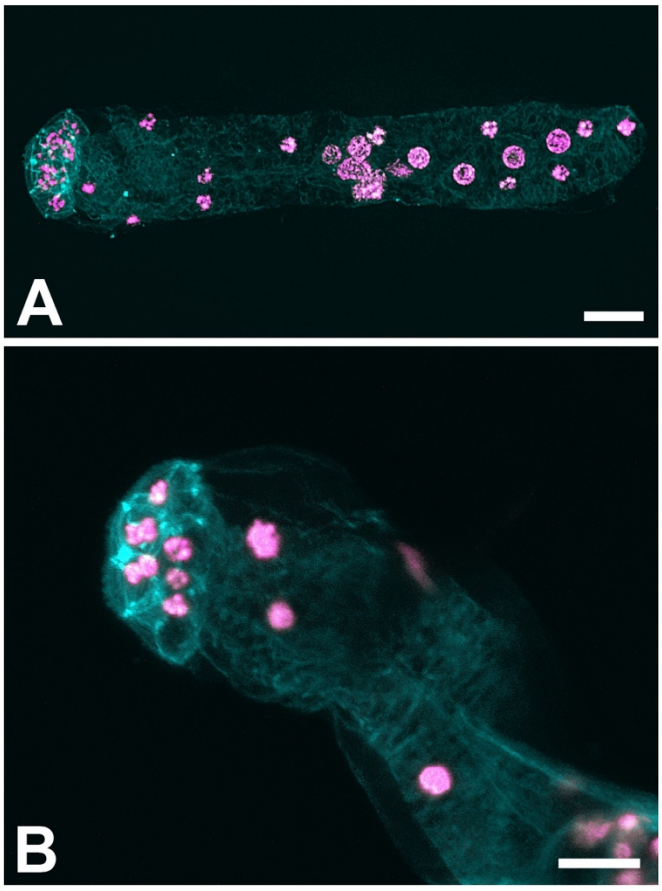


Figure 4. Dicyemid from Sepia officinalis. Confocal microscope images (DAPI/pink and phalloidin/cyan staining. (a) whole animal, (b) close up of head (calotte) with visible nuclei. The scale bar represents 10 μm.

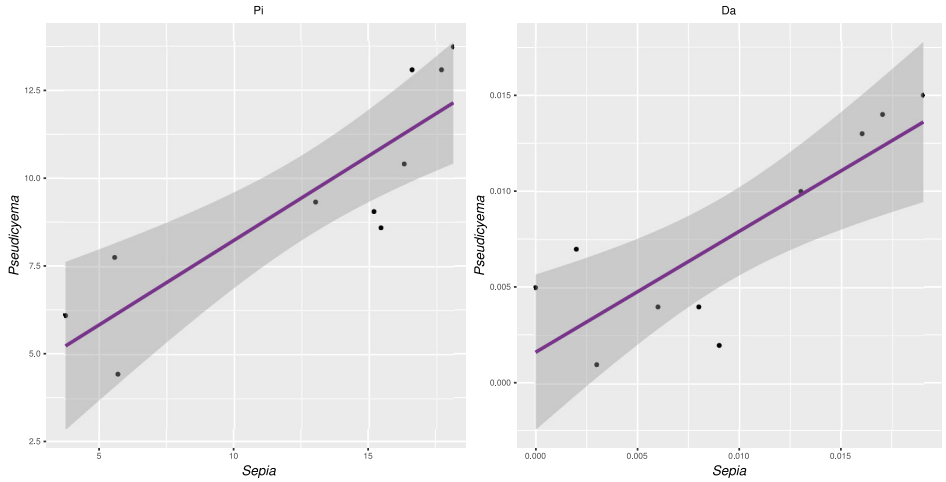


Figure 5. Correlation of pairwise population distances between the host and parasite (a) Average number of differences (Pi), adjusted R2 = 0.7127), (b) Nei's Da distance, adjusted R2 = 0.6311). Both regressions were significant at P < 0.005. Grey dashed line represents 95% regression confidence interval.

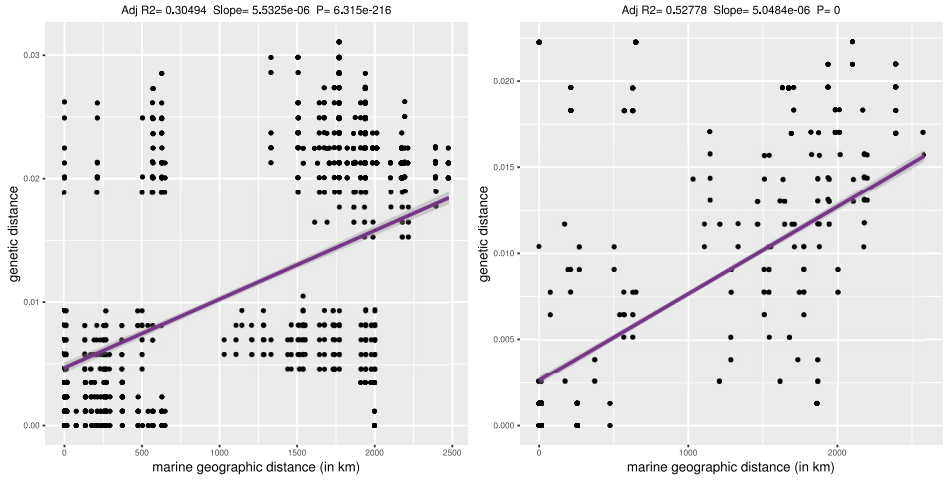


Figure 6. Correlation of pairwise individual genetic distances (computed with Tamura Nei 83 model) and geographic distances in host and parasite. (a) Host (Mantel test, observed r = 0.5, P = 0.001), (b) Parasite (Mantel test, observed r = 0.7, P = 0.001).

Regression statistics of pairwise population distances of the host and parasite showed very high level of co-diversification (Fig. 5). The slope of regression also suggests that the accumulation of genetic diversity with distance is higher in the host than in the parasite. Mantel test comparisons of individual genetic and geographic distances produced significant values in both the host and the parasite, which is pointing to IBD affecting their genetic structure (Fig. 6). There were a few population samples deviating from this pattern in both the host and parasite, showing high genetic differentiation at a small geographical scale (upper left quadrant in the plots, Fig. 6). These data points represent Sardinian population, which was genetically diverse also in other analyses (Fig. 2, Table 1). Mantel test results calculated without the distant Greek and Cyprus populations produced significant values for both host and parasite datasets (Supplementary Table S5). Despite the trend for co-diversification of the host and parasite across the Mediterranean identified in correlation analyses and in summary statistics, no pattern of congruency between individual haplotypes, or haplotype clusters, was found by mapping host genetic clusters to the parasite network in Figure 2. Dicyemid samples bearing the same haplotypes can be associated with multiple host genetic clusters.

In the case of parasites, we observed higher molecular variance between populations than within populations for the parasites in AMOVA (AMOVA: 76% variance between populations compared to 24% within populations, p=0.001, Supplementary Table S6), whereas. Contrary to parasites, hosts showed similar levels of variance between and within populations (AMOVA: 49% variance between populations and 51% within populations, p=0.001). This result is pointing to a higher degree of population fragmentation in the parasite in the Mediterranean area.

Discussion

The gross picture of the genetic diversity of *S. officinalis* across the Mediterranean Sea supports the view, already indicated by the previous studies (Perez-Losada et al., 1999, Perez-Losada et al., 2002, Perez-Losada et al., 2007), that, probably due to a low dispersion capability, this organism displays a high degree of genetic diversification among the geographically distant populations. A new strong evidence for this view was obtained by including the samples from Sardinian coast, which turned to be an area with the highest degree of diversity, possibly representing a contact zone between the East and West populations (Fig. 2 and Fig. 3). These findings demonstrate that rather than a continuous population across the Mediterranean Sea, dominantly shaped by IBD, *S. officinalis* forms a complex structure with different degree of genetic exchange among the subpopulations. In addition, we were able to demonstrate the assumed incongruencies between the host and parasite genetic structures within local populations.

While the majority of samples were collected from the central area of the Mediterranean Sea (i.e. Adriatic, Ligurian and Tyrrhenian Seas), the inclusion of several distant localities (Atlantic coast of Portugal, Greek coast and Cyprus) allowed for assessing the diversity along the whole West-East axis of the Mediterranean Sea. At this large perspective, the distant localities are genetically completely disconnected and do not share common haplotypes (e.g. Portugal vs. Adriatic vs. Cyprus). Although the number of samples is low for the east localities (Greece+Cyprus), this picture is consistent with the biological predispositions of *S. officinali*s, which is incapable of long distance dispersal, and with the previous genetic studies (Perez-Losada et al., 1999, Perez-Losada et al., 2002, Perez-Losada et al., 2007).

An interesting picture was obtained when we compared genetic composition of the Mediterranean Sea area and the Atlantic coast of Portugal. Based on the previous studies we should expect one of the

following patterns. First, the AO front serves as a major barrier, causing disconnection between the two adjacent areas. Such picture was retrieved for many organisms (for review see Patarnello et al., 2007). Second, the potential of the AO front as major determinant of population structure is suppressed by biological traits of the organisms. In such case we would expect to see either mixed population across the whole area (large dispersal capability) or strong fragmentation which would "hide" the effect of any physical barrier. However, in our data, we did not see any of these two possibilities. Instead, we found two distant genetic lineages covering disjunct geographic areas but overlapping at the Sardinian coast. In fact, the Sardinian samples proved to be genetically the most diverse. As shown in the haplotype network (Fig. 2), Sardinian samples included several haplotypes from the "Atlantic" lineage (cluster 1), several haplotypes shared with the Croatia-Italian samples (cluster 3) and even an isolated, genetically unique sample (cluster 2). In the broader analysis, which included additional *S. officinalis* publicly available sequences, the Sardinian coast seem to be located on a high diversity zone caused by an overlap of several genetic clusters (depicted in Fig. 3 and Supplementary Figure S3). Similarly to the *Sepia*, we found several unrelated haplotypes also in the Sardinian dicyemid population (nodes in dark blue, Fig. 2). Furthermore, the positively significant results of Tajima's D statistics found for this dicyemid population (Table 1) can be interpreted as a sign of past admixture between several lineages (or a bottleneck, which, however seems less probable given the distribution of Sardinian haplotypes in the network).

This scenario suggests a prominent position of the Sardinian coast as a determinant of genetic diversity, independent on the known oceanographic factors (i.e. the currents and fronts). While the current data do not allow for identification of this spot as a contact zone or source of the genetic diversity, it is interesting to note that similar picture of the diversity at Sardinian coast was obtained for the octopus (Melis et al., 2018). Although this octopus study was limited to Sardinia, and the diversity could not therefore be compared to other localities, the obtained pattern strongly

suggests presence of two distant genetic lineages (Fig. 1 in Melis et al., 2018). It might also be relevant to mention that in their analysis on connectivity among various Mediterranean areas, Andrello et al. (2013), although working with biologically very different model, detected Sardinian coasts as the localities with the highest "betweenness centrality".

The comparison between the host and the parasite population structures reveals two conspicuous patterns. First, there is no straightforward correspondence between the genetic origin of the host and the parasite, and second, the overall degree of genetic diversification is lower in the parasite than in the host (e.g. correlations in Fig. 4). Usage of Mantel tests on mtDNA data was shown to have limitations, in particular IBD is inferred erroneously when distinct regional populations are pooled (Teske et al., 2018). However, we believe this was not the case in our dataset. Based on the lack of diversity in 18S dicyemid data and the congruent pattern of increased local diversity in Sardinia for both the host and its parasite we believe that results of the correlations were not affected by artificial pooling of spatially subdivided populations. To increase the power of IBD tests it is recommended to complement the analyses with multilocus nuclear datasets to increase power of the test (Teske et al., 2018). Addition of such markers, despite difficult for the small bodied parasites nested in host tissue, would be highly beneficial for future studies.

In parallel to the processes shaping the global diversity and distribution of the haplotypes (influx of mutations, mixing by migration), the local diversities might be affected by demographic processes. Particularly, parasites may regularly undergo bottlenecks removing considerable portion of the diversity. In our data, a convincing example is provided by the best sampled population from the Adriatic Sea between the Croatian and Italian coasts. The host sample from the Croatian coast is genetically diversified, encompassing 11 different haplotypes, which are evenly distributed (1-4 individuals per haplotype) and most of them are shared with the host samples from the Italian coast (Fig. 2). In contrast, most of their parasites share a unique centrally located haplotype and the rest of the population

forms a typical star-like pattern. None of these haplotypes is shared with the parasites from other localities, although a related haplotype was sampled from the Italian coast in Liguria (depicted in light blue in Figure 2). Ligurian and Adriatic coast of Italy are in this area only three hours by car drive away making it not entirely impossible to sell freshly caught cephalopods on the other coast. Although this practice seems not to be wide spread, we paid particular attention to verification of the geographic origins of the purchased samples with the retailers to rule out such possible mixing. Moreover, in respect to the results of the presented analyses, it is important that any possible misidentification of the geographic origin in these close localities would not affect the revealed patterns.

The observed haplotype pattern, typical for the expansion after bottleneck, is also accompanied by the lowest (most negative) values of the neutrality tests statistics (Table 1). Although the values were statistically significant only for one of the tests (Fu's Fs) and could be alternatively interpreted as a sign of purifying selection (Fu & Li, 1993), the whole picture (i.e. the haplotype arrangement and the negative values of the statistics) strongly suggest a bottleneck followed by an expansion. This view is also well compatible with the narrower geographic distribution of the parasite's eastern Adriatic cluster (exclusively Croatian Coast, with the exception of the single Ligurian haplotype) in comparison with the host's genetic cluster 3, which encompasses the Croatian as well as the Italian and Sardinian samples, often with shared haplotypes. The resulting scenario thus includes a diversified host population, genetically interconnected between the Croatian and Italian coasts, and a recent recolonisation of the local Croatian population with a single genetic lineage of the parasite (i.e. bottleneck with early stage of the following diversification). This replacement/recolonisation scenario is also well compatible with the high degree of the incongruence between the geography-genetic patterns in the host and the parasites (denoted by stars in the Figure 2).

Considering possible bias of the population statistics due to the small sample sizes for some localities and a potentially low power of Mantel tests and mtDNA to uncover IBD patterns (Teske et al., 2018), we use these statistical parameters rather as an accessory evidence while the main population patterns are derived from the haplotype networks. When summarized, these patterns show that at the big scale, both the hosts and the parasites are strongly diversified and their populations are genetically rather fragmented. This is documented not only by the lack of shared haplotypes among the distant localities, but also by considerable genetic distances among the haplotypes (and many missing haplotypes). In this sense, the character of the parasite's genetic structure reflects the basic fragmentation and diversity of the host's populations. At the finer (i.e. local) scale, the parasite's genetics/geography pattern only partially reflects the distribution of the host. This shows that in local populations the extinctions and replacements take place regularly, leading to the genealogical incongruencies between the hosts and the parasites, and to the decrease of genetic diversity in the parasite. This scenario of reduced genetic diversity is in line with assumptions made by Maze-Guilmo et al. (2016) for hermaphroditic parasites with an asexual stage, which seems to be a common mode of reproduction in all dicyemids.

Acknowledgements

This study was supported by the Czech Grant Agency (project no. 15-08717S).

We would like to thank colleagues and institutions who let us use their facilities for sample processing, namely Graziano Fiorito and Giovanna Ponte (Stazione Zoologica Anton Dohrn, Naples, Italy); Andrea De Lucia, Andrea Camedda and Stefano Marra (Institute for Coastal Marine Environment, CNR, Oristano, Sardinia, Italy).

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site

Table S1. Sampling locations.

Table S2. PCR conditions and primer sequences.

Table S3. Accesion numbers of sequences used in phylogenetic tree and population network.

Table S4. Fst values computed for downsampled dataset with 5 samples maximum per population. Light grey and top row values for host populations, dark grey and left column values for parasites.

Table S5. Results of mantel test for testing of relationship between geographical and genetic distance of samples without samples originating from Greece and Cyprus because of low sample size.

Table S6. Results of AMOVA analysis

Figure S1. Maximum likelihood tree based on COI gene constructed in IQ-TREE (model auto selected-TIM2+F+I+G4, 1533bp). Pink bubble denotes the position of two sequences with 10% difference, not included in population analyses. Node supports expressed as aLRT/UFB.

Figure S2. Principal component analysis computed in R of both parasite (dicyemid) and its host (Sepia).

Figure S3. Distribution of Sepia officinalis genetic clusters in the Mediterranean Sea focusing on the diversity captured on Sardinia.

R code. Code used for computation and plotting of the Mantel test, AMOVA, regressions, PCoA and isolation by distance.

References

Andrello, M. et al. Low connectivity between Mediterranean marine protected areas: A biophysical modeling approach for the dusky grouper Epinephelus marginatus. PLoS ONE 8, e68564; 10.1371/journal.pone.0068564 (2013).

Ayata, S. et al. Regionalisation of the Mediterranean basin, a MERMEX synthesis. Prog. Oceanogr. 163, 7-20; 10.1016/j.pocean.2017.09.016 (2018).

Bandelt, H., Forster, P. & Rohl, A. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37-48; 10.1093/oxfordjournals.molbev.a026036 (1999).

Biswas, S. & Akey, J. Genomic insights into positive selection. Trends Genet. 22, 437-446; 10.1016/j.tig.2006.06.005 (2006).

Blasco-Costa, I. & Poulin, R. Host traits explain the genetic structure of parasites: a meta-analysis. Parasitology 140, 1316-1322; 10.1017/S0031182013000784 (2013).

Casabianca, S. et al. Population genetic structure and connectivity of the harmful dinoflagellate Alexandrium minutum in the Mediterranean Sea. Proc. R. Soc. Lond., B, Biol. Sci. 279, 129-138; 10.1098/rspb.2011.0708 (2012).

Catalano, S. A review of the families, genera and species of Dicyemida Van Beneden, 1876. Zootaxa 3646, 1-32; 10.11646/zootaxa.3646.1.11 (2012).

Catalano, S. First descriptions of dicyemid mesozoans (Dicyemida: Dicyemidae) from Australian octopus (Octopodidae) and cuttlefish (Sepiidae), including a new record of Dicyemennea in Australian waters. Folia Parasitol. 60, 306-320; 10.14411/fp.2013.032 (2013).

Catalano, S. R., Whittington, I. D., Donnellan, S. C. & Gillanders, B. M. Using the giant Australian cuttlefish (Sepia apama) mass breeding aggregation to explore the life cycle of dicyemid parasites. Acta Parasitol. 58, 599-602; 10.2478/s11686-013-0186-y (2013).

Catalano, S. R., Whittington, I. D., Donnellan, S. C. & Gillanders, B. M. Parasites as biological tags to assess host population structure: Guidelines, recent genetic advances and comments on a holistic approach. Int. J. Parasitol. Parasites Wildl. 3, 220-226; 10.1016/j.ijppaw.2013.11.001 (2014).

Dalongeville, A., Andrello, M., Mouillot, D., Albouy, C. & Manel, S. Ecological traits shape genetic diversity patterns across the Mediterranean Sea: a quantitative review on fishes. J. Biogeogr. 43, 845-857; 10.1111/jbi.12669 (2016).

Eshragh, R. & Leander, B. Molecular contributions to species boundaries in dicyemid parasites from eastern Pacific cephalopods. Mar. Biol. Res. 11, 414-422, doi:10.1080/17451000.2014.943241 (2015).

Excoffier, L. and H.E. L. Lischer (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10: 564-567.

Folmer, O., Black, M., Hoeh, W., Lutz R. & Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Marine Biol. Biotechnol. 3, 294-299 (1994).

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925

Fu, Y.-X. and W.-H. Li (1993). Statistical tests of neutrality of mutations. Genetics 133: 693-709.

Furuya, H., & Tsuneki, K. (2003). Biology of dicyemid mesozoans. Zoological Science, 20(5), 519–532. 10.2108/zsj.20.519

Hafner, M. et al. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science 265, 1087-1090; 10.1126/science.8066445 (1994).

Hauser, L. & Carvalho, G. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish 9, 333-362; 10.1111/j.1467-2979.2008.00299.x (2008).

Hoang, D., Chernomor, O., von Haeseler, A., Minh, B. & Vinh, L. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518-522; 10.1093/molbev/msx281 (2018).

Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403-1405; 10.1093/bioinformatics/btn129 (2008).

Kalyaanamoorthy, S., Minh, B., Wong, T., von Haeseler, A. & Jermiin, L. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587-589; 10.1038/NMETH.4285 (2017).

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281 (2014). doi:10.7717/peerj.281

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059-3066; 10.1093/nar/gkf436 (2002).

Keeney, D., King, T., Rowe, D. & Poulin, R. Contrasting mtDNA diversity and population structure in a direct-developing marine gastropod and its trematode parasites. Mol. Ecol. 18, 4591-4603; 10.1111/j.1365-294X.2009.04388.x (2009).

Laconcha, U. et al. New nuclear SNP markers unravel the genetic structure and effective population size of albacore tuna (Thunnus alalunga). PLoS ONE 10, e0128247; 10.1371/journal.pone.0128247 (2015).

Lane, H., Jones, B. & Poulin, R. Comparative population genetic study of an important marine parasite from New Zealand flat oysters. Mar. Biol. 165; 10.1007/s00227-017-3260-4 (2018).

Lapan E. A. & Morowitz H. J. The dicyemid Mesozoa as an integrated system for morphogenetic studies. J Exp Zool. 193, 147-160; 10.1002/jez.1401930204 (1975).

Leigh, J.W. & Bryant D. PopART : full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110-1116; 10.1111/2041-210X.12410 (2015).

Levin, I. I. and Parker, P. G. (2013). Comparative host-parasite population genetic structures: obligate fly ectoparasites on Galapagos seabirds. Parasitology, 140(9), 1061–1069.

Martinů, J., Hypša, V., Štefka, J. (2018). Host specificity driving genetic structure and diversity in ectoparasite populations: Coevolutionary patterns in Apodemus mice and their lice. Ecology and Evolution, 8, 10008–10022.

Maze-Guilmo, E., Blanchet, S., McCoy, K. & Loot, G. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecol. Lett. 19, 336-347; 10.1111/ele.12564 (2016).

Melis, R. et al. Genetic population structure and phylogeny of the common octopus Octopus vulgaris Cuvier, 1797 in the western Mediterranean Sea through nuclear and mitochondrial markers. Hydrobiologia 807, 277-296; 10.1007/s10750-017-3399-5 (2018).

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proc.Natl. Acad. Sci. USA 70: 3321-3323.

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

Palumbi, S. Marine speciation on a small planet. Trends Ecol. Evol. 7, 114-118; 10.1016/0169-5347(92)90144-Z (1992).

Palumbi, S. Genetic-divergence, reproductive isolation, and marine speciation. Annu. Rev. Ecol. Syst. 25, 547-572; 10.1146/annurev.ecolsys.25.1.547 (1994).

Pascual, M., Rives, B., Schunter, C. & Macpherson, E. Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. PLoS ONE 12, e0176419; 10.1371/journal.pone.0176419 (2017).

Pascual, S., Gestal, C., Abollo, E. & Arias, C. A method for the collection and preparation of Dicyemid Mesozoans. Res. Rev. Parasitol. 57, 135-138 (1997).

Patarnello, T., Volckaert, F. & Castilho, R. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? Mol. Ecol. 16, 4426-4444; 10.1111/j.1365-294X.2007.03477.x (2007).

Perez-Losada, M., Guerra, A. & Sanjuan, A. Allozyme differentiation in the cuttlefish Sepia officinalis (Mollusca : Cephalopoda) from the NE Atlantic and Mediterranean. Heredity 83, 280-289; 10.1038/sj.hdy.6885520 (1999).

Perez-Losada, M., Guerra, A., Carvalho, G., Sanjuan, A. & Shaw, P. Extensive population subdivision of the cuttlefish Sepia officinalis (Mollusca : Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. Heredity 89, 417-424; 10.1038/sj.hdy.6800160 (2002).

Perez-Losada, M., Nolte, M., Crandall, K. & Shaw, P. Testing hypotheses of population structuring in the Northeast Atlantic Ocean and Mediterranean Sea using the common cuttlefish Sepia officinalis. Mol. Ecol. 16, 2667-2679; 10.1111/j.1365-294X.2007.03333.x (2007).

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online https://www.R-project.org/. (2017).

Ramos-Onsins, S. E. and J. Rozas (2002). Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19: 2092-2100.

Robinson, A. R., Leslie, G., W., Theocharis, A. & Lascaratos, A. Mediterranean Sea Circulation in Encyclopedia of Ocean Sciences (ed. Steele, J.) 1689-1705; 10.1006/rwos.2001.0376 (Elsevier, 2001).

Rogers, A. R. and H. Harpeding (1992). Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9: 552-569.

Rozas, J., Sanchez-DelBarrio, J., Messeguer, X. & Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497; 10.1093/bioinformatics/btg359 (2003).

Souidenne, D. et al. Redescription of Dicyemennea eledones (Wagener, 1857) (Phylum Dicyemida) from Eledone cirrhosa (Lamarck, 1798) (Mollusca: Cephalopoda: Octopoda). Syst. Parasitol. 93, 905-915; 10.1007/s11230-016-9659-3 (2016).

Tajima, F. (1989). The effect of change in population size on DNA polymorphism. Genetics 123: 597-601.

Teske, P.R. Golla, T.R., Sandoval-Castillo, J., Emami-Khoyi, A., der Lingen, C.D., der Heyden, S., Chiazzari, B., Jansen van Vuuren, B., Beheregaray, L.B., Mitochondrial DNA is unsuitable to test for isolation by distance. Scientific Reports 8, 8448 (2018).

Trifinopoulos, J., Nguyen, L., von Haeseler, A. & Minh, B. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232-W235; 10.1093/nar/gkw256 (2016).

Valdivia, I., Criscione, C., Cardenas, L., Duran, C. & Oliva, M. Does a facultative precocious life cycle predispose the marine trematode Proctoeces cf. lintoni to inbreeding and genetic differentiation among host species? Int. J. Parasitol. 44, 183-188; 10.1016/j.ijpara.2013.10.008 (2014).

Watson, H., McKeown, N., Coscia, I., Wootton, E. & Ironside, J. Population genetic structure of the European lobster (Homarus gammarus) in the Irish Sea and implications for the effectiveness of the first British marine protected area. Fish. Res. 183, 287-293; 10.1016/j.fishres.2016.06.015 (2016).

Winkelmann, I. et al. Mitochondrial genome diversity and population structure of the giant squid Architeuthis: genetics sheds new light on one of the most enigmatic marine species. Proc. R. Soc. Lond., B, Biol. Sci. 280, 1759; 10.1098/rspb.2013.0273 (2013).

Zamborsky, D. & Nishiguchi, M. Phylogeographical patterns among Mediterranean sepiolid squids and their Vibrio symbionts: Environment drives specificity among sympatric species. Appl. Environ. Microbiol. 77, 642-649; 10.1128/AEM.02105-10 (2011).

Zheng, X. et al. Genetic diversity and population structure of the golden cuttlefish, Sepia esculenta (Cephalopoda: Sepiidae) indicated by microsatellite DNA variations. Mar. Ecol.-Evol. Persp. 30, 448-454; 10.1111/j.1439-0485.2009.00294.x (2009).

CHAPTER III.

Genetic analysis of dicyemid infrapopulations suggests sexual reproduction and host colonization by multiple individuals is common.

Organisms Diversity & Evolution (2021) 21, 437–446
https://doi.org/10.1007/s13127-021-00493-0

Genetic analysis of dicyemid infrapopulations suggests sexual reproduction and host colonization by multiple individuals is common.

Drábková Marie*[1,2], Tereza Flegrová[1,2], Eva Myšková[1], Václav Hypša[1], Jan Štefka[1,2]

[1]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic
[2]Institute of Parasitology, Biology Centre, CAS, v.v.i., České Budějovice, Czech Republic

* Corresponding author

Abstract

Dicyemida is a group of small-bodied marine parasites infecting cephalopods with many unknown life history details, such as their population structure and diversity, and their relation to sexual and asexual reproductive stages. To reveal (infra)population structure of *Dicyema moschatum* Whitman, 1883 in its host (*Eledone moschata* Lamarck, 1798) we isolated microsatellite sequences from a draft genome of *D. moschatum* and tested the loci for amplification success and genetic diversity. Eight microsatellite loci were selected for an analysis of *D. moschatum* populations from several octopus individuals sampled at two Mediterranean localities. The majority of microsatellite alleles were shared across the studied range, but several private alleles were also identified. Analysis of population structure identified two to four genetic clusters, mostly concordant with the geographic origin of the samples. Allelic patterns seen in individual dicyemid genotypes revealed that although dicyemids inside one host individual show low genetic variance, they do not represent genetically identical clones. These results suggest that infection is established by several dicyemid larvae within the lifetime of the host and sexual reproduction of dicyemids occurs inside the host.

Keywords

dicyemids, population genetics, parasites, cephalopods, reproduction

Introduction

Dicyemida are a little-studied group of small bodied metazoan parasites living in renal organs of benthic cephalopods (e.g. cuttlefish and octopus). Despite the fact that Dicyemida were first described over a hundred years ago (Kölliker 1881), some features of their life cycle remain still unknown. Dicyemids produce two types of embryos, a worm-like vermiform embryo resembling a miniature adult, and an infusoriform embryo, which is smaller in size, of round shape and covered in cilia. Vermiform embryos are assumed to be produced asexually, quickly multiplying dicyemid numbers and increasing density of infection in the host. Infusoriform embryos are thought to serve as a dispersal stage that infects a new host, and they are assumed to be produced sexually (Van Beneden 1876; Furuya et al. 2003). Classical methods of observation, such as light microscopy, did not provide conclusive answers regarding sexual and asexual reproduction, and interpretation of the observed details has been disputed among researchers (Stunkard 1954). Furthermore, since the exact way of transmission to a new host is unknown, the existence of other intermediate stages cannot be excluded, and the frequency of transmission is also uncertain. Some progress in this area was made by Lapan and Morowitz (1975) in the study focused on dicyemid in vitro cultivation. When cuttlefish were raised in aquaria from eggs in presence of infusoriform larvae, they became infected by dicyemids, albeit with low prevalence, suggesting that dicyemids are able to complete lifecycle without an intermediate host. Unfortunately, this short mention of a side experiment lacks details. To obtain more information, further search for intermediate hosts and the performance of infection experiments was suggested (Lapan and Morowitz 1975). However, such experiments are expensive, may prove to be difficult to carry out, or even impossible to perform under the new legislation concerning cephalopod use in experiments (at least in the UK and EU).

Another unknown is whether the vertical transfer between cephalopods is possible (passing of infection from mother to its offspring). High prevalence seen in dicyemids suggests a very effective way of transmission coupled with low virulence, or a combination of horizontal and vertical

transfer (Lipsitch et al. 1995). Vertical transmission could be realized, for example, through egg infection (as seen in vertically transmitted parasites like *Wolbachia* Hertig, 1936, Microsporidia, or in a kinetoplastid, *Leptomonas pyrrhocoris* Zotta, 1912, that parasitizes firebugs; Taylor and Hoerauf 1999, Dunn and Smith 2001, Salem et al. 2015). However, in their experiments, Catalano et al. (2013) failed to find any evidence of vertical transmission in dicyemids. Their efforts to search for dicyemid DNA by PCR amplification of the cytochrome oxidase I gene (COI) were not successful from the eggs of *Sepia apama* Gray, 1849 nor from seawater obtained during cuttlefish breeding (eDNA).

An alternative approach to unveil more information about the dicyemid life cycle is to investigate dicyemid population structure and genetic diversity inside the host (parasite infrapopulation). This could reveal whether the dicyemid population within a single host is established by infection with a single larva. It should also provide information on whether the high density of dicyemids in the host's renal organ is maintained predominantly by asexual multiplication, creating a genetically uniform population, or by sexual reproduction, leading to higher genetic diversity. If the infection is established by several larvae followed by asexual reproduction inside the host, all individuals should be representatives of a few genetic lineages (clones). Alternatively, if dicyemid infrapopulations are established by several larvae (or by repeated infections) and maintained mostly by sexual reproduction, they should show a considerable level of genetic diversity and contain only a few genetically identical individuals (clones).

To answer these questions, we selected *Dicyema moschatum,* a relatively large dicyemid species that can grow up to 6 mm (Nouvel 1947). *Dicyema moschatum* is associated with a widespread Mediterranean octopus *Eledone moschata,* often found on fish markets in Croatia and Italy. The prevalence of infection in this host population is very high, usually up to 100%. The availability of the *Eledone moschata* host species, and the size and prevalence of *D. moschatum* make them a suitable study system for understanding dicyemid biology and population genetics. *D. moschatum*

can be also rarely found in *Sepiola rondeletii*, Leach 1817 (prevalence 5% in Nouvel 1947). Occasionally, the host species *Eledone moschata* may be infected by *Dicyemennea eledones*, a dicyemid from a different genus, more commonly found in *Eledone cirrhosa*, with the possibility of mixed infection (Nouvel 1947, Souidenne et al. 2016; pers. obs.).

Due to their small body, comprising only dozens of nucleated cells, dicyemids produce very low amounts of DNA during extractions from individual specimens. Microsatellites, which require a low amount of input DNA for successful amplification and are not prone to contamination by host DNA, provide the most convenient tool for the examination of population structure and diversity in dicyemids. Here, we designed and applied the analysis of microsatellite loci in *D. moschatum* aiming to answer the questions stated above on the possibility of multiple infections and the extent of asexual vs. sexual reproduction in the dicyemid life cycle.

Methods

Octopuses of the species *Eledone moschata* were bought freshly killed, directly from local fishermen during two sampling trips in 2016 and 2017 in two localities: Naples (Italy) and Pula (Croatia). Octopuses were kept on ice until dissection. Octopuses were opened by a mid-ventral cut and renal tissue containing live dicyemids was placed in falcon tube or petri dish with artificial seawater (ASW, prepared according to Lapan and Morowitz 1975). In some cases, liquid with live dicyemids was fixed in RNAlater (Invitrogen, Carlsbad, CA). Individual live dicyemids were separated either directly from the ASW solution or RNAlater by pipetting with the use of a stereomicroscope. Every single live dicyemid from ASW was placed onto dry filter paper previously soaked in TE buffer and extraction of DNA was performed with QIAamp DNA micro kit (Qiagen, Hilden, Germany) following the protocol for dried blood spots. In the case of samples fixed in RNAlater, individual dicyemids were washed in distilled water and dried. Their DNA was extracted with QIAamp DNA micro kit (Qiagen, Hilden, Germany) following the protocol for tissue extraction.

Sample of the host muscle tissue (0.5 cm x 0.5 cm cube excised from arm without chromatophores) was fixed in pure ethanol for verification of the species identification. DNA was extracted with DNAeasy blood and tissue kit (Qiagen, Hilden, Germany). Universal primers for amplifying approximately 1000bp long fragment of the cytochrome oxidase I gene (COI) in invertebrates (F1490 Folmer 1994; H7005 Hafner 1994) were used in a 20µl PCR reaction: 0.2 µl TaqPolymerase 5 U (Top-Bio, Praha, Czech Republic), dNTPs 2.5mM each 0.5 µl, 2 µl 10x PCR blue buffer (Top-Bio, Praha, Czech Republic), forward/reverse primer 5mM each 1 µl, template DNA 1µl. The PCR protocol involved an initial denaturation period at 94°C for 5 min, 30 cycles of denaturation at 92°C for 1 min, annealing at 52°C for 1 min and elongation at 72°C for 1 min followed by final elongation period at 72°C for 5 min. PCR products were checked on 1% agarose gel and sequenced with PCR primers on ABI analyzer (service provider SEQme, Dobříš, Czech Republic or Eurofins, Ebersberg, Germany). Chromatograms were analyzed in Geneious R10 (http://www.geneious.com, Kearse et al. 2012) using both strands to assemble the sequences. Resulting sequences were compared to the GenBank database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to rule out possible misidentification of our specimens with a similarly looking *Eledone cirrhosa* present in the same areas.

Individual extracted samples of dicyemid DNA were used for amplification of the 18S marker to ascertain the presence of only one species in our dataset. PCR reaction was run in 20µl: 0.2 µl TaqPolymerase 5 U (Top-Bio, Praha, Czech Republic), dNTPs 2.5mM each 0.5 µl, 2 µl 10x PCR blue buffer (Top-Bio, Praha, Czech Republic), forward/reverse primer 5mM each 1 µl,and template DNA 1 µl. The PCR protocol involved an initial denaturation period at 94°C for 2 min, then 40 cycles of denaturation at 94°C for 45 s, annealing 50°C for 45 s, and elongation at 72°C for 2 min, followed by final elongation period at 72°C for 5 min with forward primer F3 and reverse primer R2 (Eshragh and Leander 2014) amplifying 1300 bp fragment of dicyemid 18S sequence. Chromatograms were analyzed in Geneious R8 (http://www.geneious.com, Kearse et al. 2012) using both

strands to assemble the sequence. Accession numbers of obtained sequences are MT703888-MT703900.

Primers for the amplification of dicyemid microsatellite loci were designed with QDD3 pipeline (Meglécz et al. 2014) using *Dicyema moschatum* draft genome (brief summary: 30 individual dicyemids were pooled, sequenced and assembled into draft genome, dicyemid species was determined by its host *Eledone moschata* and by a simultaneously run confocal imaging experiment, more details in Drabkova et al. in prep). Selected 45 promising primer pairs were synthesized (Sigma-Aldrich,UK) and further tested for amplification success. From the 45 candidate primers, we selected eight primer pairs that reliably amplified microsatellite loci on selected trial individual dicyemid extractions. Then the eight loci were split into two sets of fluorescently labelled primers for multiplex PCR (6fam, Ned, Vic, and Pet dyes). Primer set sequences are provided in Online Resource 1. Microsatellite loci were amplified in two 20µl multiplex PCR reactions: 0.5 µl TaqPolymerase (Top-Bio, Praha, Czech Republic), 2 µl 10x PCR blue buffer (Top-Bio, Praha, Czech Republic), dNTPs 2.5mM each 0.5 µl, primer 0.6 µl each, 1 µl Q solution (Qiagen, Hilden, Germany), template DNA 1 µl. The PCR protocol involved an initial denaturation period at 94°C for 5 min, 35 cycles of denaturation at 94°C for 30 s, annealing at 54°C for 30 s, and elongation at 72°C for 1 min, and followed by final elongation period at 72°C for 4 min). Results of PCRs were visualized on 2% agarose gel. For some samples, particular loci in the multiplex failed to amplify. In such cases, a single primer pair of the missing locus was run in an additional PCR run. Amplified fragments were analyzed with 600 LIZ size standard on ABI analyzer (service provider Macrogene, Seoul, Rep. of Korea or SEQme, Dobříš, Czech Republic).

Resulting chromatograms were manually analyzed and checked in Geneious R10 (http://www.geneious.com, Kearse et al. 2012; microsatellite plugin). Population genetic statistics describing the genetic variation of the populations, in this case host individuals, (Fst, Nm, AMOVA, PCoA) and summary of the data (allele frequencies) were

performed in Genalex (Peakall and Smouse 2006 and 2012) and Arlequin (AMOVA, HWE; Excoffier et al. 2005). Infrapopulations with data on less than five dicyemid individuals were not included in the calculations of summary statistics. Linkage disequilibrium (LD) analysis between loci was computed in Arlequin (Excoffier et al. 2005) for hosts (infrapopulations) which comprised of at least ten individuals (samples CP5, IN1). Sequential Bonferroni correction was applied to p-values to assess the significance of LD. Estimation of the number of source populations (genetic clusters) and genetic profile assignment of the individuals was done in STRUCTURE (Pritchard et al. 2000; Falush et al. 2003; Hubisz et al. 2009) using all screened specimens. STRUCTURE was run with burn-in set to 100,000 from a total of 1,000,000 steps. Fifteen repetitions were performed for each value of k from 1 to 15. Results from multiple runs were analysed and collated in CLUMPAK (Kopelman et al. 2015), and the optimal value of k was selected in Structure harvester (Earl et al. 2012).

To exclude a possibility of cross-amplicifation with *Dicyemennea eledones*, another species known to infect *E. moschata*, six *D. eledones* (and one *D. moschatum* individual used as a positive control) were tested for amplification in the two microsatellite multiplexes. Whilst both multiplexes amplified in the positive control, we obtained no PCR product in any of the *D. eledones* individuals.

Results

From the total of 45 candidate primer pairs designed in QDD3 and tested in PCRs, eight pairs were selected and used to analyze 49 individual dicyemid samples originating from seven host individuals from two localities (Pula, Croatia and Naples, Italy). Thus, multiple parasite specimens were analysed from each host individual to explore the diversity of dicyemid infrapopulations (population inside one host). Despite a thorough PCR optimization and primer testing, some microsatellite loci showed lower amplification success (e.g., 35% in DicMiSat27; Tab. 1).

Linkage disequilibrium test (LD) revealed only one pair of loci in significant linkage in one host after sequential Bonferroni correction of p-value (loci dicMiSat02 and dicMiSat07 in host IN1). Because significant LD was found only in one population, we decided to keep all loci in further analyses.

Table 1 Amplification success of microsatellite loci

	success in %
DicMiSat02	94
DicMiSat05	90
DicMiSat07	96
DicMiSat08	96
DicMiSat17	59
DicMiSat27	35
DicMiSat31	55
DicMiSat41	94

The pattern of allele distribution of microsatellite loci in dicyemid individuals shows that dicyemids inhabiting one host octopus are not represented by identical genotypes (Fig. 1), suggesting that the founding of the population by a single individual followed by clone-like asexual reproduction is improbable.

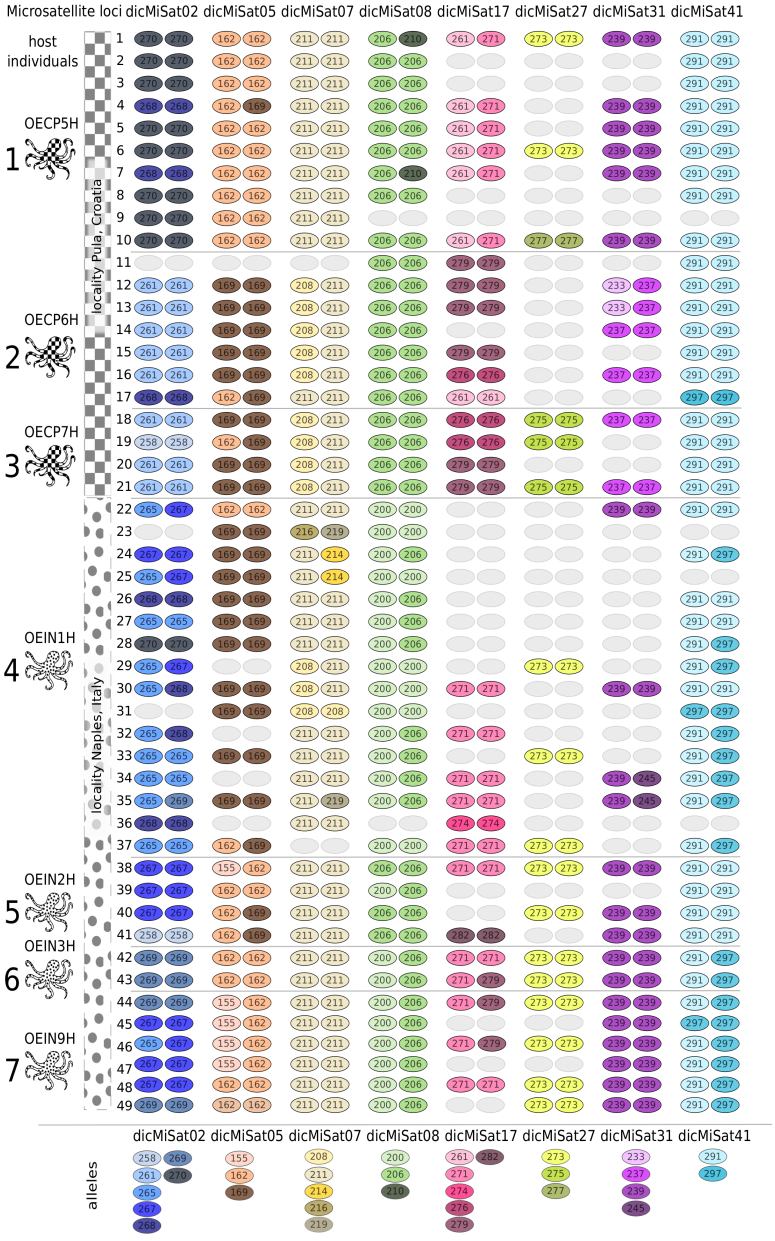


Figure 1. Pattern representing a distribution of microsatellite alleles in individual dicyemids. Dicyemids in each host individual did not share identical patterns of microsatellite alleles. Each row represents a single dicyemid individual. Microsatellite alleles are colour coded according to their length (in bp). Octopuses in checkerboard represent host individuals from Croatia, Pula (CP5, CP6, CP7); Octopuses in polka dots represent hosts from Italy, Naples (IN1, IN2, IN3, IN9). Missing data (no amplification) in grey.

Some microsatellite alleles were locally specific (private) but most were shared between localities, suggesting that gene flow is maintained among the studied populations (for example see dicMiSat02 with local allele 261 in Croatia, local allele 258 in Italy and shared allele in both locations 268; Online Resource 2).

Out of the 34 locus/population tests for Hardy-Weinberg equilibrium (HWE), we found 13 cases of loci showing significant deviation (p level 0.05) (Online Resource 3). On the contrary, most loci were significantly out of HWE (p level 0.05) when analysed across all the populations. Only loci dicMiSat08 and dicMiSat41 were not significantly out of HWE across the populations (for results see Online Resource 3). Departure from HWE is expected in substructured populations, which is consistent with our results.

In the studied infrapopulations, we found generally low level of heterozygosity (0.271 on average) and similar levels of observed and expected heterozygosities, with slightly lower expected heterozygosity in the host sample IN9 (Tab. 2, for results including also three populations with less than five individuals, see Online Resource 4). Two of the studied loci showed overall low amounts of heterozygosity (observed heterozygosity for dicMisat02 = 0.085 and dicMiSat27 = 0.063; Online Resource 4).

Genetic differentiation between dicyemid infrapopulations (hosts) from the two studied localities was relatively high, with Fst values between 0.328 and 0.462 (Tab. 3). Differentiation among infrapopulations on one locality was generally low (see Online Resource 5). However, one of the samples from Croatia (CP5) showed surprisingly high divergence from other Croatian samples (genetic distance = 0.788, Fst = 0.417; Tab. 3, Online Resource 5). This pattern is also visible in the graphical result of Structure analysis, where this Croatian sample showed a similar pattern to Italian samples, when assigned to two original populations (k=2) and showed a

distinct pattern when assigned to four original populations (k=4; Fig. 2). The same pattern also emerged in PCoA analysis, where this sample clustered with Italian samples (Online Resource 6).

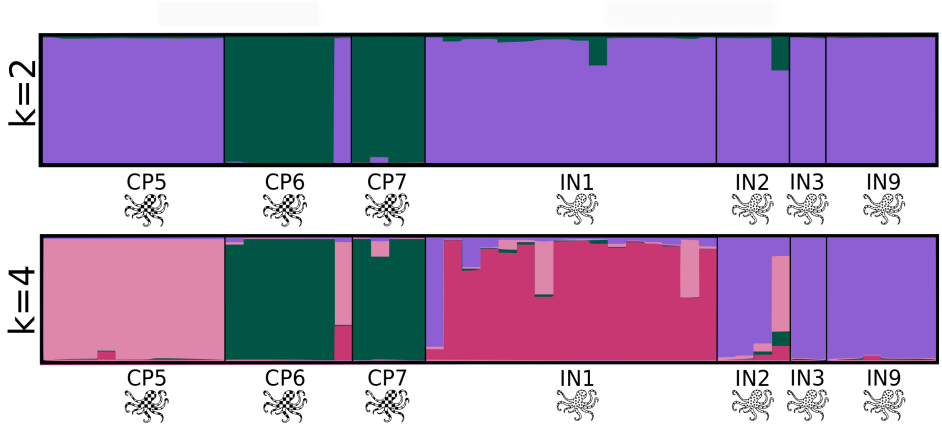


Figure 2. Graphical result of Structure analysis. Each bar represents a single dicyemid individual. Shown are plots for two original populations (k=2, selected by Evanno method) and for four original populations (k=4, highest probability k). Host population abbreviations as in Fig. 1

Structure analysis assigned individual dicyemid profiles to either two original populations (best k=2 according to Evanno method; Evanno et al. 2005) or four original populations (k=4 had the highest probability). Both clustering patterns showed a clear distinction between hosts from Italian and Croatian localities (with the exception of host CP5 at k=2, for more information, see above). A more nuanced distinction for CP5 and IN1 was seen in the results modelling four clusters (k=4; Fig. 2).

Genetic variance of dicyemids in the AMOVA test was highest on an individual level (60%). Comparison of the localities (groups=Adriatic and Tyrrhenian seas=Croatia, Pula and Italy, Naples) showed statistically significant genetic variation (27%). Infrapopulations within localities showed a low degree of variance (8.2%), which suggests, together with the low variance between individuals within infrapopulation, a high relatedness of dicyemid individuals in one host (4.3%; AMOVA results Fig. 3).

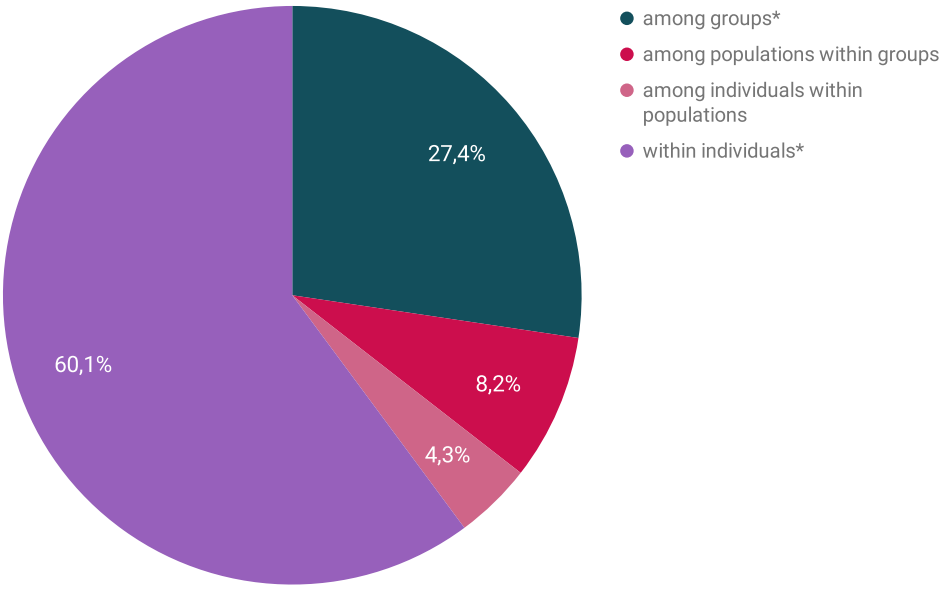


Figure 3. Percentage of variance present in the samples (based on the four-level AMOVA computed in Arlequin using Rst metrics). Groups set as two localities from Tyrrhenian (Italy, Naples) and Adriatic seas (Pula, Croatia). Hierarchical levels showing significant variation ($p < 0.05$) are marked by an asterisk.

Discussion

Questions about dicyemid life cycle and their mode of reproduction (mainly the role of asexual and sexual reproduction phases) have been debated for decades without reaching clear, unified consensus when classical methods of observation were used (Stunkard 1954; Hochberg 1990). Our first multilocus study on genetic diversity in dicyemids shows that their infrapopulations are composed of genetically distinguishable individuals, suggesting frequent outcrossing. The data also indicate the presence of at least two genetic clusters in the north-west Mediterranean, in accordance with the geographic origin of collected samples.

Infarapopulation genetic composition and life cycle of dicyemids

Analysis of multilocus genetic information performed across several populations is an effective tool to elucidate transmission route, life cycle and mating behaviour in hermaphroditic parasites (e.g. amount of outcrossing, clonality and selfing rate; for review see Criscione et al. 2005). The occurrence of multiple loci in strong linkage disequilibrium (LD) and/or presence of identical genotypes are linked to high rates of self-fertilization and low levels of transmission. This was shown for example for *Plasmodium* Marchiafava and Celli, 1885 parasites by Anderson et al. (2000). On the contrary, in the trematode *Plagioporus shawi* McIntosh, 1939 microsatellite profiles of individuals in infrapopulations showed surprisingly high diversity, low clonality, and frequent outcrossing, even though the parasite is hermaphroditic and can reproduce by self-fertilization (Criscione and Blouin, 2006). Similarly, multilocus data showed high levels of heterozygosity and low amount of selfing also in other parasites, such as tapeworms (Štefka et al. 2009; Sprehn et al. 2015). Our results showed that dicyemid individuals within one host (infrapopulation) are not genetically identical and that none of the screened loci are in LD (with one exception, see also below). These findings bring new information about dicyemid life-history traits. It shows that sexual reproduction (and outcrossing) must be common in dicyemid populations. Furthermore, it indicates that colonization of a host individual occurs multiple times during the lifetime of a cephalopod, or by multiple individuals during certain life period. Exact estimation of the proportion of clones in dicyemid infrapopulations and connected analyses were hampered by a high amount of missing data in some loci and a relatively low number of screened individuals. Nonetheless, with all limitations considered, the level of genetic diversity observed and the fact that only two loci were found in LD in one dicyemid infrapopulation point to the fact that sexual reproduction is frequently taking place in cephalopod renal organs and that self-fertilization or clonal reproduction is less frequent than previously thought (cf. Furuya and Tsuneki 2003).

Genetic diversity and structure in marine parasites

Parasites of marine organisms are often expected to have panmictic populations (Criscione et al. 2011) as was shown, for example, in *Aniskakis* spp. Dujardin, 1845 infecting sardines (Baldwin et al. 2011). However, in some parasites, the situation may be more complex. For example, three cryptic genetic clusters were observed in the trematode *Lecithochirium* Lühe, 1901 in *Conger* Bosc, 1817, but inside these clusters the populations were still panmictic (Criscione et al. 2011), in accordance with the "high mixing in aquatic habitats" hypothesis (Criscione and Blouin 2006). Sample sizes available for our study were in many host individuals limited to fewer than ten individuals, not allowing for precise estimation of population genetic statistics for the infrapopulations. Obtaining sufficient numbers of individual dicyemids for DNA extractions proved challenging due to limited resources (mainly by involved manual labour when separating individual dicyemids, which needs to be done with great precision in a short time span, while dicyemids are still alive). However, using our data we were able to infer inter-population characteristics, comparing the two studied localities in the Mediterrenean Sea.

We found differentiation between Croatian and Italian samples, showing a distinction of Adriatic and Tyrrhenian seas, based on allelic profiles, Structure analysis, AMOVA and Fst values. These results rule out existence of one panmictic population of *D. moschatum* in north-west Mediterranea. Such geographical division is in line with a previous study performed by our team on a larger geographical scale, but using only a single locus (mtDNA sequence of cytochrome oxidase I) and analysing a different dicyemid species, *Pseudicyema truncatum* Whitman, 1883, a parasite infecting cuttlefish (Drábková et al. 2019). Furthermore, the study suggested a higher level of genetic structure and population turnover in the parasite than in its cuttlefish host. Microsatellite data obtained here for *D. moschatum* indicate similar patterns (existence of infrapopulation structure), however, with such scarce sampling, we cannot properly assess dicyemid diversity in general. Further sampling on a finer scale could

provide background for answering more specific questions about dicyemid diversity and infection patterns between hosts.

Limitations and advantages of the microsatellite protocol

Analysis of microsatellite loci in dicyemids provides a feasible and informative method for exploring population relationships and life cycle in small bodied organisms. Given the very low volume of DNA material obtained from dicyemids, microsatellites are also cost-effective compared to genomic methods, which would require the costly application of single-cell genomics methods to obtain population-level data. By collecting information from multiple individuals, microsatellite analysis allowed us to explore previously little known features of a dicyemid life cycle, such as the extent of sexual vs. asexual reproduction in their hosts, or the level of infrapopulation structure. On the other hand, microsatellite amplification success that we observed was low for some loci and we were able to successfully apply only a limited number of microsatellite loci due to a high number of discarded candidates during the optimization process. Whereas a high disposal rate may be connected to the low diversity (high homozygosity) of the populations, the low amplification success rate may also possibly be caused by low amounts of input DNA extracted from such small-bodied animals. Similar issues with a low amplification success rate of microsatellite loci were also found for low quality input DNA extracted from challenging samples like faecal material (Taberlet et al. 1996; Zhu et al. 2017). Other factors influencing the microsatellite amplification success rate could also involve the storage conditions of the samples and their handling. However, quick desiccation, a method similar in principle to the handling of samples employed in this study, was found to be effective for panda faecal samples, a challenging material for DNA extraction (Zhu et al. 2017).

In conclusion, despite limitations presented by the relatively low volume of the studied material, our findings bring new information about the understudied features of dicyemid life history. Our data show that

outcrossing is a frequent mode of reproduction inside the cephalopod host and that dicyemid infrapopulations are established by at least several individuals.

Acknowledgements

The authors are grateful to colleagues who helped with field sampling (mainly Tomáš Tyml) and also to colleagues abroad, namely Graziano Fiorito and Giovanna Ponte, who allowed us to use their facilities at Stazione Zoologica Naples.

Availability of data and material

Dataset produced in this study is available on the following link https://drive.google.com/file/d/1gLuZiwtriHedxyNkIv8bjmWmLBAUUTTF/view?usp=sharing

Sequences produced in this study are available in Genbank under accession numbers MT703888-MT703900.

References

Anderson, T. J. C., Haubold, B., Williams, J. T., Estrada-Franco, J. G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I. D., Brockman, A. H., Nosten, F., Ferreira, M. U. & Day, K. P. (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Molecular Biology and Evolution, 17(10), 1467–1482. https://doi.org/10.1093/oxfordjournals.molbev.a026247

Baldwin, R. E., Rew, M. B., Johansson, M. L., Banks, M. A., & Jacobson, K. C. (2011). Population Structure of Three Species of Anisakis Nematodes Recovered From Pacific Sardines (*Sardinops sagax*) Distributed Throughout the California Current System. Journal of Parasitology, 97(4), 545–554. https://doi.org/10.1645/ge-2690.1

Bosc, L.A.G. (1817). Pisces accounts. Tome VII. Nouveau Dictionnaire d'Histoire Naturelle, Nouvelle Édition, Paris

Catalano, S. R., Whittington, I. D., Donnellan, S. C., & Gillanders, B. M. (2013). Using the giant Australian cuttlefish (*Sepia apama*) mass breeding aggregation to explore the life cycle of dicyemid parasites. Acta Parasitologica, 58(4), 599–602. https://doi.org/10.2478/s11686-013-0186-y

Criscione, C. D., Poulin, R., & Blouin, M. S. (2005). Molecular ecology of parasites: Elucidating ecological and microevolutionary processes. Molecular Ecology, 14(8), 2247–2257. https://doi.org/10.1111/j.1365-294X.2005.02587.x

Criscione, C. D., & Blouin, M. S. (2006). Minimal Selfing, Few Clones, and No Among-Host Genetic Structure in a Hermaphroditic Parasite With Asexual Larval Propagation. Evolution, 60(3), 553. https://doi.org/10.1554/05-421.1

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: Detecting cryptic microgeographic population structure in a parasite with a complex life cycle. Molecular Ecology, 20(12), 2510–2524. https://doi.org/10.1111/j.1365-294X.2011.05113.x

Drábková, M., Jachníková, N., Tyml, T., Sehadová, H., Ditrich, O., Myšková, E., Hypša, V. & Štefka, J. (2019). Population co-divergence in common cuttlefish (*Sepia officinalis*) and its dicyemid parasite in the Mediterranean Sea. Scientific Reports, 9(1), 1–12. https://doi.org/10.1038/s41598-019-50555-9

Dunn, A. M. & Smith, J. E., (2001). Microsporidian life cycles and diversity: the relationship between virulence and transmission. Microbes and Infection, 3(5), 381–388.

Earl, D. A. & von Holdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4(2), 359-361. https://doi.org/10.1007/s12686-011-9548-7

Eshragh, R., & Leander, B. S. (2014). Molecular contributions to species boundaries in dicyemid parasites from eastern Pacific cephalopods. Marine Biology Research, 11(November), 414–422. https://doi.org/10.1080/17451000.2014.943241

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 14(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L. G., Laval, & Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1, 47-50.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics, 164, 1567–1587.

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.

Furuya, H., Hochberg, F. G., & Tsuneki, K. (2003). Reproductive traits in dicyemids. Marine Biology, 142(July 2002), 693–706. https://doi.org/10.1007/s00227-002-0991-6

Furuya, H., & Tsuneki, K. (2003). Biology of dicyemid mesozoans. Zoological Science, 20(5), 519–532. https://doi.org/10.2108/zsj.20.519

Gray, J.E. (1849). Catalogue of the Mollusca in the collection of the British Museum I: Cephalopoda antepedia. London: British Museum. https://doi.org/10.5962/bhl.title.132909

Hafner, M. S., Sudman, P. D., Villablanca, F. X., Spradling, T. A., Demastes, J. W., & Nadler, S. A. (1994). Disparate rates of molecular evolution in cospeciating hosts and parasites. Science, 265(5175), 1087–1090.

Hertig, M. (1936). The Rickettsia, *Wolbachia pipientis* (gen. et sp.n.) and associated inclusions of the mosquito, *Culex pipiens*. Parasitology, 28(4), 453–486. https://doi.org/10.1017/S0031182000022666

Hochberg, F. G. (1990). "Diseases of mollusca: cephalopoda: diseases caused by protistans and metazoans," in Diseases of Marine Animals, Vol. III, Cephalopoda to Urochordata, ed. O. Kinne (Hamburg: Biologisches Anstalt Helgoland), 47–227.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources, 9(5), 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28, 1647─1649. https://doi.org/10.1093/bioinformatics/bts199.

Kopelman, N. M, Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015)

CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources, 15(5), 1179-1191. https://doi.org/10.1111/1755-0998.12387

Lamarck, J.B. (1798). Extrait d'un memoire sur les genres de la sèche, der Calmar et du Poulpe, vulgairement nomenies Polypes de mer. Bulletin de la Société philomathique de Paris, 2, 129–131.

Lapan, E. A., & Morowitz, H. J. (1975). The dicyemid mesozoa as an integrated system for morphogenetic studies. I. Description isolation and maintenance. Journal of Experimental Zoology, 193(2), 147–159. https://doi.org/10.1002/jez.1401930204

Leach, W.E. (1817). Synopsis of the orders, families and genera of the class Cephalopoda. The Zoological Miscellany; being descriptions of new or interesting animals, 3(30), 137–141. available online at http://biodiversitylibrary.org/page/28687004

Lipsitch, M., Nowak, M. A., Ebert, D., & May, R. M. (1995). The population dynamics of vertically and horizontally transmitted parasites. Proceedings of the Royal Society B: Biological Sciences, 260(1359), 321–327. https://doi.org/10.1098/rspb.1995.0099

Lühe, M. (1901). Ueber Hemiuriden. (Ein Beitrag zur Systematik der digenetischen Trematoden). Zoologischer Anzeiger. 24, 394–403, 473–488.

Marchiafava, E. & Celli, A. (1885). Nouve ricerche sulla infezione malarica, Archivo per le Scienze Mediche, Torino, 9, 311–340.

McIntosh, A. (1939). A new allocreadiid trematode, Podocotyle shawi n. sp. from the silver salmon. Journal of the Washington Academy of Sciences, 29, 379–381.

Meglécz, E., Pech, N., Gilles, A., Dubut, V., Hingamp, P., Trilles, A., Grenier, R. & Martin, J. F. (2014). QDD version 3.1: A user friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. Molecular Ecology Resources, 14 (6), 1302–1313. https://doi.org/10.1111/1755-0998.12271.

Nouvel, H. (1947). Les dicyémides 1er partie: systématique, générations vermiformes, infusorigène et sexualité. Archives de Biologie 58, 59–214.

Peakall, R., & Smouse, P. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6(1), 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics, 28(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155(2), 945–959.

Salem, H., Onchuru, T. O., Bauer, E., & Kaltenpoth, M. (2015). Symbiont transmission entails the risk of parasite infection. Biology Letters, 11(12). https://doi.org/10.1098/rsbl.2015.0840

Souidenne, D., Florent, I., Dellinger, M., Romdhane, M. S., Grellier, P., & Furuya, H. (2016). Redescription of *Dicyemennea eledones* (Wagener, 1857) (Phylum Dicyemida) from *Eledone cirrhosa* (Lamarck, 1798) (Mollusca: Cephalopoda: Octopoda). Systematic Parasitology, 93(9), 905–915. https://doi.org/10.1007/s11230-016-9659-3

Sprehn, C. G., Blum, M. J., Quinn, T. P., & Heins, D. C. (2015). Landscape genetics of Schistocephalus solidus parasites in threespine stickleback (Gasterosteus aculeatus) from Alaska. PLoS ONE, 10(4), 1–17. https://doi.org/10.1371/journal.pone.0122307

Štefka, J., Hypša, V., & Scholz, T. (2009). Interplay of host specificity and biogeography in the population structure of a cosmopolitan endoparasite: Microsatellite study of *Ligula intestinalis* (Cestoda). Molecular Ecology, 18(6), 1187–1206. https://doi.org/10.1111/j.1365-294X.2008.04074.x

Stunkard, H. W. (1954). The life history and systematic relations of the mesozoa. The Quarterly Review of Biology, 29(3), 230–244.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P., & Bouvet, J. (1996). Reliable Genotyping From Small Dna. Nucleic Acids Research, 24(16), 3189–3194. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC146079/pdf/243189.pdf

Taylor, M. J., & Hoerauf, A. (1999). Wolbachia Bacteria of Filarial Nematodes. Parasitology Today, 15(11), 437–442.

van Beneden, M. E. (1876). Recherches sur le Dicyemides, survivants actuels d'un embranchement des Mésozoaires. Bulletins de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique, 41, 1160–1205.

von Kölliker, A. (1849). Ueber *Dicyema paradoxum*, den Schmarotzer der Venenanhänge der Cephalopoden [in German]. Berichte von der Königlichen Zootomischen Anstalt zu Wurzburg, 59–66.

Whitman, C. O. (1883). A contribution to the embryology, life history, and classification of the dicyemids. Mittheilungen aus der Zoologischen Station zu Neapel, 4, 1–89.

Zhu, Y., Liu, H. Y., Yang, H. Q., Li, Y. D., & Zhang, H. M. (2017). Factors affecting genotyping success in giant panda fecal samples. PeerJ, 2017(5), 1–14. https://doi.org/10.7717/peerj.3358

Zotta G. (1912). Sur un flagelle du type Herpetomonas chez *Pyrrhocoris apterus*. Annales Scientifiques de l'Université de Jassy, 7, 211–223.

CHAPTER IV.

A perspective around cephalopods and their parasites, and suggestions on how to increase knowledge in the field.

Frontiers in Physiolgy (2018) 9, 1573

A perspective around cephalopods and their parasites, and suggestions on how to increase knowledge in the field.

Katina Roumbedakis[1,2]*, Marie Drábková[3,4,] Tomáš Tyml[3,5] and Carlo di Cristo[1]

[1]Dipartimento di Scienze e Tecnologie, Università degli Studi del Sannio, Benevento, Italy

[2]Association for Cephalopod Research, Naples, Italy

[3]Department of Parasitology, Faculty of Science, University of South Bohemia, České Budějovice, Czechia

[4]Institute of Parasitology, Biology Centre Academy of Sciences of the Czech Republic, České Budějovice, Czechia

[5]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czechia

* corresponding author

Abstract

Although interest in several areas of cephalopod research has emerged over the last decades (e.g., neurobiology, aquaculture, genetics, and welfare), especially following their 2010 inclusion in the EU Directive on the use of animals for experimental purposes, knowledge regarding the parasites of cephalopods is lacking. Cephalopods can be intermediate, paratenic, or definitive hosts to a range of parasites with a wide variety of life cycle strategies. Here, we briefly review the current knowledge in cephalopod parasitological research, summarizing the main parasite groups that affect these animals. We also emphasize some topics that, in our view, should be addressed in future research, including: (i) better understanding of life cycles and transmission pathways of common cephalopod parasites; (ii) improve knowledge of all phases of the life cycle (i.e., paralarvae, juveniles, adults and senescent animals) and on species from polar deep sea regions; (iii) exploration of the potential of using cephalopod-parasite specificity to assess population boundaries of both, hosts and parasites; (iv) risk evaluation of the potential of standard aquacultural practices to result in parasite outbreaks; (v) evaluation and description of the physiological and behavioral effects of parasites on their cephalopod hosts; (vi) standardization of the methods for accurate parasite sampling and identification; (vii) implementation of the latest molecular methods to facilitate and enable research in above mentioned areas; (viii) sharing of information and samples among researchers and aquaculturists. In our view, addressing these topics would allow us to better understand complex host-parasite interactions, yield insights into cephalopod life history, and help improve the rearing and welfare of these animals in captivity.

Keywords: Cephalopoda, parasites, pathogens, diseases, welfare

Cephalopods and Their Parasites: A Short Overview

The incidence of a given parasite in a cephalopod species depends on the presence of a potential definitive host and intermediate host(s) (in parasites with complex life cycles, i.e., those that use multiple hosts to complete their life cycle), as well as on biotic and abiotic factors (González et al., 2003). Cephalopods can be definitive hosts for protists, dicyemids, monogeneans and crustaceans, as well as intermediate or paratenic hosts for digeneans, cestodes and nematodes (summarized in Table 1; for review see also Table 1–5, Hochberg, 1990). As intermediate or paratenic hosts, cephalopods can accumulate parasites throughout their lifespan, thus increasing the chance of predation by the next host and, consequently, the probability of parasite transmission. This is especially relevant for cestodes and anisakid nematodes, which use cephalopod hosts as important vectors for transporting them to other intermediate or to definitive hosts (e.g., Pascual et al., 1995; Abollo et al., 1998; Petrić et al., 2011).

Table 1. Parasitic taxa (approximately 230 parasites identified at level species) infecting cephalopods (sorted by order) reported in the literature to date.

	Protozoa	Chromista	Dicyemida	Monogenea	Digenea	Cestoda	Acanthocephala	Nematoda	Annelida	Crustacea
Nautilida										• (1)
Spirulida		• (1)								
Sepiida		• (7)	• (31)		⚬ (2)	⚭ (6)		○ (3)		• (5)
Myopsida		• (2)	• (5)	• (2)	⚬ (3)	⚭ (9)		○ (3)	• (4)	• (5)
Oegopsida	• (1)	• (11)			⚬ (2)	⚭ (18)	⚭ (1)	○ (16)		• (4)
Octopoda		• (9)	• (59)		⚬ (3)	⚭ (2)			• (4)	• (13)
Vampyromorpha										

The role of the cephalopod host in the parasitic life cycle is indicated as: definitive •; intermediate ○; definitive, intermediate or paratenic ⚬; intermediate or paratenic ⚭; probably accidental ⚭. Number of parasite species identified for each cephalopod order is indicated in parenthesis. The current assessment reflects the original source material updated with current species information according to World Register of Marine Species (WoRMS; available at http://www.marinespecies.org/index.php).

In contrast to other molluscs, two characteristics of coleoid cephalopods (all living cephalopods besides *Nautilus* spp.) have crucial roles in their susceptibility to parasites and disease: (i) the loss of external shell, which enables the extensive neural and muscular development that allows high-speed locomotion; and (ii) the evolution of complex skin capable of sophisticated camouflage and signaling, but also prone to lesioning (Kinne, 1990). By shedding the rigid external shell of their ancestors, coleoids became more agile predators and adopted a more active lifestyle. This

likely increased the frequency of parasite transmission since, predators readily accumulate multi-host parasites that are transmitted upward through the food web (e.g., digeneans, cestodes and nematodes). Some parasites can even alter the behavior or appearance of their intermediate hosts (e.g., modifying host phenotypes) in order to increase the likelihood that they will be predated on by their definitive hosts (Lafferty, 1999; Heil, 2016), mechanisms that have yet to be explored in cephalopod hosts. In addition to the increased likelihood of transmission, the fragility of coleoid cephalopods' skin may increase the ease with which opportunistic pathogens (i.e., infection by bacteria, kinetoplastids, dinoflagellates, fungi, labyrinthulids) can invade the body (reviewed by Kinne, 1990).

To date, the most complete review of potential pathogenic agents affecting cephalopods is in "Diseases of Marine Animals" (DoMA; Kinne, 1990; chapters concerning cephalopods: Hanlon and Forsythe, 1990a,b; Hochberg, 1990). In his summary, Hochberg (1990) reported parasites for about 130 cephalopods, which represents less than a quarter of the described species at that time. Later reviews provided complementary information regarding the main viral, bacterial, fungal, parasitic, chemical and mechanical parasitic agents affecting cephalopods (see Pascual et al., 1996; Castellanos-Martínez and Gestal, 2013; Sykes and Gestal, 2014).

In the following paragraphs, we briefly overview the current knowledge on the most common parasites found in cephalopods. About 230 parasitic species of a variety of taxa (e.g., Chromista, Protozoa, Diciemyda, Monogenea, Trematoda, Cestoda, Acanthocephala, Nematoda, Annelida and Crustacea) are reported in the literature to date (Table 1 and Figure 1A). A map of the geographic distributions of cephalopod parasites is provided in Figure 1B. We emphasize that the data provided here likely over-represents tropical and temperate locations and coastal environments, since these areas are more easily and frequently sampled.

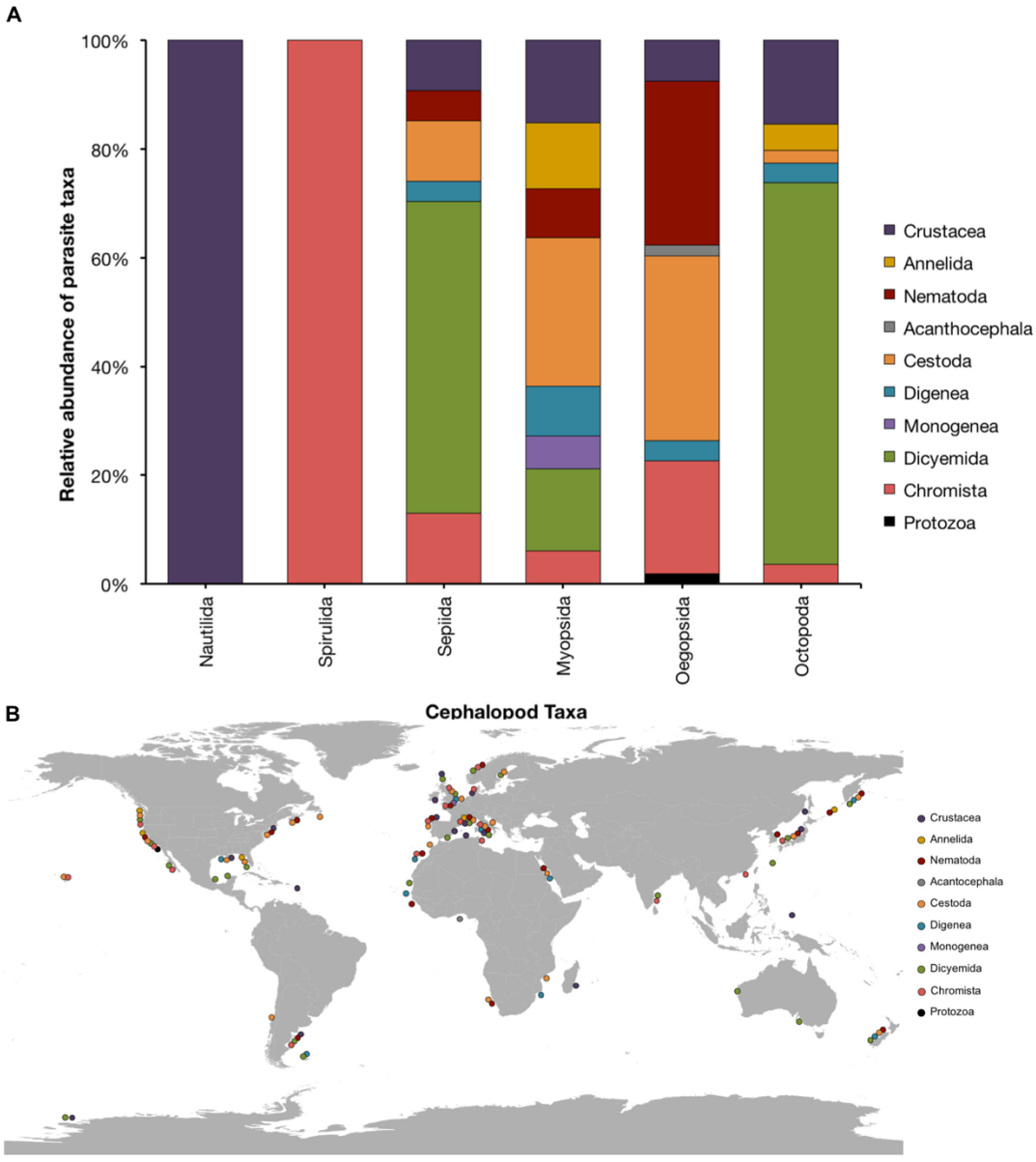


Figure 1. (A) Relative abundance of parasitic taxa affecting cephalopods. **(B)** Place of capture of the cephalopod hosts. The current assessment reflects the original source material updated with current species information according to World Register of Marine Species (WoRMS; available at http://www.marinespecies.org/index.php).

Aggregata spp.

Some of the most common parasites of cephalopods are the coccidians *Aggregata* spp. (Apicomplexa, Aggregatidae). To date, 10 species of *Aggregata* have been described parasitizing cephalopods (for review, see Gestal et al., 2010), although other (undescribed) species have also been reported (reviewed in Hochberg, 1990), so the actual diversity is likely higher. *Aggregata* spp. have complex heteroxenous life cycles, with crustaceans as intermediate hosts and cephalopods as definitive ones (Dobell, 1925; Hochberg, 1990). Most recent research (e.g., Castellanos-Martínez et al., 2013; Tedesco et al., 2017) has focused primarily on *Aggregata octopiana* and *Aggregata eberthi*, parasites of *Octopus vulgaris* and *Sepia officinalis*, respectively. This group is associated with histological and ultrastructural lesions in the digestive tract (mainly the caecum and intestine) of their cephalopod hosts (Gestal et al., 2002a), with infections of the gills, mantle, arms and mesentery also occasionally occurring (Pascual et al., 1996; Mladineo and Bočina, 2007; Tedesco et al., 2017). In addition, *Aggregata* infection can impair body growth due to "malabsorption syndrome" (Gestal et al., 2002b).

Ciliates and Dicyemids

In the renal tissue, cephalopods harbor two very unique parasitic groups, the apostome ciliates, *Chromidina* spp., and metazoans Dicyemida (= Rhombozoa). Five *Chromidina* spp. and over one hundred dicyemids have been described infecting cephalopods (Catalano, 2012; Souidenne et al., 2016). The exact impact on the hosts is still uncertain; for instance, in *O. vulgaris*, low levels of tissue abrasion caused by dicyemids could be observed by electron microscopy (Ridley, 1968), but no impact was detectable using light microscopy (Furuya et al., 2004). Consequently, these organisms may eventually come to be considered symbiotic rather than parasitic (Katayama et al., 1995; Furuya et al., 2004). Bacterial symbionts are also observed in cephalopods: for instance, the bacteria colonizing the pericardial appendage of *Nautilus* sp. (Pernice et al.,

2007; Pernice and Boucher-Rodoni, 2012) as well as the well-established association between *Euprymna scolopes* and *Vibrio fischeri* (Ruby, 1999, for review see Gerdol et al., 2018). Further studies of such symbiosis can improve not only our understanding of these complexes associations in cephalopods, but also give insights on how bacterial symbiosis occurs in mammals (Gerdol et al., 2018).

Monogeneans

A few studies have reported monogenean parasites in cephalopods (see Sproston, 1946; Palombi, 1949; Dollfus, 1958; Bychowsky, 1961). The gyrodactylid *Isancistrum subulatae* has been found in the arms and tentacles while *Isancistrum loliginis* in the mantle cavity and gills of *Alloteuthis subulata* (Llewellyn, 1984). Identifying monogeneans in cephalopods is extremely difficult due to their delicateness, small size and the thick layer of mucus in cephalopod tissues (Llewellyn, 1984), and this could be the reason for their supposed rarity. In the future, potential sites of occurrence (e.g., arms/tentacles, mantle, funnel and gills) should be thoroughly examined for a better assessment of their true prevalence.

Digeneans

The majority of information regarding digenean parasites of cephalopods is provided by Overstreet and Hochberg (1975) and Hochberg (1990), with some information added over the following decades (e.g., Shukhgalter and Nigmatullin, 2001; Nigmatullin et al., 2009), including digenean records in squid paralarvae (Vidal and Haimovici, 1999). Around 20 species have been reported from nearly 30 cephalopod hosts, usually with low prevalence of infection (Hochberg, 1990). Cephalopods do not seem to play a major role in digenean life cycles (Hochberg, 1990), though our knowledge is too limited to support this premise definitively.

Cestodes

Cephalopods are second and/or third intermediate or paratenic hosts for cestodes, acting as important vectors transporting them to other intermediate (e.g., cetaceans; Aznar et al., 2007) or definitive hosts (e.g., elasmobranchs and fishes; Hochberg, 1990). Several species have been

reported in around 60 cephalopod hosts: larval and post-larval cestodes from the orders Trypanorhyncha and Tetraphyllidea are commonly found freely in cephalopod digestive tracts, usually the stomach, caecum and intestine (Hochberg, 1990). However, they can also be found in the buccal mass (in octopus; Roumbedakis, unpublished data) or encysted in the digestive tract, mesentery and mantle cavity (Hochberg, 1990). *Phyllobotrium* spp. is the most frequently reported species (Hochberg, 1990). A general life cycle for Phyllobothriidae has recently been suggested (Klotz et al., 2018): procercoid development occurs in crustaceans (first intermediate hosts), followed by plerocercoid development in bony fish, sea turtle or squid (second intermediate host). Marine mammals can harbor both plerocercoids and merocercoids, acting as third intermediate or paratenic hosts, and sharks serve as the definitive hosts, harboring the adult parasites.

Nematodes

Larval nematodes are commonly found encysted in the viscera and musculature of cephalopods (Hochberg, 1990; Gestal et al., 1999; Abollo et al., 2001), making infected animals aesthetically unattractive for human consumption (Smith and Wootten, 1984). *Anisakis* (Anisakidae) is one of the most abundant and frequent cephalopod parasites causing important pathological effects to their hosts, such as ulceration (Abollo et al., 2001), and even castration if encysted in the gonads (Abollo et al., 1998). Transmitted through food webs, these parasites have complex life cycles involving multiple hosts: planktonic or bentho-planktonic crustaceans are the first intermediate hosts; fish and squids act as second intermediate or paratenic hosts and marine mammals (mainly cetaceans) as definitive hosts (Mattiucci and D'Amelio, 2014; Mattiucci et al., 2018). To date, a number of cephalopods (*S. officinalis*, *Ancistroteuthis lichtensteinii*, *Histioteuthis bonnellii*, *Illex coindetii*, *Todarodes sagittatus*, *T. pacificus*, *Todaropsis angolenis*, *T. eblanae*, *Nototodarus sloanii*, *Dosidicus gigas*, and *Moroteuthis ingens*) are known to be parasitized by six of the nine *Anisakis* species (*A. simplex*, *A. berlandi*, *A. nascettii*, *A. pegreffii*, *A. physeteris*, and *A. typica*) currently described (for review see Tables 2–

5, Mattiucci et al., 2018). Recent advances in anisakid biology and systematics are comprehensively summarized by Mattiucci et al. (2018). It is also worth noting that humans may also become accidental hosts if live larvae of *Anisakis* spp. are ingested through the consumption of raw or undercooked infected squid and cuttlefish. Additionally, even when ingested dead, *Anisakis* larvae can induce allergic reactions (Audicana et al., 2002; Mattiucci et al., 2013) or gastrointestinal problems (Audicana et al., 2002). Although rare, anisakiasis (the infection of a human by this parasite) is likely underdiagnosed and thus underestimated worldwide and may pose a greater threat to public health in the future (Bao et al., 2017; Mattiucci et al., 2018).

Crustaceans

Crustaceans, primarily copepods and isopods, usually parasitize the gills and mantle cavities of coleoid cephalopods (Pascual et al., 1996), but can also parasitize external surfaces, such as arms or head (Hochberg, 1990). Some attention was lately focused on tisbid copepods, parasites of deep-sea octopods. The details of the *Cholidya polypi* morphology and life cycle as well as a summary of Tisbidae infecting octopods are provided by Humes and Voight (1997), while a genus/species with an endoparasitic life stage infecting *Vulcanoctopus hydrothermalis* is described by López-González et al. (2000).

Cephalopod Parasitology: Suggestions for the Future

Despite an increase in the understanding of cephalopod parasitology during the last decades, there are still many gaps in current knowledge. Here, we briefly discuss what we believe to be the most critical issues/questions for basic and applied research that require attention.

Parasite Life Cycles and Transmission Pathways

The life cycles and transmission pathways of many cephalopod parasites are still unclear. For instance, the methods of dicyemid transmission are

completely unknown (Catalano et al., 2013), and it has been estimated that less than 5% of the life cycle of marine helminths has been fully described (Poulin et al., 2016). In the case of helminths, accurate identification of these parasites by classical methods depends on the features of adult parasites, which normally occur in vertebrates. However, the adult stages of larval helminths are frequently unknown (Aznar et al., 2007), partially due to disparity in the number of parasitological studies of invertebrates compared to vertebrates (Poulin et al., 2016). Molecular tools combined with phylogenetics can help identify trophic interactions that lead to the transmission of parasites and to a better understanding of parasite life cycles (e.g., Randhawa and Brickle, 2011). Also, our understanding of interactions between diet, feeding behavior, parasitic disease, and transmission pathways of cephalopod parasites can be improved with similar combinations of traditional approaches and modern molecular methods (e.g., Petrić et al., 2011).

Poorly Explored Life-Stages and Species From Polar and Deep Sea Regions

Most of the cephalopod parasites have been described in shallow-water species. Emerging exploration of polar and deep-sea will likely expand our knowledge about the diversity of cephalopod parasites. Similarly, the current knowledge is largely restricted to juvenile and adult cephalopod hosts, with few parasites known for paralarvae/early juveniles (Vecchione, 1987; Vidal and Haimovici, 1999) and senescent animals (Pascual et al., 2010). The extension of these limits (geographical-, life-stage-, and habitat-wise-) may be the basis for new insights into host-parasite relationships, offering important insights about the parasite diversity and complexity.

Cephalopod Parasites as Biological Tags in Population Studies

Studies of parasite distribution and host specificity can provide information about host population structure, phylogeographic distribution, migration patterns and general biology. Insights into host specificity can also help predict the likelihood of a parasite successfully establishing itself and spreading in new populations, geographical regions and hosts (Poulin and

Mouillot, 2003), a possibility which becomes increasingly important with accelerating global climate change.

Parasites are often utilized as "tags" for fisheries stock assessment, especially in small populations and limited timescales (MacKenzie, 1999; Mattiucci et al., 2015). *Anisakis* have been used as biological markers to identify sub-populations of pelagic and demersal fishes from the Mediterranean Sea (for review, see Mattiucci et al., 2015). In cephalopods, such studies are rare, mainly targeting squids (reviewed in Pascual and Hochberg, 1996; Catalano et al., 2014b). Although taxonomy within this clade is not yet well resolved (see Catalano, 2012for review), dicyemids could serve the same purpose for certain benthic cephalopods, since they are closely bound to their hosts and differ across the hosts' geographical range (Catalano et al., 2014a). Another promising taxon is *Aggregata*, which, in the Mediterranean, is differentiated into three distinct clades, potentially reflecting population differentiation of its widespread host, *O. vulgaris* (Tedesco et al., 2017).

Possible Parasite Outbreaks in Cephalopod Aquaculture

Cephalopod parasites rarely cause mortality or serious damage to wild populations. However, synergic effects between different stressors associated with captivity may favor parasites and other pathogens, making parasite outbreaks more likely in aquaculture. Coincident with the development and proliferation of aquaculture, parasites and other pathogens have proliferated (e.g., Overstreet, 1973; Lom and Dyková, 1992), many causing serious economic and environmental problems. Although our knowledge of cephalopod parasites in captivity is limited, we can extrapolate (with some caution) from knowledge obtained from other, already well-established, marine organism cultures.

In fish culture for instance, high population density is known to favor rapid spread of infections, especially those caused by parasites with direct life cycles, such as monogeneans and caligid copepods (e.g., Thoney and Hargis, 1991; Johnson et al., 2004). Both groups have already been

reported in cephalopods (e.g., Llewellyn, 1984; Pascual et al., 1996), and are thus worth monitoring particularly attentively in cephalopod aquaculture. High-density culture of hosts can also disrupt an otherwise stable parasite life-cycle scheme. For example, the myxosporeans *Enteromyxum* spp. normally alternate between two hosts (fish and annelid), but are known to be capable of direct fish-to-fish transmission in high-density conditions (Diamant, 1997). Likewise, another group of myxosporeans, *Kudoa* spp., which have been reported in wild octopus populations and are known to cause serious problems for marine fish aquaculture (Moran et al., 1999), has been suggested as a potential parasite in cephalopod culture (Yokoyama and Masuda, 2001). *Aggregata octopiana*, despite having a complex life cycle, can also impact octopus health during commercial ongrowing (Gestal et al., 2007).

In captivity, even apparently harmless symbionts, such as dicyemids and *Chromidina* spp., can become pathogens and inflict tissue damage to debilitated cephalopods (e.g., blocking the renal sacs ducts, Sykes and Gestal, 2014). At least three phylogenetically distant groups of potential eukaryotic pathogens that are capable of both a free-living and parasitic lifestyle (termed also saprophagic) can also be considered as potential pathogens of cephalopods: histophagous ciliates, known from cultured fish, crustaceans and bivalves (e.g., Cawthorn et al., 1996); amphizoic amoebae, known from cultured fish, crustaceans, bivalves and sea urchins (e.g., Dyková and Lom, 2004); and various fungal-like organisms known from cultured fish, crustaceans and molluscs (e.g., Derevnina et al., 2016). Since these pathogens are not limited by trade-offs regarding transmission or virulence because of their independent free-living stage (Kuris et al., 2014), they usually cause devastating economic impacts in aquaculture. Several ‘fungus-like organisms’ and histophagous ciliates have already been reported from cephalopods (Hanlon and Forsythe, 1990a; Tao et al., 2016) but, to date, no amphizoic amoebae have been identified.

Standardization of Parasite Sampling and Identification

Standardization of the sampling and identification methods used for cephalopods is required. Given the particular anatomy of the different cephalopod species, the publication of a guidelines, that could be used for example for parasitological and health status assessment of kept cephalopods or to determine their cause of death, would greatly facilitate research. For parasite identification, the use of classical methods (e.g., using taxonomic keys) can be extremely difficult for larval stages (Catalano et al., 2014b) or for species with high level of morphological plasticity (Poulin and Morand, 2000). In addition, some of the original parasite descriptions are not available in English (e.g., dicyemids, Nouvel, 1947, 1948; Van Beneden, 1876; Bogolepova-Dobrokhotova, 1953, 1960, 1962), are sometimes, incomplete (see Furuya, 2007), and often muddled by a variety of unresolved taxonomic and nomenclatural issues (e.g., nematodes, Smith and Wootten, 1978) which impair precise parasite identification.

The use of alternative approaches, such as search for additional morphological characters that complement classical parasite identification as suggested by Tedesco et al. (2017), the use of genetic and molecular techniques (e.g., Kopečná et al., 2006; Castellanos-Martínez et al., 2013; Souidenne et al., 2016; Tedesco et al., 2017), as well as combinations of multiple methods, is growing. Such approaches should help to better elucidate and re-evaluate the taxonomic status and host-parasite relationships, particularly where morphological plasticity might be of concern (Pascual et al., 2007). Moreover, it may clarify relationships within species complexes, such as that of *A. octopiana*infecting *O. vulgaris* in Mediterranean areas (Tedesco et al., 2017). Finally, taxonomic review of genera with morphological descriptions and molecular markers would aid research and improve assessment methods for cephalopod health and food safety in aquaculture.

The use of non- or minimally invasive methods for *in vivo* detection of cephalopod parasites should be explored in the near future. For instance, it

has been suggested that *Aggregata* infection could be diagnosed through the presence of sporocysts in the feces of living animals or through inspection of the terminal intestine by gentle retraction of the ventral mantle or by endoscopy (Sykes et al., 2017). Detection of cephalopod parasite infection using ultrasound imaging or swabbing for parasite molecular/DNA sampling might also be possible. The development of these methods would facilitate early diagnosis, ultimately preventing disease outbreaks and improving animal welfare in captivity.

Cutting Edge Molecular Methods

Transcriptomics, genomics and proteomics ("omics") are relatively new tools for understanding direct host parasite relationship on a molecular level. By enabling the study of the microbiome and metagenome of different cephalopod organs in relation to parasitic infection, the consequent pathology and immune response of hosts can be better understood (see for example Castellanos-Martínez et al., 2014a,b). Additionally, low coverage genome re-sequencing or reduced representation sequencing (RADseq methods, Davey and Blaxter, 2010) provide a tool for probing the genomic structure of populations with an unprecedented level of clarity for both host and parasites. Ultimately, such genomic information coupled with environmental data results in a "seascape genomics" approach, which can reveal both local genetic adaptations as well as the broader dynamics of gene flow (Riginos et al., 2016).

Effect of Parasites in Cephalopod Physiology and Health

Host responses to parasites may involve a variety of physiological mechanisms (e.g., neural, endocrine, neuromodulatory and immune) that can interact and alter host behavior (see review in Thompson and Kavaliers, 1994). For example, in fishes, parasitism can cause conspicuous host behavior (e.g., impaired sensory and swimming performance, increased time at water surface, etc.), increasing predation risk (Lafferty and Morris, 1996). Parasites can also affect fish performance in terms of growth and reproduction, consequently impacting their health and welfare

(Barber, 2007). Unfortunately, in cephalopods, the effects of parasitism are usually reported only at histopathological level, whereas physiological and behavioral effects are virtually unexplored. Experimental studies combining both behavioral and quantitative physiological indicators will help to better understand host-parasite systems and, hopefully, enable better assessment of cephalopod welfare. New technologies such as "omics" approaches and electron and florescent microscopy will certainly facilitate this research.

Resource Sharing

Although researchers have been able to build on previous research to some extent (e.g., through examination of collection of parasites and voucher specimens kept in museums, or gene mining in NCBI genetic database), there is much to be gained from employing a more open approach. The sharing of material through lab networks or open databases can reduce research effort and cost, maximize data use, and minimize the number of animals sampled. This is especially relevant for animals difficult to obtain, such as deep-sea cephalopods.

A database of cephalopod parasites and their cephalopod hosts available from the scientific literature, as already published for other species (e.g., Global Mammal Parasite Database, www.mammalparasites.org), possibly with extension of curated database of molecular barcodes, should be considered. In this regard, efforts are currently underway to publish a free online database of parasites and other pathogenic agents of cephalopods, the "Cephalopods' Pathogenic Agents Database (CephPAD)," which will include information on the affected tissue, anatomical-pathological findings, clinical presentation and mortality. An Atlas of Cephalopod Pathogens and Diseases is also in progress as follow-up to the activities of the COST Action FA1301. These initiatives will greatly facilitate the assessment of pathogenic agents and might facilitate early diagnosis of cephalopod pathogenic agents when they occur.

Acknowledgements

This work is considered a contribution to the COST Action FA1301 “A network for improvement of cephalopod welfare and husbandry in research, aquaculture and fisheries” (http://www.cephsinaction.org/). We would like to thank Dr. Giovanna Ponte (Chair of the COST Action FA1301) and Dr. Graziano Fiorito (Stazione Zoologica Anton Dohrn, Naples, Italy) for support and advice. We also thank Professor Lindy Holden-Dye for suggesting this contribution following the 2017 Cephs*In*Action & CIAC Meeting, “Cephalopod Science from Biology to Welfare” (Heraklion, Crete, 2017). We are also grateful to Dr. Catlin E. O’Brien for English editing of the manuscript.

References

Abollo, E., Gestal, C., López, A., González, A. F., Guerra, A., and Pascual, S. (1998). Squid as trophic bridges for parasite flow within marine ecosystems: the case of *Anisakis* simplex (Nematoda: Anisakidae), or when the wrong way can be right. *S. Afr. J. Mar. Sci.* 20, 223–232. doi: 10.2989/025776198784126575

Abollo, E., Gestal, C., and Pascual, S. (2001). Anisakis infestation in marine fish and cephalopods from Galician waters: an updated perspective. Parasitol. Res. 87, 492–499. doi: 10.1007/s004360100389

Audicana, M. T., Ansotegui, I. J., Corres, L. F., and Kennedy, M. W. (2002). Anisakis simplex: dangerous-dead and alive? Trends Parasitol. 18, 20–25. doi: 10.1016/S1471- 4922(01)02152- 3

Aznar, F. J., Agustí, C., Littlewood, D. T. J., Raga, J. A., and Olson, P. D. (2007). Insight into the role of cetaceans in the life cycle of the tetraphyllideans (Platyhelminthes: Cestoda). Int. J. Parasitol. 37, 243–255. doi: 10.1016/j.ijpara. 2006.10.010

Bao, M., Pierce, G. J., Pascual, S., González-Muñoz, M., Mattiucci, S., Mladineo, I., et al. (2017). Assessing the risk of an emerging zoonosis of worldwide concern: anisakiasis. Sci. Rep. 7:43699. doi: 10.1038/srep43699

Barber, I. (2007). Parasites, behaviour and welfare in fish. Appl. Anim. Behav. Sci. 104, 251–264. doi: 10.1016/j.applanim.2006.09.005

Bogolepova-Dobrokhotova, I. I. (1953). Materials on Dicyemids of Far-eastern Seas. Ph.D. dissertation, Academiya Nauk and Zoologicheskii Institute, Leningrad.

Bogolepova-Dobrokhotova, I. I. (1960). Dicyemidae of the far-eastern seas. I. New species of the genus Dicyema. Zool. Zhurnal 39, 1293–1302.

Bogolepova-Dobrokhotova, I. I. (1962). Dicyemidae of the far-eastern seas. II. New species of the genus Dicyemennea. Zool. Zhurnal 41, 503–518.

Bychowsky, B. E. (1961). Monogenetic Trematodes. Their Systematics and Phylogeny. Washington, DC: American Institute of Biological Sciences.

Castellanos-Martínez, S., Arteta, D., Catarino, S., and Gestal, C. (2014a). De novo transcriptome sequencing of the Octopus vulgaris hemocytes using illumina RNA-Seq technology: response to the infection by the gastrointestinal parasite Aggregata octopiana. PLoS One 9:e107873. doi: 10.1371/journal.pone.0107873

Castellanos-Martínez, S., Diz, A. P., Álvarez-Chaver, P., and Gestal, C. (2014b). Proteomic characterization of the hemolymph of Octopus vulgaris infected by the protozoan parasite Aggregata octopiana. J. Proteomics 105, 151–163. doi: 10.1016/j.jprot.2013.12.008

Castellanos-Martínez, S., and Gestal, C. (2013). Pathogens and immune response of cephalopods. J. Exp. Mar. Biol. Ecol. 447, 14–22. doi: 10.1016/j.jembe.2013. 02.007

Castellanos-Martínez, S., Pérez-Losada, M., and Gestal, C. (2013). Molecular phylogenetic analysis of the coccidian cephalopod parasites *Aggregata octopiana* and *Aggregata eberthi* (Apicomplexa: Aggregatidae) from the NE Atlantic coast using 18S rRNA sequences. *Eur. J. Protistol.* 49, 373–380. doi: 10.1016/j.ejop. 2012.11.005

Catalano, S. R. (2012). A review of the families, genera and species of dicyemida van beneden, 1876. *Zootaxa* 32, 1–32. doi: 10.11646/zootaxa.3646.1.11

Catalano, S. R., Whittington, I. D., Donnellan, S. C., and Gillanders, B. M. (2013). Using the giant Australian cuttlefish (*Sepia apama*) mass breeding aggregation to explore the life cycle of dicyemid parasites. *Acta Parasitol.* 58, 599–602. doi: 10.2478/s11686-013-0186-y

Catalano, S. R., Whittington, I. D., Donnellan, S. C., and Gillanders, B. M. (2014a). Dicyemid fauna composition and infection patterns in relation to cephalopod host biology and ecology. *Folia Parasitol.* 61, 301–310. doi: 10.14411/fp. 2014.034

Catalano, S. R., Whittington, I. D., Donnellan, S. C., and Gillanders, B. M. (2014b). Parasites as biological tags to assess host population structure: guidelines, recent genetic advances and comments on a holistic approach. *Int. J. Parasitol.* 3, 220–226. doi: 10.1016/j.ijppaw.2013.11.001

Cawthorn, R. J., Lynn, D. H., Despres, B., MacMillan, R., Maloney, R., Loughlin, M., et al. (1996). Description of Anophryoides haemophila n. sp. (Scuticociliatida: Orchitophryidae), a pathogen of American lobsters *Homarus americanus*. *Dis. Aquat. Org.* 24, 143–148. doi: 10.3354/dao024143

Davey, J. W., and Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Brief. Funct. Genomics* 9, 416–423. doi: 10.1093/bfgp/elq031

Derevnina, L., Petre, B., Kellner, R., Dagdas, Y. F., Sarowar, M. N., Giannakopoulou, A., et al. (2016). Emerging oomycete threats to plants and animals. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150459. doi: 10.1098/rstb.2015.0459

Diamant, A. (1997). Fish-to-fish transmission of a marine myxosporean. *Dis. Aquat. Org.* 30, 99–105. doi: 10.3354/dao030099

Dobell, C. C. (1925). The life history and chromosome cycle of *Aggregata eberthi* (Protozoa: Sporozoa: Coccidia). *Parasitology* 17, 1–136. doi: 10.1017/S0031182000082184

Dollfus, R. P. (1958). Copepodes, isopodes et helminths parasites de Cephalopodes de la Mediterranee et de l'Atlantique Europeen. *Faune Mar. Pyninees Orient.* 1, 61–72.

Dyková, I., and Lom, J. (2004). Advances in the knowledge of amphizoic amoebae infecting fish. *Folia Parasitol.* 51, 81–97. doi: 10.14411/fp.2004.014

Furuya, H. (2007). Redescription of two dicyemennea (Phylum: Dicyemida) From *Rossia Pacifica* (Mollusca: Cephalopoda: Decapoda). *J. Parasitol.* 93, 841–849. doi: 10.1645/GE-1003R.1

Furuya, H., Ota, M., Kimura, R., and Tsuneki, K. (2004). Renal organs of cephalopods: a habitat for dicyemids and chromidinids. *J. Morphol.* 262, 629–643. doi: 10.1002/jmor.10265

Gerdol, M., Gomez-Chiarri, M., Castillo, M. G., Figueras, A., Fiorito, G., Moreira, R., et al. (2018). "Immunity in molluscs: recognition and effector mechanisms, with a focus on bivalvia," in *Advances in Comparative Immunology*, ed. E. L. Cooper (Berlin: Springer).

Gestal, C., Abollo, E., and Pascual, S. (2002a). Observations on associated histopathology with *Aggregata octopiana* infection (Protista: Apicomplexa) in *Octopus vulgaris*. *Dis. Aquat. Org.* 50, 45–49. doi: 10.3354/dao050045

Gestal, C., De La Cadena, M. P., and Pascual, S. (2002b). Malabsorption syndrome observed in the common octopus *Octopus vulgaris* infected with *Aggregata octopiana* (Protista: Apicomplexa). *Dis. Aquat. Org.* 51, 61–65. doi: 10.3354/dao051061

Gestal, C., Guerra, A., and Pascual, S. (2007). *Aggregata octopiana* (Protista: Apicomplexa): a dangerous pathogen during commercial *Octopus vulgaris* ongrowing. *ICES J. Mar. Sci.* 64, 1743–1748. doi: 10.1093/icesjms/fsm154

Gestal, C., Pascual, S., Corral, L., and Azevedo, C. (1999). Ultrastructural aspects of the sporogony of *Aggregata octopiana* (Apicomplexa, Aggregatidae), a coccidian parasite of *Octopus vulgaris* (Mollusca, Cephalopoda) from NE Atlantic coast. *Eur. J. Protistol.* 35, 417–425. doi: 10.1016/s0932-4739(99)80051-0

Gestal, C., Pascual, S., and Hochberg, F. G. (2010). *Aggregata bathytherma* sp. nov. (Apicomplexa: Aggregatidae), a new coccidian parasite associated with a deep- sea hydrothermal vent octopus. *Dis. Aquat. Org.* 91, 237–242. doi: 10.3354/dao02264

González, A. F., Pascual, S., Gestal, C., Abollo, E., and Guerra, A. (2003). What makes a cephalopod a suitable host for parasite? The case of Galician waters. *Fish. Res.* 60, 177–183. doi: 10.1016/S0165-7836(02)00059-0

Hanlon, R. T., and Forsythe, J. W. (1990a). "Diseases of mollusca: cephalopoda: diseases caused by microorganisms," in *Diseases of Marine Animals, Vol. III, Cephalopoda to Urochordata*, ed. O. Kinne (Hamburg: Biologisches Anstalt Helgoland), 23–46.

Hanlon, R. T., and Forsythe, J. W. (1990b). "Diseases of mollusca: cephalopoda: structural abnormalitie and neoplasia," in *Diseases of Marine Animals, Vol. III, Cephalopoda to Urochordata*, ed. O. Kinne (Hamburg: Biologisches Anstalt Helgoland), 203–204.

Heil, M. (2016). Host manipulation by parasites: cases, patterns, and remaining doubts. *Front. Ecol. Evol.* 4:80. doi:10.3389/fevo.2016.00080

Hochberg, F. G. (1990). “Diseases of mollusca: cephalopoda: diseases caused by protistans and metazoans,” in *Diseases of Marine Animals, Vol. III, Cephalopoda to Urochordata*, ed. O. Kinne (Hamburg: Biologisches Anstalt Helgoland), 47–227.

Humes, A. G., and Voight, J. R. (1997). Cholidya polypi (copepoda: Harpacticoida: Tisbidae), a parasite of deep-sea octopuses in the north atlantic and northeastern pacific. *Ophelia* 46, 65–81. doi: 10.1080/00785326.1997.10432478

Johnson, S. C., Bravo, S., Nagasawa, K., Kabata, Z., Hwang, J., Ho, J., et al. (2004). A review of the impact of parasitic copepods on marine aquaculture. *Zool. Stud.* 43, 229–243.

Katayama, T., Wada, H., Furuya, H., Satoh, N., and Yamamoto, M. (1995). Phylogenetic position of the dicyemid Mesozoa inferred from 18S rDNA sequences. *Biol. Bull.* 189, 81–90. doi: 10.2307/1542458

Kinne, O. (1990). *Diseases of Marine Animals, Vol. III. Cephalopoda to Urochordata*. Hamburg: Biologisches Anstalt Helgoland.

Klotz, D., Hirzmann, J., Bauer, C., Schöne, J., Iseringhausen, M., Wohlsein, P., et al. (2018). Subcutaneous merocercoids of Clistobothrium sp. in two Cape fur seals (*Arctocephalus pusillus* pusillus). *Int. J. Parasitol. Parasites Wildl.* 7, 99–105. doi: 10.1016/j.ijppaw.2018.02.003

Kopečná, J., Jirku, M., Oborník, M., Tokarev, Y. S., Lukeš, J., and Modrý, D. (2006). Phylogenetic analysis of coccidian parasites from invertebrates: search for missing links. *Protist* 157, 173–183. doi: 10.1016/j.protis.2006.02.005

Kuris, A. M., Lafferty, K. D., and Sokolow, S. H. (2014). Sapronosis: a distinctive type of infectious agent. *Trends Parasitol.* 30, 386–393. doi: 10.1016/j.pt.2014.06.006

Lafferty, K. D. (1999). The evolution of trophic transmission. *Parasitol. Today* 15, 111–115. doi: 10.1016/S0169-4758(99)01397-6

Lafferty, K. D., and Morris, A. K. (1996). Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts. *Ecology* 77, 1390–1397. doi: 10.2307/2265536

Llewellyn, J. (1984). The biology of *Isancistrum subulatae* n.sp., a monogenean parasitic on the squid. *Alloteuthis subulata*, at Plymouth. *J. Mar. Biol. Assoc. U.K.* 64, 285–302. doi: 10.1017/S0025315400030009

Lom, J., and Dyková, I. (1992). *Protozoan Parasites of Fishes*. Amsterdam: Elsevier.

López-González, P. J., Bresciani, J., Huys, R., González, A. F., Guerra, A., and Pascual, S. (2000). Description of Genesis vulcanoctopusi gen. et sp. nov. (Copepoda: Tisbidae) parasitic on a hydrothermal vent octopod and a reinterpretation of the life cycle of cholidyinid harpacticoids. *Cah. Biol. Mar.* 41, 241–253.

MacKenzie, K. (1999). Parasites as biological tags in population studies of marine fish. *Qatar Univ. Sci. J.* 19, 117–127. doi: 10.1016/S1383-5769(98)80066-4

Mattiucci, S., Cimmaruta, R., Cipriani, P., Abaunza, P., Bellisario, B., and Nascetti, G. (2015). Integrating *Anisakis* spp. parasites data and host genetic structure in the frame of a holistic approach for stock identification of selected Mediterranean Sea fish species. *Parasitology* 142, 90–108. doi: 10.1017/S0031182014001103

Mattiucci, S., Cipriani, P., Levsen, A., Paoletti, M., and Nascetti, G. (2018). Molecular epidemiology of *Anisakis* and Anisakiasis: an ecological and evolutionary road map. *Adv. Parasitol.* 99, 93–263. doi: 10.1016/bs.apar.2017.12.001

Mattiucci, S., and D'Amelio, S. (2014). "Anisakiasis," in *Helminth Infections and Their Impact on Global Public Health*, ed. F. Bruschi (Vienna: Springer Vienna), 325–365. doi: 10.1007/978-3-7091-1782-8

Mattiucci, S., Fazii, P., De Rosa, A., Paoletti, M., Megna, A. S., Glielmo, A., et al. (2013). Anisakiasis and gastroallergic reactions associated with *Anisakis pegreffii* infection. Italy. *Emerg. Infect. Dis.* 19, 496–499. doi: 10.3201/eid1903.121017

Mladineo, I., and Boèina, I. (2007). Extraintestinal gamogony of *Aggregata octopiana* in the reared common octopus (*Octopus vulgaris*) (Cephalopoda: Octopodidae). *J. Inv. Pathol.* 96, 261–264. doi: 10.1016/j.jip.2007.05.010

Moran, J. D. W., Whitaker, D. J., and Kent, M. L. (1999). A review of the myxosporean genus Kudoa Meglitsch, 1947, and its impact on the international aquaculture industry and commercial fisheries. *Aquaculture* 172, 163–196. doi: 10.1016/S0044- 8486(98)00437- 2

Nigmatullin, C. M., Shchetinnikov, A. S., and Shukhgalter, O. A. (2009). On feeding and helminth fauna of neon flying squid *Ommastrephes bartramii* (Lesueur, 1821) (Cephalopoda: Ommastrephidae) in the southeastern Pacific. *Rev. Biol. Mar. Oceanogr.* 44, 227–235. doi: 10.4067/s0718-19572009000100023

Nouvel, H. (1947). Les dicyémides 1er partie: systématique, générations vermiformes, infusorigène et sexualité. *Arch. Biol.* 58, 59–214.

Nouvel, H. (1948). Les dicyemides 2em partie: infusoriforme, teratologie, specificite du parasitisime, affinites. *Arch. Biol.* 59, 147–223.

Overstreet, R. M. (1973). Parasites of some penaeid shrimps with emphasis on reared hosts. *Aquaculture* 2, 105–140. doi: 10.1016/0044-8486(73)90140-3

Overstreet, R. M., and Hochberg, F. G. (1975). Digenetic trematodes in cephalopods. *J. Mar. Biol. Assoc. U.K.* 55, 893–910. doi: 10.1017/S0025315400017781

Palombi, A. (1949). I trematodi d'ltalia, Parte I. Trematodi monogenetici. *Arch.Zool. Ital.* 34, 203–408.

Pascual, S., Gestal, C., Estévez, J. M., Rodríguez, H., Soto, M., Abollo, E., et al. (1996). Parasites in commercially-exploited cephalopods (Mollusca, Cephalopoda) in Spain: an updated perspective. *Aquaculture* 142, 1–10. doi: 10.1016/0044- 8486(96)01254- 9

Pascual, S., González, A., Arias, C., and Guerra, A. (1995). Helminth infection in the short-finned squid *Illex coindetii* (Cephalopoda, Ommastrephidae) off NW Spain. *Dis. Aquat. Org.* 23, 71–75. doi: 10.1645/GE-2562.1

Pascual, S., González, A., and Guerra, A. (2007). Parasites and cephalopod fisheries uncertainty: towards a waterfall understanding. *Rev. Fish Biol. Fish.* 17, 139–144. doi: 10.1007/s11160-006-9021-y

Pascual, S., González, A. F., and Guerra, A. (2010). Coccidiosis during octopus senescence: preparing for parasite outbreak. *Fish. Res.* 106, 160–162. doi: 10.1016/j.fishres.2010.05.013

Pascual, S., and Hochberg, F. G. (1996). Marine parasites as biological tags of cephalopod hosts. *Parasitol. Today* 12, 324–327. doi: 10.1016/0169-4758(96)40004- 7

Pernice, M., and Boucher-Rodoni, R. (2012). Occurrence of a specific dual symbiosis in the excretory organ of geographically distant Nautiloids populations. *Environ. Microbiol. Rep.* 4, 504–511. doi: 10.1111/j.1758-2229.2012.00352.x

Pernice, M., Wetzel, S., Gros, O., Boucher-Rodoni, R., and Dubilier, N. (2007). Enigmatic dual symbiosis in the excretory organ of *Nautilus macromphalus* (Cephalopoda: Nautiloidea). *Proc. R. Soc. B Biol. Sci.* 274, 1143–1152. doi:10.1098/rspb.2006.0353

Petric´,M., Mladineo, I., and Šifner, S.K.(2011). Insight into the short-finned squid *Illex coindetii* (Cephalopoda: Ommastrephidae) feeding ecology: is there a link between helminth parasites and food composition? *J. Parasitol.* 97, 55–62. doi: 10.1645/GE-2562.1

Poulin, R., Blasco-Costa, I., and Randhawa, H. S. (2016). Integrating parasitology and marine ecology: seven challenges towards greater synergy. *J. Sea Res.* 113, 3–10. doi: 10.1016/j.seares.2014.10.019

Poulin, R., and Morand, S. (2000). The diversity of parasites. *Q. Rev. Biol.* 75, 277–293. doi: 10.1086/393500

Poulin, R., and Mouillot, D. (2003). Parasite specialization from a phylogenetic perspective: a new index of host specificity. *Parasitology* 126, 473–480. doi: 10.1017/s0031182003002993

Randhawa, H. S., and Brickle, P. (2011). Larval parasite gene sequence data reveal cryptic trophic links in life cycles of porbeagle shark tapeworms. *Mar. Ecol. Prog. Ser.* 431, 215–222. doi: 10.3354/meps09108

Ridley, R. K. (1968). Electron microscopic studies on Dicyemid mesozoa. I. Vermiform Stages. *J. Parasitol.* 54, 975–998. doi: 10.2307/3277133

Riginos, C., Crandall, E. D., LIggins, L., Bongaerts, P., and Treml, E. A. (2016). Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Curr. Zool.* 62, 581–601. doi: 10.1093/cz/zow067

Ruby, E. G. (1999). The euprymna scolopes-*Vibrio fischeri* symbiosis: a biomedical model for the study of bacterial colonization of animal tissue. *J. Mol. Microbiol. Biotechnol.* 1, 13–21.

Shukhgalter, O. A., and Nigmatullin, C. M. (2001). Parasitic helminths of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in open waters of the central east Pacific. *Fish. Res.* 54, 95–110. doi: 10.1016/s0165-7836(01)00374-5

Smith, J. W., and Wootten, R. (1978). *Anisakis* and anisakiasis. *Adv. Parasitol.* 16, 93–163. doi: 10.1016/S0065-308X(08)60573-4

Smith, J. W., and Wootten, R. (1984). "Anisakis larvae (herringworm) (nematoda) in fish. Parasitose des poissons par les larves du nématode *Anisakis*," in *Fiches d'identification des Maladies et Parasites des Poissons, Crustaces et Mollusques*, ed. C. J. Sinderman (Silver Spring, MD: NOAA National Marine Fisheries Service), 2–4.

Souidenne, D., Florent, I., Dellinger, M., Justine, J., Romdhane, M. S., Furuya, H., et al. (2016). Diversity of apostome ciliates, *Chromidina* spp. (Oligohymenophorea, Opalinopsidae), parasites of cephalopods of the Mediterranean Sea. *Parasite* 23:33. doi: 10.1051/parasite/2016033

Sproston, N. G. (1946). A synopsis of the monogenetic trematodes. *Trans. Zool. Soc. Lond.* 25, 185–600.

Sykes, A. V., Almansa, E., Cooke, G. M., Ponte, G., and Andrews, P. L. (2017). The digestive tract of cephalopods: a neglected topic of relevance to animal welfare in the laboratory and aquaculture. *Front. Physiol.* 8:492. doi: 10.3389/ fphys.2017.00492

Sykes, A. V., and Gestal, C. (2014). "Welfare and diseases under culture conditions," in *Cephalopod Culture*, ed. R. Villanueva (Berlin: Springer), 479–489. doi: 10. 1007/978- 94- 017- 8648- 5_6

Tao, Z., Liu, L., Chen, X., Zhou, S., and Wang, G. (2016). First isolation of *Miamiensis avidus* (Ciliophora: Scuticociliatida) associated with skin ulcers from reared pharaoh cuttlefish *Sepia pharaonis*. *Dis. Aquat. Org.* 122, 67–71. doi: 10.3354/dao03067

Tedesco, P., Gestal, C., Begic´, K., Mladineo, I., Castellanos-Martínez, S., Catanese, G., et al. (2017). Morphological and molecular characterization of aggregata spp. Frenzel 1885 (Apicomplexa: Aggregatidae) in *Octopus vulgaris* Cuvier 1797 (Mollusca: Cephalopoda) from Central Mediterranean. *Protist* 168, 636–648. doi: 10.1016/j.protis.2017.08.002

Thompson, S. N., and Kavaliers, M. (1994). Physiological bases for parasite-induced alterations of host behaviour. *Parasitology* 109, S119–S138. doi: 10.1017/S0031182000085139

Thoney, D. A., and Hargis, W. J. Jr. (1991). Monogenea (Platyhelminthes) as hazards for fish in confinement. *Annu. Rev. Fish Dis.* 1, 133–153. doi: 10.1016/0959- 8030(91)90027- H

Van Beneden, M. E. (1876). Recherches sur le Dicyemides, survivants actuels d'un embranchement des Mésozoaires. *Bull. Acad. R. Sci. Lett. Beaux Arts Belgique* 41, 1160–1205.

Vecchione, M. (1987). "Juvenile ecology," in *Cephalopod Life Cycles*, ed. P. R. Boyle (London: Academic Press), 61–84.

Vidal, E. A. G., and Haimovici, M. (1999). Digestive tract parasites in rhynchoteuthion squid paralarvae, particularly in *Illex argentinus* (Cephalopoda: Ommastrephidae). *Fish. Bull.* 97, 402–405.

Yokoyama, H., and Masuda, K. (2001). *Kudoa* sp. (Myxozoa) causing a post-mortem myoliquefaction of North-Pacific giant octopus *Paroctopus dofleini* (Cephalopoda: Octopodidae). *Bull. Eur. Assoc. Fish Pathol.* 21, 266–268 *Fish Pathol.* 21, 266–268.

SUMMARY

1 Summary of results

This thesis brings new information into the area of invertebrate biology and tries to widen our knowledge about life-history traits, population structure, and phylogenetic position of two enigmatic clades of tiny marine parasites, Dicyemida and Orthonectida.

The first chapter explores the phylogenetic placement of both clades based on the phylogenomic data analyses and examines the impact of the selection of currently used methods. In all our analyses, both Dicyemida and Orthonectida were found to be firmly placed inside Lophotrochozoa and almost always joined together in monophyletic Mesozoa. However, the relationships between lophotrochozoan phyla were found to be unstable and dependent on the phylogenetic method used (e.g., Maximum likelihood, Bayesian inference, coalescence). In most of our analyses, based on a carefully constructed dataset with almost 1000 genes, Lophotrochozoa split into two main groups: Platyzoa *s.l.* (Gnathostomulida, Micrognathozoa+Rotifera, Gastrotricha, Platyhelminthes, Entoprocta+Cycliophora, and Bryozoa) and Trochozoa (Mollusca, Annelida, Nemertea, and Phoronida+Brachiopoda), with Dicyemida and Orthonectida as a monophyletic Mesozoa clade placed in Platyzoa. The exceptions, where Dicyemida and Orthonectida were seen as separate clades were the results of some of the Bayesian inference analyses (BI) with reduced representation of genes and one ML likelihood analysis with profile mixture model (similar model to CAT+GTR in BI) on 50 gene matrix, which showed similar result to BI. In BI, Dicyemida stayed in Platyzoa but Orthonectida were found in multiple positions across the trees depending on the matrix used (reduced matrices differed in the number of genes used and their characteristics). BI approach with the complex CAT+GTR model has, compared to ML and coalescent, the disadvantage of high computational demands and often does not converge on larger datasets, which limits its use to only dozens of genes. This was also generally true for our dataset where matrices with more than 50 genes did not converge on a single solution and showed variable results, unfortunately, especially regarding the position of mesozoan clades

(usually orthonectids). A similar discrepancy in results was also shown in phylogenomic datasets in the case of the Ctenophora vs. Porifera first controversy (Whelan et al. 2017; Simion et al. 2017). Discussed reasons for this discrepancy were mainly the different models of sequence evolution used by BI and maximum likelihood approaches and phylogenetic inference approaches in general but also dissimilar taxon sampling and quality of matrices (Whelan et al. 2017; Simion et al. 2017).

In a published phylogenomic study, mesozoan clades were found to be monophyletic and close to Gastrotricha and/or Platyhelminthes (Lu et al., 2017). In contrast, in the study based mostly on mitochondrial genes by Schiffer et al. (2018) Mesozoa was not found monophyletic, because Orthonectida was associated with Annelida, but Dicyemida was placed somewhere in Lophotrochozoa with uncertain affiliation. In another phylogenomic study by Zverkov et al. (2019) Mesozoa was also found not monophyletic, because Orthonectida was placed in Annelida and Dicyemida close to Platyhelminthes or other lophotrochozoan clades. This study also showed a strong dichotomy between BI and ML results regarding the position of mesozoan clades. Contrastingly, in the ML approach, Mesozoa was found monophyletic. Nonetheless, the authors strongly argue for taking into account only the BI results.

So, are Mesozoa monophyletic? What is their precise position in the Tree of Life? Results of phylogenomic studies narrowed down the options but do not seem to provide a definitive answer. These incongruous views on Mesozoan phylogeny still leave us in suspense, regarding which of the suggested scenarios is closer to the true evolutionary pathway, similarly to the Porifera vs. Ctenophora first controversy. Our results revive the classic monophyletic Mesozoa hypothesis and support their position as near Platyhelminthes or Gnathifera as part of the wider Platyzoa *s.l.* group. The key improvements that seem necessary to move the field of lophotrochozoan phylogeny forward are wide and careful taxon sampling, correct handling of the data (e.g., orthology assignment, contamination checks), and further testing of the robustness of phylogenomic methods.

The second and third chapters demonstrate the use of population studies on dicyemids. The second chapter compares the population structure of dicyemids from *Sepia officinalis* (Common Cuttlefish) with the population structure of their host, based on cytochrome oxidase gene (COI) sequences. Population structuring of organisms is affected by their motility (either as adults or larvae) and/or by their dependency on a specific habitat and its fragmentation. In general, due to the homogeneity of the marine environment, organisms living there are expected to have little structured populations. However, the actual population structure in individual cases can be unpredictable. In our study, Common cuttlefish shows substantial population structure within its range and surprisingly high genetic diversity in the seas surrounding Sardinia. Dicyemids mostly follow the pattern of their host and show the highest diversity in the same area but not to the level of individual hosts, suggesting that co-divergence of host and dicyemid haplotypes might not be as strong on the locality.

The third chapter focuses on the dicyemid population inside one host (dicyemid infrapopulation) employing metrics used in population genetics to reveal dicyemid population structure and relatedness in one host. This information can be then used to infer information about dicyemid reproduction strategies and transmission. Specifically, the study is based on eight novel dicyemid microsatellite loci and focuses on *Dicyema moschatum* inhabiting renal organs of *Eledone moschata* octopuses, sampled from two localities from two seas. Pula, Croatia represented the Adriatic Sea, and Naples, Italy, the Tyrrhenian Sea. Results showed that dicyemids inside one host are not clones, which points to the occurrence of sexual reproduction in the renal organs of its host. Dicyemid infrapopulations from the two localities were distinct, however, the sampling was too sparse to assess the overall connectivity of dicyemid infrapopulations. This study showed us that population genetic methods, such as microsatellite studies and genetic diversity statistics, may be useful in revealing additional information about life-history traits in tiny enigmatic animals such as dicyemids.

The fourth chapter summarizes what is currently known about the parasites of cephalopods and points out possible future research challenges in the area. It includes information about dicyemids and their unique position amidst other cephalopod parasites. Therein described features of dicyemids are their close association with their hosts and their high prevalence. This chapter also states that future directions in dicyemid research (and cephalopod parasites in general) should be directed into an exploration of their diversity (virtually unknown in dicyemids), systematics, elucidation of lifecycles, parasite/host coevolution and population structure and the extent of morbidity (harmfulness) of the parasites. All of the above mentioned can be useful in cephalopod stock management and aquaculture, but the last mentioned may be the most.

Overall, this thesis brings important pieces of information in the area of phylogeny, population structure, and life history for little known but very interesting marine parasites and hopefully helps us fill the gaps in the big biological puzzle called the Tree of Life.

2 Future perspectives

Despite the new pieces of information about the evolution and life of mesozoans, they still remain rather enigmatic. Especially orthonectids are very rarely seen due to their low prevalence. This unpredictability makes it difficult to collect samples for further studies. However, it also presents an opportunity for new species discovery in less known host species (e.g., the discovery of *Rhopalura xenoturbellae*, Nakano & Miyazawa, 2019).

Compared to orthonectids, dicyemids can be relatively easily obtained from a range of benthic cephalopods. However, the species boundaries in dicyemids and their distributions are not clear. Integration of molecular methods and classic morphology is crucial for a better understanding of both dicyemid and orthonectid systematics.

A promising way of increasing insight into mesozoan diversity would be a comparison of an assemblage of parasitic species in different hosts across

the world seas through analysis of genetic markers. Specifically, a usage of universal primers coupled with next-gen sequencing (alias amplicon sequencing) could provide interesting clues into mesozoan worldwide diversity, species distributions, and host-parasite interactions.

Improved molecular methods targeting dicyemids may also be potentially used for the development of tools for cephalopod stock assessment in fisheries and aquaculture. Parasites are sometimes used in fisheries as 'tags' to ascertain the origin of a given fish and dicyemids would be a good candidate for the same purpose in cephalopods (Catalano et al., 2014). If parasite species distributions are known, then the origin of an unknown host individual (for example a suspicious octopus seized on the market) can be determined. With climate change that affects marine ecosystems, the interest in research concerning cephalopods intensifies. As cephalopods are short-lived and opportunistic, their populations can react quickly to the changing environment. Diminishing of fish stocks and increase in octopus populations cause changes to ecosystems but may also drive the change in human consumption of cephalopod products, which has been already observed in France. Dicyemid tags may prove to be a useful tool in the protection of the marine environment and stock assessment in our changing seas.

Even though orthonectid and dicyemid phylogenetic position was examined with modern phylogenomic methods, their precise position in Lophotrochozoa and their sister group/s has not been pinpointed with certainty. Further phylogenomic studies with increased taxon sampling (with the inclusion of Chaetognatha in particular), careful data curation, and management will be of utmost importance for elucidation of the evolution of Lophotrochoza. The position of mesozoa relative to other lophotrochozoan groups is essential for the recovery of a reliable phylogenetic backbone of Lophotrochoza, which is a key for further research into the evolution of invertebrates.

In conclusion, after decades of research into their life histories, evolution, and genetics, orthonectids and dicyemids, still retain a few secretes and

remain mysterious. From the story of mesozoan research, we can learn that some questions are not easily answered, even with modern methods and that sometimes the answers are not as simple as expected. However, it also shows us that there are things unexplored out there and there still could be a joy found in pursuing the natural history of organisms. May there always be discoveries to be made and seas to explore.

"The Sea, once it casts its spell, holds one in its net of wonder forever."

Jacques – Yves Cousteau

3 References

Catalano, S. R., Whittington, I. D., Donnellan, S. C., & Gillanders, B. M. (2014). Parasites as biological tags to assess host population structure: Guidelines, recent genetic advances and comments on a holistic approach. International Journal for Parasitology: Parasites and Wildlife, 3(2), 220–226. https://doi.org/10.1016/j.ijppaw.2013.11.001

Lu, T. M., Kanda, M., Satoh, N., & Furuya, H. (2017). The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. Zoological Letters, 3(1), 1–9. https://doi.org/10.1186/s40851-017-0068-5

Nakano, H., & Miyazawa, H., (2019). A new species of Orthonectida that parasitizes Xenoturbella bocki: Implications for studies on Xenoturbella. Biological Bulletin, 236, 66–73. Doi:10.1086/700834

Schiffer, P. H., Robertson, H. E., & Telford, M. J. (2018). Orthonectids Are Highly Degenerate Annelid Worms. Current Biology, 28(12), 1970-1974.e3. https://doi.org/10.1016/j.cub.2018.04.088

Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D. J., Di Franco, A., Roure, B., Satoh, N., Quéinnec, E., Ereskovsky, A., Lapébie, P., Corre, E., Delsuc, F., King, N., Wörheide, G., & Manuel, M. (2017). A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals. Current Biology, 27(7), 958–967. https://doi.org/10.1016/j.cub.2017.02.031

Whelan, N. V., Kocot, K. M., Moroz, T. P., Mukherjee, K., Williams, P., Paulay, G., Halanych, K. M. (2017). Ctenophore relationships and their placement as the sister group to all other animals. Nature Ecology and Evolution, 1(11), 1737–1746. https://doi.org/10.1038/s41559-017-0331-3

Zverkov, O. A., Mikhailov, K. V., Isaev, S. V., Rusin, L. Y., Popova, O. V., Logacheva, M. D., Penin, A. A., Moroz, L. L., Panchin, Y. V., Lyubetsky, V. A., & Aleoshin, V. V. (2019). Dicyemida and Orthonectida: Two stories of body plan simplification. Frontiers in Genetics, 10(443). https://doi.org/10.3389/fgene.2019.00443

Author’s Curriculum vitae

Author's Curriculum vitae

Name: Mgr. Marie Drábková
Date of Birth: 06/09/1989
Nationality: Czech
Address: Svechynova 941, Chrudim, 537 01, Czech Republic
Tel: +420 731 301 581
E-mail: marie.drabkova7@gmail.com
ORCID: 0000-0003-1389-3605

Education

2013–now	PhD candidate in Parasitology. University of South Bohemia in Ceske Budejovice, Faculty of Science.
2011–2013	Mgr. in Biology of Ecosystems. University of South Bohemia in CB, Faculty of Science.
2008–2011	Bc. in Biology. University of South Bohemia in CB, Faculty of Science.
courses	
2013	Winterschool, Basics in biostatistics and computational biology (Brno, Czech republic)
2014	Workshop on Genomics Český Krumlov
2014	Workshop on Phylogenomics (Lund, Sweden)
2015	Workshop on Molecular Evolution Český Krumlov

Work experience

2013–2019	Research assistant, Laboratory of Molecular Ecology, Institute of Parasitology, Biology Centre CAS
2013–now	PhD student, Biology Centre CAS, Institute of Parasitology
2015–2018	Research assistant, Department of Parasitology, Faculty of Science, University of South Bohemia
2012–2013	Research assistant, Department of Biology of Ecosystems, Faculty of Science, University of South Bohemia

Internship

February – April 2017

Kocot Lab, Department of Biological Sciences, University of Alabama, Tuscaloosa, US
Supervisor: Kevin M. Kocot, Ph.D

Field experience

Leadership of field surveys of dicyemids in available cephalopod hosts (e.g. *Eledone moschata*, *Octopus vulgaris*, *Sepia officinalis*, *Callistoctopus macropus*) in Portugal, Croatia and Italy (2015-2017). Research stays at Stazione Zoologica Naples (2017, Italy) and IAS, Oristano (2016, Sardinia, Italy) exploring infra-population characteristics of dicyemids. Both carried out during project: Phylogenomics and Molecular diversity of Mesozoa (project no. GACR 15-08717S).

Teaching

2017-2018 teaching assistant, Molecular ecology, Faculty of Science, University of South Bohemia
2015 teaching assistant, Field course of marine biology, Faculty of Science, University of South Bohemia

supervision of theses:
2018 Bc. Nikola Jachníkova
Porovnání populační struktury parazita a hostitele na systému Sépie obecná - Dicyemida (in czech), Faculty of Science, University of South Bohemia
2019 Bc. Daniela Kotalová
Populační struktura chobotnice *Eledone moschata* ve Středozemním moři na základě RADseq dat (in czech), Faculty of Science, University of South Bohemia
2019 Bc. Tereza Flegrová
Analýza populační struktury *Dicyema moschatum* v hostiteli Eledone moschata (in czech), Faculty of Science, University of South Bohemia

Publications

Drábková M., Jachníková N., Tyml T., Sehadová H., Ditrich O., Myšková E., Hypša V., Štefka J. (2019) Population co-divergence in common cuttlefish (Sepia officinalis) and its dicyemid parasite in the Mediterranean Sea. Scientific Reports 9 : 14300. DOI: 10.1038/s41598-019-50555-9

Roumbedakis K. , **Drábková M**., Tyml T., di Cristo C. (2019) A Perspective Around Cephalopods and Their Parasites, and Suggestions on How to Increase Knowledge in the Field. Frontiers in Physiology 9 : 1573. DOI: 10.3389/fphys.2018.01573

Drábková, M., Flegrová, T., Myšková, E., Hypša, V., Štefka, J. (2021). Genetic analysis of dicyemid infrapopulations suggests sexual reproduction and host colonization by multiple individuals is common. Organisms Diversity & Evolution. 21. 1-10. 10.1007/s13127-021-00493-0.

International conferences

Krausová M., Husník, F., Štefka J., 2017 Genome of Dicyemids, enigmatic parasites of cephalopods assembling the genome of non-model organism with focus on host parasite separation [poster]. February 21st – February 23th 2017, BioGenomics2017, Washington D.C., US.

E-mail: marie.drabkova7@gmail.com

Evolution and Phylogeny of Mesozoa
Ph.D. Thesis Series, 2022, No. 10

Printed in the Czech Republic by Typodesign
Edition of 10 copies

University of South Bohemia in České Budějovice
Faculty of Science
Branišovská 1760
CZ-37005 České Budějovice, Czech Republic

Phone: +420 387 776 201
www.prf.jcu.cz, e-mail: sekret-fpr@prf.jcu.cz