Hidden diversity of two intracellular parasites, Microsporidia and *Rickettsia*, in New Zealand amphipod hosts: patterns and causes

Eunji Park

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago
Dunedin, New Zealand
December 2020

Abstract

Intracellular parasites, endosymbionts that are specialised to live in host cells, have evolved independently in many different lineages. Endosymbionts can be transmitted vertically and horizontally, and many of them maintain intimate associations with their hosts over space and time. For example, the mitochondrion, which is now an essential organelle in eukaryotes, is believed to have an endosymbiotic origin.

This thesis focuses on two intracellular parasite groups, Microsporidia and *Rickettsia*, in diverse amphipod hosts in New Zealand. Microsporidia and *Rickettsia* are phylogenetically distantly related, but both are ecologically and evolutionary successful in diverse groups of hosts. I aimed to understand their diversity, the factors that have shaped their current distribution, the various phylogenetic patterns they are involved in, their spatiotemporal variations and cooccurrence, and their evolutionary histories.

By starting with molecular screening of diverse amphipods collected throughout the country, I uncovered a diversity that was previously unknown in this region, a first for both groups in the Southern Hemisphere. For Microsporidia, by expanding their known geographical and host ranges, I applied phylogenetic and cophylogenetic methods to infer the evolutionary history of host-parasite associations. Based on the congruent phylogenetic and phylogeographical patterns, I provided evidence for their shared evolutionary histories.

Some vertically transmitted parasites can manipulate host reproduction, and this can have various ecological and evolutionary consequences. Certain endosymbionts are known to disrupt phylogenetic patterns of mitochondrial DNA mainly due to linkage disequilibrium. I discussed the direct and indirect impacts of *Rickettsia* infections on the use of mitochondrial DNA in barcoding, phylogenetic, and phylogeographical studies. Also, I explored the possible role of *Rickettsia* infections in accelerating host mitochondrial DNA evolution, which could result in mitonuclear discordance patterns appearing on deep time scales.

Because Microsporidia and *Rickettsia* share the same amphipod hosts and coexist in many populations, they may interact with each other within the same individual. I investigated spatiotemporal dynamics in the prevalence of both parasites, and tested whether infection by one parasite influences the probability of infection by the other. Although there were variations in prevalence among different sampling times and locations, there were no clear consistent patterns between the two parasites and their patterns of co-occurrence within the same individual hosts did not depart from random expectations.

Taken together, Microsporidia and *Rickettsia* are widespread in New Zealand amphipod hosts. It seems that both vertical and horizontal transmission have played important roles in their current distribution. The vertical transmission of endosymbionts seems to have great potential to cause profound effects on host mitochondrial DNA. However, demonstrating this will require more data from multiple systems and scales. In the future, population- to community-level studies will be especially valuable to understand the ecological interactions between hosts and parasites and among different strains/species/groups of parasites.

Acknowledgements

First of all, I would like to express my gratitude to my supervisor Robert Poulin for providing me with a great opportunity to study and live in an amazing place. Also, I thank him for his accessibility, prompt feedback, generosity, and great empathy. He allowed me to follow my intellectual curiosity, but provided timely suggestions that guided my study in the right direction. I also thank Jon Waters for his encouragement and answering my questions about New Zealand phylogeography. I appreciate Christoph Matthaei for serving as a committee member and providing advice and support, along with Jon Waters.

The happy environment contributed by staff and students in the Zoology department made my PhD more enjoyable. I especially thank Tania King for her excellent management of the genetics lab and all her technical help in the lab. I appreciate Nat, Nicky, Wendy, and Paulin for their administrative and technical support.

I thank Sophie Courjal and Eleanor Wainwright for helping during my big field trips that made my studies in this thesis possible, and Colin Décout for helping with lab work. I appreciate Dr. Jim Lowry's generous help in identifying Talitridae species.

I have been privileged to meet great people from many countries who visited and joined the Evolutionary and Ecological Parasitology Research Group of the University of Otago. I appreciate Fátima Jorge for her helpful advice and for being a great mentor. I also thank Jerusha Bennett for her amazing and kind support and encouragement during the whole time of my PhD. I also thank Bronwen Presswell for her kind encouragement and providing parasite specimens. I appreciate all the friends who made great memories with me during fun activities and trips: Marin Milotic, Antoine Filion, Laurie Emma, Alan Eriksson, Amanda, Daniela de Angeli Dutra, and Xuhong Chai. I also thank Christian Selbach and Brandon Ruehle who welcomed me when I joined the lab and for helping during my first fieldwork.

I appreciate all the people who gave me precious memories during my PhD in Dunedin. I appreciate my previous flat mates, especially Elizabeth Yarnall. I express my thanks to Moony Kim for her friendship. I also appreciate Heloise Pavanato for her friendship and moral support. She was willing to help me to collect amphipod specimens during our holiday trip, which led to me to detect the first *Rickettsia*-infected amphipod specimens.

I express my thanks to my friends in Korea for their friendship and encouragement. I especially thank Eun-Jung Yoon for our long-night talks. I thank my Master's supervisor Prof. Yong-Jin Won, who introduced me to the field of molecular ecology during my undergraduate studies, for his support throughout my early scientific career.

I appreciate my partner Jeff Doherty for his friendship and affection. Thanks to him, I made lots of exciting and great memories in New Zealand. I thank him for reading my manuscripts and providing helpful comments.

Finally, I thank my lovely and wise mother, who has always been supportive of my decisions. Also, I thank my aunties and my grandmother for their love and affection.

Table of contests

Chapter 1

General Introduction

1.1. Microorganisms living inside host cells

Parasitism is a pervasive mode of life that has evolved countless times independently within diverse lineages of life (Poulin, 2011). Some parasites are large enough to be visible, like helminths (nematodes, cestodes, and trematodes), insects, and mites. On the other hand, some are microscopic and cannot be seen by the naked eye; these include viruses, bacteria, fungi, and protists. The former are generally called 'macroparasites' and the latter '**microparasites**' (Anderson and May, 1981). Among microparasites, some can live inside the host cells; these are generally referred to as '**intracellular parasites**.' These intracellular parasites obtain (or 'steal') resources from the host cell for their own growth and reproduction, often causing mild to lethal effects on the host organism. Some intracellular parasites are highly dependent on their host for survival and reproduction, and therefore cannot survive outside of the host cell; these '**obligate intracellular parasites**' differ from 'facultative intracellular parasites' that are capable of surviving and reproducing outside of host cells. In the scientific literature, different terminology is sometimes used. For instance, '**endosymbiont**' refers to organisms that live inside another organism, whether they are intracellular or extracellular. Since 'symbiosis' can be defined as an intimate and long-term interaction between two organisms and includes mutualism, commensalism, and parasitism (Leung and Poulin, 2008), endosymbionts therefore include intracellular parasites.

Whether parasitic or not, intracellular endosymbionts exert a strong influence on their hosts, over both ecological and evolutionary time scales. The best-known example of intracellular endosymbiosis is probably that of **mitochondria** and plastids such as chloroplasts (Archibald, 2015; Dyall et al., 2004). Mitochondria and chloroplasts are essential organelles; mitochondria produce ATP and are found in almost all eukaryotes, whereas chloroplasts fix nitrogen and provide energy essential to plants (Brown, 1992; Jasid et al., 2006). Extensive studies have been conducted to elucidate the origin of these organelles and their symbiotic history with host cells (Gray, 1989; Roger et al., 2017). It is now widely accepted that a once free-living bacterium was engulfed by Asgard archaea, an archaeal group with some eukaryotic characteristics, and the resulting fusion of cells probably became the first eukaryotic cell (Zaremba-

Niedzwiedzka et al., 2017). Another line of evidence indicates that the proto-mitochondrion, the hypothetical common ancestor of all mitochondria, is related to alphaproteobacteria, now considered either a sister group or a lineage within the Rickettsiales (Andersson et al., 1998; Wang and Wu, 2015). Without intracellular endosymbionts, life on earth would have been very different from the current one.

The overall goal of this thesis is to explore the diversity, evolutionary history, and spatio-temporal ecology of host-parasite interactions between amphipods and two groups of intracellular symbionts. Also, I explore their possible role in shaping host mtDNA evolution. In the sections that follow, I provide the conceptual background underpinning the thesis, followed by an overview of the model taxa and a summary of my key objectives.

1.1. Mode of transmission and virulence

There are two modes of transmission for intracellular endoparasites: horizontal and vertical (Figure 1.1; Lipsitch et al., 1995; Poulin, 2011). **Horizontal transmission** is a common mode of transmission that occurs between different organisms of the same or different species. This may occur trophically (by eating an infected individual) or through the environment (by being exposed to the parasite's infective stages). Horizontal transmission typically involves high virulence (severe pathology, decreased fecundity, and high mortality), because growth and high replication of parasites (for a better chance of transmission) require the use of host resources (Stewart et al., 2005). On the other hand, **vertical transmission**, i.e. the direct transmission of parasites from a parent to offspring, is less common but has been reported in bacteria, viruses, protists, and helminths, as well as microsporidians (Dunn et al., 2001). Vertical transmission can occur via germ cells, milk, or through the placenta (Barthel et al., 2013; Quicke et al., 2016). The transmission success of vertically transmitted parasites is totally dependent on the host successfully producing offspring, thereby explaining the low virulence associated with this transmission mode (Dunn and Smith, 2001).

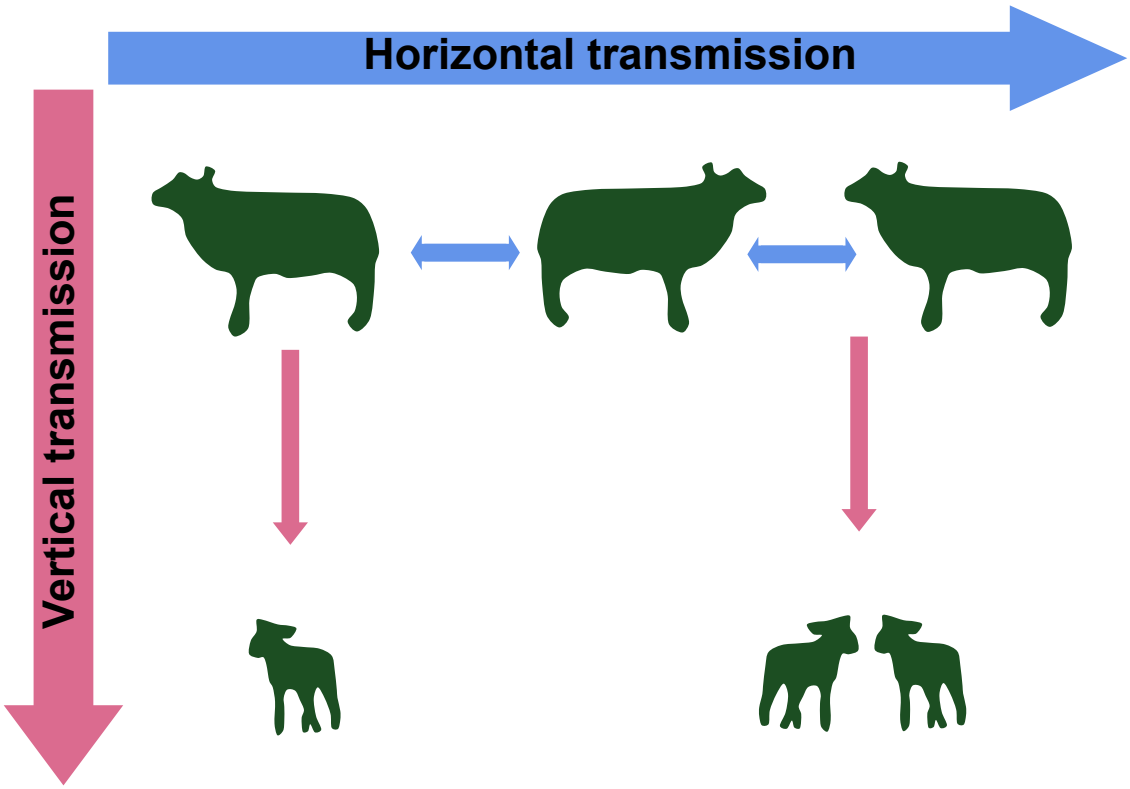


Figure 1.1 Horizontal and vertical transmission.
Modified from Farm Health Online (http://www.farmhealthonline.com).

1.2. Reproductive manipulation

In most lineages, vertical transmission occurs via the female host because of differences in gamete size (sperm cells lack cytoplasm and are too small to accommodate intracellular parasites). A male host is therefore an evolutionary dead end unless endosymbionts within a male host can be transmitted via other routes (i.e. via horizontal transmission) (Murlas Cosmides and Tooby, 1981; Watanabe et al., 2014). Interestingly, some vertically transmitted endosymbionts are capable of manipulating host sex (i.e. **sex-ratio distortion**) which results in an higher proportion of the transmitting sex (= female) in the host population (Cordaux et al., 2011). By manipulating host sex to

increase host reproduction, these sex-ratio distorters can spread rapidly through the population.

Several different mechanisms of reproductive manipulation are known, namely **cytoplasmic incompatibility**, **male killing**, **parthenogenesis induction**, and **feminisation** (Figure 1.2; Engelstädter and Hurst, 2009). Cytoplasmic incompatibility causes sterility between an infected male and an uninfected female or a female infected with a different strain (Bourtzis et al., 1996). This mechanism may lead to the removal of uninfected individuals in a population. Endosymbionts with male-killing abilities induce the death of males in the early or later stages of their development (Hurst and Jiggins, 2005). This may lead to increased resource allocation to female siblings and/or increased horizontal transmission through killed males (Dunn et al., 2001; Nakanishi et al., 2008). Parthenogenesis induction has been reported in diverse haplodiploid insects, in which sex is determined by the ploidy of the embryo. By doubling the chromosome numbers, non-transmitting males can be converted to transmitting females (Hagimori et al., 2006; Stouthamer, 1997). Although all the above mechanisms are only known in insects and arachnids, feminisation by endosymbionts has also been documented in crustaceans (amphipods and isopods) (Bouchon et al., 1998; Terry et al., 1997). Feminising endosymbionts can change genotypic males into functional phenotypic females, or even genotypic males into genotypic females.

Detailed mechanisms and the phenomenon of reproductive manipulation itself have been particularly well studied in insect hosts infected with the bacterium *Wolbachia*. However, other bacteria also manipulate host sex, for instance *Cardinium*, *Rickettsia*, *Spiroplasma*, *Flavobacteria*, and *Arsenophonus* (Cordaux et al., 2011; Duron et al., 2008). Although they are eukaryotes, species of Microsporidia are also known to manipulate host sex by inducing feminisation. *Rickettsia* and Microsporidia are the focal taxa of this thesis and their biology is presented in section 1.6.1.

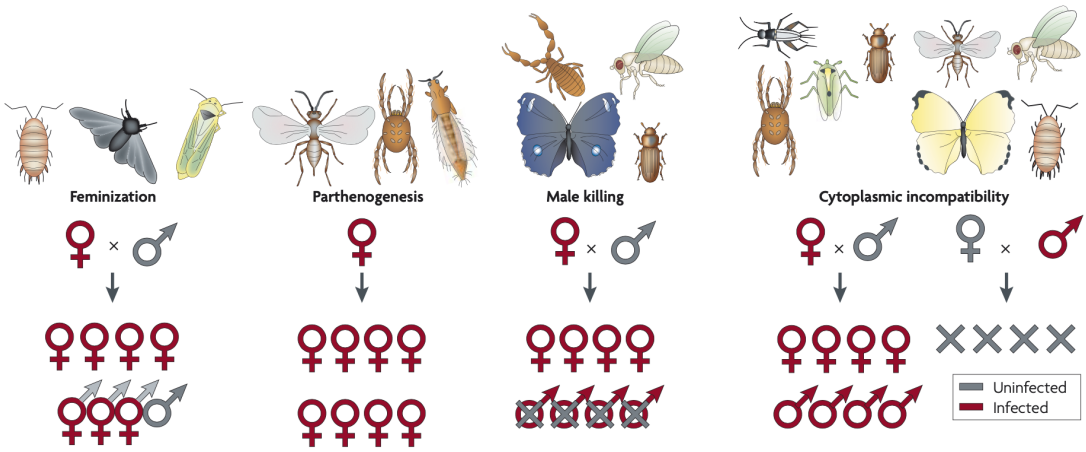


Figure 1.2 Several different phenotypes of reproductive manipulation by *Wolbachia*. Although these phenomena are best known for *Wolbachia,* other endosymbionts are also capable of reproductive manipulation. This figure was taken from Werren et al., 2008.

1.3. From ecological interactions to phylogenetic patterns

All species interact with other species. In any pair of interacting species, one may affect the other's fitness and exert selection pressure, which could lead to evolutionary changes in the affected taxon. If these pressures are reciprocal and occur both ways, '**coevolution** (=coadaptation)' will ensue. Coevolution can be defined as 'microevolution of two or more interacting species in response to reciprocal selection between them' (Janzen, 1980). On a broader, macroevolutionary scale, **codiversification** can be defined as 'correlated diversification of interacting lineages' (Janz, 2011). Codiversification can result from a combination of coevolution *sensu stricto* in combination with a range of historical and biogeographical events. Codiversification is often studied by comparing two phylogenies, i.e. that of a higher taxon of symbionts with that of their host taxon. To date, a range of **cophylogenetic methods** have been widely applied to diverse host-parasite systems.

Since parasites are highly dependent on their host for survival, their phylogenies are expected to be similar; this well-known hypothesis is called **Fahrenholz's rule**:

'parasite phylogeny mirrors that of its host' (Fahrenholz, 1913). This is expected because speciation in the host lineage may directly cause speciation of its parasites and result in **cospeciation**. However, after extensive cophylogenetic studies on diverse host-parasite associations, it has been shown that phylogenies of hosts and parasites very rarely approach perfect congruence (de Vienne et al., 2013). This is because other events such as host shifts, duplications, or extinctions, which disrupt cophylogenetic signals, can also occur during the shared evolutionary history of hosts and parasites (de Vienne et al., 2013; Page, 2003).

Several methods for cophylogenetic analysis are now available (Figure 1.3). They are broadly categorised into two types: event-based and global fit methods. Event-based methods look for the most likely history of host-parasite associations. Available event-based methods include Jungles, Tarzan, and Jane, each implemented with custom-made software (Conow et al., 2010; Merkle and Middendorf, 2005; Toews and Brelsford, 2012). On the other hand, global fit methods compare the overall congruence between the phylogenies of the two taxa. For example, ParaFit and PACo employ this approach (Balbuena et al., 2013; Legendre et al., 2002). As event-based methods rely on well-resolved phylogenies and can be computationally demanding, global fit methods are better suited for large-scale phylogenies even if they include uncertain branching patterns. Recently, a new method, Random Tanglegram Partitions (Random TaPas), has been developed (Haley, 1954). This method can more reliably identify the host-parasite associations that contribute the most to the cophylogenetic signal compared to earlier methods, and also allows the user to differentiate between cospeciation and pseudocospeciation (parasite phylogeny mimicking host phylogeny as a result of preferential host-switching among closely related host species rather than cospeciation).

The increasing availability of user-friendly cophylogenetic tools provides opportunities to achieve more robust testing of hypotheses regarding host-parasite associations and their diversification. On macroevolutionary scales, geographical barriers and dispersal abilities of organisms are important factors in determining the distribution of taxa. Although some studies have examined the role of these factors in parasite diversification (Engelbrecht et al., 2016; Sweet and Johnson, 2018; Weckstein, 2004),

studies linking hosts and parasites and their respective biogeographical patterns to the outcome of their coevolutionary association are still scarce (Nieberding et al., 2010).

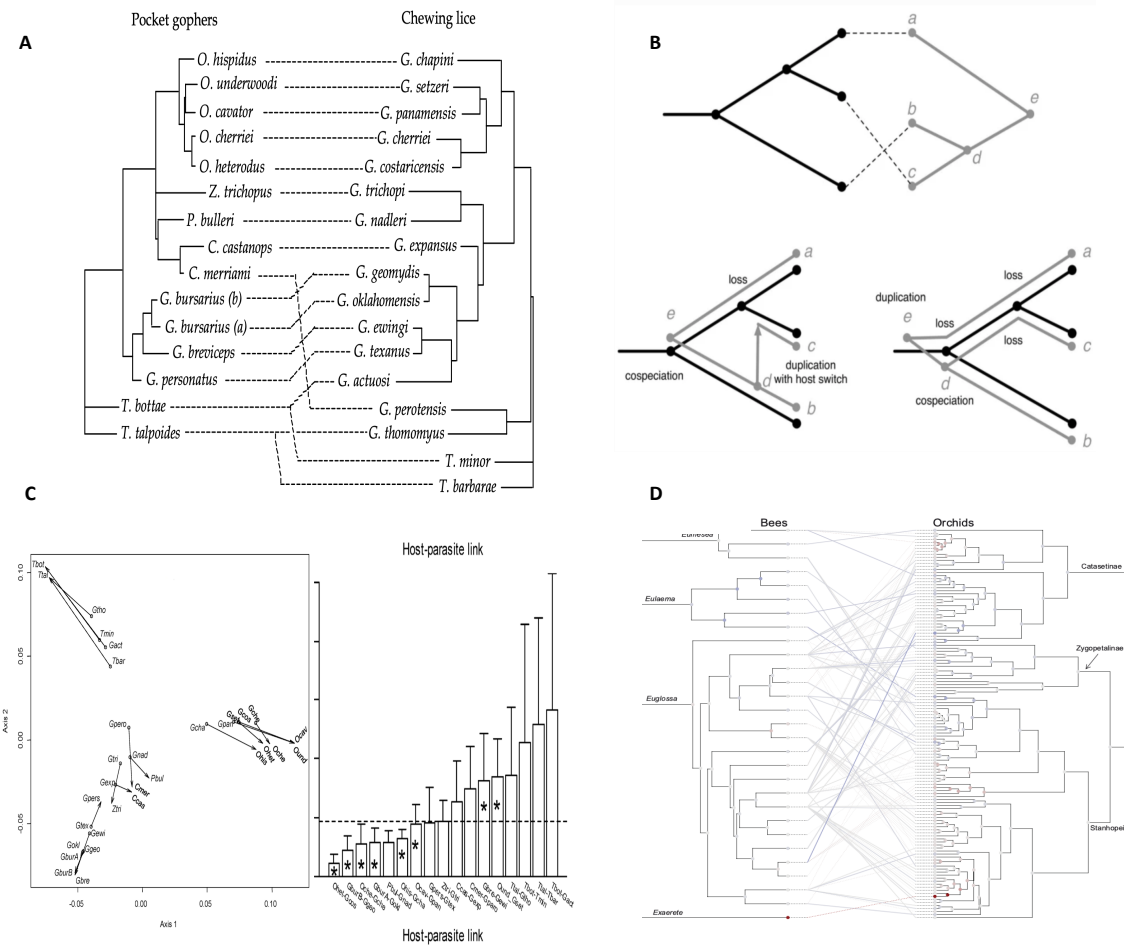


Figure 1.3 Visual outputs from several approaches for cophylogenetic studies. **A**. A simple tanglegram of the most classical example of cophylogeny between chewing lice and pocket gophers. Taken from Hafner et al., 2003. **B**. A tanglegram showing two possible reconstructions for the current host-parasite associations using an event-based method (Jane). Taken from Conow et al., 2010. **C**. PACo, a global fit method. Taken from Balbuena et al., 2013. **D**. Random Tanglegram Partitions (Random TaPas). Taken from Balbuena et al. 2020.

1.4. Co-transmission of mitochondria and some endosymbionts

Mitochondrial DNA (mtDNA) has been widely used for DNA barcoding, population genetics, and phylogenetic, phylogeographic, and cophylogenetic studies (Hajibabaei et al., 2007; Hebert et al., 2003), with several benefits. First, with a high copy number in every cell, it is easy to amplify. Second, a variable region flanked by highly conserved regions allows the use of universal primers (i.e. cytochrome c oxidase I or COI for DNA barcoding). Third, mtDNA evolves faster than nuclear genes and is likely to show population structures that nuclear genes would not reveal. Fourth, recombination is rare in mitochondria. However, numerous studies have raised concerns regarding the reliability of mtDNA for a range of reasons (Galtier et al., 2009). The first kind of problem is the amplification of pseudogenes (i.e. nuclear mitochondrial pseudogenes; numts) or sequences of untargeted organisms (e.g. endosymbionts) (Smith et al., 2012; Song et al., 2008). This type of problem is common but can be significantly reduced with certain precautions (e.g. by checking the presence of stop codons). However, even if orthologous sequences are obtained, there are other kinds of problems that may undermine the usefulness of mtDNA by violating the assumption of neutrality (Ballard and Whitlock, 2004; Galtier et al., 2009; Hurst and Jiggins, 2005).

Other problems arise due to the maternal transmission of mitochondria. In most lineages, mitochondria are transmitted vertically from mother to offspring (i.e. maternal transmission). Because not only mitochondria but also endosymbionts can be transmitted through the egg, co-transmitted mitochondria and endosymbionts in a germ cell will be in **linkage disequilibrium (LD)**. In genetics, LD refers to the non-random associations between alleles at different loci (Hill and Robertson, 1968). If one of the linked alleles is under selection, the frequency of this allele will be increased in a population just as other linked ones. This phenomenon produces a '**selective sweep**' (Kim and Neilsen, 2004). Although they do not exactly have the same meaning, these terms are also used for the non-independent association between mitochondria and endosymbionts and resulting phenomena. If endosymbionts are under positive or negative selection and their frequency in the population changes, the frequency of the associated mitochondrial haplotype will also increase or decrease, even if the

mitochondria themselves are neutral (Hurst and Jiggins, 2005; Jiggins, 2003). Furthermore, some endosymbionts can manipulate host sex and thus the frequency of mitochondrial genomes associated with endosymbionts will rapidly increase in the host population. In these cases, mtDNA can no longer be used as a neutral marker. As a result, phylogenetic or phylogeographical inferences based on these genes will not reflect the true evolutionary histories of the host taxa.

In some cases, a problem referred to as '**mitonuclear discordance**', meaning significantly different patterns observed with mitochondrial and nuclear genes, may arise due to endosymbionts (e.g. *Wolbachia*) (Toews and Brelsford, 2012). Often, two populations of the same species or different species share the same mtDNA. Hybridisation followed by mitochondrial introgression and rapid spread of endosymbionts and associated mitochondrial haplotypes can explain the described patterns (Jiggins, 2003). Because *Wolbachia* is common in insects, mitonuclear discordance patterns in insects are often attributed to the presence of *Wolbachia*. However, the diversity and distribution of other reproductive manipulators and their effects on mtDNA evolution are much less known outside of insects and mites.

1.5. Study system

This thesis focuses on two groups of obligate intracellular parasites, **Microsporidia** and ***Rickettsia***, in amphipod hosts in New Zealand.

1.5.1. Parasites

1.5.1.1. Microsporidia

Microsporidia is an extremely diverse group of obligate intracellular parasites that belongs to the eukaryotes, with genetic affinity to fungi (Edlind et al., 1996; Keeling and Doolittle, 1996). They are highly dependent on their host for metabolism; they lack mitochondria and import energy from the host cell using ADP/ATP transporters (Tsaousis et al., 2008). Their genomes are also highly reduced; for instance, the

microsporidian *Encephalitozoon* has the smallest genome among eukaryotes (Corradi et al., 2010).

Microsporidia have been known for over 150 years due to their devastating impact on species of economic importance. In the early 1800s, 'pebrine disease' in silkworms swept across many countries in Europe, and later it was shown to be caused by the microsporidian species, *Nosema bombycis* Nageli, 1857, the first-named microsporidian (Balbiani, 1882; Pasteur, 1870). Subsequently, Microsporidia have been reported from many other economically important species such as bees, crabs, lobsters, shrimps, and fishes. The collapse of the ocean pout fisheries was attributed to *Pleistophora macrozoarcides,* while *Glugea hertwigi* was responsible for the decline of the rainbow smelt fisheries in North America (Haley, 1954; Sandholzer et al., 1945). *Enterocytozoon hepatopenaei* (EHP), first discovered in 2009, is causing large economic losses in cultured shrimps in many Asian countries (Chaijarasphong et al., 2020; Tourtip et al., 2009). Beside their economic impacts, microsporidian infections, such as *Nematocida parisii* in *Caenorhabditis elegans* and *Pseudoloma neurophilia* in zebrafish, have been persistently reported from model animal species in research facilities (Sanders et al., 2012; Troemel et al., 2008).

Not only occurring in species of economic importance, microsporidians have also been found in almost all animal phyla: Acanthocephala (de Buron et al., 1990), Annelida (Larsson, 1992), Bryozoa (Desser et al., 2004), Cnidaria (Clausen, 2000), Gastrotricha (Manylov, 1999), Kinoryncha (Adrianov and Rybakov, 1992), Mesozoa (Czaker, 1997), Mollusca (Sagristà et al., 1998), Nematoda (Ardila-Garcia and Fast, 2012), Phoronida (Temereva and Sokolova, 2018), and Rotifera (Gorbunov and Kosova, 2001). These studies suggest a very high diversity of microsporidia and a broad distribution in the environment. For instance, although there have been only two reports of microsporidians in cnidarians, Clausen (2000) found heavily infected marine hydrozoans, *Halammohydra intermedia*, from two localities and suggested that the host-parasite relationship was broadly established. The rotifer *Brachionus calyciflorus* was found infected by *Microsporidium asperospora* with a prevalence of up to 40% in the Volga delta, the largest river delta in Europe (Gorbunov and Kosova, 2001). Targeted

studies with group-specific primers (either with Sanger or next-generation sequencing) have uncovered a great diversity of microsporidians in nematodes, amphipods, and from the environment (Grabner, 2017; Sapir et al., 2014; Williams et al., 2018).

Considering their wide host range and distribution, relatively few microsporidians have been reported in New Zealand. In 2010, *Nosema ceranae* was detected in honeybees in New Zealand, as part of an investigation of unusual colony losses (Frazer et al., 2015). In 2008, a new microsporidian species, *Myospora metanephrops*, was found in a marine lobster caught around the Auckland Islands, south of New Zealand (Stentiford et al., 2010). Although there have been some cases of microsporidians of commercial relevance in New Zealand, a few studies suggest that they perhaps occur more widely in various animals across the country. According to Malone et al. (1987), microsporidian infections were commonly found in *Costelytra zealandica* and *Wiseana* spp., major insect pest species in New Zealand. Microsporidian infections in freshwater crayfishes were also recorded along the Leith stream in Otago (Quilter, 1976).

While some intracellular parasites cause severe pathology, most of them do not cause evident pathology. For example, Microsporidia-infected insect larvae often exhibit melanised areas on their cuticle, so infection can be externally apparent (Becnel and Andreadis, 2014). Some aquatic arthropods have a translucent body and the presence of microsporidia can be easily observed. For example, *Fibrillanosema crangonycis* infection in the amphipod *Crangonyx pseudogracilis* can be observed, as the ovary appears opaque (Slothouber Galbreath et al., 2004). However, microsporidian infection in nature is typically chronic and cryptic, and obvious symptoms are not often visible (Franzen, 2008). In this case, genetic detection of microsporidians is the only effective approach.

1.5.1.2. *Rickettsia*

The genus *Rickettsia* belongs to the order Rickettsiales, which includes *Wolbachia*. This genus comprises diverse pathogenic species that can cause vector-borne diseases in birds and mammals including humans, as well as non-pathogenic species. Some rickettsioses with severe symptoms are well known and include Rocky Mountain

spotted fever, Queensland tick typhus, rickettsial pox, murine typhus, and epidemic typhus (Parola et al., 2005; Perlman et al., 2006; Weinert, 2015). Some species have established a mutualistic relationship within their host. For example, *Rickettsia* are necessary for oogenesis in booklice and they provide resistance to fungal infections in pea aphids (Łukasik et al., 2013, 2013). To date, at least 13 groups are known within the genus *Rickettsia*: Adalia, Bellii, Canadensis, Guiana, Helvetica, Meloidae, Mendelii, Rhyzobious, Spotted fever, Scapularis, Torix, Transitional, and Typhus. All these groups except the Torix group are exclusively associated with arthropod hosts, such as mites, fleas, ticks, and spiders. The Torix group, which is sister to all other groups, is the only group that includes non-arthropod hosts such as amoebae and leeches. In addition to these freshwater hosts, the Torix group occurs in diverse arthropod groups that spend part of their life cycle in the freshwater environment.

In New Zealand, *Rickettsia felis* from the Spotted fever group, which is widely reported from North and South America, Europe, Africa, and Australia, has also been found in cat fleas (Chandra et al., 2017). Some *Rickettsia*-like organisms (RLO) were observed in Chinook salmon and surf clams, however, none belong to the genus *Rickettsia*. Torix *Rickettsia*, which is a target study group of this thesis, have never been reported in New Zealand.

1.5.2. New Zealand phylogeography

New Zealand comprises two main islands and many small islands. New Zealand is an interesting place to study biogeography because of its current isolation and early separation from other landmasses. New Zealand was separated from Australia around 82 MYA (Kamp, 1986), leading to New Zealand being called ‘Moa’s arc’ (Bellamy, 1990; Stevens and Stevens, 1980). Although a ‘complete marine inundation’ scenario during the Oligocene challenged this view (Landis et al., 2008), recent studies suggest that New Zealand may not have been completely submerged, with some terrestrial refugia remaining throughout its history (Giribet and Boyer, 2010; Strogen et al., 2014; Wallis and Jorge, 2018).

In addition, several relatively recent geological events such as the formation of the alps (5 MYA), land emergence due to sea level changes (1-3 MYA), the last glacial maximum (LGM; 20 KYA), and volcanic eruptions (2 KYA) have largely shaped the current distribution and population structure of many extant taxa (Trewick et al., 2011). The land area (and therefore coastlines) has changed continuously over the past several million years (Figure 1.4). During the LGM of the Pleistocene, the North and South Islands were connected. The current separation of the two main islands by the Cook Strait occurred less than 500 KYA (Lewis et al., 1994).

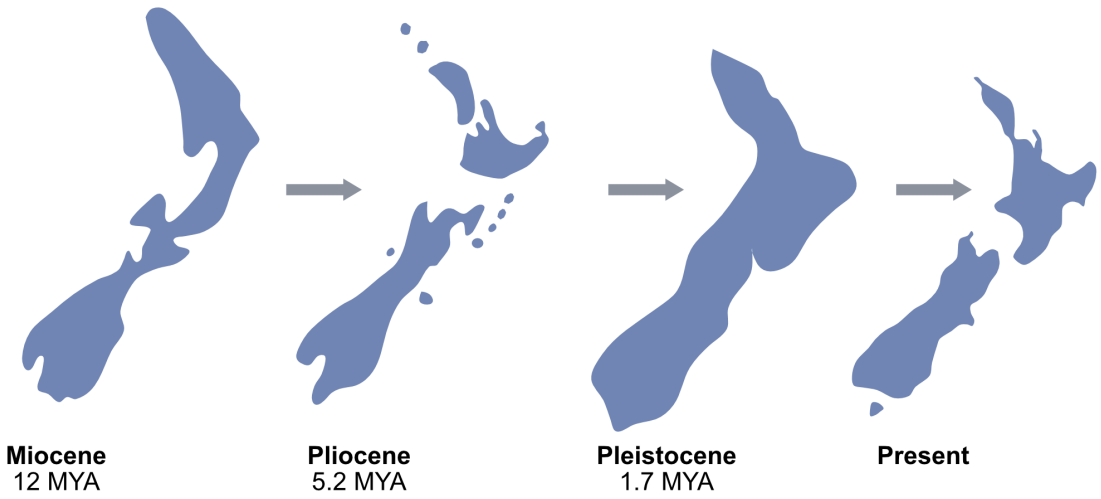


Figure 1.4 Changes of New Zealand land area (above sea level) through time, from the Miocene to the present. Modified from Fleming (1979) and (Sutherland et al., 2009).

1.5.3. Amphipod hosts

The order Amphipoda belongs to the subphylum Crustacea and comprises more than 10,000 species inhabiting diverse aquatic and terrestrial environments (Arfianti et al., 2018). Amphipods are important members of food webs because many of them are detritivores and scavengers; they are also major prey items for fishes (Conlan, 1994; Väinölä et al., 2008). Amphipods are widely used as intermediate hosts by diverse parasites such as nematodes, trematodes, and acanthocephalans (Bethel and Holmes, 1977; Friesen et al., 2019; Poulin and Latham, 2002).

According to a recent molecular analysis, amphipods originated during the late Palaeozoic (280-240 MYA), and major groups diversified between the early Cretaceous and early Paleogene (140-60 MYA) (Copilaş-Ciocianu et al., 2020). It is believed that the dispersal abilities of amphipods are largely limited due to the direct development of their eggs within a brood pouch on the underside of females, the lack of independent larval dispersal stages, and an extended parental care (Dick et al., 1998; Thiel, 1999; Väinölä et al., 2008). Amphipod diversity at the family level differs widely between the Northern and Southern Hemispheres, suggesting an important role of vicariance in shaping the current global diversity and geographical distribution of amphipods (Lowry and Myers, 2017).

Despite some pioneering studies about amphipods in New Zealand (Jerry Laurens Barnard, 1974; Chilton, 1906, 1882; Thomson, 1987), the diversity of amphipods remains largely unknown (Chapman et al., 2011; Fenwick, 2001). About 24 species from 10 genera are known from freshwater (Chapman et al., 2011). Among them, *Paracalliope* (Paracalliopiidae) is the most common genus. *Paracalliope* includes marine, brackish, and freshwater species; some of which having a wide range of tolerance to salinities (Knott, 1975). Among New Zealand freshwater amphipods, only COI sequences from *Paracalliope* and *Paracorophium* (Corophiidae) are publicly available in GenBank (Hogg et al., 2006; Lagrue et al., 2016; Stevens et al., 2006). A few amphipod families of Gondwanan origin also occur in New Zealand, such as Paraleptamphopidae and Phreatogammaridae (Fišer et al., 2013; Lowry and Myers, 2013).

1.5.3.1. Microsporidia in amphipods

Microsporidians are among the most common parasites of amphipods with more than 30 named species from 13 genera found in amphipod hosts (Bojko and Ovcharenko, 2019), and many other unnamed taxa detected from several regional and large-scale studies (Dimova et al., 2018; Grabner, 2017; Grabner et al., 2015; Kuzmenkova et al., 2008; Quiles et al., 2019; Slothouber Galbreath et al., 2004; Williams et al., 2018). Geographically, all the known diversity of microsporidians in amphipods is from the Northern Hemisphere. This may simply reflect the fact that about 83 % of known

freshwater amphipod species occur in the Northern Hemisphere (Lowry and Myers, 2013; Väinölä et al., 2008). Accordingly, microsporidian diversity has mostly been studied in amphipod diversity hotspots, including southern Europe, the Ponto-Caspian area, Lake Baikal (Asia), and southeastern USA (Bacela-Spychalska et al., 2018; Dimova et al., 2018; Quiles et al., 2019; Slothouber Galbreath et al., 2004; Väinölä et al., 2008).

1.5.3.2. *Rickettsia* in amphipod hosts

Although *Rickettsia* species are known as common pathogens or endosymbionts in arthropod hosts, these agents have never been reported in crustaceans. *Rickettsia*-like organisms (RLO) have been detected in several species of gammarids, such as *Diporeia* sp. and *Crangonyx floridanus* (Graf, 1984; Larsson, 1982; Messick et al., 2004; Winters et al., 2015). However, none of them actually belong to the genus *Rickettsia*.

1.5.4. Study sites

Research in this thesis is based on amphipods and other crustacean specimens that were collected throughout New Zealand (Figure 1.5). Throughout all chapters and published manuscripts, the same location IDs have been used. For Chapter 5, three sites (S34, S37, S40) in Otago and Southland regions were visited every two months.

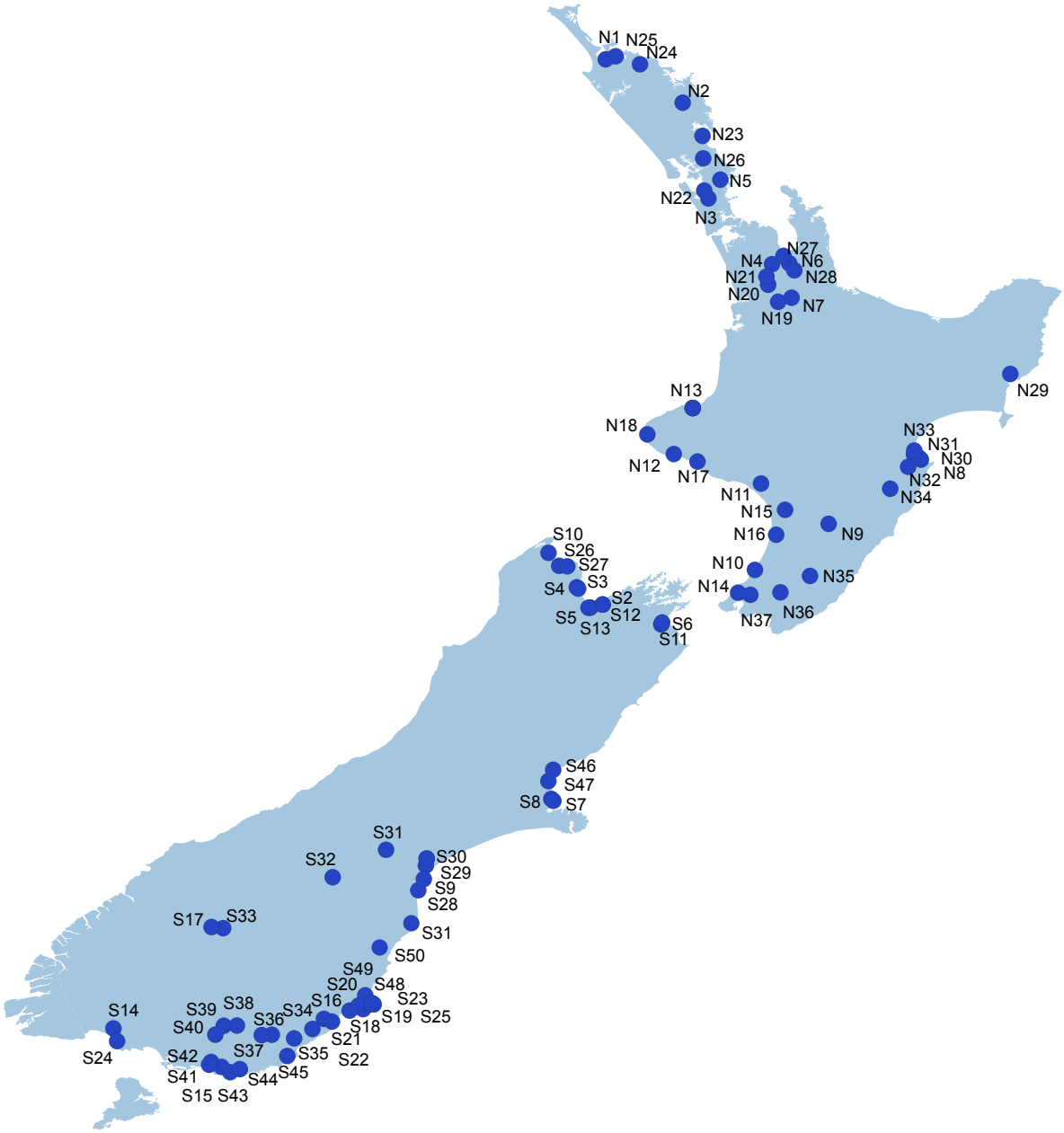


Figure 1.5. Sampling sites. Amphipods and other crustacean specimens were collected from 87 locations.

Table 1.1 A list of sampling sites

Region	Location ID	Location	Habitat type
Northland	N1	Mangatete River	River
	N2	Hukerenui	Stream
	N22	Mangakura	Stream
	N23	Ruakaka	Stream
	N24	Whangaroa	Stream
	N25	Taipa	Stream
	N26	Kaiwaka	River
Auckland	N3	Kaukapakapa	Stream
	N4	Waerenga	Stream
	N5	Wairere Reserve	Pond
	N21	Fisher Rd	Stream
Waikato	N6	Kaihere	Stream
	N7	Tauwhare	Stream
	N19	Hamilton Garden	Pond
	N20	Taupiri	Stream
	N27	Mangatarata	Stream
	N28	Tahuna	Stream
	N29	Makauri	Stream
Hawke's Bay	N8	Haumoana	Stream
	N30	Clive	Stream
	N31	Tradale	River
	N32	Pakipaki	Stream
	N33	Orotu Park	Stream
	N34	Waipaua	River
Taranaki	N12	Manaia	Stream
	N13	Tirorangi	Stream
	N17	Mokoia	Stream
	N18	Rahotu	Stream
Malborough	S6	Grovetown Lagoon	Stream
	S11	Riversidepark	River
Manawatu-Wanganui	N9	Woodville	Stream
	N11	Wanganui	Stream
	N15	Manawatu River	River
	N16	Bulls	Pond
Wellington	N10	Waihanae River	River
	N14	Porirua stream	Stream
	N35	Masterton	River
	N36	South Featherston	Stream
	N37	Lower Hutt	River
Tasman	S3	Motueka River	River
	S4	Riwaka	Stream
	S5	Waimea River	River
	S10	Collingwood	Estuary
	S13	Waimea inlet	Inlet
	S26	Takaka River	River
	S27	Pohara	Stream
Nelson	S2	Matai River	River
	S12	Queensgarden	Pond

Canterbury	S1	Waitaki River	River
	S7	Avon Hathcote Estuary	Estuary
	S8	Avon River	Reservoir
	S9	Saltwater creek	Creek
	S28	Otaio River	River
	S29	Opihi River	River
	S30	Winchester	Stream
	S31	Fairlie	Stream
	S32	Lake Benmore	Stream
	S33	Crown Range	Stream
	S46	Amberley	River
	S47	Waikuku	Stream
Otago	S16	Lake Waihola	Lake
	S17	Lake Hayes	Lake
	S18	Tomahawk beach	Beach
	S19	Tomahawk lagoon	Lagoon
	S20	Rosyln, Dunedin	Garden
	S21	Kaikorai estuary	Estuary
	S22	Taieri mouth	Estuary
	S23	Hoopers inlet	Inlet
	S25	Portobello	Bay
	S34	Tokomairiro	Stream
	S35	Clutha River	River
	S36	Waiwera South	Stream
	S37	Clinton	Stream
	S45	Owaka	Stream
	S48	Waitati	Stream
	S49	Nichols Creek	Pond
	S50	Herbert	Stream
Southland	S14	Waiau river	River
	S15	Tokanui	Stream
	S24	Te waewae bay	Beach
	S38	Pukerau	Stream
	S39	Gore	River
	S40	Waimumu	Stream
	S41	Waimahaka	Stream
	S42	Fortrose	River
	S43	Haldane	Stream
	S44	Niagara Falls	Stream

1.6. Pilot study and thesis development

This thesis has been gradually developed throughout my PhD period. I initially started my study with a focus on microsporidians. Because the diversity of microsporidians was largely unknown in New Zealand, I conducted a pilot survey to gain a general idea of how diverse they are, where they are distributed and in what hosts, and how prevalent they are. Several major groups of crustaceans (amphipods, cladocerans, ostracods, copepods, and shrimps) and some insects were collected from various localities in the Otago region. Despite extensive screening effort, only 18 individuals were positive for microsporidians among 1,864 individuals screened. Among positive individuals, 17 were amphipods and one was an insect (Chironomidae sp.). Most microsporidians from amphipods were closely related to *Dictyocoela* sp., also the most common microsporidian species in amphipods in Europe. Therefore, I decided to genetically characterise newly found microsporidian species and conduct cophylogenetic studies to infer the intimacy and antiquity of these host-parasite associations (**Chapter 2**). While identifying amphipod hosts using universal primers for the COI region, I unexpectedly obtained rickettsial sequences from several localities. This prompted me to explore the diversity and prevalence of *Rickettsia* in diverse amphipod groups, and thus **Chapter 3** was conceived. In addition, because amplification of untargeted rickettsial sequences seemed to be common, I designed and tested blocking primers to obtain targeted host COI sequences. COI sequences of *Paracalliope* amphipods are extremely divergent (Hogg et al. 2006; Sutherland et al. 2009; Lagrue et al. 2013; and my observation). Both Microsporidia and *Rickettsia* are known as sex-ratio distorters and therefore they may have affected mtDNA evolution. **Chapter 4** was therefore designed to test if these endosymbionts are responsible for divergent COI sequences. I learned from Chapters 2 and 3 that the prevalence of the parasites varies among populations, and may possibly also undergo seasonal fluctuations. In this context, **Chapter 5** was designed to quantify the temporal variation of two parasites and their co-occurrence throughout the year. Finally, the number of nucleotide sequences of Microsporidia is increasing in public databases, and many important studies regarding the evolution of microsporidians and their relatives have been published in the last decade. Therefore, I made good use of my

new data and the publicly available sequences to produce an updated global phylogeny of canonical microsporidians in **Chapter 6**.

1.7. The aims of this thesis

The overall goal of this thesis is to explore the diversity, evolutionary history, and spatio-temporal ecology of host-parasite interactions between amphipods and two groups of intracellular parasites. More specifically, this thesis aims to 1) uncover hidden diversity (**Chapters 2 and 3**), 2) infer the evolutionary history of host-parasite associations using cophylogenetic methods (**Chapter 2**), 3) identify and resolve problems caused by endosymbionts when studying host phylogeny and phylogeography (**Chapters 3 and 4**), 4) quantify the temporal variation of both parasites and their interactions (**Chapter 5**), and 5) explain the diversity of Microsporidia and *Rickettsia* detected in this study within a broader geographical context (**Chapters 2, 3, and 5**).

1.8. Thesis outline

This thesis includes five core chapters. Each chapter was written for journal publication and therefore, some parts may be repeated between chapters. I am the first author of all the manuscripts, which reflects the fact that I conducted most of the work (fieldwork, lab work, data analyses, and writing), with some help from colleagues and input from co-authors for writing (Dr. Fátima Jorge and Prof. Robert Poulin).

Chapter 1. General Introduction

Chapter 2. Shared geographic histories and dispersal contribute to congruent phylogenies between amphipods and their microsporidian parasites at regional and global scales
Published as Eunji Park, Fátima Jorge, and Robert Poulin (2020) in *Molecular Ecology* 29: 3330-3345

Chapter 3. Widespread Torix *Rickettsia* in New Zealand amphipods and the use of blocking primers to rescue host COI sequences
Published as Eunji Park and Robert Poulin (2020) in *Scientific Reports* 10: 16842

Chapter 4. Extremely divergent COI sequences of *Paracalliope* species complex (Amphipoda) due to interplay of various factors including *Rickettsia* infections
Under review in *Molecular Ecology*

Chapter 5. Sharing the same host: temporal variation in the prevalence of Microsporidia and *Rickettsia* within and among amphipod populations in New Zealand
Under review in *Parasitology*

Chapter 6. Revisiting the phylogeny of Microsporidia
Accepted for a publication in *International Journal for Parasitology*

Chapter 7. General Conclusion

Chapter 2

Shared geographic histories and dispersal contribute to congruent phylogenies between amphipods and their microsporidian parasites at regional and global scales

2.1. Abstract

In parasites that strongly rely on a host for dispersal, geographic barriers that act on the host will simultaneously influence parasite distribution as well. If their association persists over macroevolutionary time it may result in congruent phylogenetic and phylogeographic patterns due to shared geographic histories. Here, I investigated the level of congruent evolutionary history at a regional and global scale in a highly specialised parasite taxon infecting hosts with limited dispersal abilities: the microsporidians *Dictyocoela* spp. and their amphipod hosts. *Dictyocoela* can be transmitted both vertically and horizontally and is the most common microsporidian genus occurring in amphipods in Eurasia. However, little is known about its distribution elsewhere. I started by conducting molecular screening to detect microsporidian parasites in endemic amphipod species in New Zealand; based on phylogenetic analyses, I identified nine species-level microsporidian taxa including six belonging to *Dictyocoela*. With a distance-based cophylogenetic analysis at the regional scale, I identified overall congruent phylogenies between *Paracalliope*, the most common New Zealand freshwater amphipod taxon, and its *Dictyocoela* parasites. Also, hosts and parasites showed similar phylogeographic patterns suggesting shared biogeographic histories. Similarly, at a global scale, phylogenies of amphipod hosts and their *Dictyocoela* parasites showed broadly congruent phylogenies. The observed patterns may have resulted from covicariance and/or codispersal, suggesting that the intimate association between amphipods and *Dictyocoela* may have persisted over macroevolutionary time.

2.2. Introduction

Cophylogenetic analyses can be used to infer the evolutionary history of associations between two interacting taxa (Page, 2003). Congruent phylogenetic patterns at a macroevolutionary scale may arise due to adaptive processes, but also mainly due to other processes such as shared biogeographic histories (Clayton et al., 2015; Weckstein, 2004). Concomitant occurrence of speciation (=cospeciation) of two interacting taxa can promote congruent phylogenies (but see de Vienne et al. 2007 for cases of preferential host-shifts). For instance, in host-parasite associations, speciation in the host lineage can directly cause speciation of its parasites and result in cospeciation. However, phylogenies of hosts and parasites are seldom perfectly congruent due to other events that disrupt cophylogenetic patterns such as host shift, duplication, or extinctions (de Vienne et al., 2013; Page, 2003). In fact, according to an extensive review of cophylogenetic studies, host-shift speciation seems to be the dominant mechanism in parasite diversification (de Vienne et al., 2013). Although data are scarce, some systems show congruent phylogenies of host and parasite including the classical example of pocket gophers and their chewing lice (Hafner et al., 1994). In this example, a combination of several factors such as the solitary life-style of the host species, allopatric species distributions of hosts, and limited dispersal abilities of parasites were suggested as contributing factors which may have lowered the chances of host-shift, resulting in congruent host-parasite phylogenies (Clayton & Johnson, 2003; de Vienne et al., 2013; Nieberding, Jousselin, & Desdevises, 2010).

Dispersal is a fundamental biological process that acts on multiple evolutionary scales (Nathan, 2001). From the parasite's perspective, there are broadly two kinds of dispersal: host-dependent and host-independent. By their nature, parasites spend at least a part of their life within or on hosts and therefore rely on the host for dispersal to various degrees depending on lifecycle characteristics and transmission type (Blouin et al., 1995; Clayton et al., 2015). Many parasites have multiple hosts, as well as a free-living stage during which independent dispersal could occur. On the other hand, parasites that spend their whole lifespan within/on hosts probably rely on the host for dispersal (e.g. chewing lice on birds; Clayton et al., 2015). In such cases, host dispersal

is crucial for parasite dispersal, potentially leading to congruent evolutionary histories. Similarly, vertically transmitted microparasites are likely to follow the evolutionary trajectories of their hosts (Althoff et al., 2014), and it is thus unsurprising that clear cases of cospeciation typically involve parasites that have vertical transmission (de Vienne et al., 2013).

In a system where a parasite is highly reliant on its host for dispersal over the long term, these shared biogeographic histories alone may be sufficient to explain congruent phylogenies (Althoff et al., 2014), without requiring any adaptive explanation. Vicariance or dispersal events that impact host evolutionary history can simultaneously affect the parasites' evolutionary history. Therefore, the degree of host-parasite associations and dispersal capabilities of hosts and parasites can influence cophylogenetic patterns. Although some studies have underscored the role of host dispersal ability (Moon et al., 2019; Norte et al., 2020), parasite dispersal ability (Engelbrecht et al., 2016; Sweet and Johnson, 2018), and geographic barriers (Larose and Schwander, 2016; Weckstein, 2004) in parasite diversification, studies linking life history traits of hosts and parasites and their respective biogeographical patterns to the outcome of their co-evolutionary association are still scarce (Nieberding et al., 2010).

Microsporidian parasites, which can be transmitted effectively both vertically and horizontally, are common in amphipod hosts (Bojko and Ovcharenko, 2019; Lipsitch et al., 1995). Among more than 30 named species and many other unnamed taxa that were found in amphipod hosts, *Dictyocoela* is the most common genus with about 10 known species (Bojko & Ovcharenko, 2019; Dimova et al., 2018; Grabner, 2017; Grabner et al., 2015; Kuzmenkova, Sherbakov, & Smith, 2008; Quiles et al., 2019; Slothouber Galbreath, Smith, Terry, Becnel, & Dunn, 2004; Williams, Hamilton, Jones, & Bass, 2018; Supplementary Table 2.1). However, the known diversity of *Dictyocoela* seems to be restricted both geographically and in terms of host range. Geographically, all known microsporidians in amphipods are from the Northern Hemisphere, including southern Europe, the Ponto-Caspian area, Lake Baikal (Asia) and southeastern USA (Bacela-Spychalska et al., 2018; Dimova et al., 2018; Quiles et al., 2019; Slothouber Galbreath et al., 2004; Väinölä et al., 2008). In terms of host range, gammarids are the

best-studied amphipod hosts for microsporidians, although some *Dictyocoela* species were found parasitizing species of Talitridae, Melitidae, and Hyalellidae (Terry et al., 2004).

Given the limited knowledge on their distribution and diversity, the antiquity and strength of associations between *Dictyocoela* species and their hosts remain poorly understood. A recent study revealed some degree of host specificity and overlapping geographical distributions between microsporidian parasites and their amphipod hosts, suggesting their ancient associations (Quiles et al., 2019). In a pilot study conducted across a few locations on New Zealand's South Island, we detected microsporidian species similar to *Dictyocoela* in several endemic amphipod species. The presence of *Dictyocoela* in New Zealand amphipods provides an opportunity to ask questions regarding the evolutionary history of their association and codiversification patterns. How diverse and widely distributed are *Dictyocoela* and other microsporidian parasites in New Zealand amphipods? What are the phylogenetic relationships between *Dictyocoela* in New Zealand and from other parts of the world? Can we observe congruent phylogenies in the amphipod-*Dictyocoela* system? What are the underlying ecological and geological factors influencing the degree of congruence between their phylogenies? Can we infer the duration and intimacy of associations between *Dictyocoela* parasites and their amphipods hosts?

Some ecological and geographical factors make our study system highly suitable for investigating patterns of codiversification and the potential underlying roles of dispersal and geographic barriers. Amphipod dispersal abilities are highly limited due to the lack of planktonic larval stages (Kristjánsson and Svavarsson, 2007; Myers, 1993). Therefore, it is believed that vicariance may have played important roles in amphipod diversification and their biogeographical patterns reflect historical events (Copilaş-Ciocianu et al., 2020; Hou and Sket, 2016). New Zealand's geological history is relatively well known and is reflected in the unique fauna and the phylogenetic structure of a diverse range of organisms. New Zealand separated from Australia around 82 MYA (Kamp, 1986), with some lineages of archaic vicariant origin (McGlone, 2005; Stevens and Stevens, 1980). Also, several relatively recent geological events such as the

formation of the Southern Alps, shifting climatic conditions and sea-levels, and volcanic eruptions have strongly influenced the current phylogeographic structure of many extant taxa (Trewick et al., 2011). The role of vicariance, for example, has been demonstrated in New Zealand's endemic freshwater (*Paracalliope* species complex) and estuarine (*Paracorophium*) amphipods (Hogg et al., 2006; Knox et al., 2011; Sutherland et al., 2009). These studies uncovered highly divergent lineages within taxa, probably due to prolonged isolation and the presence of several cryptic species.

I conducted a regional scale cophylogenetic analysis focusing on *Paracalliope*-*Dictyocoela* associations in New Zealand. *Paracalliope* is the most common and widely distributed amphipod taxon and its phylogeographic structure largely reflects historical events (Hogg et al., 2006; Sutherland et al., 2009). I compared phylogenetic and phylogeographic structures between *Paracalliope* and their *Dictyocoela* parasites to assess the degree to which patterns overlap and the prevalence of shared co-differentiation/evolutionary histories. I then extended the cophylogenetic analyses to a global scale. The comparison of phylogenies and geographic patterns can provide interesting insights into the duration and the intimacy of host-parasite associations (Clayton et al., 2015; Garrick et al., 2017). The highly limited dispersal abilities of amphipods have resulted in largely different diversity patterns between the Northern and the Southern Hemispheres (Barnard, 1974; Lowry & Myers, 2017). Accordingly, the presence of *Dictyocoela* in amphipods in both hemispheres can be explained by either an ancient origin or recent dispersal of parasites. Considering patterns of codiversification within a geographic and ecological context, I infer their intimate and ancient associations.

To answer the questions posed above, I conducted nationwide molecular screening on diverse New Zealand amphipods for the presence of microsporidians, covering phylogenetically diverse amphipods (from population to family level). Specifically, I aimed at 1) quantifying the diversity, distribution, and prevalence of microsporidians in New Zealand, and for the first time in the Southern Hemisphere; 2) elucidating the phylogenetic positions of newly discovered microsporidian species within the phylum, and the phylogenetic relationships among *Dictyocoela* species; 3) assessing the degree

of congruence between host and parasite phylogenies and phylogeographic structures at both local and global scales; 4) inferring the intimacy and the duration of the association between *Dictyocoela* and their amphipod hosts; and 5) discussing the role of dispersal and geological barriers in explaining codiversification.

2.3. METHODS

2.3.1. Collection of specimens

Amphipods were collected from 69 sites throughout both the South and North Islands of New Zealand between August 2017 and April 2019 (Figure 2.1, and Supplementary Table 2.2). Specimens were collected with fine-mesh hand nets (< 0.2 mm) and then preserved in 96 % ethanol on site. Our main target taxon was the *Paracalliope* species complex, the most common and widely distributed freshwater amphipod species in New Zealand. *Paracalliope* spp. were obtained from 59 locations. At 63 locations, only one amphipod species was found and collected, while at 6 locations (S16, S21, S33, S39, N17, N36), two or more species were collected. Most of the specimens were found around weed beds in slow-flowing lowland streams and rivers. Some rare amphipod species were found in mountain streams and estuaries.

2.3.2. Identification of amphipods

Initial identification of collected specimens was done based on gross morphology (Chapman et al., 2011; Fenwick, 2001). Morphologically similar amphipods may be genetically distant due to the presence of cryptic species and/or morphological conservatism (Fišer et al., 2018; Murphy et al., 2009). Therefore, genomic DNA was obtained from several appendages per individual for further genetic identification. Mitochondrial COI and nuclear 28S regions were sequenced for each morphospecies per location. The sequences obtained were deposited in GenBank (Accession ID: MT465134-MT465172, MT466574-MT466580). Based on both morphological and genetic data, amphipod specimens collected in this study were ascribed to 7 families: Paracalliopidae (*Paracalliope* species complex), Paraleptamphopidae (*Paraleptamphopus* sp.), Phreatogammaridae (*Phreatogammarus* sp.), Corophiidae

(*Paracorophium excavatum*), Melitidae (*Melita awa*), Talitridae (*Parorchestia tenuis*), and one undescribed family that belongs to the suborder Senticaudata.

2.3.3. DNA extraction from pooled specimens

In order to maximize detectability while lowering the cost and time needed for molecular screening for microsporidians, I used pooled host specimens instead of individual specimens for DNA extraction, for amphipod species with small body sizes (< 4 mm). This approach allowed us to detect microsporidians even in a host population with low prevalence, with relatively low effort. I used the same number of host individuals for each location when I had enough specimens to compare relative prevalence: 12 pooled samples of 4 individuals for each location (= 48 individuals) were used for DNA extraction for most populations (Supplementary Table 2.2). For each pooled sample, the whole bodies of 4 individual amphipods were washed with distilled water, cut into small pieces and pooled into a tube. Then, 400 µl of Chelex solution and 3 µl of proteinase K were added to each tube, which was then incubated at 55 °C overnight. The next day, tubes were incubated at 90 °C for 8 minutes and then run in a centrifuge for 10 minutes at 14,000 rpm. For *Parorchestia tenuis* which has a large body size (> 10 mm), pereonites 5 to 7 were dissected (which include gonads) and used for DNA extractions without pooling specimens.

2.3.4. Detection of microsporidia by PCR

A partial small subunit ribosomal DNA (*SSU* rDNA) sequence was amplified to detect microsporidian infections. Either a primer pair of 18F (CACCAGGTTGATTCTGCC) and 1492R (GGTTACCTTGTTACGACTT), or V1f (CACCAGGTTGATTCTGCCTGAC) and MC3R (GATAACGACGGGCGGTGTGTACAA) were used to amplify 1,248 bp and 1,163 bp, respectively (Ovcharenko et al., 2010; Vossbrinck and Debrunner-Vossbrinck, 2005; Weiss and Vossbrinck, 1999; Zhu et al., 1993). For PCR reactions, 12.3 µl of distilled water, 4 µl of reaction buffer, 0.8 µl of each forward and reverse primers, 0.1 µl of MyTaq (Bioline), and 2 µl of DNA were used. For each set of PCR reactions, both negative and positive controls were included with water and DNA obtained from initial screening, respectively. PCR conditions for the primer pair of 18F and 1492R were the

following: 94°C initial denaturation for 3 min, 35 cycles of 94°C for 30 s, 50°C for 60 s, 72°C for 60 s, final extension for 10 min at 72°C. For the primer pair of V1f and MC3R, a touchdown PCR was conducted under the following conditions: initial denaturation at 94 °C for 3 min, 7 cycles of denaturation at 94 °C for 30 s, annealing at 62 °C for 30 s (decreasing 1 °C/cycle) and extension at 65 °C for 80s, 25 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s and extension at 65 °C for 80 s, with a final extension at 65 °C for 5 min. Then, 2µl of PCR product from each PCR reaction were run on a 1.5 % agarose gel. For a subset of samples, representing each species-level taxon, a primer pair of HG4f (GCGGCTTAATTTGACTCAAC) and 580R (GGTCCGTGTTTCAAGACGG) was additionally used to obtain a full SSU, ITS, and LSU sequence (a total length of ~1,760 bp) with the same PCR conditions as for V1f /MC3R (Bacela-Spychalska et al., 2018; Gatehouse and Malone, 1998; Weiss et al., 1994).

2.3.5. PCR from individual specimens and sequencing

Because I used pooled samples, the risk of contamination due to multiple microsporidian strains of one species or several species in one sample was expected to be higher than when non-pooled samples were used. Therefore, I extracted and sequenced DNA from individual specimens from populations with high infection rates (Supplementary Table 2.2). When no microsporidians were detected from eight individual samples, or when no amphipod specimens were available after the initial screening, PCR products were directly obtained from pooled samples assuming low prevalence (= single microsporidian species per tube). PCR products were purified with MEGAquick-spin™ Total Fragment DNA Purification Kit (iNtRON biotechnology) according to the manufacturer's instructions. Purified PCR products were sent to either Genetic Analysis Services at the University of Otago, New Zealand or Macrogen, Korea, for Sanger sequencing. Raw nucleotide sequences were trimmed with the trim function in Geneious prime 2019.0.4 (https://www.geneious.com) with the default settings, and then ambiguous sites were carefully examined and corrected by eye. Some multi-peaks were identified only in two short hypervariable regions within SSU. This could be due to multiple infections or intragenomic variation among rRNA copies (Ironside 2013). In our case, the pooling method may have increased the chance of

contamination by including more than one infected individual in the same tube. However, this was not likely a major factor because sequences with multiple peaks were evenly distributed among both individual and pooled samples. For these sequences, the International Union of Pure and Applied Chemistry (IUPAC code) was used to avoid possible errors in delineating strains or species (Alperi et al., 2008).

2.3.6. Species delimitation

Haplotypes were identified by using the package *pegas* (Paradis, 2010) in R version 3.5.2 (R Core Team 2013). Haplotypes that diverged by less than 1% were grouped into a putative species, following Terry et al. (2004)'s criteria. A tree-based (mPTP) method and a distance-based (ABGD) species delimitation method, along with morphological assessment, confirmed the validity of the '1% rule' (Bacela-Spychalska et al., 2018). A formal description of these species would require an integrative approach with morphological, ecological, and phylogenetic data (Stentiford, Feist, Stone, Bateman, & Dunn, 2013). Therefore, our newly discovered microsporidian putative 'species' will remain as candidates until full description. In this study, I assigned them provisional names for convenience (Supplementary Table 2.3).

2.3.7. Phylogenetic analyses

Six phylogenetic trees were assembled for four different purposes: 1) to place newly found putative species within the phylum Microsporidia; 2) to resolve the phylogenetic relationships among all dictyocoelan species, as most of our sequences belong to this genus; 3) to be used in cophylogenetic analysis between *Dictyocoela* and their *Paracalliope* hosts at a regional scale; 4) to be used in cophylogenetic analysis between *Dictyocoela* and their amphipod hosts at a global scale. The following procedures were applied to all datasets: all sequences were aligned in Geneious prime with the MAFFT algorithm (Katoh and Standley, 2013) using consistency-based iterative refinement methods (E-INS-i or G-INS-i). Ambiguous sites were then eliminated in Gblocks with the least restrictive setting (Castresana, 2000). The best-fitting model of nucleotide evolution for each dataset was determined based on the corrected Aikake information criterion (AICc) using jModelTest v2.1.6 (Darriba et al., 2012), which was conducted through the CIPRES Science Gateway v3.3 (Miller et al., 2010). For all analyses of

microsporidians, the General Time Reversible (GTR) model of nucleotide substitution along with Gamma distributed rate variation across sites (G) and the proportion of invariable sites (I) were used for Bayesian tree inference in MrBayes 3.2.7 (Ronquist et al., 2012). For the host phylogeny based on 28S sequences, GTR+G was used for tree reconstruction. For all datasets, two independent runs, consisting of four chains each, were simultaneously conducted for 2,000,000 generations with a sampling frequency of 1,000. A stop rule was applied to terminate the MCMC generations as soon as the standard deviation of split frequencies fell below 0.01. The initial 25% of samples were discarded. Maximum Likelihood trees were reconstructed in RAxML with GTRGAMMA+I as a model of nucleotide evolution. A rapid bootstrap analysis was conducted with 1,000 replicates. The resulting trees were visualized in FigTree v1.4.4. No major differences between Bayesian and ML trees were found for all the datasets (see Supplementary Figures for the ML trees).

2.3.7.1. Phylogeny of the phylum Microsporidia

A full SSU, a full ITS, and partial LSU sequences of representative species from the major clades of microsporidians (clade 1-5; Vossbrinck, Debrunner-Vossbrinck, & Weiss, 2014) and several sequences that are similar to our sequences (>88 %), based on a BLAST search, were obtained from GenBank (Supplementary Table 2.4). Two species that belong to the 'expanded Microsporidia', *Nucleophaga amoebae* (JQ288099), *Paramicrosporidium saccamoebae* (JQ796369), and one aphelid species, *Amoeboaphelidium protococcarum* (JX507298), were used as outgroups (Bass et al., 2018). An alignment of 1,115 bp of 93 sequences was used for tree reconstruction after eliminating ambiguous sites in Gblocks as described above.

2.3.7.2. Phylogeny of the genus *Dictyocoela*

Representative sequences of each dictyocoelan species were included for the analysis. A full SSU, a full ITS, and partial LSU sequences were used to resolve deeper relationships within the genus (Bacela-Spychalska et al., 2018). Ten dictyocoelan species (*D. duebenum*, *D. muelleri*, *D. roeselum*, *D. berillonum*, *D. dipoereiae*, *D. gammarellum*, *D. cavimanum*, *D. deshyesum*, *D.* sp. N1, and *D.* sp. N4) known from Eurasia and the USA were included along with our seven newly identified, species-level

taxa (see results) of the *Unikaryon-Dictyocoela* group. *Glugea anomala* (AF044391), *Pleistophora mulleri* (FN434084), *Spraguea lophii* (AF104086), *Nosema granulosis* (AJ011833), and *Enterocytozoon bieneusi* (L07123) were included as outgroups.

2.3.7.3. Cophylogeny on a regional scale: *Dictyocoela* and *Paracalliope*

A fine-scale tree (population-species level) was made for all haplotypes belonging to the dictyocoelan species (*Dictyocoela* sp. NZ1-3) discovered from the *Paracalliope* species complex to test for congruent phylogeny between parasite and host species. The two sequences of *Dictyocoela* sp. NZ4, obtained from *Paracorophium excavatum* and *Melita awa*, were used as outgroups. For host phylogeny, nuclear 28S sequences obtained from each *Paracalliope* population were included in the ingroup. Three sequences with the highest similarity in GenBank were used as outgroups, based on the BLAST search. Several methods are available for cophylogenetic analyses. A tanglegram is commonly used to visually represent congruence between two phylogenies (Page, 2003; but see de Vienne, 2019 for criticism). Therefore, a tanglegram was drawn manually on a vector graphics editor, Affinity Designer (https://affinity.serif.com/). Additionally, overall congruence between parasites and hosts was quantified using Procrustean Approach to Cophylogeny (PACo), one of the commonly used distance based global-fit methods for cophylogenetic analysis (Balbuena et al., 2013). PACo computes a goodness-of-fit statistic from the residual sum of squares of the Procrustean fit as a measure of congruence between parasite and host phylogenies, with its significance established by randomization of the host-parasite association matrix. It also allows for the assessment of the contribution of each individual host-parasite association to the overall global fit. PACo provided several advantages for our study. Firstly, this method does not require fully resolved trees and allows multiple parasite-host associations for analysis. Secondly, PACo is especially appropriate for study systems where one phylogeny is expected to depend upon another. Assuming inherently high dependence of *Dictyocoela* upon their hosts (i.e. the former being an obligate intracellular parasite), I hypothesized that the phylogeny of *Dictyocoela* should mirror that of its *Paracalliope* hosts, by showing a significant degree of congruence with amphipod phylogeny. Three data matrices were used as input: two phylogenetic trees of hosts and parasites, and a binary matrix of parasite-host

associations. The two trees were transformed into matrices of patristic distances, and then the parasite matrix was rotated and scaled to fit the host matrix by Procrustean superimposition. A residual sum of squares was obtained as a global goodness-of-fit statistic; its significance was established by assigning hosts randomly to parasites in the parasite-host matrix with 100,000 permutations. The sum of squared residuals and the upper 95% confidence intervals of each parasite-host link were obtained using a jackknife method, and used to assess the contribution of each link to the overall goodness-of fit. A significance level of 0.05 was applied for all the analyses.

2.3.7.4. Cophylogeny at a global scale: *Dictyocoela* and amphipods

To evaluate the level of cophylogenetic congruence at a global scale between *Dictyocoela* and their respective amphipod hosts, the distance-based PACo was again used. Only one sequence per species of *Dictyocoela* was included for tree inference, because including multiple sequences of the same species could overestimate the degree of phylogenetic congruence (Refrégier et al., 2008). For amphipod hosts, I inferred a genus-level phylogenetic tree, since the family Gammaridae was not recovered as a monophyletic group in Copilas et al. (2020). I used 18S, 28S, and COI sequences available in GenBank and obtained in this study (Supplementary Table 2.5). Most of the *Dictyocoela* were found from freshwater amphipods; because freshwater amphipods evolved independently multiple times from marine groups (Lowry and Myers, 2017), most freshwater amphipod families are distantly related and comprise a small portion of the diversity across all amphipods. However, unresolved trees would not affect our inferences since PACo estimates overall congruence of the two phylogenies based on the patristic distances which measure the amount of genetic divergence accounting for the divergence time among taxa (Balbuena et al., 2013). Using these two trees, PACo analysis was conducted as described above.

2.4. RESULTS

2.4.1. Microsporidians are widespread in diverse New Zealand amphipods

Microsporidians were widely distributed in freshwater and estuarine amphipods in New Zealand (Figure 2.1). Also, a putative microsporidian species was detected from a freshwater-terrestrial amphipod species (*Parorchestia tenuis*). Among 69 locations, microsporidians were detected from 51 sites (73.9 %). Because I used pooled samples, the actual prevalence in a population could not be estimated. Relative prevalence varied from low (1/12 pooled samples) to high (12/12 pooled samples) among populations and sites (Supplementary Table 2.2). Among 7 identified host taxa, 5 harbored microsporidians: *Paracalliope*, *Paraleptamphopus*, *Paraorchestia*, *Paracorophium*, and *Melita*. Forty-six of 59 *Paracalliope*, 2 of 4 *Paraleptamphopus*, 1 of 6 *Parorchestia*, 2 of 3 *Paracorophium*, and 2 of 2 *Melita* populations were positive for microsporidian infections. In total, 169 of 724 pooled samples (23.3 %) tested positive for microsporidians.

2.4.2. Placing species-level taxa within the phylum

A total of 71 SSU sequences was obtained from the 51 sampled locations. In total, 31 haplotypes were identified from 71 sequences, which were delimitated as 9 species-level taxa. A BLAST search against GenBank showed that 28 of these haplotypes were genetically similar to sequences of the genus *Dictyocoela* (86~94 % uncorrected sequence similarity). Of the remaining 3 haplotypes, one was 96.4% identical to the sequence of *Unikayon legeri* (KX364285), a hyperparasitic species (a parasite whose host is also a parasite) infecting the digenean trematode *Meiogymnophallus minutus*, which is a parasite of cockles (Stentiford et al., 2017). The Bayesian tree of the phylum shows the phylogenetic positions of the newly identified species (Figure 2.2). *Dictyocoela* and *Unikaryon* are closely related, forming a monophyletic clade (PP=1). Two haplotypes were located outside of the *Unikaryon-Dictyocoela* clade (Figure 2.2). One of these two haplotypes (GenBank ID: MT462181) was obtained from both *Paracalliope* and *Paracorophium* in Lake Waihola (S16), and is similar to sequences obtained from other amphipods forming a monophyletic clade (Figure 2.2). The other

haplotype (GenBank ID: MT462180) found from a divergent lineage of *Paracalliope* in Kaingaroa (N1) was 91.27 % identical to *Facilispora margolish* (HM800849) previously reported from a parasitic copepod in the northeast Pacific Ocean (Jones et al., 2012).

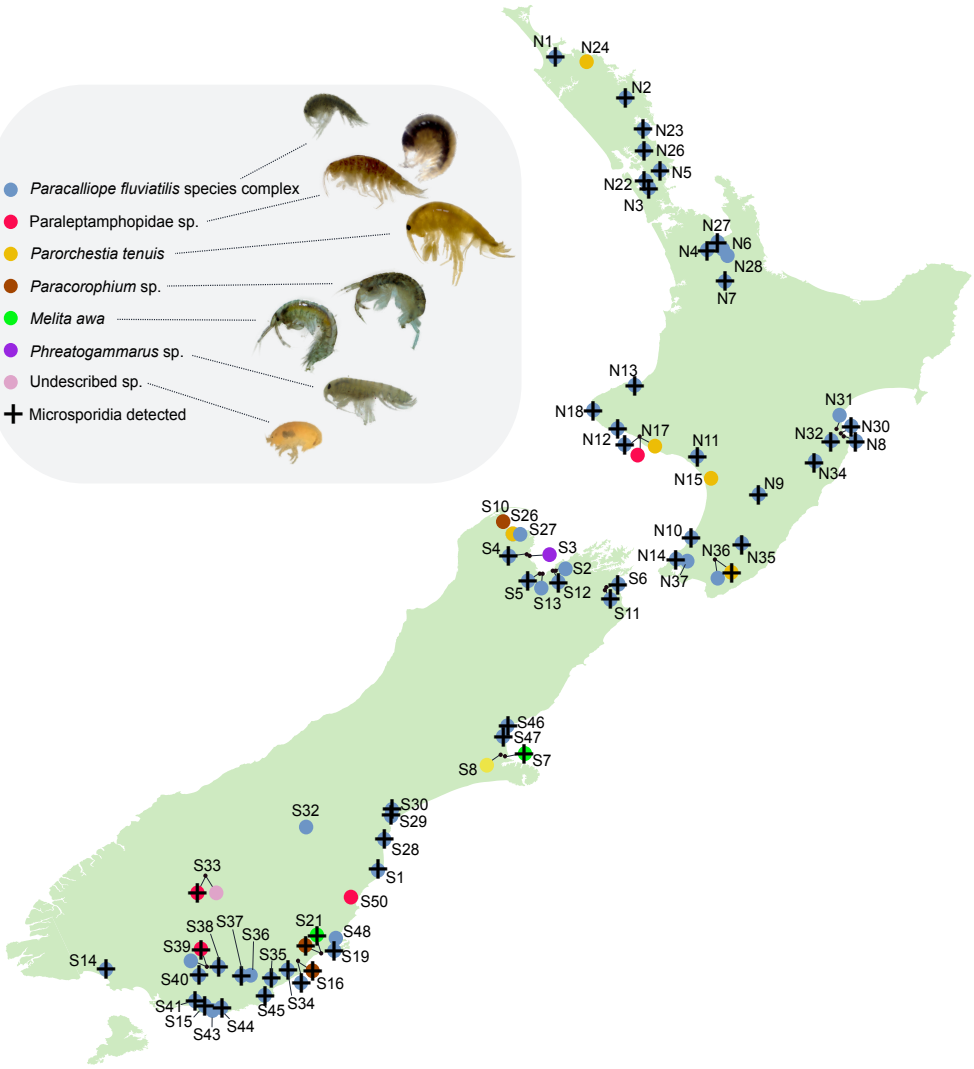


Figure 2.1 Map of New Zealand showing the 69 sampling sites with circles. Seven different families of amphipods are marked with circles of different colours. The sites where microsporidians were detected are marked with +.

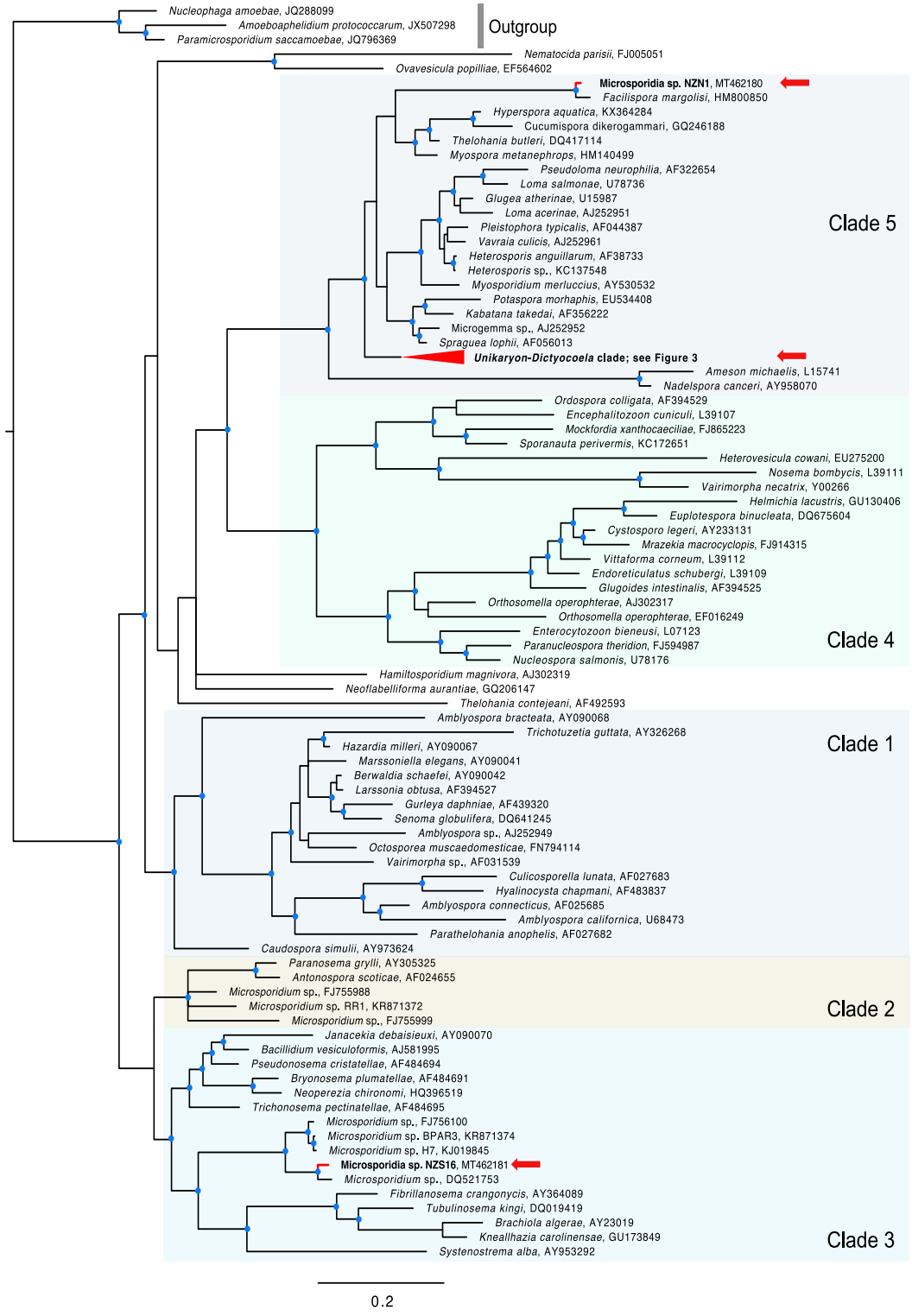


Figure 2.2 Bayesian trees showing phylogenetic positions of microsporidians obtained from this study within the phylum Microsporidia. Nodes with posterior probability higher than 0.9 are shown with blue circles. Major clades defined from Vossbrinck et al. (2014) are marked. All our newly discovered sequences (red arrows) belong to either Clade 3 or 5

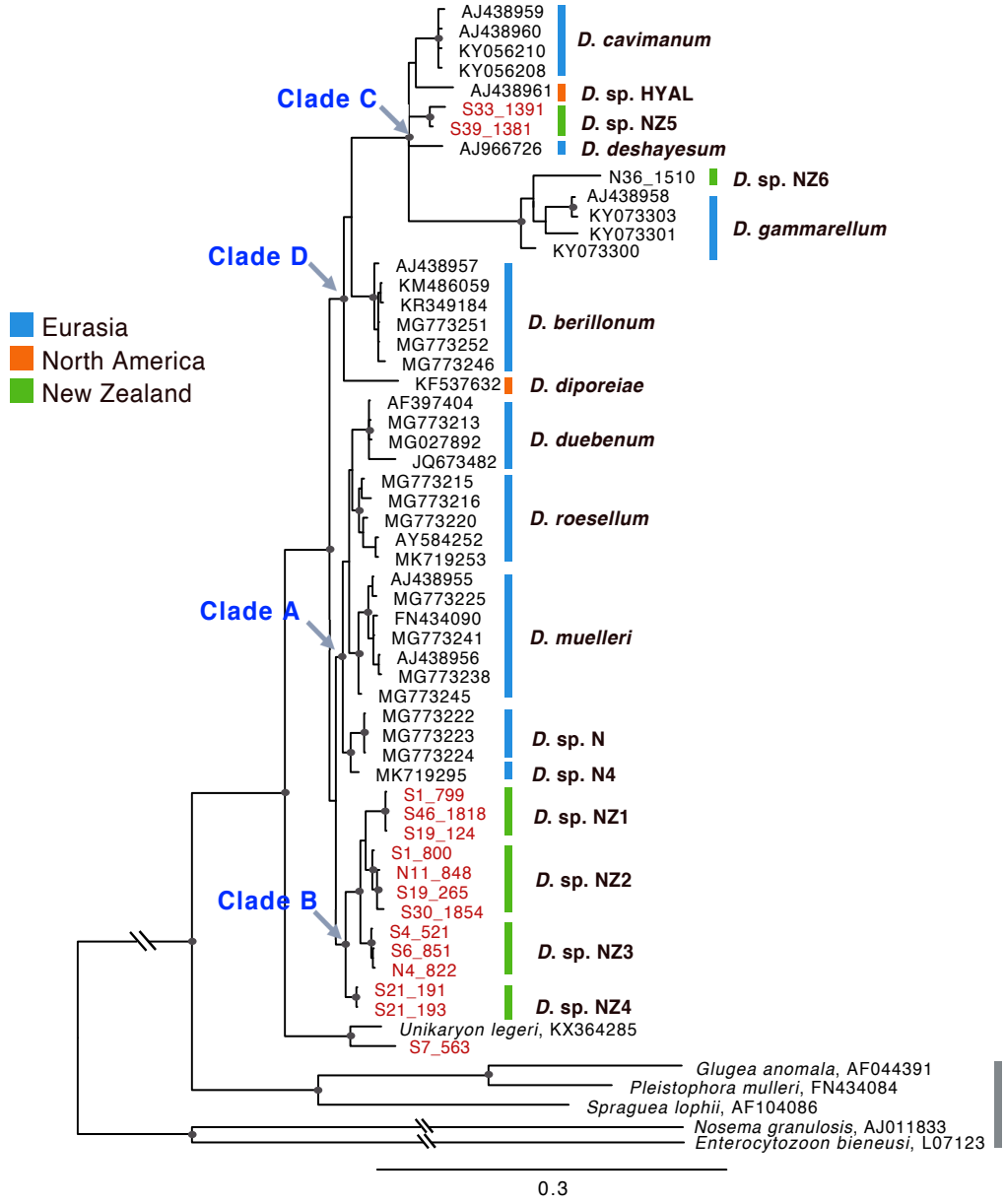


Figure 2.3 Bayesian tree of the relationships among species within *Dictyocoela*. Each coloured bar represents a species and their region of origin is shown with different colours. Well-supported major groups are marked (Clade AD)

2.4.3. Six new molecular species identified within *Dictyocoela*

Figure 2.3 shows phylogenetic relationships among the dictyocoelan species obtained globally, including six species-level taxa (*Dictyocoela* sp. NZ1-6) obtained in this study. Several well-supported clades were identified. Clade A (PP=0.97) contains all the dictyocoelan species obtained from gammarids in Eurasia, except for *D. berillonum*. Clade B (PP=0.98) includes all the dictyocoelan species obtained from *Paracalliope*, *Paracorophium*, and *Melita* (*Dictyocoela* sp. NZ1-4), all obtained from this study. Clade A and B were grouped together, but their sister relationship was weakly supported (PP=0.74). Clade C includes all the microsporidian species from talitrids from Europe, New Zealand, and the USA. *Dictyocoela* sp. NZ5 from Paraleptamphopidae was the exception, being the only species in this clade not found from the superfamily Talitroidea. In addition to Clade C, *D. berillonum* and *D. diporeiae* were clustered together forming a highly supported Clade D (PP=1).

2.4.4. Host specificity of New Zealand microsporidian species

Dictyocoela sp. NZ1-3 were exclusively detected from *Paracalliope* amphipod hosts despite their large geographic ranges, suggesting strong host fidelity. Some evidence for host specificity was observed at several sites with co-occurring species. Two or three amphipod species of different families were sampled from 6 locations (see Figure 2.1). In 4 of 6 of these locations, a microsporidian species was detected from only one species. However, some evidence of horizontal transmission among distantly related hosts was observed in 2 locations. Specifically, *Paracorophium excavatum* and *Melita* sp., in Kaikorai estuary, harbored dictyocoelan sequences that were genetically very similar to each other (*Dictyocoela* sp. NZ4; > 99.4% similarity). Similarly, *Paracorophium excavatum* and *Paracalliope* sp. shared the same microsporidian species (Microsporidia sp. NZS16; 100% identical sequences) in Lake Waihola.

2.4.5. Cophylogeny at a regional scale

Figure 2.4 shows a tanglegram of associations between *Dictyocoela* and their hosts from the *Paracalliope* species complex (see also supplementary Figure 2.3). Based on 28S sequences, host populations were largely divided into two groups showing geographic structure: Northern and Central (NC) and Southern (S) (Figure 2.4B). Interestingly, the

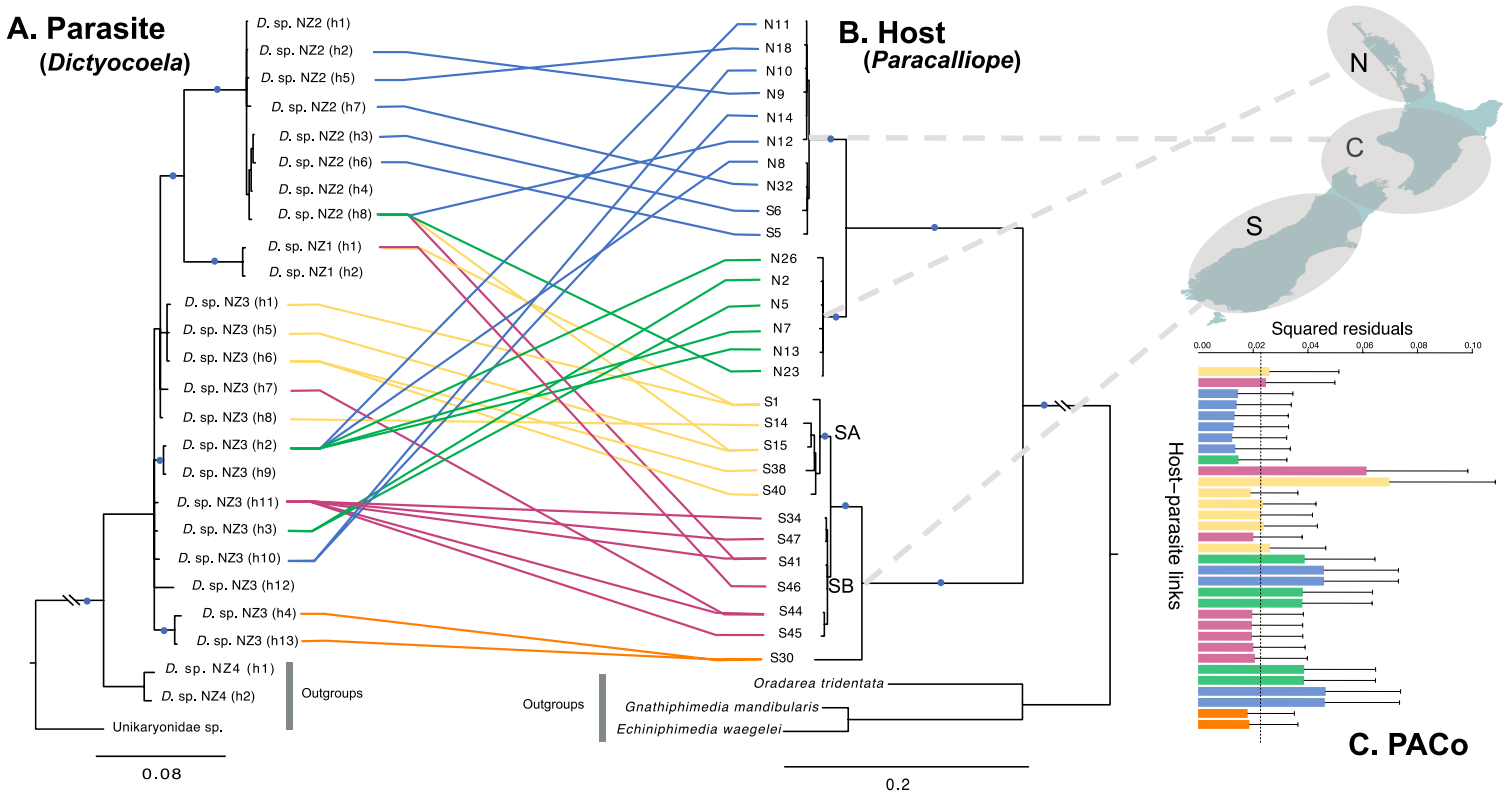


Figure 2.4 Tanglegram illustrating associations between parasites (*Dictyocoela*) and their hosts (*Paracalliope*). Lines between the two trees represent individual parasite-host links. Colours of the lines and bars are marked based on the host group (see below). (a) Bayesian tree of the 26 haplotypes of SSU rDNA sequences obtained from D. sp. NZ1-3. (b) Bayesian tree (28S rDNA sequences) showing the relationships among *Paracalliope* populations. Four main groups are defined (N, C, SA and SB). The map shows geographic distributions of the host groups (upper right). (c) Residual bars of each parasite-host link are shown. The dotted line shows the median residual value and error bars show the upper 95% confidence intervals (see Figure S3 for identity of each parasite-host link)

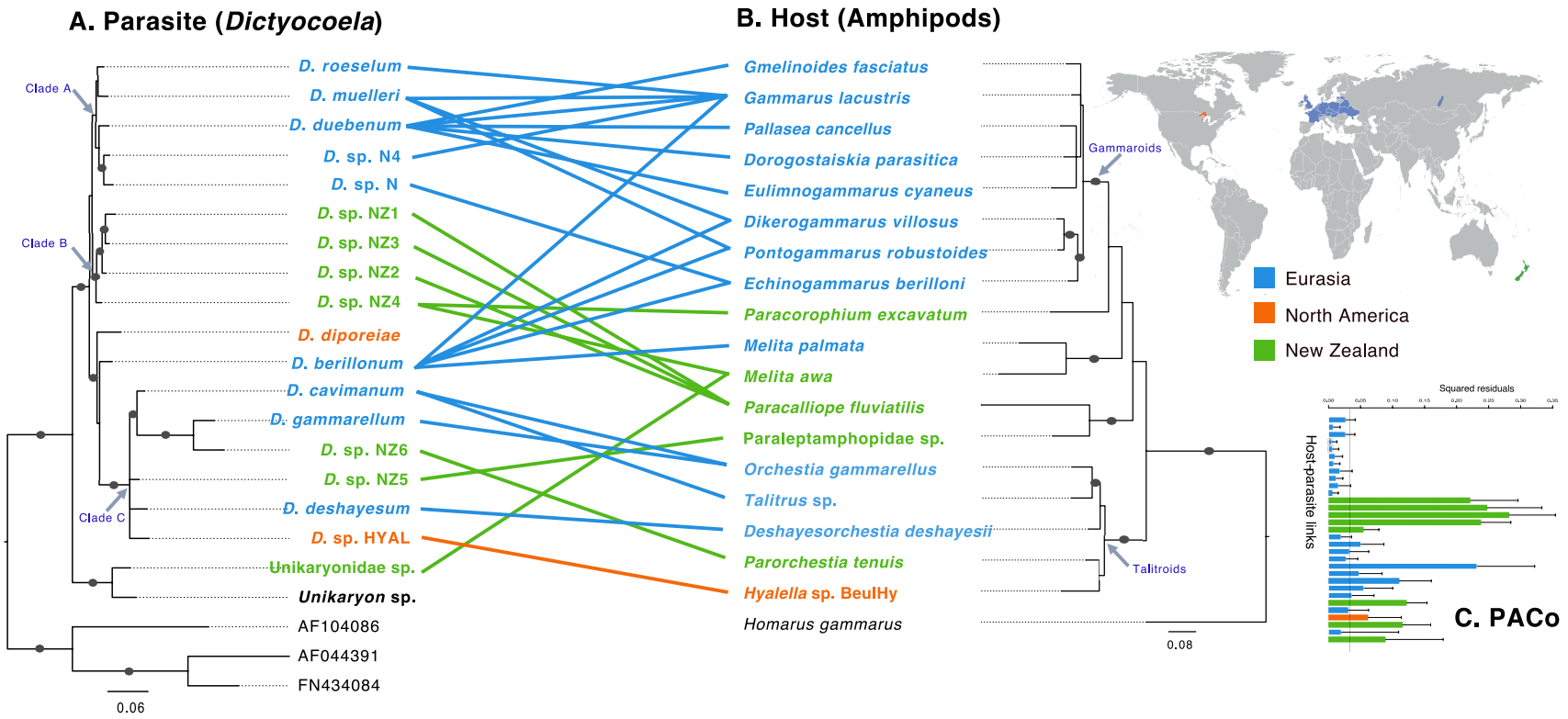


Figure 2.5 Tanglegram showing associations between the parasites (*Dictyocoela*) and their amphipod hosts. The same colour code from Figure 3 was used. (a) The species level Bayesian tree of the genus *Dictyocoela* based on the SSU and LSU sequences. (b) The genus level Bayesian tree of all the amphipods which harbor Dictyocoela based on the concatenated data set (18S, 28S and COI sequences). (c) Residual bars of each parasite-host link are shown. The dotted line shows the median residual value and error bars show the upper 95% confidence intervals (see Figure S4 for identity of each parasite-host link)

Southern group was further divided into 2 main subgroups (SA and SB) and a divergent lineage (S30), with *Dictyocoela* haplotype (rather than geography) being a strong predictor of host genotypic groups, i.e. host populations of subgroups SA and SB harbored different haplotypes of *Dictyocoela*. *Dictyocoela* sp. NZ3 was found throughout the country but some haplotypes of this species were associated only with the SA, SB, or NC groups (Figure 2.3). Within *Dictyocoela* sp. NZ2, six of eight haplotypes were found in group C. Also, this species included a widespread haplotype that was found throughout the country. The rare species, *Dictyocoela* sp. NZ1, was only found at a few locations in the South Island. Overall, some congruent patterns between parasite and host phylogenies were observed from visual inspection. In addition, I tested for a significant congruent pattern between parasite and host phylogenies using PACo including all the three species of *Dictyocoela* (*Dictyocoela* sp. NZ1-3) found in New Zealand. I rejected the null hypothesis of random association (m^2= 0.8683122, P= 0.01559), in favour of the alternative hypothesis that overall *Dictyocoela* phylogeny is constrained by that of their amphipod hosts.

2.4.6. Cophylogeny at a global scale

Figure 2.5 shows the tanglegram between the genus *Dictyocoela* and its amphipod hosts. PACo analysis provided evidence for rejection of the null hypothesis that host phylogeny does not predict the parasite ordination (m^2= 1.93561, P<0.0001). Therefore, I opted for our alternative hypothesis that the host-parasite associations show some degree of phylogenetic congruence. Several host-parasite association patterns were observed. Firstly, clade A of *Dictyocoela* was found only in gammarid hosts, with the respective squared residuals of each of their individual links contributing the most to the congruence of the two phylogenies (see also supplementary Figure 2.4) and could potentially represent codivergence links. Secondly, Clade B was found only in New Zealand amphipods but in genetically distantly related families (Paracalliopidae, Corophiidae, and Melitidae). Squared residuals associated with these links contributed the least to the level of congruence of the phylogenies. Thirdly, all the species from clade C, except for *Dictyocoela* sp. NZ5, were found in talitroid hosts distributed across different continents. Generally, the host-parasite links of clade C seem to represent

incongruent coevolutionary links given the high associated residuals (Figure 2.5C; see also supplementary Figure 2.4).

2.5. Discussion

Our study uncovered the diversity of microsporidian parasites in amphipods in New Zealand and, for the first time, across the Southern Hemisphere. I also investigated patterns of codiversification by comparing phylogenies between *Dictyocoela* microsporidians and their amphipod hosts at both local and global scales, providing new insights into processes that may have shaped their current diversity and distribution. In addition, I inferred the duration and intimacy of amphipod-*Dictyocoela* associations in the context of codiversification history, dispersal limitation, and historical geological events.

It is important to note that I am not explicitly testing coevolutionary diversification, or linking coevolution with codiversification, as I am not assessing reciprocal natural selection pressures or the resulting microevolutionary changes between host and parasite (Althoff et al., 2014; de Vienne et al., 2013). Coevolutionary diversification is the process by which coevolution between two or more taxa increases net diversification in at least one of them (Althoff et al. 2014). Instead, I hypothesise a pattern of codiversification (correlated diversification between interacting lineages; Clayton et al. 2015) as inferred from the level of cophylogenetic and cophylogeographic congruence. I discuss some ecological traits (mode of parasite transmission, dispersal abilities of host and parasite, the presence of genetically similar taxa within the same area) as well as shared biogeographic histories (e.g. covicariance and codispersal) to explain phylogenetic congruence at macroevolutionary scales.

2.5.1. Phylogenetic and phylogeographic congruence between *Paracalliope-Dictyocoela* suggests their shared phylogeographic history

Considering the overall congruent phylogenetic and geographic patterns, and the known geological history of New Zealand, I infer that the phylogeographic pattern of

Dictyocoela may in part reflect the colonization history of *Paracalliope*. Based on the sequence divergence among *Paracalliope* lineages within New Zealand, it was estimated that dispersal of *Paracalliope* from Australia to New Zealand occurred during the Miocene (~17MYA) (Sutherland et al., 2009). During the Pliocene (~5 MYA), New Zealand was divided into several large and small islands due to rising sea levels. The S group may have been isolated from the NC groups since that time. During the Pleistocene (3-1 MYA), the sea level decreased, and land emerged in the southern North Island (Trewick et al., 2011), and *Paracalliope* may have (re)colonized the newly available area like other invertebrates. The North and South Islands were connected until about 500 KYA (Fleming, 1979; Lewis et al., 1994), which may explain why the populations of the southern North Island and the northern South Island are genetically homogeneous. Meanwhile, the presence of the two main host subgroups in the South Island (SA and SB) may have resulted from different colonization events. (Re)colonizations may have occurred due to sea-level changes and occasional flooding and tsunami events in lowland streams, which are common events in the South Island (Scrimgeour and Winterbourn, 1989). When (re)colonization occurred, the associated parasite haplotypes may have co-dispersed with their hosts into the new habitat. Based on their high abundance in the host populations, overlapping geographic regions, host fidelity, and associated population structures, I infer that dictyocoelan parasites were present in the most recent common ancestor of *Paracalliope* and have been maintaining their relationship mainly by vertical transmission.

2.5.2. Codiversification between *Dictyocoela* and their amphipod hosts

The presence of *Dictyocoela* in several ancient lineages of major amphipod families could explain their present-day occurrence in numerous extant amphipod taxa. Gammarids in Europe, Ponto-Caspian, and Lake Baikal are genetically closely related and are believed to have originated from their the most recent common continental ancestors (Barnard & Barnard, 1983; Macdonald, Yampolsky, & Duffy, 2005; Väinölä et al., 2008). This may explain why similar microsporidians were found in diverse gammarid hosts across those regions, despite geographic distances (Ironside and Wilkinson, 2018). On the other hand, frequent horizontal transmission among genetically similar hosts was assumed based on the common presence of *Dictyocoela,*

and their lack of clear host specificity among some gammarids in Eurasia (Ironside & Wilkinson, 2018; Quiles et al. 2019). Meanwhile, the New Zealand amphipod fauna has low species-level diversity but a broad taxonomic range, with 24 described species from 10 genera and 8 families (Chapan et al., 2011). It seems that dictyocoelan species can switch relatively easily between con-familiar species but overcoming the barrier between species of different families is much more difficult. Interestingly, although strong host fidelity of *Dictyocoela* in *Paracalliope* and overall family-level host specificity were observed in New Zealand, host-switching events among genetically distantly related hosts were also inferred (see 3.8). The strongly supported monophyly of clade B (Figures 2.4 and 2.5) suggests that host-shift speciation contributed to the diversity of *Dictyocoela* in New Zealand.

2.5.3. Inferring the evolutionary history of amphipod-*Dictyocoela* associations

Our findings significantly expand the known geographic and host range of *Dictyocoela*. Because *Dictyocoela* is highly specialized in amphipod hosts, its present-day global distribution could be explained either by an ancient origin followed by prolonged association with the host, or perhaps by more recent, host-independent dispersal events. Given the life history characteristics of microsporidians and the strong association between *Dictyocoela* and their amphipod hosts, I suggest three possible scenarios to explain the transoceanic distribution of *Dictyocoela* (Figure 2.6).

Vicariance

An ancient origin of *Dictyocoela* predating the split of the supercontinent Pangaea would be the most parsimonious scenario for the occurrence of *Dictyocoela* species in most freshwater amphipod species whose dispersal abilities are significantly limited (Figure 2.6). In the absence of effective indirect means of transfer, *Dictyocoela* must have been vertically transmitted or horizontally transmitted between other species in the same habitat. It is believed that the breakup of Pangaea and the formation of Laurasia and Gondwana supercontinents played an important role in the diversification of many amphipod groups, although this hypothesis still remains to be tested (Copilaş-Ciocianu et al., 2020). Given that the northern and southern amphipod fauna are largely different

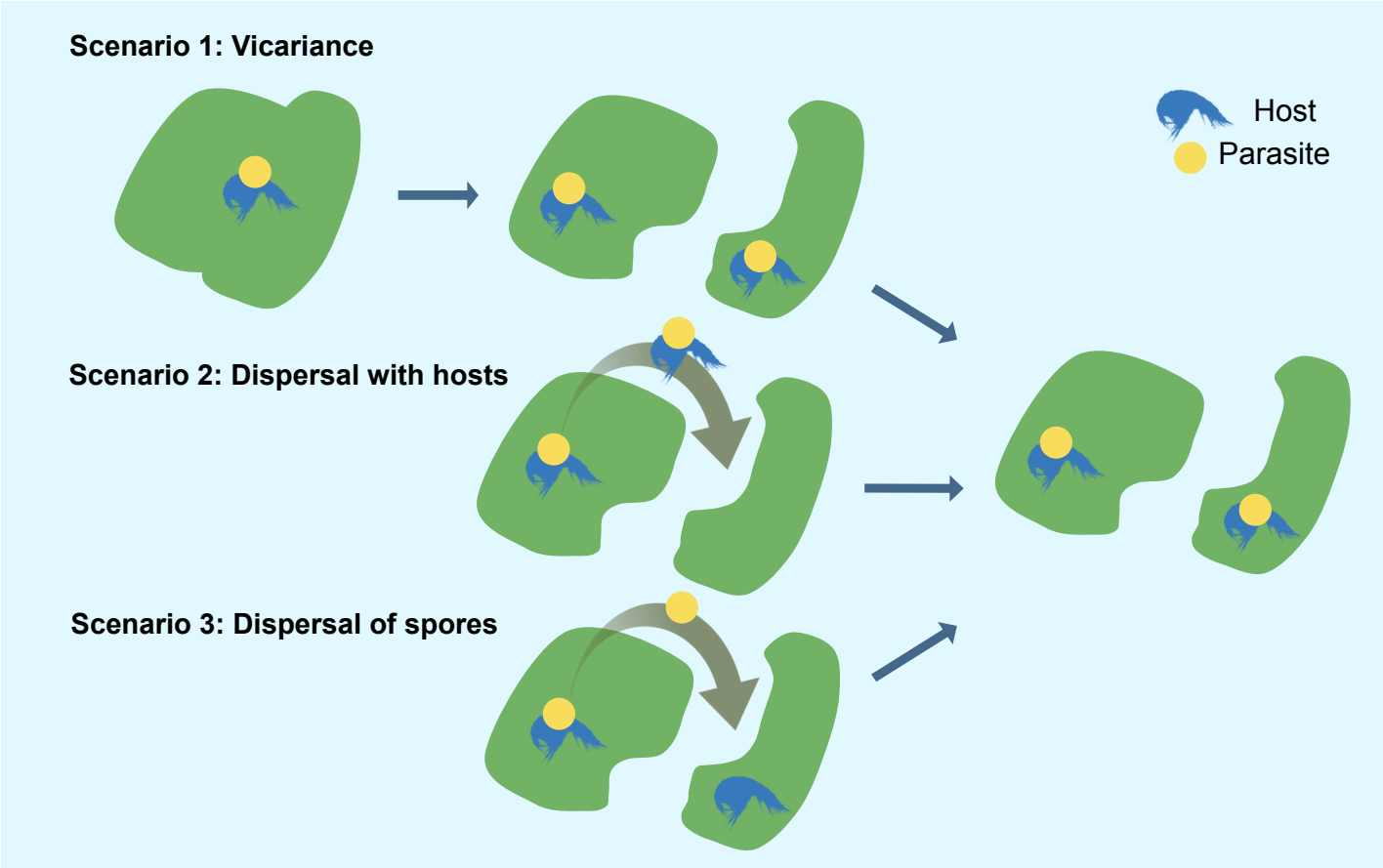


Figure 2.6 Three possible scenarios to explain transoceanic (inter-continental) distribution of parasites. **Scenario 1.** Vicariance. Split of parasite lineages in different continents may be due to vicariance events. **Scenario 2.** Host-dependent long-distance dispersal (LDD). Parasites that are highly dependent on hosts may have travelled with them, if hosts were capable of LDD. **Scenario 3.** Host-independent parasite LDD. Spores may have dispersed via water, air, driftwood, or birds, to other remote area. In all cases, current host-parasite associations are not necessarily assumed. Accordingly, the amphipod host and the parasite in the figure represent hypothetical species

(Barnard, 1974; Lowry & Myers, 2013), the origin of *Dictyocoela* could potentially date back to the split of Pangaea (~180 MYA), before the major diversification of amphipods and/or when they were separated into different lands.

Host dependent Long Distance Dispersal: Dispersal with hosts

Long distance dispersal (LDD) of parasites within or on their hosts (Figure 2.6) could play an important role for parasite dispersal, as shown in ticks on their penguin hosts and tick-borne bacterial pathogens in birds (Moon et al., 2019; Norte et al., 2020). This scenario can explain especially the global distribution of talitroid hosts and their associated *Dictyocoela* parasites. Talitridae is the only family that includes terrestrial species that can disperse by several means (Fenwick & Webber, 2008; Friend & Richardson, 1986; Wildish, 2012). Interestingly, all *Dictyocoela* species detected from talitrids belong to Clade C, suggesting codiversification (Figure 2.5). Also, talitrids and their *Dictyocoela* parasites are distributed globally, consistent with a codispersal scenario.

Host independent parasite LDD: Dispersal of spores

Dictyocoela may have traveled long distances independently as spores. If this is the case, then the origin of *Dictyocoela* does not necessarily need to be ancient. Microsporidians produce spores that are resistant to the external environment, but little is known about their dispersal potential over long distance. Spores may be able to travel in the air and water and explain the wide distributions of opportunistic microsporidians such as *Encephalitozoon* and *Enterocytozoon* (Stentiford et al., 2016). However, this mode is more plausible for generalists and is incompatible with the host specificity observed in *Dictyocoela* in amphipod hosts on remote islands and seen in this study. Nevertheless, because we know very little about the dispersal ability of spores, and the lack of occurrence data of microsporidians in amphipods over large areas, I cannot rule out the possibility of this scenario.

Distinguishing among these scenarios requires more evidence. Firstly, targeted screening of microsporidians of marine and additional freshwater amphipods covering a larger geographic area would provide valuable information for inferring historic host-

parasite associations. Secondly, only with time calibrated phylogenies can we possibly discern between a vicariant origin vs other more recent origin of *Dictyocoela* in New Zealand. Thirdly, assessing the dispersal abilities of microsporidian spores could support the LDD by spore scenario.

2.5.4. Host-parasite associations across scales: from mode of transmission to macroevolutionary patterns

Vertical and horizontal transmission occur within a short time frame and across small geographic scales. Although vertical transmission is often expected to produce congruent patterns with hosts on larger scales (Althoff et al., 2014), the impact of vertical transmission on macroevolutionary patterns has rarely been shown by empirical data. Biological interactions can affect the distribution of species, but this is a scale-dependent process and its role over large scale patterns remains controversial (Araújo and Rozenfeld, 2014; McGill, 2010). A mathematical model predicts that parasites can co-occur with their hosts across geographical scales according to their dependency on the hosts (Araújo and Rozenfeld, 2014). Our study provides empirical evidence that vertically transmitted parasites show similar phylogenetic and geographic patterns to their hosts across spatial scales. However, this does not undermine the role of horizontal transmission (which could lead to host-shift speciation) in parasite diversification. A large body of evidence suggests that host-shift speciation is a common process even for specialized symbionts or vertically transmitted parasites (de Vienne et al. 2013; Bailly-Bechet et al., 2017; Doña et al., 2017). Microsporidians show similar patterns. Despite apparent vertical transmission, frequent host-switching events have been inferred in studies of microsporidians in Eurasia (Ironside and Wilkinson, 2018; Quiles et al., 2019). In New Zealand, even though tight host-parasite associations were inferred based on the congruent spatial-genetic structure in *Paracalliope*, horizontal transmission (or host-shift) among species of different families was also inferred. Therefore, it seems that host-switching may be common from local to regional scales. When comparing Europe and New Zealand, it seems that geological barriers that simultaneously acted on both *Dictyocoela* and their amphipod hosts, coupled with vertical transmission as the main transmission mode, could in part explain their congruent phylogeographies.

In summary, evidence of both vertical and horizontal transmission can be seen at small scales, and both modes may have played pivotal and far-arching roles; however, parasite distribution at larger scales could be mainly explained by host distribution and geographical processes. Our study underscores that considering multiple processes operating at different scales is necessary to explain parasite distribution and its connection to host associations.

2.6. Conclusions

Our study confirmed the worldwide distribution of *Dictyocoela* in many different lineages of aquatic amphipods. Based on their strong reliance on the host for dispersal as an intracellular parasite as well as the limited dispersal capabilities of amphipod hosts, I inferred their intimate association that may have persisted over macroevolutionary time, by comparing phylogenetic and phylogeographic patterns. Both vertical and horizontal transmission may have played substantial roles in the evolution of the parasites. However, at a macroevolutionary scale, host range and geological processes can primarily explain parasite distribution. Our study highlights that considering multiple processes operating at different scales is necessary to explain codiversification of hosts and their parasites. Also, our study shows that uncovering parasite diversity in new host taxa and geographic regions can provide novel insights into the evolutionary history of host-parasite associations. Further studies of diverse host-parasite systems with varying ecological traits and known biogeographic histories will be important to further investigate patterns of codiversification and underlying mechanisms.

Chapter 3

Widespread Torix *Rickettsia* in New Zealand amphipods and the use of blocking primers to rescue host COI sequences

3.1. Abstract

Endosymbionts and intracellular parasites are common in arthropod hosts. As a consequence, (co)amplification of untargeted bacterial sequences has been occasionally reported as a common problem in DNA barcoding. While identifying amphipod species with universal COI primers, I unexpectedly detected rickettsial endosymbionts belonging to the Torix group. To map the distribution and diversity of *Rickettsia* species among amphipods hosts, I conducted a nationwide molecular screening of seven families of New Zealand freshwater amphipods. In addition to uncovering a diversity of Torix *Rickettsia* species across multiple amphipod populations from three different families, our research indicates that: 1) detecting Torix *Rickettsia* with universal primers is not uncommon, 2) obtaining '*Rickettsia* COI sequences' from many host individuals is highly likely when a population is infected, and 3) obtaining 'host COI' may not be possible with a conventional PCR if an individual is infected. Because *Rickettsia* COI is highly conserved across diverse host taxa, I was able to design blocking primers that can be used in a wide range of host species infected with Torix *Rickettsia*. I propose the use of blocking primers to circumvent problems caused by unwanted amplification of *Rickettsia* and to obtain targeted host COI sequences for DNA barcoding, population genetics, and phylogeographic studies.

3.2. Introduction

The cytochrome c oxidase subunit 1 gene (COI), a partial fragment of mitochondrial DNA, is the marker of choice for DNA barcoding, and is also widely used for population genetics and phylogeographic studies (Bucklin et al., 2011; Hajibabaei et al., 2007; Hebert et al., 2004). A variable region is flanked by highly conserved regions; this allowed for the design of a pair of universal primers and their application to various organisms (Folmer et al., 1994; Hebert et al., 2003). With the advancement of fast and cost-effective next-generation sequencing technologies, which enables metabarcoding (Elbrecht and Leese, 2015; Taberlet et al., 2012), the number of COI sequences is increasing rapidly in public databases such as GenBank and The Barcode of Life DataSystems (BOLD) (Porter and Hajibabaei, 2018). However, quality control is often an issue due to the presence of questionable "COI-like" sequences (Buhay, 2009) or nuclear mitochondrial pseudogenes (numts) that are often coamplified with orthologous mtDNA (Song et al., 2008). Bacterial sequences are also often coamplified with universal primers. Indeed, there have been reports of the amplification of untargeted sequences of endosymbiotic bacteria such as *Wolbachia* and *Aeromonas* during DNA barcoding with universal primers and their misidentification as those of invertebrate hosts during deposition in databases (Mioduchowska et al., 2018; Smith et al., 2012).

The bacterial genus *Rickettsia* is another of these endosymbiotic taxa. This genus belongs to the order Rickettsiales along with *Wolbachia*, and comprises diverse pathogenic species that can cause vector-borne diseases in birds and mammals including humans, as well as non-pathogenic species. Some rickettsioses with severe symptoms are well known, and include Rocky Mountain spotted fever, Queensland tick typhus, rickettsial pox, murine typhus and epidemic typhus (Parola et al., 2005; Perlman et al., 2006; Weinert, 2015). To date, at least 13 groups are known within the genus *Rickettsia*: Adalia, Bellii, Canadensis, Guiana, Helvetica, Meloidae, Mendelii, Rhyzobious, Spotted fever, Scapularis, Torix, Transitional, and Typhus (Binetruy et al., 2020; Hajduskova et al., 2016; Weinert et al., 2009). All these groups except the Torix group are exclusively associated with arthropod hosts, such as mites, fleas, ticks, and spiders. The Torix group, which is sister to all other groups, is the only group that includes non-arthropod

hosts such as amoeba and leeches (Galindo et al., 2019; Kikuchi et al., 2002; Kikuchi and Fukatsu, 2005). In addition to these freshwater hosts, the Torix group occurs in diverse arthropod groups that spend part of their life cycle in the aquatic environment (e.g. Coleoptera and Diptera) (Küchler et al., 2009; Pilgrim et al., 2017; Weinert et al., 2009).

Although *Rickettsia* species are known as common pathogens or endosymbionts in arthropod hosts, these agents have never been reported in crustaceans. *Rickettsia*-like organisms (RLO) have been reported in some groups of crustaceans including crabs, crayfish, lobsters, shrimps, and amphipods (Gollas-Galván et al., 2014). However, most reports of these RLOs were based on morphological similarity with *Rickettsia* and were rarely confirmed by molecular data. In amphipods, RLOs were reported in several species of gammarids, as well as other taxa (e.g., *Crangonyx floridanus* and *Diporeia* sp.) (Graf, 1984; Larsson, 1982; Messick et al., 2004; Winters et al., 2015). 16S rRNA (rrs) sequences of RLOs are available for *Diporeia* sp. and some gammarids, but none of them belong to the genus *Rickettsia* (Bojko et al., 2017; Winters et al., 2015).

While identifying amphipod species with DNA barcoding for cophylogenetic analyses between microsporidians and their amphipod hosts (Park et al., 2020), I obtained a suspicious COI sequence from *Paracalliope fluviatilis*, the most common freshwater amphipod species in New Zealand. According to a blast search in GenBank, this sequence obtained from a stream in the Southland region (S15 in Figure 3.1) was ~99.5% identical to 12 sequences of the same 'amphipod' species available in GenBank from a previous population genetic study (Lagrue et al., 2016). These sequences were obtained from two different locations (Waikouaiti River and Waitaki River) in the Otago region. However, all sequences were highly divergent from that obtained from other populations (~57%) of the same host group (*Paracalliope* species complex). DNA from the individual amphipod was extracted from its legs (i.e., low chance of contamination due to gut contents).I obtained a clear chromatogram with no ambiguous peaks. Furthermore, these sequences were similar to other COI sequences obtained from diverse insects (Coleoptera, Diptera, Hemiptera, Hymenoptera, Odonata) in GenBank with sequence similarity ranging from 80 to 99 %. However, these sequences were also

similar (~92%) to sequences of rickettsial endosymbionts of insects and spiders that have been recently registered in GenBank. Because such highly conserved COI sequences among distantly related arthropod groups are unlikely, I assumed that these sequences were actually obtained from their rickettsial endosymbionts. I independently confirmed the presence of *Rickettsia* in our amphipod hosts using three genetic markers that were designed to be specific to *Rickettsia*.

In fact, this phenomenon of the amplification of non-targeted COI sequences with widely used DNA barcoding primers has already been reported. Řezáč et al. (2014) obtained a rickettsial COI sequence from a spider species in a study with a different purpose. Ceccarelli et al. (2016) obtained COI sequences of Torix *Rickettsia* from six individuals of a spider species while conducting DNA barcoding, and these authors formally discussed the presence of misidentified COI sequences in GenBank. However, despite these early reports, the deposition of misidentified sequences to GenBank has continued until recently. A very recent survey on BOLD reported that 0.41 % of the barcode submissions in BOLD are actually from *Rickettsia*, which is higher than that from *Wolbachia* (0.17 %) (Pilgrim et al., 2020).

Because *Rickettsia* species are endosymbionts within host cells, DNA extracts from infected host tissue will inevitably include DNA of endosymbionts as well. If binding sites for 'universal primers' are conserved in both hosts and their endosymbionts, PCR products obtained from these mixed templates may result in mixed signals in chromatograms, or in the amplification of endosymbiont instead of host sequences (Ceccarelli et al., 2016). Using primers that are designed to bind uniquely and specifically to host templates would reduce this problem. However, designing group-specific primers is not always possible, especially when reference sequences are scarce or not available. Also, finding conserved regions across a given taxonomic group may not be achievable. Alternatively, blocking primers can be used to prevent the amplification of unwanted or dominant sequences among DNA templates (Vestheim and Jarman, 2008). For example, this method has been successfully applied to identify prey items (by suppressing the amplification of predator DNA in gut contents), or to obtain rare mammal sequences from ancient DNA (by blocking the amplification of

human DNA) (Boessenkool et al., 2012; Vestheim et al., 2011). Because COI sequences of Torix *Rickettsia* are highly conserved in diverse host groups, I were able to design blocking primers that are intended to specifically block the amplification of Torix *Rickettsia* but allow amplification of the COI region of (any) host mtDNA.

In this study, I first screened rickettsial infections in diverse amphipods collected throughout New Zealand to determine their prevalence and distribution. Secondly, I characterized the genetic diversity of the newly found *Rickettsia* species in relation to other Torix *Rickettsia* using 4 distinct markers, namely *rrs*, *gltA*, *atpA*, and COI, and expand current understanding of *Rickettsia* phylogeny. Thirdly, I demonstrate that unwanted amplification of rickettsial COI sequences during DNA barcoding is a common problem, and that such sequences have been frequently reported and misidentified in GenBank. Fourthly, I suggest that using blocking primers in addition to universal primers for PCR is an effective solution to obtain targeted host COI sequences. Finally, I discuss the implications of these pseudo-sequences in public databases, ways to reduce this problem, and possible applications of blocking primers for similar problems.

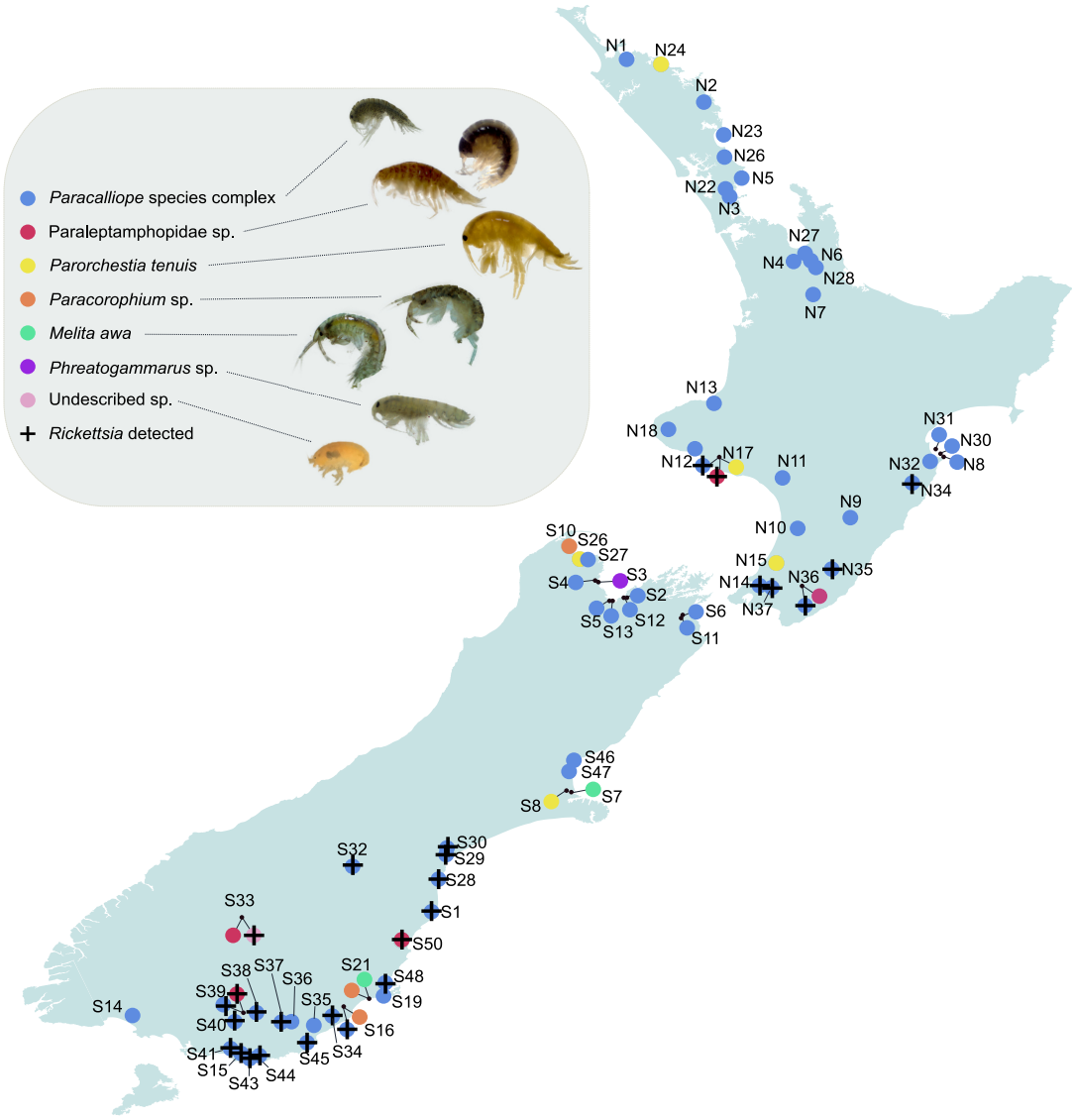


Figure 3.1 Map of sampling sites. Map of New Zealand showing the sixty-nine sampling sites with circles. Seven different families of amphipods are marked with different colours in the circles. The sites where *Rickettsia* was detected are marked with +. Site codes correspond to those in Supplementary Table3.1.

3.3. Materials and methods

3.3.1. Confirming *Rickettsia* infections by PCR

The presence of *Rickettsia* was confirmed by amplification of three different markers (*rrs*, *gltA*, and *atpA*) using *Rickettsia*-specific primer pairs (Küchler et al., 2009; Pilgrim et al., 2017) (Table 3.2). The DNA sample (S15_470) from which I obtained presumably 'rickettsial COI' was used as DNA templates. Seven additional DNA samples from the same population were also included for PCR detection to compare the efficiency of the primer sets in order to select the best marker for molecular screening. For PCR reactions, 12.3 μl of distilled water, 4 μl of reaction buffer, 0.8 μl of each forward and reverse primers, 0.1 μl of MyTaq (Bioline), and 2 μl of DNA were used. PCR reactions were conducted under the following conditions: 95°C initial denaturation for 5 min, 35 cycles of 94°C for 30s, 54°C for 30s, 72°C for 120s, final extension for 7 min at 72°C. Then, 2μl of PCR product from each PCR reaction was run on a 1.5 % agarose gel.

3.3.2. Molecular screening of *Rickettsia* in New Zealand amphipods

I obtained extracted DNA samples of diverse amphipod specimens from Park et al. (2020; Chapter2), in which the authors investigated the diversity of microsporidian parasites, a group of obligate intracellular eukaryotic parasites of amphipod hosts. Seven families of amphipods (Melitidae, Paracalliopidae, Paraleptamphopidae, Phreatogammaridae, Talitridae, Paracorophiidae, and an undescribed family of Senticaudata) were collected from 69 locations throughout both the South and North Islands (Figure 3.1 and Supplementary Table 3.1). A total of 724 pooled DNA samples obtained from 2,670 individuals (mostly 4 individual amphipods per pool) were screened for *Rickettsia* by PCR by amplifying the *rrs* region under the PCR conditions and procedures described above.

Table 3.1 List of primers used in this study. All blocking primers were modified with C3 spacer at the 3' end. Five deoxyinosines (I) were added in the middle of HCO_DPO primer as a linker.

Gene	Primer	Sequence (5'-3')	References
16S rRNA	Ri170_F	GGGCTTGCTCTAAATTAGTTAGT	Kuchler et al. 2009
	Ri1500_R	ACGTTAGCTCACCACCTTCAGG	Kuchler et al. 2009
gltA	RiGltA405_F	GATCATCCTATGGCA	Pilgrim et al. 2017
	RiGltA1193_R	TCTTTCCATTGCCCC	Pilgrim et al. 2017
atpA	RiAtpA327_F	GTCGGTAAAGCATTGCTTGGT	Pilgrim et al. 2017
	RiAtpA1309_R	ATTGATCCTGCTTCAATA	Pilgrim et al. 2017
CoxA	LCO1490	GGTCAACAAATCATAAAGATATTGG	Vrijenhoek 1994
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	Vrijenhoek 1994
CoxA (Blocking primers)	COI_BlocF	GATATTGGCRTAATGTATAT	This study
	COI_BlocR	AAATCAAAATAAATGCTGA	This study
	BlocF_2	TTGGCRTAATGTATATTATATTYGCC	This study
	HCO_DPO	GGRTGACCAAAAAAYCAAAATAAATGCTGRIIIIIYACTGGATCTCC	This study

3.3.3. Sequencing

rrs, *gltA*, and *atpA* sequences were obtained from populations that had been found positive for *Rickettsia* infections by PCR screening (Supplementary Table 3.1). PCR products were purified with MEGAquick-spinTM Total Fragment DNA Purification Kit (iNtRON Biotechnology) following the manufacturer's instructions. Purified PCR products were sent to Macrogen, Korea, for Sanger sequencing. Raw sequences were aligned in Geneious Prime 2019.0.4 (https://www.geneious.com) and ambiguous sites were carefully examined by eye. Haplotypes were identified by using the package *pegas*(Paradis, 2010) in R version 3.5.2 (R Development Core Team, 2011). I obtained some rickettsial COI sequences as a byproduct during host identification of amphipods with universal primers. I included these COI sequences along with nucleotide sequences from other genes for further analyses.

3.3.4. BLAST search

Blast search was done on GenBank with the *rrs*, *gltA*, *atpA*, and COI sequences obtained in this study. Based on the result of BLAST searches, all sequences that were considered as the Torix group (sequences with similarity to the query sequence higher than that between the query sequence and the Bellii group *Rickettsia*) were downloaded from GenBank (see Supplementary Tables 3.3-5) for further phylogenetic tree reconstructions. *rrs* and *gltA* sequences from other *Rickettsia* groups were also obtained and included for tree reconstruction (Supplementary Table 3.6). In addition, *rrs* sequences from recently discovered close relatives to *Rickettsia* (*Candidatus* Trichorickettsia, *Candidatus* Gigarickettsia, and *Candidatus* Megaira), and *Orientia tsutsugamushi* were included as outgroups.

3.3.5. Phylogenetic analyses

For each gene set, all sequences were aligned in Geneious Prime with the MUSCLE algorithm. Ambiguous sites were then eliminated in Gblocks with the least restrictive setting (Castresana, 2000). The best-fitting model of nucleotide evolution for each dataset was determined based on the corrected Aikake information criterion (AICc) using jModelTest v2.1.6 (Darriba et al., 2012), which was conducted through the CIPRES Science Gateway v3.3 (Miller et al., 2010). For all datasets, the General Time

Reversible (GTR) model of nucleotide substitution along with Gamma distributed rate variation across sites (G) and the proportion of invariable sites (I) were chosen as the best model. Bayesian trees were inferred in MrBayes 3.2.7a (Ronquist et al., 2012). For each dataset, two independent runs, which consisted of four chains each, were simultaneously conducted for 10,000,000 generations with a sampling frequency of 1000. The initial 25% of samples were discarded. The resulting trees were visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

3.3.6. Design of blocking primers

Blocking primers were designed following the guidelines of Vestheim et al. (2011). In order to design blocking primers, COI sequences of the Torix group and some species belonging to other groups of *Rickettsia*, and COI sequences of New Zealand amphipods were aligned in Geneious Prime (Figure 3.2). I designed four different annealing inhibiting blocking oligos which were intended to compete with universal primers. All met the following criteria: First, the blocking primers should overlap with one of the universal primers. Second, the blocking primers should specifically bind to the unwanted DNA templates (i.e. *Rickettsia*) but not to our target DNA templates (i.e. amphipod hosts). Third, 3'-end was modified so that it does not prime amplify (here, all with C3 spacer). Initially, two primers were designed: Bloc_F and Bloc_R (Table 3.2). However, GC contents of these primers were too low (27.5% and 21%, respectively), which resulted in a low expected melting temperature (Tm) of 43.2°C and 42.5 ºC, respectively. Ideally, Tm of a blocking primer should be higher than that of the competing primers (Vestheim et al., 2011). I, therefore, designed a longer primer, BlocF_2, to increase Tm to 51.7 ºC (Table 3.2). A fourth primer was designed with a dual priming oligonucleotide (DPO): HCO_DPO (Table 3.2). A DPO can be used when it is impossible to find an appropriate binding site for a blocking primer adjacent to a binding site of a universal primer (Vestheim and Jarman, 2008). A DPO primer consists of two separate segments connected with five deoxyinosines, and with C3 spacer modification at the 3'end. The total length of a typical DPO primer is long but it does not suffer from high Tm because a deoxyinosine linker, which assumes a bubble-like structure, allows the two segments to act independently (Chun et al., 2007). All

synthesized primers were purified with polyacrylamide gel purification (PAGE) to increase binding specificity by removing under-synthesized oligos.

3.3.7. Validation of blocking primers

I applied fragment analysis to test and compare the effectiveness of our blocking primers (Vestheim and Jarman, 2008). Fragment analysis of fluorescently labeled PCR products on capillary electrophoresis can separate fragments in different sizes and can be used as a semi-quantitative method. When amplified with the universal LCO1490 and HCO2198 primers, the expected lengths of PCR products were different for amphipod hosts and *Rickettsia* COI, because Rickettsial COI is 6 bp longer. The FAM dye was attached to the 5' end of the LCO1490 primer. This fluorescently labeled forward primer was added to the PCR mixture instead of the normal (unlabeled) LCO1490 primer. Various factors can affect PCR success with blocking primers: Tm of primers, the concentration of primers (relative ratio between blocking primer and regular primer), the amount of the DNA templates in a PCR mixture (concentration of DNA), and the number of PCR cycles (Vestheim and Jarman, 2008). To optimize PCR conditions, PCR reactions were conducted under several different PCR conditions (Table 3.2). Fragment analyses were carried out with a 1,200 LIZ size marker on an ABI 3730xl System (Applied Biosystems) at Macrogen (Korea). Results were analyzed with Peak Scanner Software 1.0 (Applied Biosystems; https://www.thermofisher.com/).

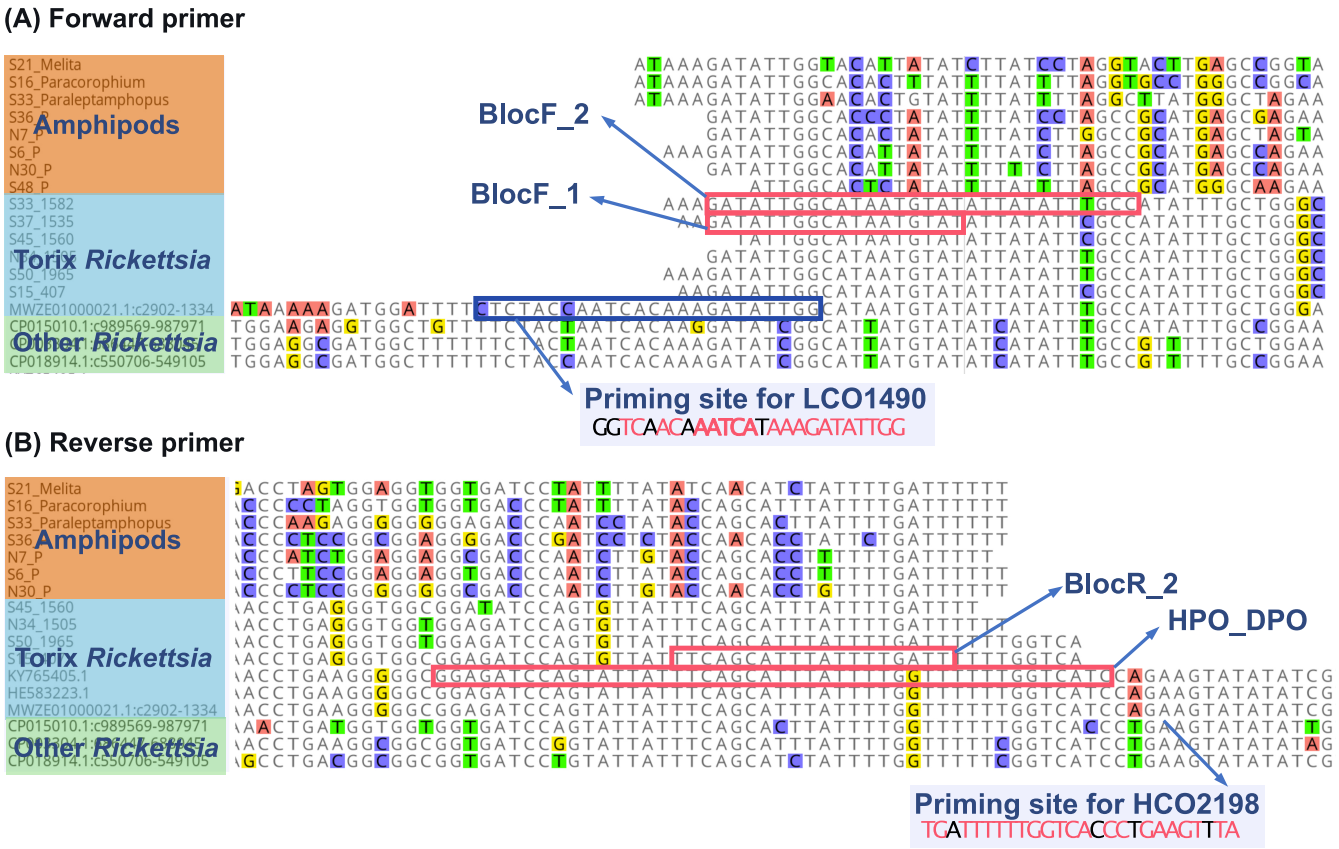


Figure 3.2. COI sequence alignments showing priming sites for both forward and reverse primers. Conserved regions to which universal primers bind are highlighted with blue squares. Positions at which nucleotides are the same as in universal primers are highlighted with pink texts in the primer sequences. COI sequences are highly divergent among amphipods whereas COI sequences from Torix *Rickettsia* from diverse host groups are highly conserved, which allowed the blocking primers (binding regions are highlighted with pink squares).

Table 3.2 PCR conditions and results of fragment analyses. PCR products obtained under different PCR conditions (different primers with different concentrations, annealing temperature, different DNA templates, number of PCR cycles) were run on capillary electrophoresis. The number of amplicons of host COI is followed by the number of amplicons of *Rickettsia* COI separated by (/), and their ratio is shown in parenthesis. The effects of blocking primers are highlighted in the grey boxes. Universal to blocking primer ratios that were higher than 0.8 are highlighted in red. NA: fragment analysis was conducted but the result was not available.

	Condition 1				Condition 2				Condition 3			Condition 4
Annealing temperature	44 °C				48 °C				48 °C			48 °C
DNA templates	DNA extract 1				DNA extract 1				DNA extract 2			DNA extract 1
Number of cycles	38				38				38			48
Primer ratio (universal:blocking)	0	2	5	10	0	2	4	8	1	2	3	2
-	NA	-	-	-	448/7010 (0.0601)	-	-	-	NA	-	-	-
COI_BlocF	NA	7126/3426 (0.6753)	9257/1669 (0.8472)	1809/303 (0.8565)	NA	NA	NA	959/4955 (0.1622)	-	-	-	-
COI_BlocR	NA	7428/5660 (0.56754)	NA	1751/1450 (0.5470)	NA	NA	406/5975 (0.0636)	NA	-	-	-	-
COI_BlocF_2	-	-	-	-	NA	404/2028 (0.1783)	631/772 (0.4498)	618/332 (0.6505)	NA	779/2676 (0.2255)	NA	-
HCO_DPO	-	-	-	-	NA	407/0 (1)	156/0 (1)	NA	445/144 (0.7555)	463/91 (0.8357)	175/0 (1)	18488/196 (0.9895)

3.4. Results

3.4.1. Distribution of *Rickettsia* in amphipod hosts in New Zealand

Rickettsia was detected in 26 of 69 locations (37.7%) from 3 families of amphipods: *Paracalliope* species complex (24/59 populations; 40.7%), *Paraleptamphopus* sp. (3/5 populations; 60%), and one undescribed family of Senticaudata sp. (1/1 population; 100%) (Figure 3.1 and Supplementary Table 3.1). Because pooled samples were used, accurate prevalence in each population could not be obtained. However, a relative comparison was possible among the populations in which the same number of individuals per sample and the same total number was used (i.e. populations with a total of 48 individuals, with 12 samples each containing 4 individuals) (Supplementary Table 3.1). With parsimonious interpretation, among 18 populations, 12 populations showed at least 10% prevalence (>5 positive tubes/12 tubes tested). Seven populations showed at least 20% prevalence (>10 positive tubes/12 tubes tested). And three populations showed 100% positive tubes (12/12), with prevalence thus possibly ranging from 25-100%. Although *Rickettsia* was detected in both the North and South Islands, its distribution was confined to the southern parts of both islands (Figure 3.1).

3.4.2. Genetic characterization of *Rickettsia* sequences

At least one *rrs*, one *gltA*, or one *atpA* sequence was obtained from each of the population/species that were positive in the initial molecular screening (Supplementary Table 3.1). Specifically, 24 sequences of *rrs*, 14 sequences of *gltA*, and 19 sequences of *atpA* were obtained. Also, 8 sequences of COI were added to our dataset (Supplementary Table 3.2). Fourteen genotypes were identified using *rrs* sequences. All rrs showed higher similarity to each other (>99.4%). All *gltA*, *atpA*, and COI sequences of Torix *Rickettsia* from amphipods showed high similarity to each other: >95%, >94%, >95%, respectively.

3.4.3. Compiling molecular data on Torix *Rickettsia* from GenBank

A total of 183 nucleotide sequences of Torix *Rickettsia* were obtained from GenBank (Supplementary Tables 3.3-5). Specifically, 51 *rrs* sequences were available from Amoeba, Annelida, Arachnida, Coleoptera, Diptera, Hemiptera, Hymenoptera, Psocoptera, Megaloptera, and an environmental sample representing 18 studies. A total

of 68 sequences of *gltA* were obtained from Arachnida, Coleoptera, Diptera, Hemiptera, Hymenoptera, and Siphonaptera representing 12 studies. A total of 64 COI sequences were available from Amphipoda, Arachnida, Coleoptera, Diptera, Hemiptera, Hymenoptera, Megaloptera, and Odonata, representing 17 studies. Among these COI sequences, 42 sequences from 11 studies assigned rickettsial COI sequences to their invertebrate hosts. Since the very first misassignment in 2013, these mislabeled sequences have been deposited every year. Eight of these studies (representing 26 sequences) were involved with DNA barcoding and therefore voucher specimens (Supplementary Table 3.5).

3.4.4. Phylogeny of *Rickettsia*

All three trees inferred by *rrs*, *gltA*, and COI sequences clearly show two lineages within the genus *Rickettsia*: one clade of Torix *Rickettsia* and the other clade including all other 12 recognized groups within *Rickettsia* (Figures 3.3-5). A Bayesian tree based on *rrs* sequences (Figure 3.3) shows that all sequences obtained from New Zealand amphipods belong to the Torix group of *Rickettsia*. Even when the *rrs* conserved marker was used, all sequences obtained from amphipods (except S39_1542), were grouped in the same clade and distinct from other sequences, although this clade was not strongly supported (PP=0.84). Several subgroups were identified in the *gltA* tree (Figure 3.4). Most *Rickettsia* from amphipods were grouped within the same clade, similar to that revealed in the *rrs* tree. Also, a Bayesian tree inferred from COI sequences (Figure 3.5) shows that *Rickettsia* from amphipods and some insects obtained in New Zealand are closely related, and the clade containing them is strongly supported (PP=0.96).

3.4.5. Testing and validating blocking primers

The ratio of the amplicons of host COI to *Rickettsia* COI (based on their expected fragment sizes) was calculated to compare the efficiency of each primer under different conditions (Table 3.2). Not all fragment analyses were successful, but I was able to compare some effects. Except for Bloc_R, which always resulted in amplifying an excess of *Rickettsia* COI (ratio 0.06~0.57), all primers showed some blocking effects. Specifically, Bloc_F, Bloc_F2, and HCO_DPO primers showed increased blocking

effects (higher ratio of hostCOI:RickettsiaCOI) when more blocking primers were added. An increased number of PCR cycles resulted in a high number of host COI fragments. However, this increased somewhat the amplification of *Rickettsia* COI as well. Overall, HCO_DPO showed the highest efficiency among tested primers by preventing the amplification of *Rickettsia* COI, even at low concentration.

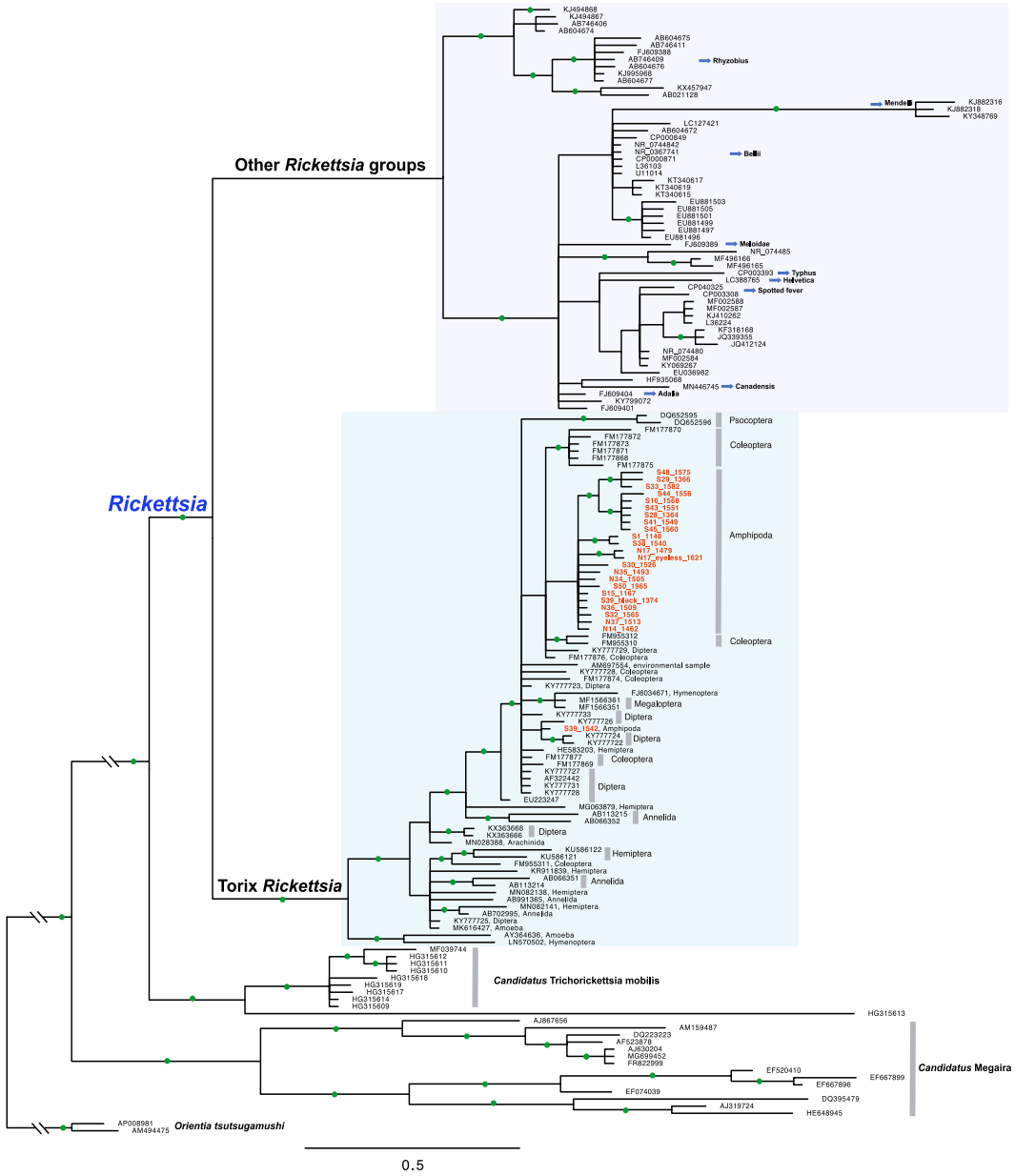


Figure 3.3 Bayesian tree of the genus *Rickettsia* based on 16S rRNA sequences. An alignment of 1,198 bp of 158 taxa was used. Two well-supported clades are shown within the genus *Rickettsia*. One is the Torix group which includes endosymbionts of diverse hosts (host taxa indicated on the right), and the other clade includes all other 12 recognized groups of *Rickettsia*. Nodes with a posterior probability higher than 0.9 are shown with green circles. Sequences obtained in this study are highlighted in orange colour.

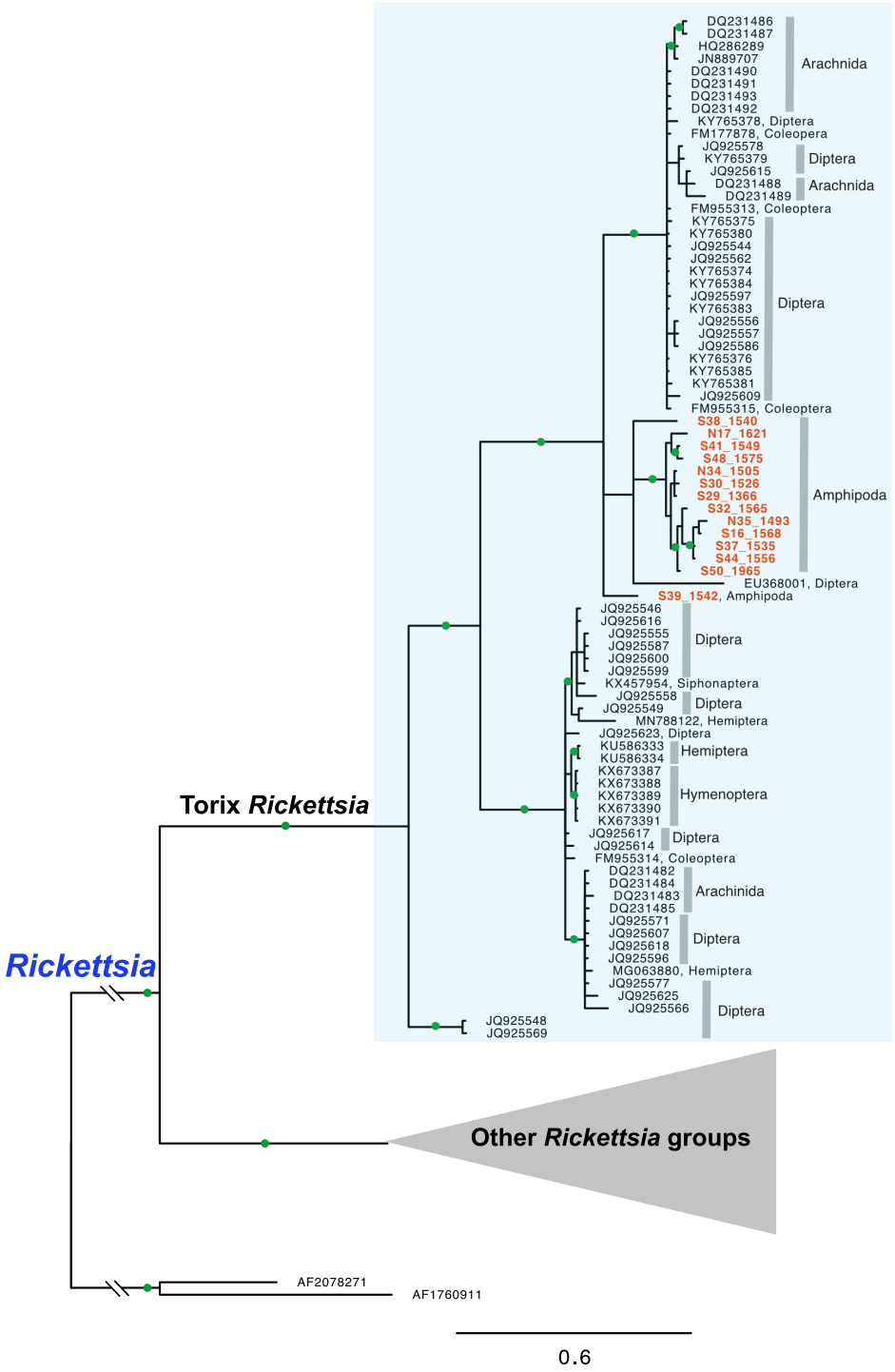


Figure 3.4 Bayesian tree of the genus *Rickettsia* based on an alignment of 765 bp of *gltA* sequences of 130 taxa (host taxa indicated on the right). Nodes with a posterior probability higher than 0.9 are shown with green circles. Sequences obtained in this study are highlighted in orange colour.

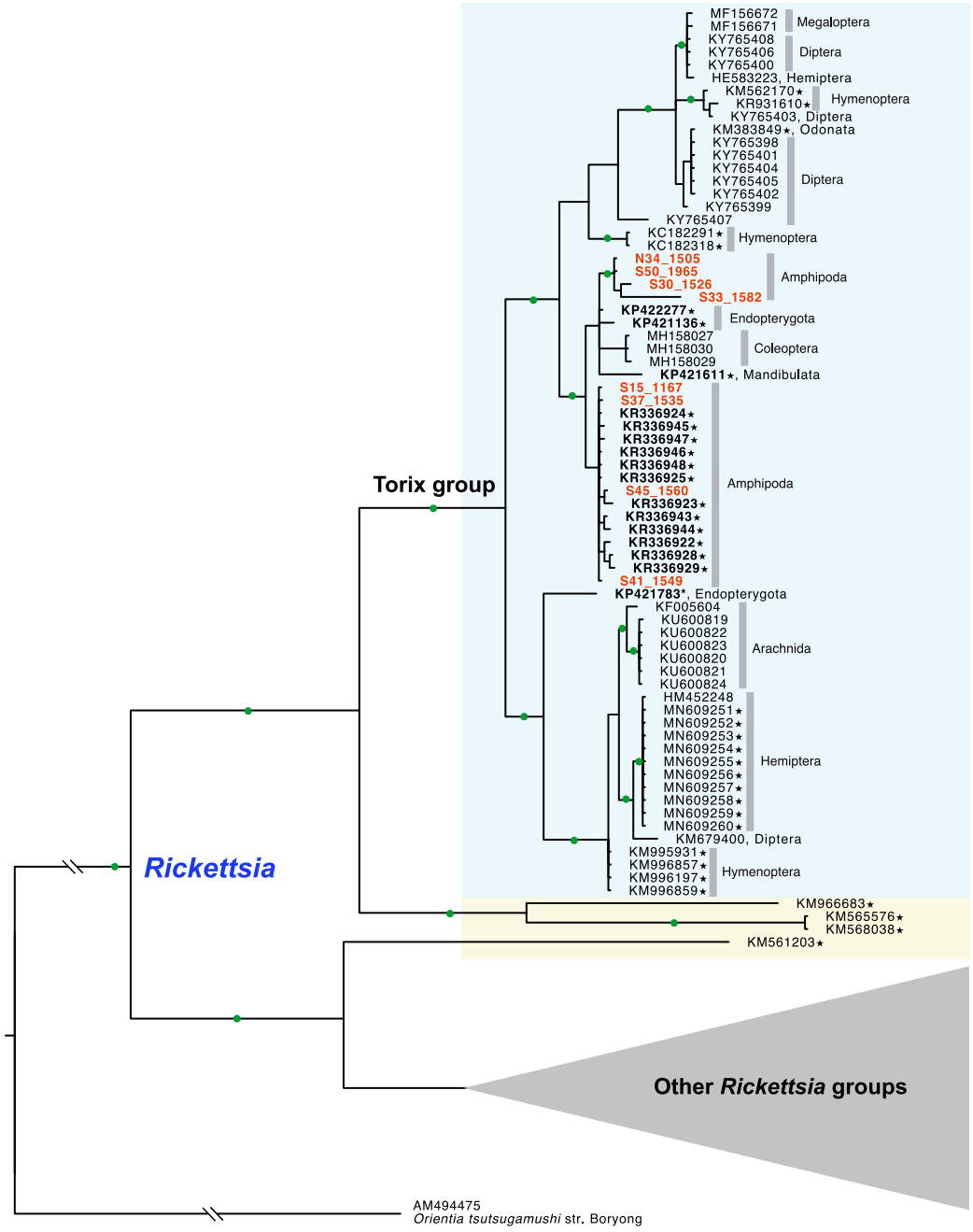


Figure 3.5 Bayesian tree of the genus *Rickettsia* based on an alignment of 559 bp of COI sequences of 114 taxa (host taxa indicated on the right). Nodes with a posterior probability higher than 0.9 are shown with green circles. Sequences that were misidentified as COI from invertebrate hosts, and initially not as rickettsial endosymbionts, in GenBank are highlighted with ★. Sequences obtained in this study are highlighted in orange colour. Sequences obtained in New Zealand are highlighted with bold text. Sequence similarity among Torix *Rickettsia* in the blue box is above 89%. Sequences that are similar to the other Torix *Rickettsia* but with lower similarity (80~82%), are highlighted within the yellow box.

3.5. Discussion

Our nationwide molecular screening results show that the Torix group of *Rickettsia* is widespread in freshwater amphipod hosts in New Zealand, and is to our knowledge the first report of *Rickettsia* in crustacean hosts. Because of the lack of information on *Rickettsia* infections in other groups of amphipods in other parts of the world, when and how these bacteria colonised and spread among New Zealand amphipods remain in question. Because freshwater amphipods have limited dispersal abilities (Myers, 1993), the widespread distribution of Torix *Rickettsia* in New Zealand amphipods may be explained by an ancient acquisition followed by vertical transmission, or by many independent events of recent horizontal transmission from other organisms. Several lines of evidence indicate that both horizontal and vertical transmission may have played roles in spreading and maintaining these bacteria in New Zealand amphipod hosts. The monophyletic relationships of most *Rickettsia* from New Zealand amphipod populations inferred by *rrs* and *gltA* sequences (Figures 3.3 and 4) seem to support the ancient acquisition scenario. Also, genetically closely related *Paracalliope* populations harbored the same genotype of *Rickettsia*, which also strongly supports their long-lasting relationship probably maintained by vertical transmission. Meanwhile, sharing of *Rickettsia* genotypes between sympatric amphipod species of different families suggests host shifts among genetically distant host species within the same order. Such a complex evolutionary history involving both vertical transmission and horizontal transfers has been reported for other insect/endosymbiotic systems (Duron et al., 2014). The Bayesian tree obtained with COI sequences provides some hints for horizontal transmission among amphipods and other arthropods (Figure 3.5). *Rickettsia* sequences from darkling beetles (*Pimelia* sp.) obtained in Europe were highly similar to the sequences identified in New Zealand amphipods (96~98% similarity) (López et al., 2018). Moreover, *Rickettsia* sequences obtained from several unspecified arthropod species (Mandibulata sp., Endopterygota sp., and Formicidae sp.) in New Zealand (although these were originally identified as invertebrate COI sequences) (Drummond et al., 2015) are closely related to those of New Zealand amphipods (96~99% similarity), providing strong evidence of recent horizontal transmission among them. Unfortunately, details regarding the host specimens and local origins of these sequences in New Zealand are not available. Direct detection of *Rickettsia* from these arthropod

species and multi-gene analyses will be necessary to elucidate their transmission routes. Interestingly, with a larger dataset with multigene data, Pilgrim et al. (2020) have recently inferred frequent horizontal transmissions of Torix *Rickettsia* among distantly related hosts. This supports the recent horizontal transmission of *Rickettsia* among amphipods and insects in the shared habitat, warranting further investigation.

Our findings support the early observation that the Torix group of *Rickettsia* may be highly associated with aquatic and damp environments (Weinert et al., 2009), which was based on the detection of this group in leeches, amoeba, Diptera, and Coleoptera (Dyková et al., 2003; Kikuchi et al., 2002; Küchler et al., 2009; Perlman et al., 2006). Pilgrim et al. (2017) provided support for this view by detecting *Rickettsia* from 38 % of *Culicoides* species tested and hypothesized that Torix *Rickettsia* may be dominant in insects with aquatic larval stages, and the 'aquatic hotspot' hypothesis has recently been strongly supported by a comparison of the incidence of Torix *Rickettsia* between terrestrial and aquatic hosts (Pilgrim et al., 2020). Horizontal transmission of Torix *Rickettsia* among genetically distantly related but spatially co-occurring species may have occurred frequently (Weinert et al., 2009). The high prevalence of Torix *Rickettsia* and their stable association with their hosts suggest negligible pathogenic effects of this group (Dyková et al., 2003; Kikuchi et al., 2002; Küchler et al., 2009; Wang et al., 2020). Some Torix *Rickettsia* may even be beneficial for their hosts. For example, infected leeches can have larger body sizes than uninfected individuals, although the possibility that larger individuals are more likely to acquire *Rickettsia* via horizontal transmission cannot be ruled out (Kikuchi and Fukatsu, 2005; Perlman et al., 2006). Ecological impacts of Torix *Rickettsia* on their hosts, and direct evidence of horizontal transmission among aquatic host groups, could be better answered with targeted community-level studies.

With the advancement of molecular techniques, our knowledge of the diversity of the sister groups of *Rickettsia* is also increasing and changing rapidly. Earlier studies focused mainly on the pathogenic and medically important species in arthropod hosts (Azad and Beard, 1998; Raoult et al., 2001). Until 2005, only two genera, *Rickettsia* and *Orientia*, were known within the family Rickettsiaceae, which now contains seven more

genus-level taxa (Castelli et al., 2016; Sabaneyeva et al., 2018). All these new genera are exclusively found in aquatic environments, mostly within ciliate hosts. It seems that adaptation to the use of arthropod hosts occurred several times independently within the family Rickettsiaceae (Castelli et al., 2016). In addition, the nomenclatural status and relationships among *Rickettsia* groups are also changing. The Hydra group, which was once considered to be an ancient group within the genus *Rickettsia*, along with the Torix and Belli groups, is now regarded as a separate genus: *Candidatus* Megaira (Schrallhammer et al., 2013). Another recently recognized group, *Candidatus* Trichorickettsia, is now believed to be a sister clade to the genus *Rickettsia* (Sabaneyeva et al., 2018).

The Torix group is largely different from the other groups of *Rickettsia* in many respects, including host range and habitat. The Torix group includes not only endosymbionts of diverse aquatic invertebrates (that are more complex than ciliates), but also diverse terrestrial arthropod hosts. Also, the Torix group is genetically distinct from other groups of *Rickettsia*, which all are sister to Torix *Rickettsia*. Specifically, the genetic similarities between the Torix and the Bellii groups are 96% in *rrs*, 78% in *gltA*, and 76% in COI sequences. The delimitation criteria I used for the Torix group in this study were 98.1% in *rrs*, 87.6% in *gltA*, and 89% in COI (broadly 80%; see Figure 3.5). Two genome sequences of Torix *Rickettsia* recently became available (Pilgrim et al., 2017; Wang et al., 2020). These genomes have the typical characteristics of *Rickettsia* (e.g. reduced genome size, and biosynthetic and catabolic capacity) but also have unique characteristics different from the other groups of *Rickettsia* (e.g. the presence of non-oxidative PPP, methionine salvage pathway, and glycolysis). It would be interesting to see how these two sister lineages, one mainly pathogenic and the other nonpathogenic, evolved and diverged from their common ancestor.

Interestingly, COI sequences revealed a diversity of Torix *Rickettsia*, as much as other popular markers for Rickettsiales such as *rrs* and *gltA*, even though COI has rarely been used as a marker of choice. Among 17 studies that have generated COI sequences of *Rickettsia*, only 3 were specifically intended to obtain COI sequences from *Rickettsia* (Gerth et al., 2017; MacHtelinckx et al., 2012; Pilgrim et al., 2017). The remaining 14

studies obtained *Rickettsia* COI sequences as a byproduct of other research objectives (i.e. host identification, population genetic studies, or DNA barcoding) with PCR using universal primers. *rrs* is a widely used marker but may be too conserved to resolve phylogenetic relationships among closely related species, while the *gltA* gene shows more variability. Only 6 studies (including the present one) produced both *rrs* and *gltA* sequences (Küchler et al., 2009; Noda et al., 2012; Pilgrim et al., 2017; Reeves et al., 2008; Wang et al., 2020). *gltA* sequences are not available for most Torix *Rickettsia* including endosymbionts of leeches and amoeba, as the reports of Torix *Rickettsia* from these groups precede the first use of *gltA* sequences to study Torix *Rickettsia* (Dyková et al., 2003; Goodacre et al., 2006; Kikuchi et al., 2002; Kikuchi and Fukatsu, 2005). Conversely, only *gltA* sequences are available for some species found in spiders and dipterans, which made it difficult to resolve the phylogenetic position of these rickettsial endosymbionts along with other *Rickettsia* (Goodacre et al., 2006; Perlman et al., 2006). 'Limoniae' and 'Leech' groups were used within Torix *Rickettsia* in some studies based on the *gltA* gene and concatenated sequences of *gltA* and *rrs* genes, although the 'Leech' group was found not to be monophyletic (Gualtieri et al., 2017; Küchler et al., 2009). Our *gltA* tree showed two main lineages which correspond to the clades found in previous studies. Similarly, two main linages were identified in the Bayesian tree inferred from COI sequences. However, whether the clade containing all endosymbionts from spiders represents the 'Leech' group could not be confirmed without direct multi-gene data from the same group. It is likely that there are many more subgroups, given the limited number of targeted studies available to date.

Endosymbionts and vertically transmitted intracellular parasites are common in arthropod hosts (Rousset et al., 1992; White et al., 2013). In the context of the growing recognition of the 'Holobiont' concept (Minard et al., 2013; Thompson et al., 2014), obtaining bacterial sequences from DNA extracts from host tissue is not surprising. Most bacterial 'contaminations' are filtered out during processing of metabarcoding data (Leray et al., 2013; Siddall et al., 2009). The frequent recovery of COI sequences from Torix *Rickettsia* can be partly explained by their nucleotide sequence similarity with mitochondria. The Proto-mitochondrion (the hypothetical common ancestor of all mitochondria) is often recovered as a sister to the Rickettsiales or within the

Rickettsiales (Roger et al., 2017). The alignment of COI sequences among several lineages of *Rickettsia* shows the high similarity between priming sites and the sequences used for universal primers (Figure 3.2). The priming site for the forward primer is 80% (20/25 nucleotides) identical to the LCO1490 sequences, and the priming site for the reverse primer is 84.6% (22/26 nucleotide) identical to the HCO2198 sequences. In addition, priming sites for universal primers are not conserved in many groups, which necessitates the need for group-specific or degenerate markers (Geller et al., 2013; Ward et al., 2005). However, this does not explain the more frequent reports of Torix COI in GenBank, because priming sites are also highly conserved in other groups of *Rickettsia* (Figure 3.2). Pilgrim et al. (2020) proposed that the lack of SNP near the 3' end of the priming site of Torix *Rickettsia* may be responsible for this bias. Additionally, I hypothesize that overall high prevalence of Torix *Rickettsia* compared to other groups of *Rickettsia* in an infected host population can partly explain the bias, even though a formal comparison of prevalence between Torix and non-Torix groups was not made (but see Weinert et al., 2015) for prevalence of diverse groups of *Rickettsia* in host populations). Therefore, several individuals from a given population may all be infected and could yield rickettsial COI, as illustrated in some previous studies (Ceccarelli et al., 2016; Lagrue et al., 2016), and in the current study.

These problems can be managed, as they are with *Wolbachia* (Smith et al., 2012). As mentioned earlier, Torix COI is highly conserved across diverse hosts. Therefore, comparing newly obtained (and suspicious) COI sequences with known Torix *Rickettsia* COI sequences can be easily done to distinguish Torix *Rickettsia*. Comparing sequences from the same taxon or genetically closely related groups could be useful. Checking for the presence of stop codons could largely decrease this problem, as for numts (Song et al., 2008). Bacterial sequences will show stop codons with the translation table for invertebrates, but will be in an open reading frame with the translation table for Bacteria. In addition to the high prevalence of Torix *Rickettsia* in many populations, high copy numbers of *Rickettsia* in host cells also make it difficult to obtain genuine host COI sequences, once a population is infected. In this case, applying blocking primers is a practical solution. Unfortunately, using blocking primers for *Rickettsia* does not always guarantee the amplification of host COI because other symbionts or parasites

might still be amplified. Nevertheless, blocking primers can be widely used for any host groups that are infected by *Rickettsia*, and for both next-generation-sequencing as well as Sanger sequencing. These sequences should not be confounded with those of hosts, yet these 'unwanted sequences' or 'contaminations' can provide useful information about their endosymbionts and parasites. Our current study provides an example, as I confirm the presence of Torix *Rickettsia* after discovering contaminated sequences from a previous study (Lagrue et al., 2016). Similarly, the COI sequence obtained from a damselfly suggested the presence of *Rickettsia* in this host group (order Odonata); this has recently been confirmed by a study with a targeted screening (Thongprem et al., 2020). Targeted studies are likely to uncover a huge but under-detected diversity of Torix *Rickettsia*, and with more data, we will be able to answer questions regarding transmission, host switching, and the evolution of pathogenicity. Furthermore, detailed research on a finer scale is needed to elucidate the impact of these widespread endosymbionts on their diverse hosts.

Chapter 4

Extremely divergent COI sequences of *Paracalliope* species complex: a possible role for *Rickettsia* infections?

4.1. Abstract

Deep divergence in mitochondrial DNA (mtDNA) with little or no variation in nuclear DNA (nuDNA) is a type of mitonuclear discordance. Although this may arise due to causes other than recent divergence of lineages, this issue has received little attention. *Paracalliope*, the most common freshwater amphipod genus in New Zealand, is known to have extremely divergent COI sequences (~24%) despite a lack of differentiation in morphology. In Chapter 2, I showed that 28S rRNA nuclear sequences well reflect geographical histories, with little genetic variation within each group (N, C, SA, and SB). Here, I investigate patterns and causes of mitonuclear discordance in *Paracalliope*. I start by contrasting phylogeographical patterns between mitochondrial (COI) and nuclear (28S rRNA and ITS1-5.8S rRNA-ITS2) sequences. I examine whether Microsporidia and/or *Rickettsia* could be responsible, because some endosymbionts can disrupt mtDNA patterns. Results indicate that *Rickettsia* is associated with mitonuclear patterns: *Rickettsia* was detected in the clades where discordance occurred; *Rickettsia* was in a linkage disequilibrium with a host COI haplotype in one of four populations tested; *Rickettsia* was prevalent in the SA group where substitution saturation was detected. I examine 'ancient lineage' and 'accelerated rate' scenarios and explore possible underlying mechanisms. I suggest that endosymbionts and host-specific life-history traits (environmental sex determination, variation in fecundity, and short generation time) may act synergistically to accelerate substitution rates and result in substitution saturation. In light of our findings, I discuss several important implications for the use of molecular clocks and the detection of cryptic species. Considering the hidden diversity of endosymbionts, endosymbiont-related mitonuclear discordance may be much more common than currently recognized across a wide range of taxa.

4.2. Introduction

Mitonuclear discordance, a phenomenon that can be generally defined as 'a significant difference in the patterns of differentiation between mitochondrial and nuclear markers' (Toews and Brelsford, 2012), has been widely recognized across various groups of vertebrates and invertebrates (Hinojosa et al., 2019; Linnen and Farrell, 2008; Morales et al., 2015). Although most studies deal with incongruent topologies (branching patterns), substantial differences in branch lengths can also be regarded as a type of mitonuclear discordance (Toews and Brelsford, 2012; Zhang et al., 2019). Patterns of deeply divergent mtDNA with little or no variation in nuDNA have been observed in diverse organisms including reptiles, birds, annelids, insects, and crustaceans (Bernardo et al., 2019; Giska et al., 2015; Hinojosa et al., 2019; Kvie et al., 2013; Zhang et al., 2019). This issue can have important implications because deep splits are often interpreted as a result of prolonged isolation (hence low or no gene flow between populations is inferred), and/or they provide a basis for species delimitation (Hebert et al., 2003; Ratnasingham and Hebert, 2013). When deeply diverged mtDNA lineages are not supported by nuclear markers and cannot be distinguished morphologically, mitonuclear discordance is associated with 'cryptic species' (Bickford et al., 2007).

The patterns of mitonuclear discordance that arise over shallow and deep time scales may be caused by different underlying mechanisms, however this issue has received little attention. Several possible causes have been suggested to explain mitonuclear discordance including incomplete lineage sorting (ILS), introgression followed by hybridization, sex-biased dispersal, adaptation, and the presence of endosymbionts (e.g. *Wolbachia*) (Toews and Brelsford, 2012), of which the first two are the most frequently invoked. More fundamentally, discordance patterns arise due to the different nature of inheritance of mitochondrial and nuclear genomes: haploidy and uniparental inheritance of mtDNA in contrast to diploidy and biparental inheritance of nuDNA (Ballard and Whitlock, 2004; Wolff et al., 2014). Mitochondrial DNA has a 4-fold smaller effective population size than that of nuDNA. As a result, coalescent events are expected to occur faster in mtDNA than in nuDNA (Hudson and Turelli, 2003). Therefore, if two lineages have split only recently, nuDNA may fail to coalesce while

mtDNA does, which is the cause of ILS (Funk and Omland, 2003; McGuire et al., 2007). On the other hand, secondary contact of long-isolated populations or closely related species can result in hybridization (Funk and Omland, 2003). If this is followed by the introgression of mtDNA from one to another, individuals from previously isolated populations or closely related species will share the same mitochondrial haplotypes. This scenario may explain why some closely related (and sometimes distantly related) species harbor the same mitochondrial haplotype despite having divergent nuDNA (Forsman et al., 2017; Raupach et al., 2016).

However, the ILS scenario does not generally agree with deeply divergent lineages because long branches mean a considerable accumulation of substitutions in two lineages since their most recent common ancestor. The scenario of hybridization followed by mitochondrial introgression may explain the presence of two or more deeply divergent mtDNA lineages within a population. However, this scenario requires secondary contact, and the differentiation level between populations may be reduced but not increased. A few explanations have been provided specifically for deep divergence. For example, in a bird species complex, deeply divergent mtDNA lineages, despite a lack of differentiation in the nuclear genome, were explained by 'ghost introgression', a mitochondrial introgression due to ancient hybridization events with extinct lineages (Zhang et al., 2019). Ancient hybridization events were also suggested as a cause of mitonuclear discordance in other taxa (Brennan et al., 2016; Dupuis and Sperling, 2020; Tóth et al., 2017). In cnidarians and copepods, substitution saturation (hence providing little phylogenetic information) was suggested as a responsible factor for incongruent topologies (Pratlong et al., 2017; Thum and Harrison, 2009). However, explanations for mitonuclear discordance arising in deep time are still scarce.

Another possible cause for mitonuclear discordance is the presence of vertically-transmitted endosymbionts and intracellular parasites capable of manipulating host reproduction (Hurst and Jiggins, 2005; Jiggins, 2003). These parasites can spread quickly within a population in several different ways (Engelstädter and Hurst, 2009; Werren et al., 2008). Because both mitochondria and endosymbionts are maternally transmitted, they are not independent but in linkage disequilibrium (LD). As a result,

endosymbionts along with the associated mitochondrial haplotype can increase in frequency within a population. This phenomenon, i.e. a selective sweep, could happen rapidly and result in reduced mtDNA diversity in an infected population (Hurst and Jiggins, 2005). If multiple strains are present in a population, multiple haplotypes will be maintained, resulting in increased mtDNA diversity within a population (James and Ballard, 2000; Schulenburg et al., 2002). Such patterns arising due to LD are well known in diverse insect and isopod hosts of the bacteria *Wolbachia* (Bouchon et al., 1998; Werren et al., 2008). In fact, there are at least six bacterial endosymbionts associated with reproductive manipulation: *Arsenophonus*, *Cardinium*, Flavobacteria, *Rickettsia*, *Spiroplasma*, and *Wolbachia* (Cordaux et al., 2011; Duron et al., 2008). Although they are eukaryotes, Microsporidia are also reproductive manipulators capable of feminizing their host. However, the effect of endosymbionts on host mtDNA evolution is largely unknown outside of *Wolbachia*-insect and *Wolbachia*-isopod associations.

Amphipods provide a great study system to investigate deep divergence and associated mitonuclear discordance. Amphipods are generally known to have high intra- and interspecies genetic divergence (Bradford et al., 2010; Witt et al., 2006). Also, it is believed that cryptic species and morphological stasis are prevalent in amphipods (Fišer et al., 2018; Havermans et al., 2011; Witt et al., 2006). *Paracalliope*, the most common freshwater amphipod genus in New Zealand, is also known to comprise several cryptic species based on deeply divergent COI sequences (~24 %; Hogg et al. 2006; Sutherland et al. 2009). However, no morphological differences have been identified among genetically divergent lineages (Hogg et al., 2006; Sutherland et al., 2009). While identifying amphipod species using the COI marker, I also observed extremely divergent COI sequences even among adjacent populations (~ 22%) far exceeding the general degree of intraspecific genetic divergence in other taxa (Hebert et al., 2003; Raupach and Radulovici, 2015). However, I found little or no genetic divergence among these populations based on 28S rRNA (28S) sequences (Chapter 2). 28S sequences are traditionally believed to evolve slowly, although this marker also includes highly variable regions which are used as barcoding markers for some invertebrate taxa (Shylla et al., 2013; Zhao et al., 2020). *Paracalliope* 28S sequences clearly showed two main

lineages (NC and S) which is concordant with a known geographical barrier (Chapter 2).

In previous chapters (Chapters 2 and 3), Microsporidia and *Rickettsia*, two known manipulators of host reproduction, were detected in multiple *Paracalliope* populations. Microsporidians are among the most common eukaryotic parasites in amphipods (Bojko and Ovcharenko, 2019). Although their feminizing abilities have been investigated (Terry et al., 1998), their possible impact on host mtDNA evolution remains unknown. Moreover, sex-ratio distortion and its consequences for mtDNA evolution are known in insect hosts of *Rickettsia* (Hagimori et al., 2006; Schulenburg et al., 2002). Because the presence of *Rickettsia* was confirmed in amphipod hosts just recently in Chapter 3, their impact on host mtDNA evolution is unknown.

Here, I investigate patterns and causes of mitonuclear discordance in *Paracalliope*. I started by testing patterns of mitonuclear discordance using COI, 28S, ITS1-5.8S rRNA-ITS2 (ITS) sequences. I determine whether Microsporidia or *Rickettsia* may be responsible for mitonuclear discordance patterns. Even though endosymbionts can explain mitonuclear discordance patterns to some extent, the seemingly widespread occurrence of cryptic species in amphipods may not be fully explained by endosymbionts; this may also have intrinsic explanations. Therefore, I additionally explored other potential causes for the deep divergence of mtDNA and mitonuclear discordance in the *Paracalliope* species complex. I focused on two aspects, time and rate, because branch lengths in the phylogenetic tree are determined by the combined effects of these two components. I examined the 'ancient lineage' scenario in which long branches indeed reflect the evolutionary time. Then, I considered the 'accelerated substitution rates' in mtDNA scenario. I propose endosymbionts, some amphipod lineage-specific life-history traits, and their synergistic effects as potential factors driving the accelerated substitution rates in amphipod mtDNA.

4.3. Materials and methods

4.3.1. Sample collection and DNA extraction

Amphipod specimens and extracted genomic DNA samples of the *Paracalliope* species complex including their endosymbionts, which were collected from 67 locations throughout New Zealand, were obtained from our previous studies (Chapters 2 and 3).

4.3.2. COI, 28S, and ITS sequences

In order to compare phylogenetic patterns between mtDNA and nuDNA, COI, 28S, and ITS sequences were obtained. ITS sequences were only obtained from a subset of samples. Although 28S sequences were already obtained from 28 locations (one sequence per location) from Chapter 2, sequences from 11 further locations were additionally obtained for the present study, for better geographical coverage. PCR reactions were conducted as described in Chapter 2. Since I obtained untargeted rickettsial sequences using universal COI primers, blocking primers were added to securely amplify host COI sequences (Chapter 3). For PCR reactions for the COI region, 11.9 ml of distilled water, 4 ml of reaction buffer, 0.8 ml of each forward and reverse primers, 0.4 ml of blocking primer (DPO_HCO), 0.1 ml of MyTaq (Bioline), and 2 ml of DNA were used. PCR conditions were the following: 95°C for the initial denaturation of 3 min, 5 cycles at 94°C for 60s, 45°C for 90s, 72°C for 40s, 35 cycles at 94°C for 60s, 48°C for 90s, 72°C for 40s, the final extension for 5 min at 72°C. ITS sequences were obtained using ITS_amp_F (5'-AACCTGCGGAAGGATCATTA-3') and ITS_amp_R primers (5'-ATGCTTAAGTTCAGCGGGTAGTCCC-3'). For PCR reactions for the ITS region, 12.3 ml of distilled water, 4 ml of reaction buffer, 0.8 ml of each forward and reverse primers, 0.1 ml of MyTaq (Bioline), and 2 ml of DNA were used with the following PCR conditions: 95°C for the initial denaturation of 5 min, 35 cycles at 94°C for 30s, 55°C for 30s, 72°C for 80s, and the final extension for 7 min at 72°C. After PCR reactions, 2 ml of PCR product from each PCR reaction were run on a 1.5 % agarose gel.

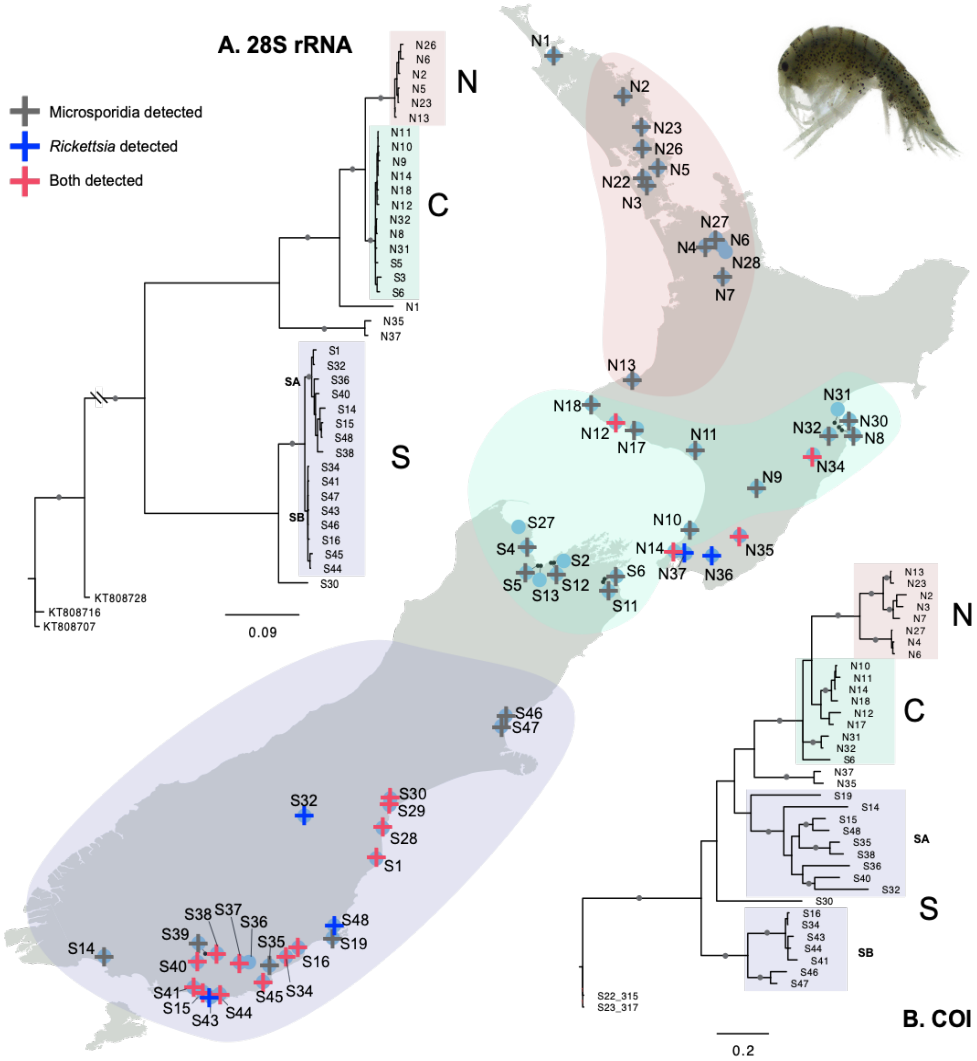


Figure 4.1 Map of New Zealand showing sampling sites of *Paracalliope* with circles. The sites where parasites were detected are marked with crosses of different colours. Also, Bayesian trees of (A) 28S and (B) COI sequences are shown on the left and right sides of the map, respectively. Geographical regions are marked with different colors of shades on the map and the trees. The C and S groups are not monophyletic in the COI tree.

4.3.3. LD between host mitochondria and intracellular parasites

In order to see if host mtDNA and parasites are in LD within a population, four populations were chosen (S15, S30, S40, and S44) among locations with both Microsporidia and *Rickettsia* (Figure 4.1). For each population, 2-4 appendages were obtained from 12-24 individuals to purify host genomic DNA. The rest of the body was separately used for DNA extraction to be used for parasite detection. DNA was extracted using the Chelex method as described in Chapter 2 or with the Qiagen DNeasy Blood & Tissue kit following the manufacturer's protocol. Microsporidia and *Rickettsia* were detected as described in Chapters 2 and 3, respectively. The presence or absence of both parasites was recorded for each individual amphipod. After conducting PCR reactions for the COI region as described above, approximately the same number of infected and uninfected individuals were chosen for sequencing.

4.3.4. Sequencing

All PCR products obtained for both phylogeographic and LD patterns were purified with MEGAquick-spinTM Total Fragment DNA Purification Kit (iNtRON Biotechnology) according to the manufacturer's instructions. Purified PCR products were sent to Genetic Analysis Services at the University of Otago, New Zealand. Raw nucleotide sequences were carefully examined and corrected by eye and the existence of nuclear mitochondrial pseudogenes (numts; Song et al., 2008) was checked by translating the nucleotide sequences into protein sequences.

4.3.5. Phylogenetic trees and networks

4.3.5.1. Phylogeography

Phylogenetic trees of COI and 28S sequences were inferred using MrBayes 3.2.7 (Ronquist et al., 2012) through CIPRES Science Gateway v3.3 (Miller et al., 2010). For all datasets, two independent runs, consisting of four chains each, were simultaneously conducted with GTR+G+I as a model of nucleotide substitution for 2,000,000 and 5,000,000 generations for 28S and COI sequences, respectively, with a sampling frequency of 1,000. The initial 25% of samples were discarded. The resulting trees were visualized in FigTree v1.4.4. Also, phylogenetic networks of COI and 28S, as well as of ITS sequences, were produced based on HKY85 distances using a Neighbor-Net

method implemented in Splitstree5 (Bryant and Moulton, 2002; Huson and Bryant, 2006), to visually and effectively compare overall phylogenetic patterns. Sequences that were too short were excluded from the analyses because these increased ambiguities which were represented by many boxes in the networks.

4.3.5.2. Linkage disequilibrium

Another Bayesian COI tree was inferred to better observe LD between endosymbiont infection status and mtDNA (i.e. COI haplotypes). For this, 64 sequences that I obtained from the four populations (S15, S30, S40, S44) were added to the dataset of the phylogeographical patterns. COI sequences from Hogg et al. (2006) and Lagrue et al. (2016) were also included, producing a final dataset of 128 taxa. Two independent runs, consisting of four chains each, were simultaneously conducted with GTR+G+I as a model of nucleotide substitution for 1,000,000 generations with a sampling frequency of 1,000. The initial 25% of samples were discarded. The resulting trees were visualized in FigTree v1.4.4. Additionally, median-joining haplotype networks were drawn for each of the four populations (S15, S30, S40, and S44) using PopArt (Leigh and Bryant, 2015) to better observe associations between *Rickettsia* and host COI haplotypes.

4.3.6. Intrapopulation variation and neutrality tests

The number of haplotypes, number of segregating sites, and number of synonymous and non-synonymous mutations were counted in DnaSP 6 (Rozas et al., 2017). Haplotype diversity and nucleotide diversity were calculated in Arlequin v3.5.2.2 (Excoffier and Lischer, 2010). Tajima's D and Fu's Fs tests for neutrality were also conducted using the same program.

4.3.7. Pairwise genetic distance

Uncorrected pairwise genetic distances of 28S and COI sequences were calculated in Mega7 using alignments used for phylogeographical patterns in *4.3.5.1.* Highly variable regions within 28S sequences were useful to distinguish groups but also included indels which made some parts of the alignment ambiguous. Therefore, I obtained pairwise genetic distances using two different datasets: one with only a highly conserved region (432 bp) and another with the full length of the amplicon (~1,760 bp).

4.3.7. Substitution saturation

Substitution saturation was tested using Xia's method implemented in DAMBE7 (Xia et al., 2003). Because substitution saturation often occurs in the third codon position (Breinholt and Kawahara, 2013), the first and second codon positions, and the third codon position were tested separately. The tests were conducted for all the data, and each group separately because the N, C, SA, and SB clades may represent cryptic amphipod species that followed different evolutionary trajectories. Separate analyses for each group also allowed us to verify which group is likely responsible for the overall substitution saturation. In addition, substitution saturation was visually assessed by plotting transitions and transversions against K2P distance, also using DAMBE7.

4.4. Results

4.4.1. Phylogeographical patterns and mitonuclear discordance

In total, 38 COI and 39 28S rRNA sequences of *Paracalliope* were used to infer phylogeographical trees (Figure 4.1). The phylogeographical patterns shown with 28S sequences well reflect known geographical events (e.g. marine transgression, recolonization of the newly available land area, and the last glacial maximum that have shaped the geographical distributions of many other taxa in New Zealand (see Chapter 2). A clear split between the NC and S groups with large genetic distances suggests that these two main lineages diverged a long time ago from their most recent common ancestor. The 28S tree supports several independent freshwater invasions because freshwater lineages in the North and South Islands are not monophyletic but they are closely related to brackish lineages (e.g. N1, S30), respectively. The marine species (*Paracalliope novizealandiae*; S22, S23) is sister to all other freshwater and brackish lineages (Figure 4.1). The SA and SB clades may represent different lineages of colonization from marine or brackish to the freshwater environment. The SB lineage has only been found near coastal areas in contrast to the SA group that was found in both inland and coastal areas (Figures 4.2).

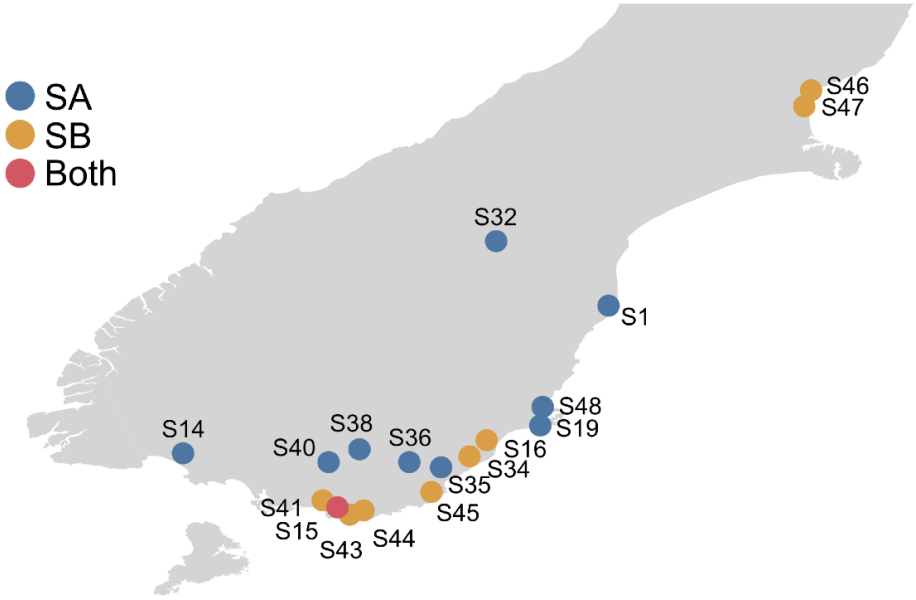


Figure 4.2 A New Zealand map of the southern part of the South Island. Circles represent sampling locations. The SB group was found only in the lower streams near the coastline, or a lake with saltwater intrusion. On the other hand, the SA group was found in the upper streams as well as near the coastlines.

Although the COI tree also generally shows the same major groupings (N, C, SA, and SB), the C and S groups were not recovered as monophyletic (Figure 4.1). Also, branch lengths were considerably longer within the SA group compared to other groups. Some deep splits were also seen in the N, C, and SB groups, despite little variation in 28S sequences (Figure 4.1). I only obtained six clear sequences of ITS due to many repeats and indels. Although most of the ITS sequences I obtained were discarded because of low quality, the clear distinction among different major groups, which was similar to the pattern of 28S sequences, was observable. ITS was used for network analyses (see below).

The discordance patterns are also clearly seen in the phylogenetic networks (Figure 4.3). The COI network shows extremely divergent lineages especially within the SA group (Figure 4.3A). Deep splits were seen in all other groups, but to a lesser extent (N, C, and SB). In contrast, both 28S and ITS networks show similar patterns of the deep split between NC and S groups (Figure 4.3B and C).

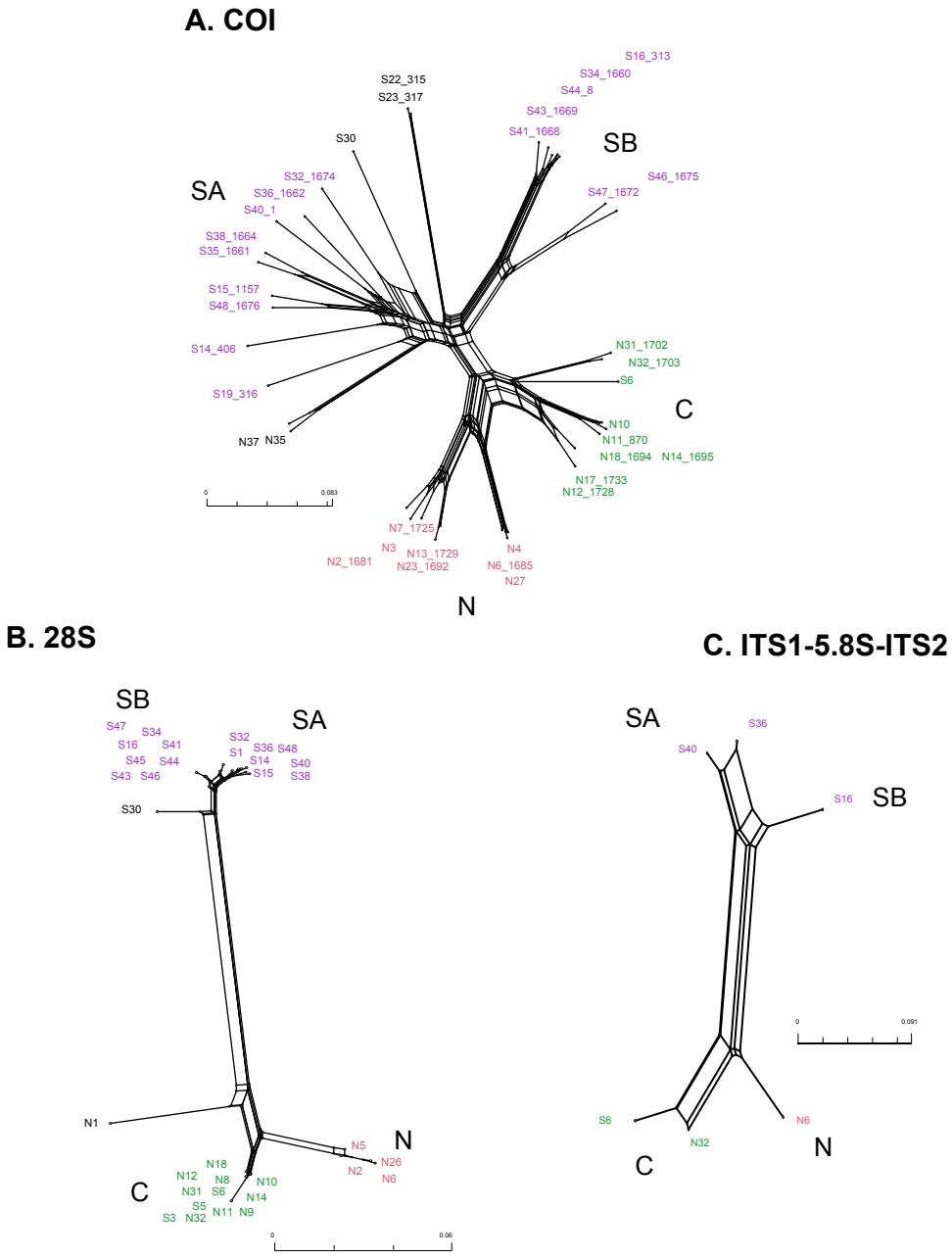


Figure 4.3 Phylogenetic networks of (A) COI, (B) 28S, and (C) ITS sequences. COI sequences are highly divergent especially in the SA group showing long terminal branches. 28S and ITS networks show a clear divide between the NC and S (SA and SB) groups.

4.4.2. Pairwise genetic distance among populations

Mitonuclear discordance was also observed by comparing patterns of genetic distances between 28S and COI sequences (Table 4.1-3). With the full length of 28S sequences which include highly variable sites, four main groups were distinguished (Table 4.1). The difference between the NC and S groups with the full 28S sequences was large (15~20 %). While the genetic divergence within each group was very small (0~1.2 %), the genetic divergences between N and C groups were 4.5~5.9 %, and between SA and SB groups were 0.9~2.8 % (Table 4.1). Using only the conserved region of 28S sequences, the genetic distances between the NC and S groups were 6~8 % (Table 4.2). Unlike 28S sequences, which showed very little within-group genetic distance, COI sequences show high intra- and intergroup genetic distances (Table 4.3). Notably, the SA group showed generally high genetic divergence (13-25 %) compared to that of other groups (2-16 %). Considering that the SA group was collected with the most-dense sampling, the genetic divergence among adjacent populations is significantly high in this group. Interestingly, the genetic divergence shown with 28S and COI sequences was not proportional. Based on the conserved region of 28S sequences, the genetic distances of 28S sequences between SA and SB groups (0.2-0.7 %) were shorter than those between N and C groups (1.2-1.4 %), but the genetic distances of COI sequences were higher between SA and SB groups (20.6-23.6 %) than between N and C groups (16-18.6 %).

4.4.3. Parasite distribution and mitonuclear discordance patterns

The distribution of Microsporidia and *Rickettsia* is shown in Figure 4.1 (data obtained from Chapters 2 and 3). Microsporidia were detected from all the major groups of *Paracalliope* (N, C, SA, and SB). However, *Rickettsia* was not found from the N group.

Table 4.1. Uncorrected pairwise distance of 28S sequences of *Paracalliope* populations. A full length of amplicon was used.

	N26	N6	N2	N5	S3	S5	N31	N32	N8	N11	N10	N9	N18	N14	N12	S6	N1	S1	S32	S34	S41	S47	S43	S46	S16	S45	S44	S36	S40	S14	S15	S48	S38	S30
N26																																		
N6	0.4																																	
N2	0.2	0.2																																
N5	0.6	0.5	0.2																															
S3	5.4	5.9	5.1	4.7																														
S5	5.4	5.6	5.9	5.4	0.1																													
N31	5.1	5.7	5.1	4.5	0.7	0.1																												
N32	5	5.7	5.1	4.5	0.8	0.1	0.2																											
N8	5.7	6	5.2	5.2	0.4	0.1	0.1	0																										
N11	5	5.6	4.9	4.5	1	0.4	0.4	0.3	0.3																									
N10	5	5.4	4.9	4.6	0.3	0.5	0.3	0.3	0.3	0																								
N9	5.2	5.6	4.9	4.5	1.2	0.4	0.6	0.5	0.3	0.2	0																							
N18	4.9	5.6	4.9	4.4	0.8	0.4	0.4	0.4	0.4	0.2	0	0.4																						
N14	5	5.6	4.9	4.5	1	0.5	0.4	0.3	0.3	0	0	0.2	0.2																					
N12	5.2	5.9	5.1	4.5	0.9	0.3	0.4	0.5	0.4	0.3	0.1	0.5	0.3	0.3																				
S6	5.4	6	5.1	4.8	0.7	0.1	0.5	0.5	0.3	0.7	0.3	1	0.7	0.7	0.7																			
N1	10	10.1	7.7	9.7	7	7.6	7	7	7	7.3	7.3	7.3	7.3	7.3	7.2	7																		
S1	22.7	21.4	15.7	20.3	20.9	18.6	20	20	18.7	20.1	16	21.1	20.2	20	21	20.2	16.6																	
S32	18	17.9	15.4	16.6	15.6	17.6	15.6	15.6	16	15.8	15.8	15.8	15.8	15.8	15.6	15.6	17	0.3																
S34	22	20.6	14.9	19.8	20.2	18.1	19.7	19.6	17.9	19.7	15.5	20.4	19.7	19.6	20.2	19.8	16.7	1.5	0.8															
S41	17.6	17.6	15.1	16.5	15.6	17.3	15.6	15.6	15.8	15.7	15.6	15.7	15.7	15.6	15.6	15.6	16.7	1.1	0.8	0														
S47	22	20.6	14.9	19.8	20.2	18.1	19.6	19.6	17.9	19.7	15.5	20.4	19.7	19.5	20.2	19.8	16.7	1.5	0.8	0.1	0													
S43	22	20.6	14.9	19.8	20.2	18.1	19.7	19.6	17.9	19.7	15.5	20.4	19.7	19.6	20.2	19.8	16.7	1.5	0.8	0.1	0	0.1												
S46	21.9	20.6	14.9	19.7	20.1	18.1	19.7	19.5	17.9	19.6	15.5	20.3	19.6	19.5	20.2	19.7	16.7	1.4	0.8	0.1	0	0.1	0.1											
S16	17.7	17.7	15.1	16.4	15.4	17.3	19.7	15.4	15.7	15.6	15.6	15.6	15.6	15.6	15.4	15.4	17.4	1.2	0.9	0	0	0	0	0										
S45	21.3	20.2	14.8	20.1	19.5	17.6	19.7	19.5	17.6	19.7	15.1	19.7	19.7	19.5	19.6	19.8	16.2	2.2	0.9	0.4	0.4	0.4	0.4	0.4	0.2									
S44	21.5	20.5	15.9	21.1	20.4	17.9	19.7	20.5	18.3	20.7	16.2	20.7	20.7	20.5	20.6	20.7	16.5	2.3	0.8	0.2	0.1	0.2	0.2	0.2	0	0.2								
S36	22.4	20.9	15.5	20.3	20.9	18.3	19.7	20.2	18.5	20.3	15.8	21.1	20.4	20.2	21	20.4	16.6	1	0.9	1.8	1.5	1.7	1.8	1.6	1.5	1.9	2.2						
S40	23	21.7	15.7	20.4	21.2	18.8	19.7	20.1	19	20.2	16.1	21.4	20.3	20.1	21.3	20.3	17.1	1.1	1.2	1.6	1.7	1.5	1.6	1.5	1.7	2.2	2.5	0.9						
S14	22.3	20.9	15.4	20.1	20.9	18.5	19.7	19.9	18.3	20	15.9	20.8	20.1	19.9	20.6	20.2	16.8	1.5	1.4	1.9	2.1	1.9	1.9	1.8	2	2.4	2.8	1.2	1					
S15	22.4	20.9	15.5	20.1	20.9	18.5	19.7	19.8	18.5	20	16.1	21	20.1	19.9	20.9	20.1	17	1.2	1.4	1.7	1.8	1.6	1.7	1.6	1.8	2.1	2.4	1	0.8	0.8				
S48	22.4	20.9	15.6	20	20.9	18.6	19.7	19.8	18.5	19.9	16	21	20	19.8	20.8	20.1	16.9	1.2	1.4	1.7	1.8	1.6	1.7	1.6	1.8	2.1	2.4	0.9	0.8	0.6	0.2			
S38	22.1	21.3	15.4	20.8	20.4	18.6	19.7	20.5	18.7	20.7	16	20.7	20.7	20.5	20.6	20.6	16.9	1.6	1.1	2.1	1.5	2.1	2.1	2.1	1.5	2.1	2.4	0.9	1.1	1.2	0.7	0.6		
S30	21.6	20.9	14.4	19.6	19.8	18.5	19.7	19.2	18	19.2	15.7	19.8	19.4	19	19.6	19.4	16.2	6.2	3.1	5.7	3.5	5.6	5.7	5.6	2.5	5.2	5.4	6.2	6.1	6.3	6.2	6.3	6.4	

Table 4.2. Uncorrected pairwise genetic distance of 28S sequences of *Paracalliope* populations. Only a conserved region was used.

	N26	N6	N2	N5	N23	S3	S5	N31	N32	N8	N11	N10	N9	N18	N14	N12	S6	N7	N1	S1	S32	S34	S41	S47	S43	S46	S16	S45	S44	S36	S40	S14	S15	S48	S38	S30	S22
N26																																					
N6	0.0																																				
N2	0.0	0.0																																			
N5	0.3	0.3	0.2																																		
N23	0.0	0.0	0.0	0.2																																	
S3	1.4	1.2	1.2	1.4	1.2																																
S5	1.4	1.2	1.4	1.7	1.4	0.0																															
N31	1.4	1.2	1.2	1.4	1.2	0.0	0.0																														
N32	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0																													
N8	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0																												
N11	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0																											
N10	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0																										
N9	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0																									
N18	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0																								
N14	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0																							
N12	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0																						
S6	1.4	1.2	1.2	1.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0																					
N7	0.0	0.0	0.0	0.2	0.0	1.2	1.4	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2																				
N1	2.2	1.5	2.1	2.3	2.1	2.3	2.6	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.1																			
S1	7.7	6.2	6.8	7.1	6.8	7.5	8.9	7.5	7.5	7.8	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.0	8.1																		
S32	7.4	5.8	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	7.8	0.2																	
S34	7.4	6.2	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5																
S41	7.4	6.2	6.7	6.9	6.7	7.4	8.6	7.4	7.4	7.5	7.4	7.4	7.4	7.4	7.4	7.4	7.4	6.8	8.3	0.7	0.5	0.0															
S47	7.4	6.2	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.0	0.0														
S43	7.4	6.2	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.0	0.0	0.0													
S46	7.4	6.2	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.0	0.0	0.0	0.0												
S16	7.4	6.2	6.6	6.9	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.0	0.0	0.0	0.0	0.0											
S45	7.4	6.2	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.0	0.0	0.0	0.0	0.0	0.0										
S44	7.4	6.2	7.2	7.5	7.2	8.0	8.6	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	7.2	8.3	0.8	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0									
S36	7.1	5.8	6.4	6.7	6.4	7.1	8.4	7.1	7.1	7.3	7.1	7.1	7.1	7.1	7.1	7.1	7.1	6.6	8.1	0.5	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3								
S40	7.1	5.8	6.4	6.6	6.3	7.1	8.4	7.1	7.1	7.3	7.1	7.1	7.1	7.1	7.1	7.1	7.1	6.6	8.1	0.5	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.0							
S14	7.1	5.8	6.4	6.6	6.3	7.1	8.4	7.1	7.1	7.3	7.1	7.1	7.1	7.1	7.1	7.1	7.1	6.6	8.1	0.9	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.5	0.5						
S15	7.4	5.8	6.6	6.9	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.2	0.2					
S48	7.4	5.8	6.6	6.8	6.6	7.3	8.6	7.3	7.3	7.5	7.3	7.3	7.3	7.3	7.3	7.3	7.3	6.8	8.3	0.7	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.2	0.2	0.0				
S38	7.1	5.8	6.4	6.6	6.3	7.1	8.4	7.1	7.1	7.3	7.1	7.1	7.1	7.1	7.1	7.1	7.1	6.6	8.1	0.5	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.0	0.0	0.5	0.2	0.2			
S30	6.1	5.9	5.3	5.5	5.2	6.0	7.3	6.0	6.0	6.2	6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.4	6.1	3.1	2.9	2.6	2.7	2.6	2.6	2.6	2.6	2.6	2.9	2.9	2.9	2.4	2.6	2.6	2.9		
S22	11.6	11.8	10.4	10.1	10.8	10.1	11.9	10.1	10.1	10.5	10.1	10.1	10.1	10.1	10.1	10.1	10.1	11.2	12.3	9.2	9.5	9.2	9.4	9.2	9.2	9.2	9.3	9.2	10.1	9.6	9.5	9.2	9.6	9.5	9.5	8.3	

Table 4.3. Uncorrected pairwise genetic distance of COI sequences among *Paracalliope* populations

	N10	N14	N11	N18	N12	N17	N31	N32	S6	N13	N23	N2	N3	N7	N27	N4	N6	N37	N35	S19	S16	S34	S43	S44	S41	S46	S47	S14	S15	S48	S35	S38	S36	S40	S32	S30	S22	S23
N10																																						
N14	0.4																																					
N11	1.7	1.5																																				
N18	3.5	2.8	3.9																																			
N12	9.5	9.7	9.7	9.5																																		
N17	8.0	8.0	8.8	8.3	4.7																																	
N31	12.6	12.1	13.0	12.9	12.7	12.3																																
N32	11.7	11.4	12.1	12.3	12.3	11.5	2.7																															
S6	12.0	12.8	12.4	12.6	13.2	12.3	13.3	12.1																														
N13	18.0	17.9	17.7	17.4	17.6	16.2	17.4	16.5	17.1																													
N23	18.5	18.6	18.5	17.8	18.0	16.3	18.0	17.4	18.0	0.9																												
N2	17.1	17.7	17.0	16.5	17.1	16.1	17.3	16.4	17.1	7.9	8.3																											
N3	17.6	17.9	17.6	17.1	17.7	17.3	16.7	16.5	18.0	6.8	7.4	4.7																										
N7	17.0	17.3	17.3	17.1	18.3	16.7	17.3	16.4	16.7	6.5	7.4	5.3	4.7																									
N27	16.6	15.7	16.6	17.0	14.4	14.7	16.6	15.8	17.6	12.8	13.1	13.8	14.4	12.9																								
N4	17.0	16.0	17.0	17.4	14.5	15.2	17.3	16.7	17.7	13.6	14.0	14.4	15.0	13.8	0.8																							
N6	17.3	16.0	17.3	17.7	15.2	15.5	17.7	17.0	18.3	13.5	13.8	14.8	15.5	13.9	0.9	0.8																						
N37	19.5	19.7	20.3	21.1	18.8	19.2	20.0	19.5	19.1	22.7	23.4	22.6	22.9	21.2	21.1	21.4	21.8																					
N35	20.3	20.5	21.2	21.5	19.8	18.9	20.5	20.0	19.1	22.4	23.2	22.6	22.6	21.2	21.3	21.5	22.0	4.8																				
S19	20.5	20.5	20.9	21.5	22.0	20.8	20.3	19.7	20.0	21.7	22.5	20.3	21.8	19.8	21.7	21.8	22.0	17.3	18.2																			
S16	21.1	20.9	21.8	21.4	22.0	20.8	22.1	22.1	20.2	21.7	22.3	22.4	22.6	22.9	23.9	23.8	23.9	23.5	23.2	23.3																		
S34	20.8	20.5	21.5	21.1	21.8	20.6	22.0	22.0	20.3	21.5	22.3	22.6	22.7	23.0	24.0	23.9	24.1	23.6	23.3	23.5	0.5																	
S43	21.2	21.0	21.7	21.1	21.7	20.5	22.1	21.8	20.8	22.0	23.2	23.3	23.5	23.3	23.4	23.3	23.5	23.8	23.3	24.4	3.3	2.9																
S44	20.6	20.7	21.1	20.5	21.6	19.8	21.6	21.8	20.8	21.9	22.3	22.7	22.6	22.9	23.3	23.2	23.4	22.9	22.7	23.8	2.2	1.9	3.0															
S41	21.4	21.0	21.4	21.1	21.2	20.9	22.0	21.7	21.4	23.2	24.1	23.8	23.8	24.7	23.7	23.6	23.8	24.4	24.2	24.1	5.0	4.5	4.7	4.2														
S46	20.9	21.2	21.7	21.4	20.8	20.2	20.9	20.5	21.1	23.0	24.0	22.0	22.9	23.0	22.6	22.9	23.3	21.7	21.4	23.5	15.3	15.5	16.8	16.0	16.7													
S47	20.6	20.5	21.1	20.8	20.3	19.7	20.6	19.7	20.5	22.6	22.9	22.6	23.6	23.3	23.1	23.3	23.8	22.0	22.6	22.9	15.0	14.8	15.6	15.4	15.3	6.7												
S14	22.0	21.6	22.7	22.3	23.3	22.6	21.8	20.9	22.3	21.1	21.2	21.5	21.5	20.3	23.1	23.3	23.5	19.1	19.7	18.0	23.0	23.2	22.9	22.9	23.6	22.7	22.0											
S15	20.8	20.5	21.2	21.2	21.2	20.8	20.8	20.5	22.6	22.4	23.6	21.1	21.8	21.4	21.9	22.4	22.6	18.0	18.6	17.1	20.6	20.8	21.4	21.3	20.9	21.7	20.8	16.4										
S48	20.7	20.7	21.0	21.0	21.8	21.0	21.5	21.1	22.4	21.0	21.7	19.4	20.2	19.7	22.0	22.6	22.8	18.4	18.2	18.2	21.1	21.5	22.4	21.9	21.1	21.6	21.3	15.1	7.8									
S35	20.6	20.4	20.0	20.8	23.3	21.4	21.4	21.6	21.4	22.6	23.7	23.7	23.5	23.0	22.9	23.5	23.7	20.6	20.0	19.8	22.2	22.2	21.8	22.0	21.4	24.1	23.7	16.3	13.2	13.6								
S38	21.4	21.0	20.9	21.5	21.8	20.9	20.9	21.1	22.6	20.8	22.3	21.5	21.1	21.5	22.8	23.0	23.2	21.4	21.4	19.7	22.7	22.4	22.3	22.7	22.1	23.8	23.5	17.6	14.8	14.3	6.4							
S36	20.3	19.9	20.8	20.3	21.5	21.7	21.7	20.6	21.7	22.4	22.7	21.4	21.7	21.7	21.1	21.8	22.1	18.8	17.9	18.9	22.6	22.7	22.6	22.4	21.7	22.4	23.2	18.0	13.9	15.0	17.1	17.3						
S40	21.9	22.2	22.8	22.0	22.5	21.9	21.7	21.1	24.1	23.2	23.4	22.2	23.0	23.8	23.1	23.5	23.6	18.4	19.2	17.6	23.6	23.6	23.2	22.6	23.6	22.2	22.7	18.7	14.7	16.4	16.0	15.5	14.2					
S32	22.3	21.8	22.7	22.0	21.5	21.4	20.3	20.2	21.7	23.3	24.5	21.5	22.6	22.1	22.2	22.3	22.4	19.5	20.6	17.9	23.0	23.0	22.6	22.7	22.9	22.4	22.1	18.5	17.3	16.4	17.1	17.7	17.7	14.7				
S30	21.7	22.0	21.9	21.7	21.9	21.7	21.1	21.7	21.1	22.8	23.4	23.0	23.8	23.1	23.2	22.8	23.1	19.5	20.7	21.2	23.4	23.4	23.3	22.7	22.8	24.1	23.0	22.2	20.1	20.4	22.2	21.2	21.9	20.8	17.9			
S22	23.3	23.5	23.8	22.7	23.6	24.1	22.7	22.1	22.7	23.5	24.1	23.5	23.3	23.5	22.3	22.9	23.3	22.9	22.6	22.0	23.2	23.3	23.2	23.2	22.7	24.1	24.2	24.2	21.5	21.1	23.3	23.0	20.3	23.3	22.1	23.0		
S23	23.4	23.6	24.0	22.9	24.0	24.1	22.6	22.4	22.9	23.6	23.8	23.4	23.3	23.3	22.8	23.4	24.0	23.1	23.1	22.0	23.6	23.8	23.8	23.8	23.3	25.0	24.5	24.8	22.6	21.0	23.7	22.9	20.6	24.0	22.7	23.4	0.5	

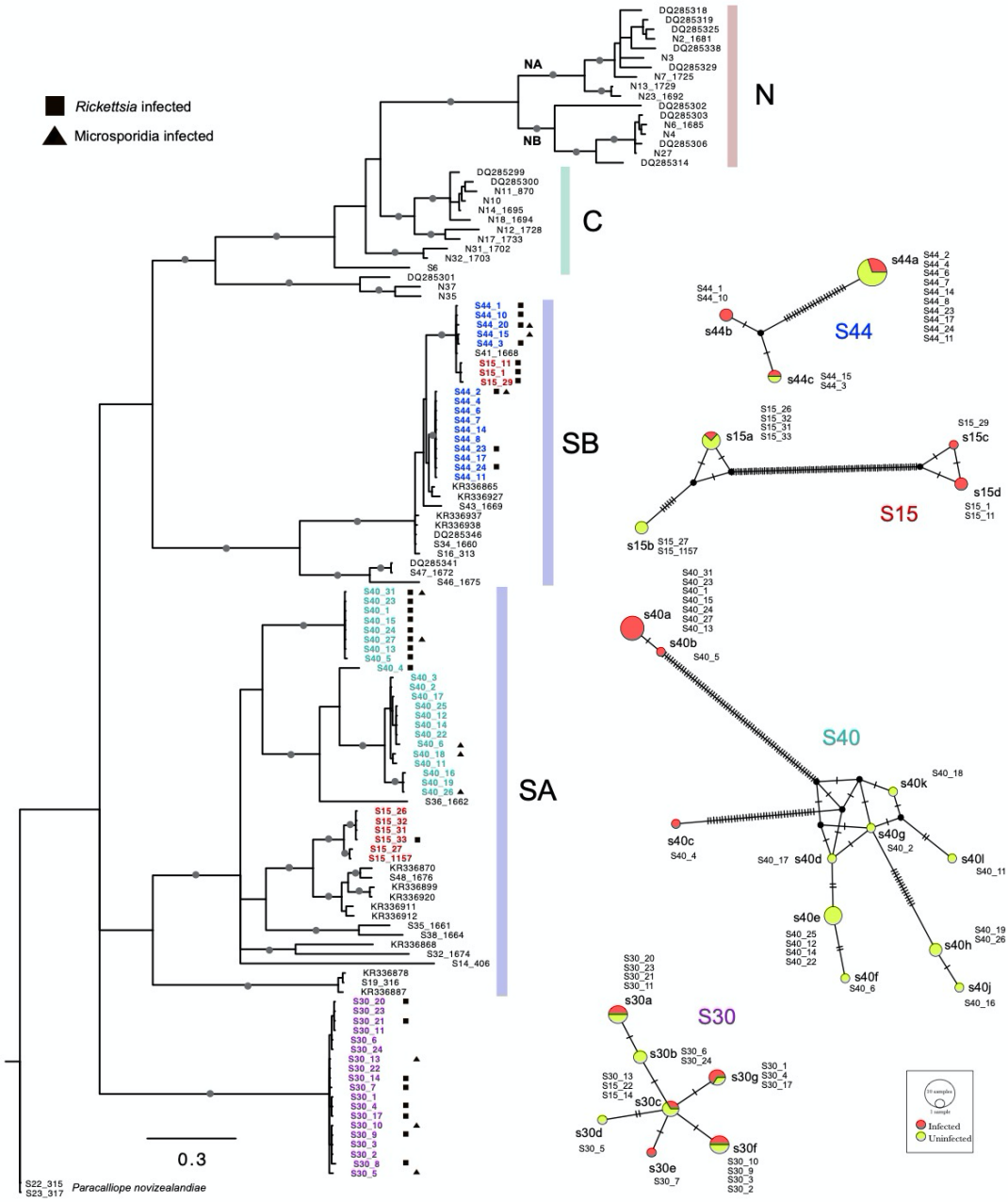


Figure 4.4 A Bayesian tree of COI sequences to show the linkage disequilibrium (LD) between parasites (Microsporidia and *Rickettsia*) and host COI haplotypes. The data on parasite infection status and host haplotype were obtained from four populations (S15, S40, S40, S44). Microsporidians do not show a clear pattern of LD, but *Rickettsia* shows clear LD in the S40 population. Haplotype networks are also shown for each population. Colours in the haplotype networks represent infection status (Pink; infected, lime; uninfected).

4.4.4. Linkage disequilibrium within a population

In total, 64 COI sequences (=64 individual amphipods) were obtained from 4 populations (S15, S30, S40, and S44). The infection status by each parasite (Microsporidia and *Rickettsia*) in each individual is shown (Figure 4.4). There were no noticeable associations between Microsporidia and COI haplotypes. On the other hand, a clear pattern of LD between *Rickettsia* and COI haplotypes was shown in S40 (Figure 4.4). Among 9 infected individuals sequenced, 7 individuals were associated with the same haplotype (s40a), and one with another haplotype (s40b), but only with one SNP different from s40a. One infected individual (S40_4) was associated with a haplotype distinct from all other haplotypes (s40c). All the uninfected individuals were associated with other variant haplotypes (s40d~l). Because both the SA and SB lineages were found in S15 (Figure 4.4), I additionally obtained 28S rRNA sequences of a few individuals of both lineages to see if this is a result of hybridization followed by introgression of mtDNA from one to another. Based on the sequencing result, COI lineages correspond to the 28S lineages (i.e. both were either SA or SB), which means that the two lineages coexist in the same habitat and there is no evidence for hybridization. Interestingly, individuals belonging to the SB group within S15 were closely related to individuals from S41 and S44, most of which harbour *Rickettsia*. S30 shows no pattern of LD.

4.4.5. Intra-population variation and neutrality tests

Some standard population genetic indices were obtained (Table 4.4). The nucleotide diversity was highest in S40 and lowest in S30. Both Tajima's D and Fu's Fs were below 0 in S30. Negative values could mean recent selective sweeps or bottleneck events. However, none of these tests were statistically significant, which is probably due to low sample numbers.

Table 4.4 Standard population genetic indices and results of Tajima's D and Fu's Fs tests for four populations.

Populaiton	No. of sequences	No. of haplotypes	No. of segregating sites	Haplotype diversity	Nucleotide diversity	Tajima's D	P-value	Fu's Fs	P-value
S15									
S30	18	7	7	0.8758 +/- 0.0396	0.002919 +/- 0.001967	-0.34196	0.374	-1.70804	0.117
S40	21	11	94	0.8667 +/- 0.0586	0.083088 +/- 0.041777	2.18384	0.997	10.39079	1
S44	14	3	25	0.4835 +/- 0.1425	0.018597 +/- 0.010122	1.4992	0.954	10.90255	1

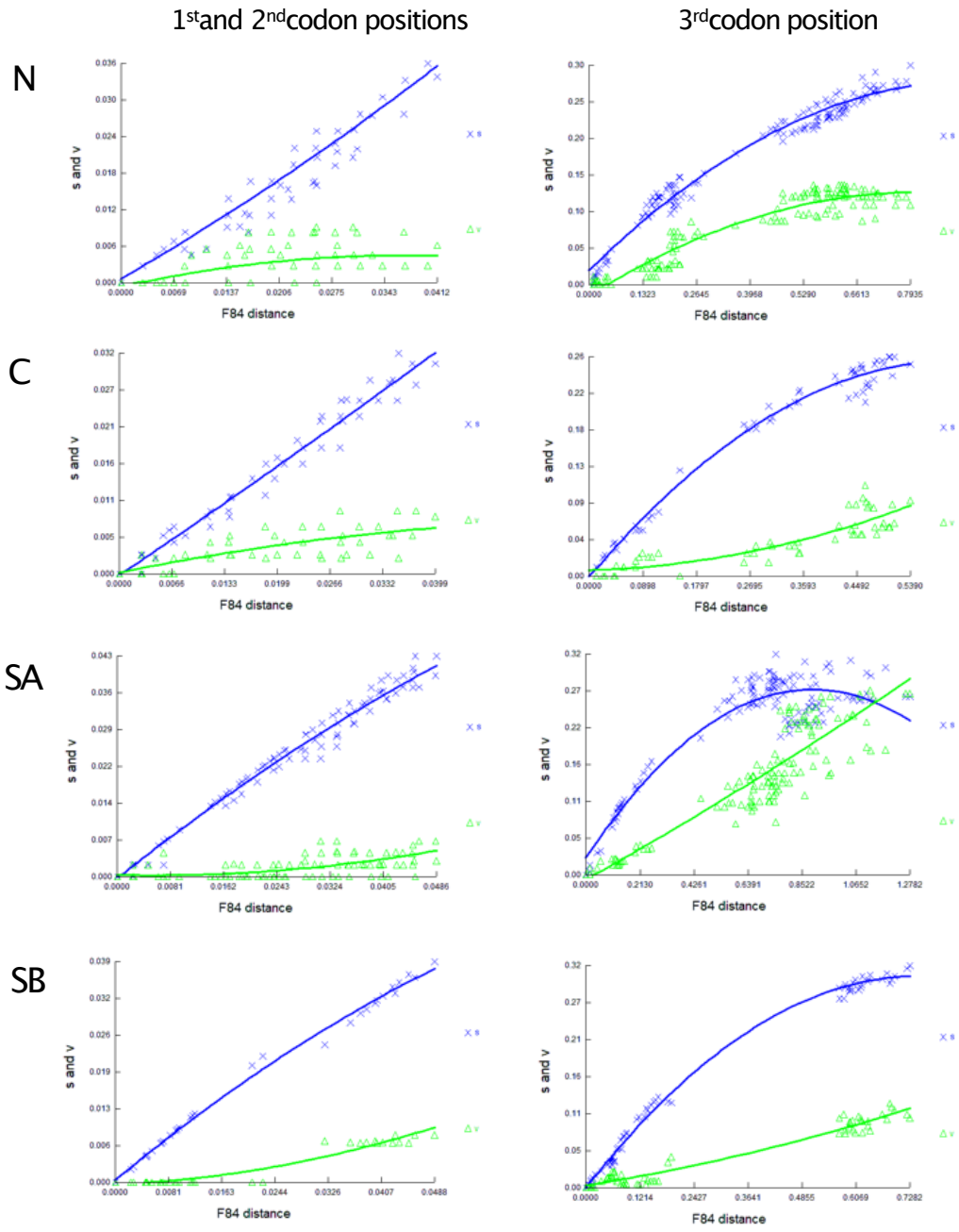


Figure 4.5 Plots of transitions and transversions against F84 genetic distance to visually diagnose substitution saturation. The plots were drawn for each group of *Paracallope*, and for the first and the second codon positions, and the third codon position, separately. The third codon position of the SA group is saturated with transversions catching up transitions. S and V in the y-axis represent transitions and transversions, respectively.

4.4.6. Substitution saturation

Plots of transitions and transversions versus F84 distance show no saturation on the first and second codon positions in all the main lineages, with transitions higher than transversions (Figure 4.5). The third codon position of N, C, and SB groups showed no saturation, however in the SA group transversions caught up with transitions. This means that the same sites were likely to be affected by multiple hits, and therefore this position has little phylogenetic information. The results of Xia's tests show that the third codon position is significantly saturated in New Zealand *Paracalliope* (Iss > Iss.c, p=0), therefore this position is useless for phylogenetic inference (Table 4.5). Although Iss is not larger than Iss.c in Xia's test for the third codon position of the SA group, they are not significantly different (p=0.11), suggesting significant saturation (Xia et al., 2003). Xia's test suggests the third codon position of the N group is also saturated, but separate tests for the NA and NB groups show no saturation (See Figure 4.4 for NA and NB). Xia's tests also show that the first and the second positions are not saturated in any of the groups.

Table 4.5 Xia's test of substitution saturation shown separately for the first and second codon positions, and for the third codon position, in the main lineages of the amphipod *Paracalliope*.

	1st and 2nd			3rd		
	Iss	Iss.c	Probability	Iss	Iss.c	Probability
N	0.2979	0.7016	0	0.5999	0.6367	0.6047
NA	0.2872	0.7655	0	0.4147	0.7365	0
NB	0.2683	0.7238	0	0.4472	0.7033	0
C	0.1532	0.7181	0	0.4004	0.6924	0
SA	0.125	0.7016	0	0.5721	0.6367	0.1137
SB	0.1258	0.7096	0	0.3121	0.6717	0
All	0.247	0.695	0	0.786	0.686	0.0393

4.5. Discussion

With more accessible and affordable molecular tools for sequencing, reports of mitonuclear discordance patterns have been increasing in the recent literature (Després, 2019; Gompert et al., 2008; Hinojosa et al., 2019; Toews and Brelsford, 2012). Patterns of mitonuclear discordance related to deeply divergent mitochondrial lineages are relatively rare compared to those for recently diverged lineages, but the former may be more common in certain taxa including amphipods. Here, I examine the patterns of mitonuclear discordance in *Paracalliope* and suggest several potential underlying causes for these patterns considering both extrinsic and intrinsic factors. Mitonuclear discordance problems often directly influence discrimination among of cryptic species and phylogeographical inferences, therefore I discuss some related issues and implications.

Mitonuclear discordance on deep time scales and substitution saturation

Our phylogenetic trees and networks show apparent mitonuclear discordance in both branching patterns (monophyly vs paraphyly) and branch lengths (little divergence vs extreme divergence) (Figures 4.1 and 4.3). However, the pattern of non-monophyly in this study is different from those usually observed over shallow time scales. In other words, discordance patterns did not arise due to the sharing of the same mtDNA haplotypes between distinct nuDNA lineages. Rather, COI sequences were highly divergent among populations (<24 %), and even within a single population (e.g. ~13 % in S40). The most striking difference between 28S and COI trees was branch lengths within each group. 28S sequences showed little variation within each group, but showed clear distance gaps between groups (Figure 4.1 and Tables 4.1-2). Branching patterns of 28S and COI trees were incongruent due to the non-monophyly of the C and S groups in the COI tree, which was caused by substitution saturation in the SA group (Figure 4.5). The third codon position of the SA group was substantially saturated, and contained little phylogenetic information (Figure 4.5 and Table 4.5). The genetic distances among populations of the SA group were generally higher than those of other groups; these were also represented as long terminal branches on the phylogenetic tree and the network (Figure 4.1 and 4.3).

Underlying causes for divergent COI

Several lines of evidence suggest a potential association between *Rickettsia* and mitonuclear discordance in *Paracalliope*: *Rickettsia* was found in populations of the S and C groups, where non-monophyly occurred (Figure 4.1); *Rickettsia* was the most prevalent in the SA group, in which COI sequences were extremely divergent and substitution saturation was observed (Figure 4.2 and 4.5); an apparent LD between *Rickettsia* and a host COI haplotype was observed in a population with divergent lineages (S40; Figure 4.4). However, phylogeographical patterns are determined by various factors including geographical and demographical histories. The long branches and substitution saturation in the SA group can be explained by several different processes. I propose two hypotheses: 'ancient lineage' and 'accelerated substitution rates', considering the fact that branch length reflects both time and rate components (Sanderson, 2003; Takezaki et al., 1995).

'Ancient lineage' hypothesis

Substitution saturation and overall high genetic distances among populations of the SA group may indeed reflect the old age of this group. The Otago region has been geologically and environmentally stable for a long time compared to other regions in the South Island (Apte et al., 2007; Craw et al., 2017). For example, Canterbury, Nelson, and southern coastal areas did not exist before the Pliocene (5 MYA). Accordingly, genetic diversity is high in the Otago region in some taxa, some of which having originated in this region (Buckley et al., 2001; Burbidge et al., 2003; Waters and Wallis, 2001). The origin of the genus *Paracalliope* is uncertain. However, *Paracalliope* has been discovered in freshwaters in India, the Philippines, New Caledonia, and Tasmania (Chilton, 1920; Iannilli and Ruffo, 2007; Knott, 1975), suggesting their ancient origin possibly dating back to Gondwana. Chilton (1920) reported that he could not find any morphological differences between amphipod specimens of *Paracalliope* from New Zealand, India, and the Philippines. If this genus indeed has an ancient origin, this would be a striking example of 'morphological stasis' (Lee and Frost, 2002; Sturmbauer and Meyer, 1992). The paleolake Manuherikia, formed during the early Miocene and lasted until the late Miocene, covered a large area (~56,000 km^2) which is now part of the Otago region (Reichgelt et al., 2015). If *Paracalliope* existed in the South Island

during that time, the SA group could be a descendant of the old lineage that flourished in the paleolake and then became isolated in different streams and rivers. However, the very limited or lack of divergence in 28S and ITS sequences among populations of the SA group does not support the simple 'ancient lineage' scenario, because considerable evolutionary changes would have accumulated in nuclear markers as well. However, if frequent hybridization events have homogenized nuDNA but not mtDNA (Fuertes Aguilar et al., 1999), the 'ancient lineage' scenario could still be plausible. Although freshwater amphipods have limited dispersal abilities, historical connections among different catchments due to changes in drainage geometry (Carrea et al., 2013) or occasional flooding may have allowed gene flow among *Paracalliope* populations.

'Accelerated substitution rates' hypothesis

COI sequences are extremely divergent among amphipod populations, especially in the SA group, but they are divergent in other groups (N and C) as well. It has also been suggested that substitution rates are elevated in some groups of amphipods (Verheye et al., 2016). Therefore, I hypothesize that substitution rates may be generally high in *Paracalliope*, but the presence of endosymbionts may have elevated substitution rates even higher in some lineages.

Endosymbionts

Endosymbionts can disrupt population and phylogeographical patterns in mtDNA in various ways (Engelstädter and Hurst, 2009; Hurst and Jiggins, 2005; Jiggins, 2003). Several authors have suggested a potential role for *Wolbachia* and other symbionts in promoting speciation (Brucker and Bordenstein, 2012; Shoemaker et al., 1999; Werren, 1998). Here, I argue that *Rickettsia* (and/or other endosymbionts) may have played some roles in accelerating substitution rates in mtDNA in amphipods. I suggest two possible mechanisms for this. First, reproductive manipulation caused by endosymbionts may affect gene flow among individuals and lead to reproductive isolation. It is well known in *Wolbachia*-host associations that cytoplasmic incompatibility caused by endosymbionts can generate genetic barriers between infected and uninfected individuals, or between males and females infected with different strains (Bourtzis et al., 1996). The presence of divergent lineages within a population and the

clear LD between *Rickettsia* and a lineage within the S40 population support this possibility. Also, despite the SA and SB groups coexisting in some locations (S15 and Lagrue et al. 2016) without any evidence of hybridization, the fact that the two groups are associated with different strains of *Dictyocoela* (Microsporidia) and *Rickettsia* (Chapters 2 and 3), suggests a genetic barrier between SA and SB groups possibly maintained by endosymbionts.

Second, recurrent selective sweeps may accelerate substitution rates in a population, although this hypothesis has not yet been theoretically and empirically tested. For a neutral site, the possibility for fixation of a mutation is equal to the mutation rate (Kimura, 1977). However, positive selection can quickly fix a mutation in a population (Fitzpatrick et al., 2010). It is well known that a selective sweep can rapidly increase the frequency of an allele in a population (Kim and Neilsen, 2004). The same applies to the cases of reproductive manipulators and associated mitochondrial haplotypes (Hurst and Jiggins, 2005). For example, a rare haplotype was rapidly fixed in *Drosophila* in the USA (Turelli et al., 1992). This has resulted in fixation of a rare haplotype in a very short period, which would have taken a long time without a selective sweep. If another selective sweep occurs in the same population with another rare haplotype, and if this process is repeated, recurrent selective sweeps can make two populations diverge rapidly. However, this hypothesis needs to be investigated further.

Another complicating factor is that *Rickettsia* is not alone. There could be other endosymbionts that are not known yet and their complex interactions within a population could generate various patterns. Different combinations of endosymbionts exert different selection pressures on different populations, which can result in a highly heterogeneous selective landscape and divergence among populations. Torix *Rickettsia* are common in diverse invertebrates and frequent horizontal transmission has been inferred (Chapter 3; Pilgrim et al., 2020). Therefore, recurrent selective sweeps by different strains are highly likely.

Host life-history traits

Genetic variation within a population changes over time as a result of mutation, gene flow (migration), natural selection, and genetic drift. In other words, mutation generates new variants, new alleles can arrive from other populations through migration, and evolutionary forces or stochastic processes can change the allele frequency in a population (Slatkin, 1987). If mitonuclear discordance associated with deep branching is more common in certain taxa than others, there may be some lineage-specific traits behind the elevated substitution rates. Because most substitutions in *Paracalliope* COI occurred on the third codon position (hence no changes in amino acid composition), fixation due to direct positive selection is unlikely. Effective population size (N_e) is one of the most important parameters underpinning population dynamics and the effectiveness of selection relative to drift (Charlesworth, 2009; Nei and Tajima, 1981). A population will be more susceptible to genetic drift when the effective population size is small. Therefore, I focus on several life-history traits in amphipods that may contribute to lowering the effective population size by affecting population structure. First, fluctuating sex-ratio throughout the year may lower the effective population size. In amphipods, environmental sex determination (ESD) is believed to play an important role (Bachtrog et al., 2014). ESD is known in *Gammarus* and *Echinogammarus*; it is believed to be adaptive as it allows males born earlier than females to grow larger, which is advantageous for mating success (Guler et al., 2012; McCabe and Dunn, 1997; Nomura, 2002). Therefore, sex-ratio periodically changes in an amphipod population with ESD. The effective population size can be estimated as $N_e = 4N_mN_f / (N_m + N_f)$, where N_m is the number of males and N_f is the number of females (Nomura, 2002). There are many ways to estimate N_e (see Ryman et al., 2019), though it is always lower when the sex ratio is not 1:1. Second, fecundity varies considerably among individual females and throughout the year (Bella and Fish, 1996; Cunha et al., 2000). The variation in fecundity also lowers the effective population size (Vucetich et al., 1997). Third, amphipods may have high mutation rates due to short generation times. Amphipods reach maturity within a few months with some variation (27-210 days; Welton and Clarke, 1980). Every generation, new mutations accumulate within a population, meaning more diversity on which natural selection and genetic drift can act (Kimura, 1977).

Synergistic effects

Endosymbionts and taxon-specific life-history traits are not mutually exclusive as causes of mitonuclear discordance, but may also act synergistically to accelerate substitution rates and result in substitution saturation. For example, endosymbionts can influence host fecundity and change host sex-ratio (Dunn et al., 2001). In other words, sex-ratio change caused by endosymbionts could exacerbate the sex-ratio imbalance caused by ESD. In addition, the selective sweep caused by endosymbionts often reduces genetic diversity and lowers the effective population size within an infected population (Hurst and Jiggins, 2005; Johnstone and Hurst, 1996). Therefore, after a selective sweep, the population will be more likely to be affected by genetic drift, which can lead to the fixation of a rare mutation.

Implications

Molecular clock

Molecular clock rates of 1.4-2.6 %/MY have been widely used for a wide range of taxa especially when fossil data are scarce or not available (Hipsley and Müller, 2014; Knowlton and Weigt, 1998). However, if some amphipods including *Paracalliope* indeed have higher substitution rates than other invertebrates, applying substitution rates obtained from other taxa may result in overestimation of divergence times. Furthermore, if recurrent selective sweeps play a major role in mtDNA evolution in amphipods, the substitution rates will be highly influenced by the frequency and the extent of selective sweeps rather than inherently high mutation rates.

Cryptic species

Deep genetic divergence is often interpreted as evidence for the existence of cryptic species (Bickford et al., 2007; Fišer et al., 2018). There are numerous examples where mtDNA is highly divergent suggesting the presence of cryptic species, while nuDNA, behavioral, and morphological data do not support cryptic species (Giska et al., 2015; Hinojosa et al., 2019; Pazhenkova and Lukhtanov, 2019). In the biological species concept, species are defined as ‘groups of interbreeding natural populations that are reproductively isolated from other such groups’ (Mayr, 2000). Sutherland et al. (2010) conducted an interesting mating experiment to test if mate discrimination proportionally

increases with genetic distances in *Paracalliope*, using 7 divergent populations from both the North and South Islands. Despite high COI divergence (19.5 %) between two populations (Hamilton and Napier; both from North Island), individuals from these populations paired and produced eggs. However, individuals from North and South populations (divergence >21.5 %) tended not to pair when in the presence of each other; the few individuals that did pair did not produce eggs, suggesting behavioural and genetic isolation between North and South Island populations. It is likely that Hamilton and Napier populations belong to the N and C groups, respectively. In fact, the 28S phylogeny explains the result of the mating experiment much better than the COI phylogeny. Females from the Wellington population in Sutherland et al (2009), which is genetically similar to the N35 population in our study, were able to produce eggs when mated with males from the Hamilton (N group) population. The N35 population is sister to all other N populations in our 28S tree. This mating experiment suggests that there is a genetic barrier between NC and S. Cryptic species indeed exist and are probably common in amphipods due to frequent 'morphological stasis'. However, mtDNA is likely to overestimate the number of cryptic species. Similar cases are also known in other amphipods. For instance, divergent COI lineages were not supported by morphology among Australian chiltoniid amphipods (King et al., 2012). Mitonuclear discordance between COI and ITS sequences was observed in *Niphagus* in Austria (Stoch et al., 2020). In Iceland, five distinct COI clades were identified in a groundwater amphipod *Crangonyx* (Kornobis et al., 2010), but ITS sequences and nuclear genomic data (ddRAD) showed different patterns and did not support cryptic species (Eme et al., 2017; Kornobis and Pálsson, 2011).

In conclusion, mtDNA has been a source of useful information to resolve the evolutionary history of organisms. However, mtDNA and nuDNA have inherently different natures and are likely to show different phylogenetic patterns. Also, because mtDNA is maternally transmitted, it is prone to be affected by co-transmitted endosymbionts. In this chapter, I examined various factors that may have shaped the current phylogeographical patterns of *Paracalliope* amphipods in New Zealand. The mitonuclear discordance patterns I observed are likely due to a combination of various factors. Further investigation of certain unresolved issues would be valuable. First, the

incidence and diversity of endosymbionts in other amphipods need to be quantified to see if endosymbionts have indeed played some roles in host mtDNA evolution. Second, genomic scale data should be used to estimate demographical histories which may also have caused mitonuclear discordance. Third, understanding how different strains/species/groups of endosymbionts interact at a population level would inform on further conclusions regarding their evolutionary consequences. Because of widespread morphological stasis and homoplasy, using multiple markers will be especially important in amphipods, and any mtDNA pattern should be interpreted with caution.

Chapter 5

Two parasites in one host: spatiotemporal dynamics and cooccurrence of Microsporidia and *Rickettsia* in an amphipod host

5.1. Abstract

Biological interactions can greatly influence the abundance of species. This is also true for parasitic species that share the same host, especially if they occupy the same tissues within the host. Depending on their mode of transmission, their virulence, and whether their interests are aligned or not, various outcomes are possible. Both Microsporidia and *Rickettsia* are common in populations of *Paracalliope*, the most common freshwater amphipods in New Zealand. Although both parasites coexist in many populations, it is unclear whether they interact with each other. In this brief chapter, I investigated spatial-temporal dynamics and co-occurrence of the two parasites, Microsporidia and *Rickettsia*, in *Paracalliope* hosts, across one annual cycle and in three different locations. Prevalence of both Microsporidia and *Rickettsia* changed over time. However, while the prevalence of *Rickettsia* varied significantly between sampling times, that of Microsporidia did not change significantly and remained relatively low. The two parasites therefore followed different temporal patterns. Also, the prevalence of both parasites differed among locations, though the two species reached their highest prevalence in different locations. Lastly, there was no evidence for positive or negative associations among the two parasite species; in other words, they did not co-infect the same individual hosts more or less often than expected by chance. Although the presence of one parasite in an individual host does not appear to influence the probability of infection by the other parasite, due to environmental heterogeneity across locations, their respective prevalence may follow different patterns among populations on larger a spatial scale.

5.2. Introduction

The abundance of any species in any habitat varies over time. Environmental factors such as temperature and precipitation drive the abundance in many organisms (Pollard et al., 1999; White et al., 2000), and so do biological interactions (Martins and Haimovici, 1997; Woodin, 1974). For example, the density of prey and predators in the habitat (Arditi and Ginzburg, 1989), competition for resources (Robertson, 1996), and parasites and diseases can all contribute to shape the abundance and dynamics of animal populations (Poulin, 1999; Scott and Dobson, 1989). The prevalence of parasites themselves is also governed by various factors. Because parasites are highly dependent on their host for survival, all environmental and biological factors that influence host abundance can, directly and indirectly, also affect the prevalence of their parasites (Arneberg et al., 1998). More specifically, changes in host behavior, host immune response, and fluctuating host births and deaths themselves can alter the prevalence of parasites in a host population (Grassly and Fraser, 2006).

Another important factor that determines the prevalence of parasites is the interaction among parasites that use the same hosts (Rigaud et al., 2010). Depending on their mode of transmission, virulence, and their ability to manipulate the hosts, various outcomes are possible (Haine et al., 2005; Poulin, 2011). For example, vertically transmitted parasites often have little or no effect on their host's fitness because the hosts' reproductive success is crucial for the parasite as well (Dunn and Smith, 2001). On the other hand, some trophically (=horizontally) transmitted parasites are capable of changing host behavior, which ultimately leads to predation of the intermediate host by the final host (Thomas and Poulin, 1998). Parasites with these contrasting modes of transmission sharing the same host are therefore in conflict. For example, in an amphipod host, the coinfection of vertically transmitted microsporidia has been shown to weaken the behavioral alteration induced by the trophically transmitted acanthocephalans (Haine et al., 2005). In contrast, parasites with the same transmission mode that share the same host may have interests that are well aligned. In an amphipod host, two vertically transmitted parasites, one microsporidian and one paramyxean, were shown to co-occur more frequently than expected by chance, which suggested that one

parasite was hitchhiking with another capable of feminizing the host (Short et al., 2012), or that both can feminize their hosts (Arundell et al., 2015; Pickup and Ironside, 2018).

Amphipods host diverse macroparasites (trematodes, acanthocephalans, and nematodes) as well as diverse microparasites (viruses, bacteria, and protists) (Bethel and Holmes, 1977; Bojko and Ovcharenko, 2019; Friesen et al., 2019; Poulin and Latham, 2002). Amphipods are used as intermediate hosts by many parasites (Dezfuli et al., 2000), and a single amphipod individual can be infected by several different groups of parasites at the same time (Haine et al., 2005; Short et al., 2012). I detected Microsporidia and the bacteria *Rickettsia* in multiple New Zealand populations of *Paracalliope* amphipods (see Chapters 2 and 3). The two parasites often coexist in the same host population (Chapter 4). However, since they are both intracellular parasites, could occur in gonadal cells for vertical transmission, they may compete for space within the same individual host, which could lead to the prior infection by one parasite leading to the exclusion of the other parasite from the same individual host. On the other hand, they may also be positively associated. For example, since microsporidians can feminize their host (Dunn et al., 2001) and therefore improve their chances of transmission (vertical transmission is only possible from female hosts to their offspring), *Rickettsia* may benefit from associating with microsporidian-infected host individuals. The question is then, do they tend to coinfect the same individual amphipod more frequently than expected by chance, avoid each other, or are there no associations between them?

Because both microsporidians and *Rickettsia* include species of economic importance and serious pathogens in humans, livestock, and companion animals, their seasonal dynamics are relatively well known in those groups. For example, microsporidian keratitis peaks during the rainy season in several countries (Reddy et al., 2011; Tham and Sanjay, 2012). Rocky Mountain spotted fever caused by *Rickettsia rickettsia* peaks during seasons when vector species (i.e. ticks) are abundant (Walker, 1995). The prevalence of parasites follows different temporal trends in different host systems. A recent meta-analysis showed that there is no universal pattern in the seasonal dynamics of aquatic parasites; instead seasonal variation in infection levels depends on taxa and habitat (Poulin, 2020). Microsporidians also show various temporal trends driven by

different factors. For example, *Octosporea bayeri* in *Daphnia* hosts showed clear cyclic prevalence patterns increasing in summer and decreasing in winter. This was related to the host lifecycle (i.e. diapause) rather than external temperature (Lass and Ebert, 2006). Other microsporidians, also in *Daphnia* but in different locations, showed more-or-less constant prevalence over space and time (Wolinska et al., 2011). Microsporidian species in *Artemia* displayed a clear pattern of seasonality although this was affected by the presence of other host species (Lievens et al., 2019). Therefore, different host-parasite associations can be characterised by various patterns of temporal prevalence fluctuations.

Here, I investigate patterns of temporal variations in prevalence and in the co-occurrence of Microsporidia and *Rickettsia* in *Paracalliope* host individuals and populations. I ask several specific questions: Does the prevalence of micrsoporidians and *Rickettsia* change throughout the year? If so, do they have similar temporal patterns in different locations? Are their temporal variations in prevalence associated with host population dynamics? Do microsporidians and *Rickettsia* tend to co-infect the same individual hosts, or not? In order to answer these questions, I sampled *Paracalliope* specimens across an entire annual cycle from three different locations. I quantify seasonal infection dynamics of both parasites as well as seasonal changes in host demographic parameters, and I test whether the co-occurrence of the two parasites among individual hosts departs from random.

5.3. Materials and methods

5.3.1. Field sampling

Three sampling sites (S34, S37, S40) were chosen among sites with both Microsporidia and *Rickettsia* based on the screening results from Chapter 2 (Figure 5.1). These sites were visited every two months between February 2019 and February 2020 (a total of seven sampling times). Individuals of *Paracalliope fluviatilis* were collected with dipnets and fine sieve nets; samples were collected among littoral macrophytes, in a standardized manner across localities and sampling times. Samples were stored in containers with 96% ethanol upon collection and then brought to the lab.

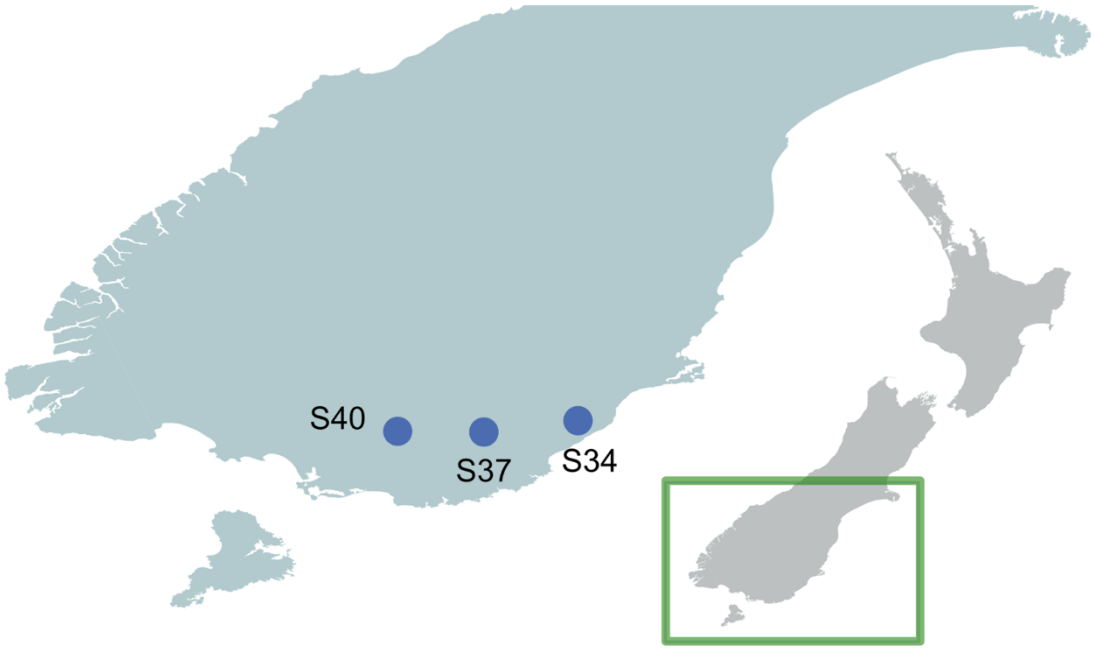


Figure 5.1 Map of New Zealand's South Island showing sampling locations. *Paracalliope fluviatilis* specimens were collected from the three sites in the Otago and Southland regions within one-year span.

5.3.2. Sample preparation

For each population and for each sampling time, 24 individuals were randomly chosen for molecular screening for parasite detection. They were sexed under a microscope and then were photographed using a DP25 camera mounted on a microscope, and the Olympus DP2-BSW application software. These photos were later used to measure the body size of each amphipod individual. The distance from the base of the first antennal segment to the base of the telson (Asochakov, 1994), was recorded as a measure of body size using ImageJ software (http://imagej.nih.gov/ij/). The brood size (=number of eggs in a brood pouch) was recorded for each ovigerous female. After being washed with distilled water, the whole body was used for DNA extraction for each individual amphipod.

5.3.3. Parasite detection by PCR

The presence of Microsporidia and *Rickettsia* for each amphipod individual was detected by PCR, using the methods described in Chapters 2 and 3, respectively.

5.3.4. Data analysis

The temporal fluctuations in the prevalence of each parasite and their co-occurrence in three different locations were visually represented as stacked bar graphs (Figure 5.2A), whereas temporal variations in amphipod body sizes (males and females separately) and the brood size of females were plotted as boxplots (Figure 5.2B-D). All plots were generated using *the ggplots2* package in R (version 3.5.2; R Core Development Team, 2018).

I used two generalized linear models (GLM), one for Microsporidia and one for *Rickettsia*, to evaluate the influence of several factors on the occurrence of each parasite in individual amphipods. The presence of the parasite was used as response variable (binomial distribution: uninfected=0, infected=1). I assessed several fixed factors: presence of the other parasite in the amphipod (absent=0, present=1), sampling time, location, sex, and amphipod body size for their effects on the focal parasite's occurrence. 'Sampling time' had 7 levels (Feb-19, Apr-19, Jun-19, Aug-19, Oct-19, Dec-19, and Feb-20), 'location' had three levels (S34, S37, and S40), and 'sex' had two

levels (female and male). The GLMs were conducted using "*glm*" function in the "*lme4*" package (Bates et al., 2015). Pairwise comparisons across different 'locations' and 'sampling times' were performed using the *glht* function in the *multcomp* package (Hothorn et al., 2008). All statistical analyses were conducted in the R environment (version 3.5.2; R Core Development Team, 2018).

5.4. Results

5.4.1. Host population dynamics

Parcalliope populations persisted throughout the year in all three locations (Figure 5.1), although I failed to collect specimens in the S34 population from one sampling time (Dec-19). *Paracalliope* demographic parameters, i.e. body size and brood size, showed clear temporal variations in all locations (Figure 5.2B-D). It seems that *Paracalliope* is most productive during the spring (September to November), based on our observation of the highest brood size in females from the October samples, across all three populations. Brood sizes appear to decrease during the summer and early winter (December-June). In June, females harboring eggs were very rare in all locations (Figure 5.2B). The body size of the females and males also varied greatly throughout the year, showing similar patterns with that of brood size. It seems that mature females and males are mostly found during the late winter to spring (August-October). During autumn to winter (February-June), the populations consisted mostly of small, immature individuals (Figure 5.2C-D).

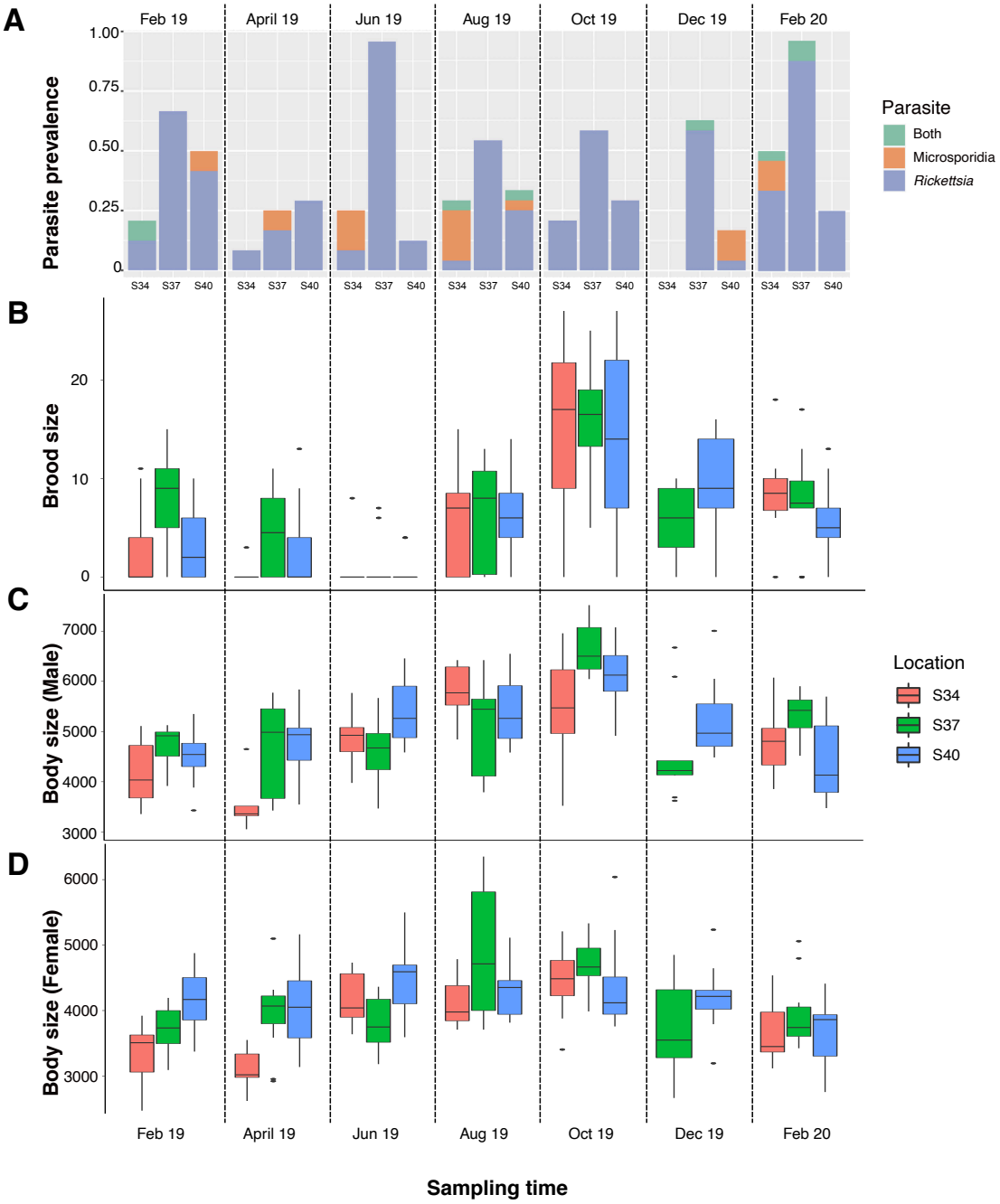


Figure 5.2 The temporal variations of (A) parasite prevalence in three locations, (B) and the brood size (=number of eggs) of amphipod females, the body size (μm) of (C) males and (D) females. For box plots, medians (central lines), 25–75 percentiles (boxes), non-outliers (whiskers), and outliers (dots) are shown. All these traits are aligned together (with dashed lines) for better comparison of the temporal trends.

5.4.2. Spatiotemporal variations in parasite prevalence

Both Microsporidia and *Rickettsia* were found from all three locations, but their prevaence changed throughout the year (Figure 5.2A). *Rickettsia* was found from all locations at all sampling times, but Microsporidia were not found at some sampling times. The prevalence of *Rickettsia* showed ranges of 8.3-37.5 %, 16.7-95.8 %, and 4.2-41.7 % in the S34, S37, and S40 populations, respectively. On the other hand, the prevalence of Microsporidia was generally low, i.e. 0-8.3 % and 0-12.5 % in the S37 and S40 populations, respectively, although its prevalence reached up to 25 % in winter at the S34 population. The GLM results supported the lack of temporal variation in the prevalence of Microsporidia (Table 5.1). On the other hand, the effect of sampling time (temporal effect) on *Rickettsia* infections was supported by GLM results (Table 5.2). Compared to the first sampling time, the April samples showed a significantly lower prevalence of *Rickettsia*. Also, locations had effects on the prevalence of both Microsporidia and *Rickettsia*, indicating that different locations tend to have a different prevalence of both parasites. The S37 population has the highest prevalence of *Rickettsia* ($z=8.23$, $p< 2e-16$). The pairwise comparisons of the effects of different sampling times and locations, respectively, on infections by each parasite, are shown in Tables 5.3 and 5.4.

5.4.3. Co-occurrence of the two parasites

Individuals simultaneously harboring both Microsporidia and *Rickettsia* were found in all locations and from several different sampling times (Figure 5.2A). According to the GLM results, there was no effect of the infection by one parasite to the presence or absence of the other parasite (Tables 5.1 and 5.2). In other words, the presence of Microsporidia did not predict the presence of *Rickettsia* at the individual host level ($z=0.10$, $p=0.92$), and vice versa ($z=0.22$, $p=0.82$).

Table 5.1 Generalized linear model results showing the effects of various factors on Microsporidia infections in amphipod individuals. Significant effects are shown in **bold**.

Fixed effects	Estimate	Standard error	Z-value	P-value
(Intercept)	-1.11E+00	1.30E+00	-0.854	0.39335
Rickettsia presence	1.16E-01	5.22E-01	0.223	0.82365
Body size	-2.02E-04	3.46E-04	-0.584	0.55916
Sampling2_April-19	-2.95E-01	9.11E-01	-0.323	0.74636
Sampling3_Jun-19	2.56E-01	7.75E-01	0.33	0.74109
Sampling4_Aug-19	1.03E+00	7.35E-01	1.396	0.16278
Sampling5_Oct-19	-1.65E+01	1.20E+03	-0.014	0.98904
Sampling6_Dec-19	1.36E+00	8.23E-01	1.65	0.09896
Sampling7_Feb-20	5.13E-01	6.95E-01	0.738	0.46066
LocationS37	-1.87E+00	6.60E-01	-2.83	**0.00466**
LocationS40	-1.42E+00	5.31E-01	-2.672	**0.00754**
Sex (male)	-4.72E-01	5.36E-01	-0.88	0.37899

Table 5.2 Generalized linear model results showing the effects of various factors on *Rickettsia* infections in amphipod individuals. Significant effects are shown in **bold**.

Fixed effects	Estimate	Standard error	Z-value	P-value
Intercept	0.0653594	0.676144	0.097	0.92299
Microsporidia presence	0.0511338	0.4990964	0.102	0.9184
Body size	-0.0003432	0.0001677	-2.047	**0.04062**
Sampling2_April-19	-1.5958317	0.4540489	-3.515	**0.00044**
Sampling3_Jun-19	-0.0306738	0.3996621	-0.077	0.93882
Sampling4_Aug-19	-0.4487632	0.4228839	-1.061	0.2886
Sampling5_Oct-19	-0.081639	0.4364869	-0.187	0.85163
Sampling6_Dec-19	-1.0378586	0.4504977	-2.304	**0.02123**
Sampling7_Feb-20	0.55359	0.3826242	1.447	0.14795
LocationS37	2.6256841	0.3192275	8.225	**< 2e-16**
LocationS40	0.7101636	0.312914	2.27	**0.02324**
Sex (male)	-0.1995707	0.2717154	-0.734	0.46265

Table 5.3 The pairwise comparisons of the effects between 'sampling times' and 'locations' on Microsporidia infections from the GLM results.

	Fixed effects	Estimate	Standard error	Z-value	P-value
Sampling time	2_April-19 - 1_Feb-19 == 0	-0.2946	0.9109	-0.323	1
	3_Jun-19 - 1_Feb-19 == 0	0.2562	0.7754	0.33	1
	4_Aug-19 - 1_Feb-19 == 0	1.0264	0.7354	1.396	0.77
	5_Oct-19 - 1_Feb-19 == 0	-16.4709	1198.8638	-0.014	1
	6_Dec-19 - 1_Feb-19 == 0	1.3585	0.8234	1.65	0.603
	7_Feb-20 - 1_Feb-19 == 0	0.5125	0.6947	0.738	0.987
	3_Jun-19 - 2_April-19 == 0	0.5508	0.9365	0.588	0.996
	4_Aug-19 - 2_April-19 == 0	1.3211	0.9062	1.458	0.732
	5_Oct-19 - 2_April-19 == 0	-16.1763	1198.864	-0.013	1
	6_Dec-19 - 2_April-19 == 0	1.6532	0.9509	1.738	0.541
	7_Feb-20 - 2_April-19 == 0	0.8071	0.8801	0.917	0.962
	4_Aug-19 - 3_Jun-19 == 0	0.7702	0.6696	1.15	0.893
	5_Oct-19 - 3_Jun-19 == 0	-16.7271	1198.8638	-0.014	1
	6_Dec-19 - 3_Jun-19 == 0	1.1024	0.8054	1.369	0.786
	7_Feb-20 - 3_Jun-19 == 0	0.2563	0.7072	0.362	1
	5_Oct-19 - 4_Aug-19 == 0	-17.4973	1198.8637	-0.015	1
	6_Dec-19 - 4_Aug-19 == 0	0.3321	0.7405	0.449	0.999
	7_Feb-20 - 4_Aug-19 == 0	-0.5139	0.6482	-0.793	0.982
	6_Dec-19 - 5_Oct-19 == 0	17.8294	1198.8639	0.015	1
	7_Feb-20 - 5_Oct-19 == 0	16.9834	1198.8638	0.014	1
	7_Feb-20 - 6_Dec-19 == 0	-0.846	0.7646	-1.106	0.909
Locations	S37 - S34 == 0	-1.8676	0.66	-2.83	**0.0128**
	S40 - S34 == 0	-1.4181	0.5307	-2.672	**0.0203**
	S40 - S37 == 0	0.4495	0.667	0.674	0.7772

Table 5.4 The pairwise comparisons of the effects between 'sampling times' and 'locations' on *Rickettsia* infections from the GLM results.

	Fixed effects	Estimate	Standard error	Z-value	P-value
Sampling time	2_April-19 - 1_Feb-19 == 0	-1.59583	0.45405	-3.515	**0.00782**
	3_Jun-19 - 1_Feb-19 == 0	-0.03067	0.39966	-0.077	1
	4_Aug-19 - 1_Feb-19 == 0	-0.44876	0.42288	-1.061	0.93845
	5_Oct-19 - 1_Feb-19 == 0	-0.08164	0.43649	-0.187	1
	6_Dec-19 - 1_Feb-19 == 0	-1.03786	0.4505	-2.304	0.2397
	7_Feb-20 - 1_Feb-19 == 0	0.55359	0.38262	1.447	0.77391
	3_Jun-19 - 2_April-19 == 0	1.56516	0.46448	3.37	**0.01313**
	4_Aug-19 - 2_April-19 == 0	1.14707	0.47767	2.401	0.19558
	5_Oct-19 - 2_April-19 == 0	1.51419	0.4933	3.07	**0.03461**
	6_Dec-19 - 2_April-19 == 0	0.55797	0.49267	1.133	0.91716
	7_Feb-20 - 2_April-19 == 0	2.14942	0.456	4.714	**< 0.001**
	4_Aug-19 - 3_Jun-19 == 0	-0.41809	0.41669	-1.003	0.95276
	5_Oct-19 - 3_Jun-19 == 0	-0.05097	0.42184	-0.121	1
	6_Dec-19 - 3_Jun-19 == 0	-1.00718	0.45974	-2.191	0.29772
	7_Feb-20 - 3_Jun-19 == 0	0.58426	0.39391	1.483	0.75262
	5_Oct-19 - 4_Aug-19 == 0	0.36712	0.4144	0.886	0.97438
	6_Dec-19 - 4_Aug-19 == 0	-0.5891	0.47154	-1.249	0.87313
	7_Feb-20 - 4_Aug-19 == 0	1.00235	0.41535	2.413	0.19126
	6_Dec-19 - 5_Oct-19 == 0	-0.95622	0.48791	-1.96	0.43737
	7_Feb-20 - 5_Oct-19 == 0	0.63523	0.42539	1.493	0.74664
	7_Feb-20 - 6_Dec-19 == 0	1.59145	0.45235	3.518	**0.00791**
Location	S37 - S34 == 0	2.6257	0.3192	8.225	**<1e-04**
	S40 - S34 == 0	0.7102	0.3129	2.27	0.0595
	S40 - S37 == 0	-1.9155	0.2624	-7.3	**<1e-04**

5.5. Discussion

I investigated the spatiotemporal variations in the prevalence of two parasites, Microsporidia and *Rickettsia*, sharing the same *Paracalliope* amphipod species complex. There was no clear temporal variation in the prevalence of Microsporidia, but *Rickettsia* showed some temporal variations in the prevalence among different sampling times. The prevalence of both parasites varied across locations, however, the patterns were different: Microsporidia was more common in the S34 population, and *Rickettsia* was more common in the S37 and S40 populations (Figure 5.2 and Table 5.1-4). Here, I discuss whether the observed temporal variation in the prevalence of *Rickettsia* may be associated with host population dynamics. Then, I discuss various factors that influence the temporal infection patterns within a population as well as spatial patterns at a larger scale. I also discuss the possible causes for the lack of co-occurrence of the two parasites at the individual host level.

Although temporal trends in the prevalence of *Rickettsia* were not significant, the April samples showed low prevalence in all populations. On the other hand, the February samples showed higher prevalence. The difference in the prevalence of *Rickettsia* between the April and February samples is also supported by the pairwise comparison of means (Table 5.2). Because I only have data for one year, I do not know if the same patterns are consistent across years. The decrease in the prevalence in April could have been due to seasonal environmental change, stochastic processes, or it may be related to the change of the amphipod cohort. It may also be the product of the relatively modest sample sizes used in the analyses. The body size of both females and males was lowest during April, which means that the previous cohorts probably died and most individuals in the population at that time were a new, young cohort. As an obligate parasite, the strong dependence of *Rickettsia* on the host is assumed (Sibley, 2004), but more long-term data will be needed to better understand if there are indeed seasonal fluctuations in infection prevalence, and if so, what factors drive those patterns.

Host-parasite associations are shaped by various biotic and abiotic factors (Anderson and Sukhdeo, 2010). These factors are not the same among localities. The compositions of the host and parasite species differ among localities, as do climatic and other environmental factors. Therefore, the absence of general temporal patterns can be due to biotic and abiotic heterogeneity among habitats. Although I did not find any evidence of non-random coinfection (or avoidance) patterns between microsporidians and *Rickettsia* at the individual level in our three study sites, it is still possible that their distributions on a larger spatial scale are not mutually independent. Indeed, there is a difference in the spatial distribution of the two parasites across New Zealand. Although Microsporidia have been found throughout the country, *Rickettsia* was found only in the southern part of North Island and the Southern part of the South Island (C and S groups; see Chapters 2 and 3). *Rickettsia* was not found in the northern region of the North Island (N group of *Paracalliope*). The reason for the absence of *Rickettsia* in the northern parts of both Islands is not understood yet but could be due to phylogeographic or environmental processes, or simply due to sampling artifacts; this remains to be studied.

Even if both Microsporidia and *Rickettsia* infect the same individual, the exact host organs and tissues they target (i.e. tissue tropism), and how they interact with the host cell may be different (Sahni and Rydkina, 2009; Tamim El Jarkass and Reinke, 2020). In this case, they might not need to compete for the same resources, although they must both be present in the eggs of female hosts for successful vertical transmission. Also, if both have little or no pathogenic effects on the host, there would be no obvious conflict between them (Rigaud et al., 2010). If one of them affects the host immune system, which one infects the host first can be an important factor (Karvonen et al., 2019). Intracellular endosymbionts have generally shown low virulence to their hosts (Dunn et al., 2001), therefore they may be able to coexist in the same individual host without strong conflicts. On the other hand, these endosymbionts may have conflicts of interests with horizontally (including trophically) transmitted parasites. *Paracalliope* is an important prey item for fishes in New Zealand and is used as intermediate host by several different helminth parasites including trematodes and acanthocephalans (Friesen et al., 2019; Lagrue et al., 2016). Complex interactions, on both ecological and

evolutionary time scales, may therefore exist between microsporidians, *Rickettsia*, and helminths, and these may be involved in generating observed patterns in the spatial distribution of the two focal parasites in this study.

From an evolutionary perspective, the coexistence of microsporidians and *Rickettsia* in amphipod hosts is interesting. Canonical Microsporidia have several unique characteristics that make them different from their relatives (see Chapter 6). One of the unique characteristics is the presence of ADP/ATP translocators (Dean et al., 2016), which allows microsporidians to effectively 'steal' energy from the host cell. These translocators are highly similar to those of *Rickettsia* or *Chadymia*, and therefore it is believed that these translocators were horizontally obtained from these bacteria or their ancestors (Dean et al., 2016). Physical proximity must be a requirement for horizontal gene transfer. Therefore, it is likely that the ancestors of long-branch microsporidia and *Rickettsia* shared the same host cell. Their coexistence within the same host cell may have provided the evolutionary novelty which has led their joint evolutionary success.

Chapter 6

Revisiting the phylogeny of Microsporidia

6.1. Abstract

Canonical microsporidians are a group of obligate intracellular parasites of a wide range of hosts comprising ~1,300 species of >220 genera. Microsporidians are related to fungi, and many characterized and uncharacterized groups closely related to them have been discovered recently, filling the knowledge gaps between them. These groups assigned to superphylum Opisthosporidia have provided several important insights into the evolution of diverse intracellular parasitic lineages within the tree of eukaryotes. The most studied among opisthosporidians, canonical microsporidians, were known to science more than 160 years ago. However, the classification of canonical Microsporidia has been challenging due to common morphological homoplasy, and accelerated evolutionary rates. Instead of morphological characters, SSU rRNA sequences have been used as the primary data for the classification of canonical microsporidians. Previous studies have produced a useful backbone of the microsporidian phylogeny, but provided only some node supports, causing some confusion. Here, I reconstructed phylogenetic trees of canonical microsporidians using Bayesian and Maximum Likelihood inferences. I included rRNA sequences of 126 described/named genera, by far the broadest taxon coverage to date. Overall, our trees show similar topology and recovered four of the five main clades demonstrated in previous studies (Clades 1, 3, 4, and 5). Family-level clades were well-resolved within each major clade, but many were discordant with the recently revised classification. Therefore, revision and some reshuffling, especially within and between Clade 1 and 3 are required. I also reconstructed phylogenetic trees of Opisthosporidia to better integrate the evolutionary history of canonical microsporidians in a broader context. I discuss several traits shared only by canonical microsporidians that may have contributed to their striking ecological success in diverse metazoans. More targeted studies on the neglected host groups will be of value for a better understanding of the evolutionary history of these interesting intracellular parasites.

6.2. Introduction

'Canonical (=classical/derived/higher)' microsporidians are a monophyletic group of highly specialized intracellular parasites that infect a wide range of hosts. As the name implies, canonical microsporidians include lineages that were first discovered and ones that share common characteristics with those lineages, differentiating them from 'microsporidia-like organisms' that are morphologically and genetically similar to them but distinct. All canonical microsporidians share several common characteristics including a compact genome (with some variability; see Wadi and Reinke, 2020), highly reduced mitochondria (mitosomes), the presence of ADP/ATP transporters and well-developed polar tubes (Dean et al., 2016; Tamim El Jarkass and Reinke, 2020; Vossbrinck et al., 2014).

The existence of Microsporidia has been known since the 19th century due to their devastating impact on animals of economic importance such as silkworms and fishes (Naegeli, 1957; Pasteur, 1870; Sandholzer et al., 1945). During the last 160 years, more than 1,300 species have been described from >220 genera (Becnel et al., 2014; Franzen, 2008), including at least 17 human-infecting species (Stentiford et al., 2016). Microsporidians are common in arthropod and chordate hosts, but they have been found from almost all animal phyla (Becnel et al., 2014; Snowden, 2014; Stentiford et al., 2013b). Although this is rare, some canonical microsporidians have been found in Ciliophora, which is the only non-metazoan host group known for these parasites (Foissner and Foissner, 1995; Fokin et al., 2008). Canonical microsporidians are extremely divergent and have successfully colonized diverse ecological niches, but several common characteristics are shared by all canonical microsporidians including a compact genome (with some variability; see Wadi and Reinke, 2020), highly reduced mitochondria (mitosomes), and the presence of ADP/ATP transporters and complex polar tubes (Dean et al., 2016; Tamim El Jarkass and Reinke, 2020; Vossbrinck et al., 2014).

Despite their well-recognized diversity, resolving the phylogenetic position of canonical microsporidians within the tree of life has been challenging especially due to fast

evolutionary rates in SSU gene(s) that often cause a long-branch attraction (LBA) problem (Lartillot et al., 2007). Their phylogenetic affinity to Fungi has now been widely accepted (Edlind et al., 1996; Keeling and Doolittle, 1996), but their phylogenetic placement in relation to their close relatives and also with or within Fungi still remain to be fully resolved. About a decade ago, it was shown that a clade containing *Rozella* (parasites of Chytridiomycetes, Blastocladiomycetes, and Oomycetes) and many unidentified environmental sequences formed a monophyletic group closely related to Fungi, and this group has been referred to as Rozellida (Lara et al., 2010). Later, the phylum Rozellomycota (=Cryptomycota) was proposed for this group (Corsaro et al., 2014b; Jones et al., 2011). After the characterization of Aphelida (parasites of algae), the Superphylum 'Opisthosporidia' was proposed to encompass Aphelida, Rozellida, and Microsporidia (so-called ARM clade), which are closely related and branched near the base of the fungal radiation (James et al., 2013; Karpov et al., 2014).

Some Microsporidia-like organisms such as *Paramicrosporidium*, *Mitosporidium*, and *Nucleophaga* have been both morphologically and genetically characterized, providing important insights into the specialization and evolutionary trait reduction within Opisthosporidia (Corsaro et al., 2014a; Galindo et al., 2018; Haag et al., 2014). Bass et al. (2018) proposed the concept of 'expanded Microsporidia' to include all these groups along with canonical Microsporidia, which are in a robust monophyletic sister group to *Rozella*. These Microsporidia-like organisms branched between *Rozella* and canonical Microsporidia, having short branches in the SSU trees, and therefore they have been referred to as short-branch Microsporidia ('SB-Microsporidia') in contrast to canonical Microsporidia which have long branches (therefore canonical Microsporidia were named 'LB-Microsporidia' in Bass et al., 2018). Metchnikovellids and *Chytridiopsis*, which have also long been known, but only recently genetically characterized, were confirmed as the closest relatives of canonical Microsporidia (Corsaro et al., 2019; Galindo et al., 2018; Mikhailov et al., 2017). Bass et al. (2018) included only canonical Microsporidians as LB-Microsporidia, but in this study, we also include Metchnikovellids and *Chytridiopsis* within LB-Microsporidia, thus slightly extending

the inclusion border because they also have considerably long branches in SSU rRNA trees.

The ribosome is essential to all life and therefore core units are still conserved in canonical Microsporidia, even though they are significantly reduced in size and are highly divergent (Bowman et al., 2020; Peyretaillade et al., 1998). Canonical microsporidians have a prokaryote-like ribosomal RNA with a fused LSU-5.8S rRNA (Vossbrinck and Woese, 1986). A comparison between the secondary structure of the rRNA of microsporidians and their relatives shows that extreme reduction occurred only in the lineage of canonical microsporidians (Corsaro et al., 2019). Metchnikovellids and *Chytridiopsis* also have long branches in SSU rRNA trees, but their ribosomal DNAs are similar to that of other eukaryotes in structure and size (Corsaro et al., 2019). Despite the differences in size, these highly conserved orthologous fragments of ribosomal RNA provide valuable information on the phylogenetic relationships among these divergent groups. Within canonical microsporidians, SSU rRNA sequences have been used as primary data for higher classification as morphological characters traditionally used for identification show common homoplasy (Stentiford et al., 2013b; Vossbrinck et al., 2014).

Previous phylogenetic studies based on SSU rRNA sequences have played an important role in the classification of canonical Microsporidia. Vossbrinck and Debrunner-Vossbrinck (2005) constructed neighbour-joining (NJ) and maximum parsimony (MP) trees based on SSU rRNA sequences of 125 species from 56 genera, proposing three classes based on the dominant host habitat of each group: Terresporidia, Aquasporidia, and Marinosporidia. Later, Vossbrinck et al. (2014) inferred a maximum likelihood (ML) and MP trees with an improved taxon sampling (71 species from 63 genera) showing five main clades: Clades 1 to 5. Clades 1 and 3 correspond to Aquasporidia, Clades 2 and 4 correspond to Terresporidia, and Clade 5 corresponds to Marinosporidia. Although ecological heterogeneity across major lineages was demonstrated using environmental sequences (Williams et al., 2018), the 5-clade system has been widely used for the classification of canonical Microsporidia. Indeed, the taxonomy of canonical Microsporidia has recently been revised to accommodate the 5 major clades

(by Tokarev and Issi in Wijayawardene et al., 2020). In the revised classification, orders Neopereziida, Ovavesiculida, and Amblyosporida, were newly established to accommodate Clades 1, 2, and 3, respectively, and orders Nosematida and Glugeida were revised for Clades 4 and 5, respectively.

At present, reporting the SSU rRNA sequence (or a sequence of a longer region of rRNA; SSU-ITS-LSU) is regarded as an essential part of species description (Stentiford et al., 2013b; Vossbrinck et al., 2014), and the number of rRNA sequences in GenBank has been increasing. Here, I make use of an expanded number of rRNA sequences to better understand the phylogeny of canonical microsporidians. For this purpose, I reconstruct the most up-to-date and most comprehensive genus-level phylogeny of canonical Microsporidia with available sequences. Also, I infer the phylogeny of Opisthosporidia to discuss the evolution of canonical Microsporidia within a broader context to provide insights into the origin and diversification of this interesting parasite group. Bayesian and ML trees in this study recovered all five of the major clades except for Clade 2 from Vossbrinck et al. (2014). This result suggests the need for revising the current classification of canonical Microsporidia. However, because our current knowledge of the diversity of microsporidians is far from complete, more major lineages might be uncovered and the relationship among them will be better resolved in the future.

6.3. Materials and Methods

6.3.1. Compiling genetic data

6.3.1.1 Microsporidia

I aimed at reconstructing phylogenetic trees of canonical Microsporidia that are as complete as possible at the genus level. Becnel et al. (2014) listed 200 formally described generic names within canonical Microsporidia. Among these, rRNA sequences (SSU or SSU-ITS-LSU sequences) of 104 genera were available in GenBank (Supplementary Table 6.1). For these genera, I included at least one rRNA sequence per

genus. Although they have not yet been formally described, I also included *Paranucleospora* and *Visvesvaria* in our analyses because rRNA sequences were also available for these genera in GenBank. In addition, rRNA sequences of 20 newly described genera since Becnel et al. (2014) were also included (*Alternosema*, *Apotaspora*, *Cambaraspora*, *Conglomerata*, *Dictyocoela*, *Fibrillaspora*, *Globulispora*, *Hyperspora*, *Jirovecia*, *Myrmecomorba*, *Nematocenator*, *Obruspora*, *Pancytospora*, *Paradoxium*, *Parahepatospora*, *Percutemincola*, *Pseudoberwaldia*, *Pseudokabatana*, *Rugispora*, and *Trichotosporea*; see Supplementary Table 6.1 for references). Finally, rRNA sequences under the names '*Microsporidium* sp.' and 'Microsporidia sp.' were also added to our dataset since many sequences obtained from diverse hosts are provisionally registered under these names, and these may represent distinct and previously unrecognized lineages within the tree of Microsporidia. For these two provisional groups, sequences that were deemed too short (< 500 bp) were excluded and only sequences with known hosts were included for further analyses. Also, highly similar sequences generated by a single study (e.g., many sequences obtained from amphipods from Lake Baikal) were reduced to one or a few representative sequences. Thus, 60 rRNA sequences of *Microsporidium* sp. and 16 sequences of Microsporidia sp. were added to the dataset. As a result, a total of 220 rRNA sequences including 126 described/named canonical microsporidian genera were used for further analyses. In addition, 13 rRNA sequences of close relatives within opisthosporidians were included as outgroups.

6.3.1.2. Opisthosporidia

SSU rRNA sequences were also compiled for the phylogenetic tree of microsporidians and their relatives (Opisthosporidia). Sequences representing canonical Microsporidia, Metchnikovellida, *Chytridiopsis*, *Paramicrosporidium*, *Mitosporidium*, *Nucleophaga*, and some representative groups of environmental sequences of SB-microsporidia in Bass et al. (2018), Aphelida, *Rozella*, and some fungi were included. SSU rRNA sequences of Holozoa and Nucleariidae were also included as outgroups. In total, 94 rRNA SSU sequences were used in our analyses (Supplementary Table 6.2).

6.3.2. Phylogenetic analysis

For canonical microsporidians, both SSU and LSU rRNA sequences were used when the LSU region was available. The ITS region was excluded because it was too divergent across microsporidians. SSU and LSU sequences were aligned respectively with the MAFFT algorithm in Geneious prime (Katoh and Standley, 2013). Removing ambiguous sites can reduce LBA problems (Qu et al., 2017; Ranwez and Chantret, 2020). Therefore, ambiguous sites were eliminated in Gblocks with the least restrictive setting (Castresana, 2000), and then SSU and LSU were concatenated. For the tree of Opisthosporidia, only the SSU region was used for analyses and the same alignment and refining procedures described above were applied. The best-fitting model of nucleotide evolution for each dataset (canonical Microsporidia and Opisthosporidia) was determined based on the corrected Akaike information criterion (AICc) using jModelTest v2.1.6 (Darriba et al., 2012), which was conducted through the CIPRES Science Gateway v3.3 (Miller et al., 2010). The General Time Reversible (GTR) model of nucleotide substitution along with Gamma distributed rate variation across sites (G) and the proportion of invariable sites (I) were used for Bayesian tree inference in MrBayes 3.2.7 (Ronquist et al., 2012). Two independent runs, consisting of four chains each, were simultaneously conducted for 20,000,000 generations with a sampling frequency of 2,000 for canonical Microsporidia, and for 10,000,000 generations with a sampling frequency of 1,000 for Opisthosporidia. The initial 25% of the samples were discarded. Maximum Likelihood trees were reconstructed in RAxML with GTRCAT approximation with 25 rate categories following the developer's recommendation (Stamatakis, 2014). A rapid bootstrap analysis was conducted with 1,000 replicates. The resulting trees were visualized in FigTree v1.4.4 (https://tree.bio.ed.ac.uk/software/figtree/).

6.3.3 Compiling data on host and habitat of canonical microsporidians

Information on host species (and higher taxonomic groups) and habitat of each microsporidian species that were actually used in our analyses was compiled for canonical microsporidians. Notably, some genera include many species that may each infect a different group of hosts and inhabit different habitats (e.g., *Encephalitozoon*). Some species are generalists infecting distantly related taxa (e.g., *Nosema* and

Vairimorpha) or alternate between different groups of hosts during their life cycle (e.g., *Amblyospora*). Because uneven taxon sampling could affect our interpretation of the phylogeny (Heath et al., 2008), only one (or a few) representative sequences per genus were included. Habitats were categorized into 5 groups; Marine (M), Marine-Freshwater (=Brackish; MF), Freshwater (F), Freshwater-Terrestrial (FT), and Terrestrial (T). FT groups include species that spend considerable portions of their life cycle both in freshwater and on land, such as parasites of many dipteran insect species, whose larval stage must develop in freshwater (Becnel and Andreadis, 2014).

Figure 6.1 (next page) A Bayesian phylogenetic tree of canonical microsporidia inferred from concatenated SSU and LSU rRNA gene sequences. Nodes with strong support in both Bayesian (posterior probability > 95) and ML (Bootstrap support >90) analyses were annotated with grey circles (●). Nodes that were strongly supported only in Bayesian analysis were annotated with grey triangles (▲). Major clades of Vossbrinck et al. (2014) and some family names are shown with red and orange colours, respectively. Name of host species and higher taxonomic classification are shown next to the tree. Habitats of host-parasite are categorized into five groups and marked with different colours.

Figure 6.1a A part of the Bayesian tree of canonical microsporidians showing clade 5.

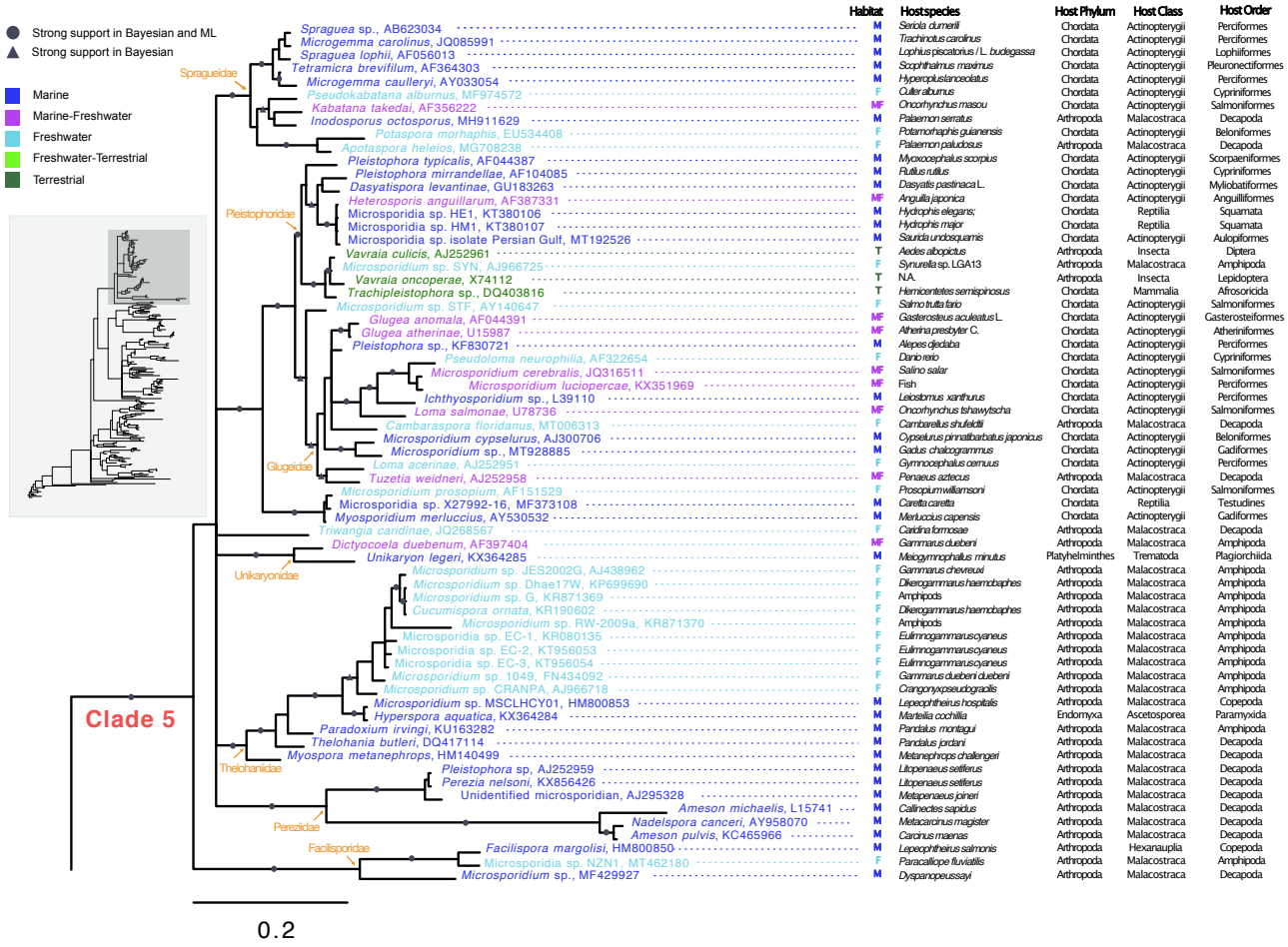


Figure 6.1b A part of the Bayesian tree of canonical microsporidians showing clade 4.

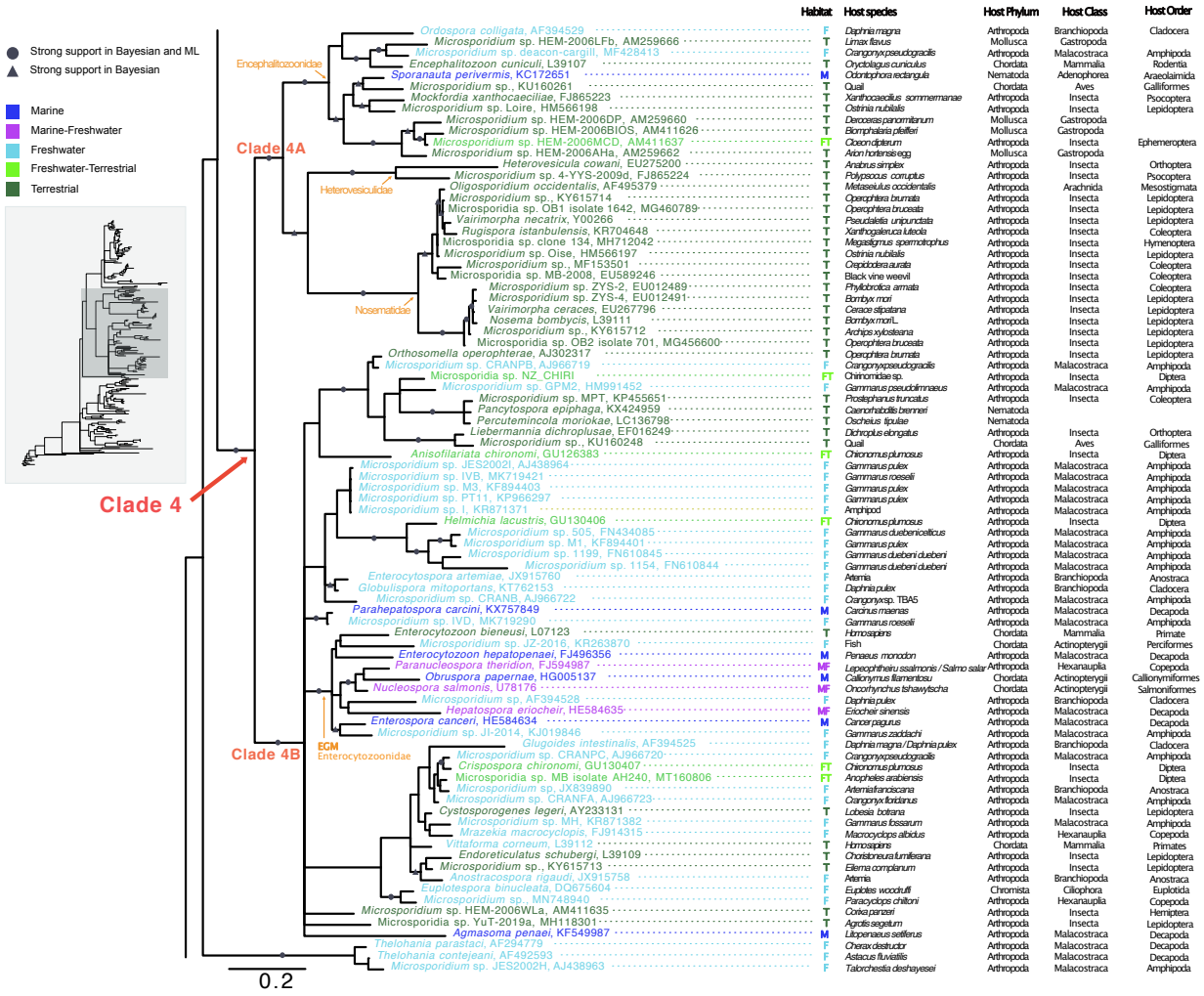
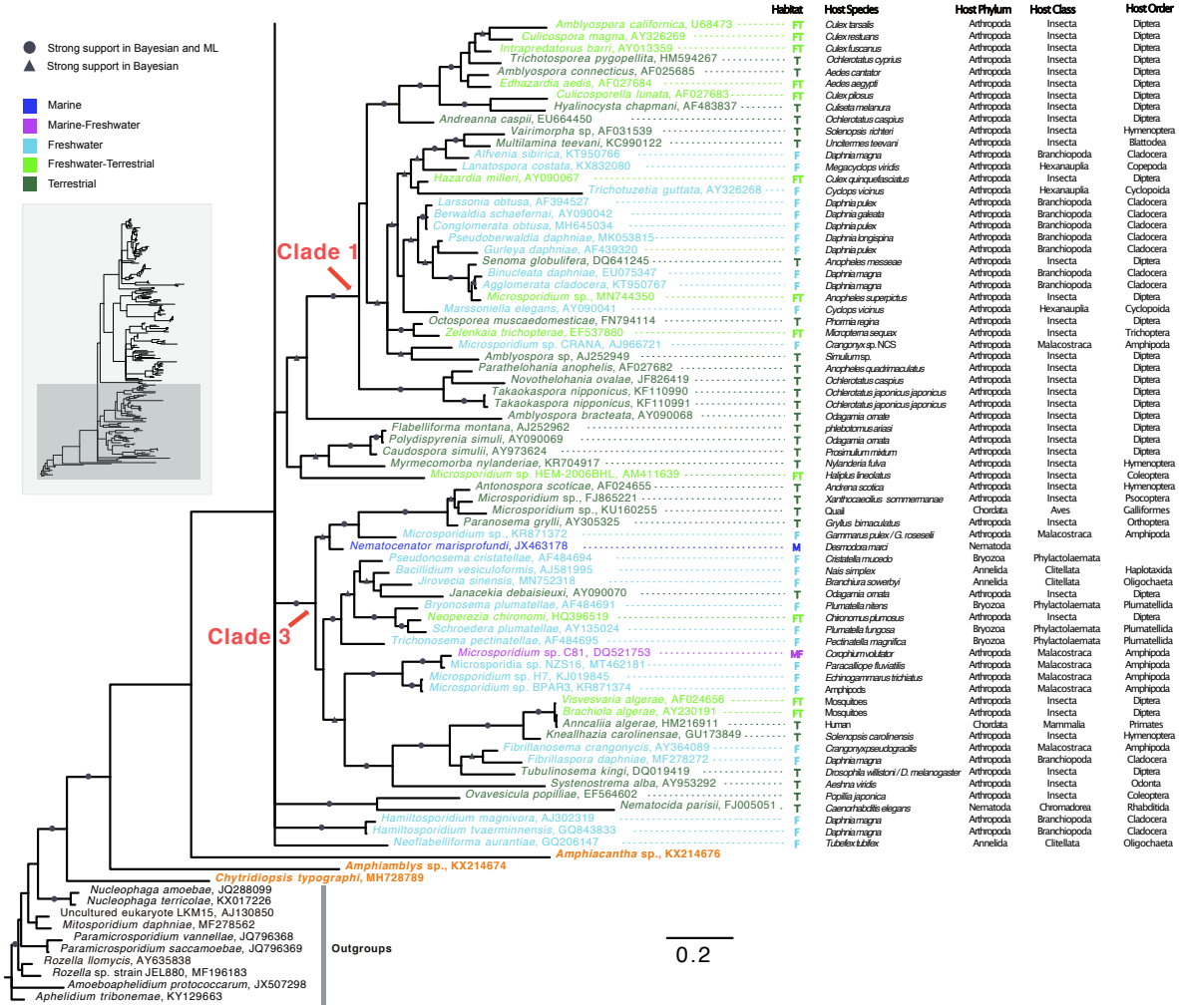


Figure 6.1c A part of the Bayesian tree of canonical microsporidians showing clades 1 and 3.



6.4. Results and discussion

Canonical microsporidians are a highly successful group parasitizing almost all animal phyla. Notably, microsporidians have been detected from Acanthocephala (de Buron et al., 1990), Annelida (Larsson, 1992), Bryozoa (Desser et al., 2004), Cnidaria (Clausen, 2000), Gastrotricha (Manylov, 1999), Kinoryncha (Adrianov and Rybakov, 1992), Mesozoa (Czaker, 1997), Mollusca (Sagristà et al., 1998), Nematoda (Ardila-Garcia and Fast, 2012), Phoronida (Temereva and Sokolova, 2018), Platyhelminthes (Levron et al., 2005), and Rotifera (Gorbunov and Kosova, 2001). Unfortunately, some of these reports were based only on morphological descriptions without SSU rRNA sequences, and therefore those could not be included in our analyses. Based on current knowledge, microsporidians seem to be most common in arthropod and chordate hosts, especially in insects, crustaceans, and fishes. However, this may represent a biased research effort (see below).

6.4.1. The phylogeny of canonical Microsporidia

Our Bayesian and ML trees reconstructed with 220 sequences from 126 named genera represent more than half of the known diversity of canonical microsporidians at the genus level (Figures 6.1a-c; full trees in Supplementary figures 6.1 and 6.2). Overall, both trees were highly congruent and recovered four of the five main clades from Vossbrinck et al. (2014); Clades 1, 3, 4, and 5, although the phylogenetic relationships among them were not well resolved in our analyses. Clade 2 was not recovered as a monophyletic group. Therefore, our Bayesian and ML trees support all the suggested orders except for Ovavesiculida from the recently revised classification (Wijayawardene et al., 2020). Several well-supported family-level clades were identified within each main clade (see figure 6.1).

Clade 5: fish and crustacean hosts from aquatic habitats

Clade 5 comprises mostly aquatic (marine, marine-freshwater, and freshwater) species (Figure 6.1a). Also, this clade includes most of the fish-infecting species. Species infecting diverse crustaceans are also included in this clade. All the subclades correspond well to the established families within the order Glugeida (by Tokarev and

Issi in Wijayawardene et al., 2020). Only families Spragueidae, Pleistophoridae, and Glugeidae include fish-infecting species. Family Unikaryonidae includes *Dictyocoela*, the most common microsporidian genus infecting amphipods globally (Bacela-Spychalska et al., 2018; Drozdova et al., 2020; Park et al., 2020). Families Thelohaniidae, Pereziidae, and Facilisphoridae are largely associated with crustaceans (amphipods, decapods, and copepods). Interestingly, two hyperparasite species (*Hyperspora aquatica* and *Unikaryon legeri*) are included in Clade 5. A possible vectoring role of hyperparasitism in the transmission of microsporidian parasites between crustacean and mollusk hosts has been suggested before (Stentiford et al., 2017).

Clade 4: a large clade containing diverse hosts from all realms

Two robust subclades were identified within Clade 4 (Clades 4A and 4B; Figure 6.1b). Clade 4A consists mostly of species from terrestrial insects (lepidopterans and coleopterans), but also includes some species infecting freshwater crustaceans. The genus *Nosema*, which includes many species that infect economically important insect species (*N. bombycis* in silkworms and *N. ceranae*, *N. bombi*, and *N. apis* in bees), belongs to this clade. Clade 4A does not include species from crustacean or dipteran hosts. On the other hand, Clade 4B includes species found in diverse crustacean hosts including amphipods, decapods, copepods, anostracans, and cladocerans. Also, Clade 4B includes some parasites of dipterans suggesting that some lineages within this clade may be largely associated with freshwater environments. The phylogenetic relationships among lineages within Clade 4B were poorly resolved with some polytomies. Many OTUs belonging to this clade were recovered in a study of environmental sequences, suggesting that Clade 4B may be highly under-sampled (Williams et al., 2018). Notably, the *Enterocytozoon* Group of Microsporidia (EGM) belongs to this clade. This robust monophyletic EGM group includes parasites of many important marine species and human-infecting species; some food- and water-borne microsporidiosis outbreaks were related to this group and possible transmission through the human food chain has been suggested (Stentiford et al., 2019, 2016).

Clades 1 and 3: Parasites with many *freshwater and terrestrial hosts*

Two well-supported clades corresponding to clades 1 and 3 were recovered in our analyses (Figure 6.1c). Species belonging to these clades were mostly obtained from freshwater, freshwater-terrestrial, and terrestrial habitats. Clade 1 includes only species of arthropod hosts (insects and crustaceans), whereas Clade 3 includes species from a broader range of hosts (insects, crustaceans, bryozoans, nematodes, and chordates).

Unrecovered Clade 2, and other lineages branching near the base of the tree of canonical microsporidians

I also identified some minor lineages that diverged near the base of canonical microsporidian radiations, which do not belong to any of the major groups. In fact, these minor lineages were assigned to either Clade 2 or Clade 3 in previous studies (Vossbrinck et al., 2014; Vossbrinck and Debrunner-Vossbrinck, 2005). To be specific, only four species (*Antonospora locustae = Paranosema locustae*, *Antonospora scoticae*, *Nematocida parisii*, *Ovavesicula papillae*) were included in Clade 2 in Vossbrinck et al. (2014). In my analyses, *Paranosema grylli* (96% identical to *A. locustae*) and *A. scoticae* belong to Clade 3, but *N. parisii* and *O. papillae* do not belong to any of the major clades (Figure 6.1C). In fact, Clade 2 was poorly supported in the ML tree (Fig 6.3a in Vossbrinck et al. 2014) in the previous study with a low bootstrap value (BS=72). Several recent studies with genomic scale data, which included both *A. locustae* and *N. parisii*, also did not recover them as a monophyletic group (Galindo et al., 2018; Mikhailov et al., 2017). Overall, clades 1-3 were not strongly supported (i.e. low bootstrap values, or bootstrap values were not provided for most of the nodes) and therefore the borders between major clades were not clearly defined in the previous studies (Vossbrink et al. 2014). For example, *Amblyspora bracteata* and *Caudospora simuli* belonged to Clade 1 in their ML tree, but these two species were not grouped with the rest of Clade 1 in the MP tree (Vossbrinck et al., 2014), which shows unstable phylogenetic positions of some lineages. These two species do not belong to any of the major clades in our analyses (Figure 6.1C). The reason for the formation of Clade 2 in the previous studies (in contrast to ours where it was not recovered) was probably due to a long-branch attraction artifact (Lartillot et al., 2007). Also, our improved taxon

sampling (the inclusion of more sequences similar to those of species in Clade 2) may have helped to break up unnatural groupings.

The need for revision of classification within canonical Microsporidia

The four major clades shown in our study have important implications since the current classification is largely based on the SSU rRNA phylogeny. Our Bayesian and ML trees support all the suggested orders except for Ovavesiculida from the recently revised classification (Wijayawardene et al., 2020). Several well-supported family-level clades were identified within each main clade (see Figure 6.1A-C). Especially, major family-level clades within Clade 5 and Clade 4A correspond well to revised families (Figure 6.1A and B, Wijayawardene et al. 2020). There are also well-defined family-level groups within Clades 1, 3, and 4B (Figure 6.1B and C), but these are largely discordant with the recently revised classification (Wijayawardene et al., 2020), which was mainly based on the previous phylogenetic studies. Therefore, reshuffling of some families and genera within and among the newly established orders Amblyosporida (Clade 1), Neopereziida (Clade 3), and Nosematida (Clade 4) is needed, and Ovavesiculida (Clade 2) should be dissolved.

Another obvious problem with the current taxonomy is the presence of para- and polyphyletic groups such as *Nosema*, *Vairimorpha*, *Plestophora*, and *Amblyospora* (Figure 6.1). This issue has been continuously discussed and has been mostly attributed to classification based on morphological characters (e.g. number of nuclei, the process of spore formation, the spore shape, and the number of polar filament coils) which could change rapidly and are therefore unreliable for classification (Baker et al., 1994; Stentiford et al., 2013a; Vossbrinck et al., 2014). The presence of numerous para- and polyphyletic lineages illustrates that morphological similarities do not necessarily mean evolutionary relatedness but could be due to the convergent evolution of those morphological characters (Vossbrinck et al., 2014). With the increasing use of molecular data for classification, many species within these genera were transferred to another genus or new genera were established (Franzen et al., 2006; Slamovits et al., 2004; Tokarev et al., 2020; Vavra et al., 2006). Likewise, even if two microsporidians are morphologically distinct, this does not necessarily mean that they are

phylogenetically distantly related. In our trees, some genera with distinct morphological and developmental features are grouped, closely related to each other, suggesting that they could be treated as congeneric species (Figure 6.1). These include *Agglomerata-Binucleata-Senoma* (Sokolova et al., 2016), *Larssonia-Berwaldia-Conglomerata* (Vávra et al., 2018), and *Spraguea-Microgemma-Tetramicra* (Casal et al., 2012; Scholz et al., 2017); see referred studies for detailed discussion. Even more strikingly, the extreme polymorphism of a single species shown in *Ameson pulvis*, which has both *Ameson*-like and *Nadelspora*-like lineages infecting the musculature of marine crabs, emphasizes the need for classification primarily based on the molecular data (Stentiford et al., 2013a, 2013b; Vossbrinck et al., 2014).

Because of difficulties in describing species, reporting sequences of newly discovered lineages under the names of Microsporidia sp. or *Microsporidium* sp. is increasingly common. In fact, sequences under these provisional names represent a large portion of diversity within canonical Microsporidia as shown in our tree (Figure 6.1). SSU rRNA sequences can be useful to assign new species into order and family. The use of SSU rRNA along with the additional marker(s), especially fast-evolving ones, could be used for the genus-species level in the future, but what markers can be used needs to be investigated further and general agreement among researchers would be also needed.

Cociversification and host switching

Many canonical microsporidian species are thought to be host specific and are associated with a single host species or related groups (e.g., con-generic or con-familiar species). Hosts and parasites that are intimately associated may show congruent phylogenies, but cophylogenetic patterns are rarely seen because of frequent host switching events (de Vienne et al., 2013). Within Microsporidia, some studies that focused on a specific host-parasite system have revealed patterns of codiversification (i.e., congruent host-parasite phylogenies at macroevolutionary scales). Such systems include *Amblyospora* and other genera in mosquitos, *Nosema* in bees, and *Dictyocoela* in amphipods (Andreadis et al., 2012; Baker et al., 1998; Park et al., 2020; Shafer et al., 2009). These studies also inferred that host switching is more likely among closely related hosts, but also suggested that transmission among distantly related hosts may

also occur. On the other hand, some genera such as *Nosema*, *Encephalitozoon*, and *Enterocytozoon* include species that infect distantly related hosts, suggesting common and frequent host switching. In fact, the most extreme example of transmission among distantly related hosts has been shown in human-infecting species (Stentiford et al., 2019). Interestingly, human-infecting taxa emerged within all the major clades. Among them, the EGM group, which contains the most common human infecting microsporidian species *Enterocytozoon bieneusi*, was not known until the 1980s but now includes pathogens of diverse companion animals and livestock (Stentiford et al., 2019). Considering their fast-evolutionary rate and frequent host switching across distantly related hosts and different habitats, more extensive exploitation of wild animals, habitat destruction, human encroachment into wild habitats, intensive animal farming, and environmental stress may promote these kinds of novel host-parasite associations.

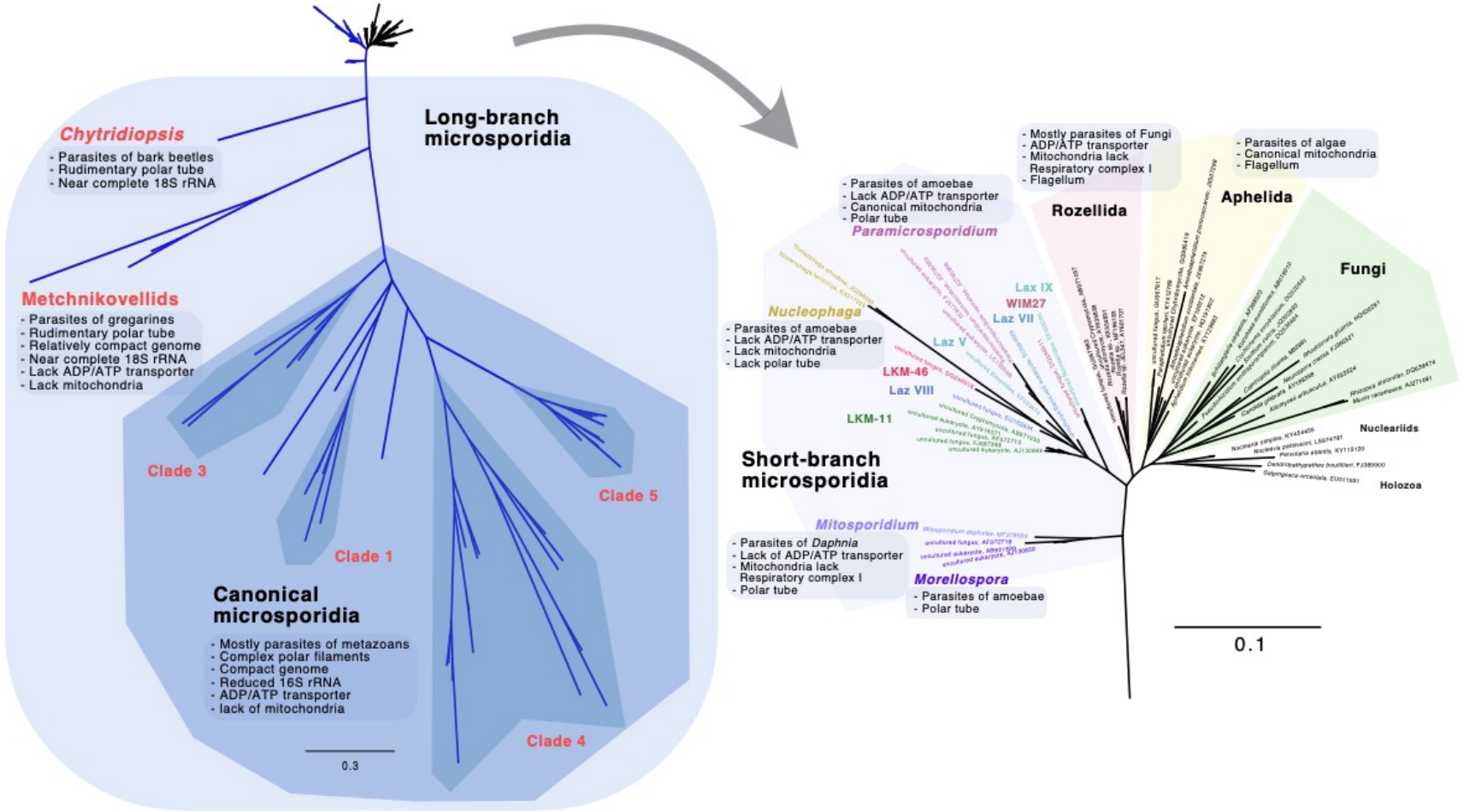
6.4.2. Phylogenetic relationships among Microsporidia and their relatives

An unrooted Bayesian tree of opisthosporidians has been constructed to highlight the genetic distance between canonical microsporidians and SB-Microsporidia and other relatives (Figure 6.2). Both Bayesian and ML trees were similar in overall topology (Figure 6.2 and Supplementary figures 6.3-5). Metchnikovellids and *Chytridiopsis* also have long branches and diverged before the last common ancestor of canonical Microsporidia. The clade containing canonical Microsporidia, Metchnikovellids, and *Chytridiopsis* is strongly supported in both Bayesian and ML trees. Bass et al. (2018) used the term ‘LB-Microsporidia’ for the first time in contrast to ‘SB-Microsporidia’. In their analyses, Metchnikovellids were also shown as sister group to canonical microsporidians, but whether they should be regarded as ‘LB-Microsporida’ was not explicitly discussed. Here, we also include Metchnikovellids and *Chytridiopsis* within LB-Microsporidia, thus slightly extending the inclusion border because they also have considerably long branches in SSU rRNA trees. Although the monophyly of LB-Microsporidia is robust, their sister group is incongruent between the Bayesian and ML trees in this study (Figure 6.2 and Supplementary Figures 6.3 and 6.4). In our Bayesian tree, the LB-Microsporidia clade is sister to a clade of *Mitosporidium* and *Morellospora*, but their sister relationship is poorly supported. Meanwhile,

Nucleophaga was grouped with *Paramicrosporidium* (Figure. 6.2). Although this is consistent with the early morphological observations that '*Nucleophaga* is similar to *Paramicrosporidium* in its infective stage, by having non-flagellated walled spores penetrating amoebae through host cell phagocytosis' (Corsaro et al., 2016, 2014a; Michel et al., 2000), this was also poorly supported. On the other hand, *Nucleophaga* is sister to LB-Microsporidia in the ML tree (Supplementary Figures 6.3 and 6.4), the same as shown in Bass et al. (2018). In both of the Bayesian and ML analyses in this study, the clade containing LB-Microsporidia, SB-Microsporidia, and *Rozellida* is sister to Aphelida, and these ophistosporidian clades are sister to the monophyletic Fungi.

In fact, the phylogenetic placement of SB-Microsporidia is far from stable, probably due to rate heterogeneity among taxa and undersampling. In cases like this, the resulting topology can be greatly affected by taxon sampling, data refinement, and the choice of the model of molecular evolution (Lartillot et al., 2020; Philippe et al., 2011; Ranwez and Chantret, 2020). I acknowledge our imperfect knowledge of these groups. Genome-scale data from additional representative groups within Opisthosporidia and the use of appropriate models of molecular evolution may allow greater resolution in future studies.

Figure 6.2 (next page) A Bayesian tree showing the inferred phylogenetic relationship among microsporidians and their relatives. SSU rRNA gene sequences of opisthosporidians (canonical microsporidia, Metchnikovellids, *Chytridiopsis*, short-branch-Microsporidia, *Paramicrosporidium*, *Mitosporidium*, *Nucleophaga*, Rozellida, Aphelida) were included; these names and other group names of environmental sequences are marked. Some major traits of each group are also shown in grey boxes. (Left) Unrooted tree. Branches of 'Expanded microsporidia' from Bass et al. (2018) are highlighted with blue colour. (Right) Enlarged part of the tree showing 'short-branch microsporidia' and other groups.


Long-branch microsporidia
Chytridiopsis
- Parasites of bark beetles
- Rudimentary polar tube
- Near complete 18S rRNA
Metchnikovellids
- Parasites of gregarines
- Rudimentary polar tube
- Relatively compact genome
- Near complete 18S rRNA
- Lack ADP/ATP transporter
- Lack mitochondria
Clade 3
Clade 1
Clade 5
Clade 4
Canonical microsporidia
- Mostly parasites of metazoans
- Complex polar filaments
- Compact genome
- Reduced 16S rRNA
- ADP/ATP transporter
- lack of mitochondria
0.3
- Mostly parasites of Fungi
- ADP/ATP transporter
- Mitochondria lack Respiratory complex I
- Flagellum
- Parasites of algae
- Canonical mitochondria
- Flagellum
Aphelida
Rozellida
Fungi
- Parasites of amoebae
- Lack ADP/ATP transporter
- Canonical mitochondria
- Polar tube
Paramicrosporidium
Lax IX
WIM27
Laz VII
Nucleophaga
- Parasites of amoebae
- Lack ADP/ATP transporter
- Lack mitochondria
- Lack polar tube
Laz V
LKM-46
Laz VIII
LKM-11
Short-branch microsporidia
Nucleariids
Holozoa
Mitosporidium
- Parasites of Daphnia
- Lack of ADP/ATP transporter
- Mitochondria lack Respiratory complex I
- Polar tube
Morellospora
- Parasites of amoebae
- Polar tube
0.1

6.4.3. What differentiates canonical microsporidians from their relatives?

Although the higher classification of microsporidians has been changing dramatically for reasons similar to those mentioned above, canonical microsporidians show distinct characteristics both morphologically and genetically, making them distinct from other groups. Many comparisons of important traits among microsporidians and their relatives have been conducted recently; these traits include energy metabolism (Timofeev et al., 2020), mechanisms of host invasion, proliferation, and exit (Tamim El Jarkass and Reinke, 2020), the structure of rRNA (Corsaro et al., 2019), and genome architecture (Wadi and Reinke, 2020). These studies highlight that traits important for these intracellular parasites have become specialized or reduced in each lineage differently. The innovative traits that canonical microsporidians acquired are complex polar tubes and ADP/ATP transporters (Alexander et al., 2016; Tsaousis et al., 2008; Vávra and Larsson, 2014). Microsporidians have lost many genes involved with DNA repair pathways, which partly explains the accelerated evolutionary rates in canonical microsporidians (Galindo et al., 2018). Traits present only in canonical microsporidians may have contributed to their successful colonization of diverse metazoans, by promoting efficient host invasion, proliferation, and adaptation to diverse ecological niches (=wide range of hosts in various habitats).

Although other opisthosporidians shared several traits with canonical microsporidians, none of them have all the traits described above. To be specific, the presence of a well-developed polar tube (commonly consists of three sections including a straight part), which allows efficient penetration of host cells upon infection, is a defining character of canonical microsporidians, although the length and thickness vary among different species (Vávra and Larsson, 2014). Metchnikovellids, sister to canonical Microsporidia, have a rudimentary polar tube (short and thick, without the straight part) and for this reason, they used to be called ‘primitive’ Microsporidia (Larsson, 2014). *Paramicrosporidium* and *Nucleophaga* invade the host cell through host phagocytosis even though *Paramicrosporidium* has a polar filament (Corsaro et al., 2014a; Scheid, 2007). All canonical Microsporidia lack canonical mitochondria, which generate ATP (Tsaousis et al., 2008). Microsporidians proliferate (produce spores to complete their life cycle) within the host cell and this is an energy-consuming process (Tamim El

Jarkass and Reinke, 2020). Canonical Microsporidia obtain ATP using ADP/ATP transporters, which are believed to have been obtained horizontally from bacteria (Tsaousis et al., 2008). The presence of ADP/ATP transporters was also identified in *Rozella* (Heinz et al., 2014). LB-Microsporidia have long branches in both the SSU tree and genome tree, which means elevated rates of molecular evolution. However, metchnikovellids did not go through genome reduction as extensively as canonical microsporidians did. It is believed that the reduction of regulatory genes and non-coding regions resulted in rapid evolutionary rates (Galindo et al., 2018). This accelerated evolutionary rates may have produced diverse traits that promoted adaptation to diverse hosts and niches. Although other opisthosporidians share several traits with canonical microsporidians, none of them have all the traits described above.

6.4.4. Canonical microsporidians in metazoan hosts: due to evolutionary adaptation or biased screening effort?

Canonical microsporidians have successfully colonized metazoan hosts. In fact, canonical microsporidians are almost exclusive to metazoan hosts. It was often said that microsporidians could infect amoebae, but these were later identified as *Paramicrosporidium*, which belongs to SB-Microsporidia (Corsaro et al., 2014b; Michel et al., 2000; Scheid, 2007). On the other hand, no opisthosporidian has been detected from metazoan hosts except for canonical microsporidians, *Chytridiopsis*, and *Mitosporidium*. Berbee et al. (2017) suggested that the divergence among Aphelida, Rozellida, and Microsporidia precedes the major diversification of multicellular organisms based on the fact that Aphelida and Rozellida lack the ability for effective intrusion and ramification to multiple cells. Although canonical Microsporidia form a monophyletic group with SB-microsporidians, and both share certain morphological similarities, the genomes of characterized SB-microsporidians are much more similar to those of *Rozella* and canonical Fungi (Haag et al., 2014; Quandt et al., 2017). Also, considering the fact that most SB-microsporidians parasitize amoebae, the divergence between LB-Microsporidia and SB-Microsporidia may have occurred a long time ago. Multiple lineages of SB-Microsporidia have recently been discovered with environmental sequencing (Bass et al., 2018; Lacerda et al., 2020). According to a recent network analysis investigating potential hosts of SB-Microsporidia, it has been

suggested that SB-microsporidians may be associated with Apicomplexa, Cercozoa and Fungi, as well as some Metazoa (Doliwa et al., 2020). Co-occurrence does not necessarily mean an actual host-parasite relationship, but we are starting to learn more about these long unknown groups. SB-microsporidians are far less known than canonical microsporidians to date, but their actual diversity may be very high. LB- and SB-microsporidians diverged from a common ancestor (probably a long time ago) and may have adapted to different host groups with different specialization and reduction of traits. However, without more occurrence data and screening efforts from a wide range of hosts and habitats, we cannot exclude the possibility that microsporidia-like organisms were simply not characterised or detected within metazoan hosts. Also, canonical microsporidians may be common in microscopic hosts.

Undoubtedly, further genetic characterization of microsporidians and their relatives from underexplored host groups, environments, or new geographic areas that have never been explored will provide valuable insights into the evolutionary history of these extremely diverse groups. These could result in adding more major groups to the tree that we are presenting here. Although our trees show four major clades, this does not mean that there are only four major clades within canonical Microsporidia. Also, targeted studies on microscopic hosts may recover the hidden diversity of canonical microsporidians as well as SB-microsporidians. Notably, primers targeting canonical microsporidians do not amplify SB-microsporidians, and vice versa (Williams et al., 2018). Considering these factors, the use of different primer sets will reveal more diversity.

6.5. Conclusion

In this study, I inferred the phylogeny of canonical Microsporidia with an improved dataset and method. The discrepancy between our trees and those from previous studies highlight our imperfect knowledge of the diversity of these parasites. Microsporidians provide an excellent system to study host-parasite associations from the cellular to the ecosystem level. Although it is only a short fragment of the genome, the SSU rRNA region is still useful for species identification and classification at the family-order level, and for detecting microsporidians from unknown hosts and environment. In addition to SSU rRNA, genomic data could be also used for resolving relationships among families and orders, and fast-evolving genes for genus-species level classification. Also, we emphasize that canonical microsporidians are distinct from 'Microsporidia-like organisms' or the rest of the 'expanded Microsporidia' despite some morphological similarities and genetic affinities among them. Only more data (genetic, morphological, and ecological) would fill our knowledge gap and provide insights into the evolutionary relationships among these extremely diverse and successful intracellular parasites.

Chapter 7

General Conclusion

All species interact with other species, and these interactions can generate occurrence and distribution patterns over space and time (Elith and Leathwick, 2009). Therefore, by using the information on the distribution and diversity of interacting taxa, we can infer the strength and evolutionary duration of the interactions between them (Clayton et al., 2015). In this thesis, I focused on two groups of obligate intracellular parasites, Microsporidia and *Rickettsia*, in diverse amphipod hosts in New Zealand. Considering the mode of transmission of both parasites, host ecology, and the geographical settings of New Zealand, I asked several questions regarding the diversity, ecological interactions among parasites and their hosts, and the evolutionary history of parasites as well as host-parasite associations.

This study started with molecular screening, which uncovered a previously unknown diversity of both parasite groups in New Zealand. By screening diverse endemic amphipod hosts for the presence of both parasite groups using group-specific primers, I revealed the widespread distribution and the phylogenetic diversity of both groups in New Zealand amphipod hosts (**Chapters 2 and 3**). The presence of these parasites in this new geographical region (for Microsporidia) and new host taxa (for *Rickettsia*) provided useful information for further investigations. **Chapter 2** confirmed that *Dictyocoela* is the most common microsporidian genus in amphipod hosts. Intimate associations between *Dictyocoela* and their amphipod hosts were inferred based on their congruent phylogenies at both regional and global scales. I emphasised that cophylogenetic patterns may have been determined by various processes operating at multiple scales. In **Chapter 3**, I focused on the Torix group of *Rickettsia* whose presence was confirmed for the first time in amphipod hosts in this study. Torix *Rickettsia* are relatively less known than other groups because they have been discovered more recently; they often cause low pathogenicity (Weinert, 2015). It is surprising that COI sequences of Torix *Rickettsia* have been unintentionally sequenced from diverse groups of hosts from many different countries around the world; this suggests that they are widely distributed. I used these 'mislabelled' sequences deposited in public databases to resolve the phylogeny of Torix *Rickettsia*. My findings illustrate that the unrecognised diversity of many endosymbionts may be greater than what is currently known. Whereas **Chapter 3** dealt with untargeted amplification of rickettsial

sequences as a direct result of infection, **Chapter 4** dealt with indirect and possibly profound impacts of infection. It is well known that mtDNA sequences are often non-neutral and do not necessarily reflect the evolutionary history of organisms (Ballard and Whitlock, 2004). Mitonuclear discordance is a commonly observed pattern across diverse groups of organisms. Despite an increasing number of reports of this phenomenon, only a small number of studies have investigated the possible causes (Després, 2019; Streicher and Day, 2020). In this chapter, I focused on a special case of mitonuclear discordance arising from deep time scales and suggested several possible underlying mechanisms. In **Chapter 5**, I quantified the spatiotemporal patterns in the prevalence of Microsporidia and *Rickettsia*, and tested if they tend to co-occur in the same host individual more or less frequently than expected by chance alone. Finally, **Chapter 6** provides the most up-to-date phylogeny of canonical microsporidia. Here, I highlighted several unique characteristics of canonical microsporidians which differentiate them from their close relatives.

Most of the chapters in this thesis are about the patterns and causes of '**co-**'s: '**co**phylogeny', '**co**amplification', '**co**transmission', and '**co**occurrence'. Indeed, cophylogenetic patterns between *Dictyocoela* (Microsporidia) and their amphipod hosts were observed at both regional and global scales. Cotransmission of *Rickettsia* and mtDNA through vertical transmission were shown to cause direct and indirect problems in DNA barcoding, phylogenetic, phylogeographical, and cophylogenetic studies. Although they coexist in the same habitats and host populations, there was no pattern of non-random cooccurrence between Microsporidia and *Rickettsia* in the same host individual. Below, I discuss some remaining questions that show much promise for future studies.

Molecular screening has proven to be useful in detecting microscopic parasites (Grabner, 2017; Stensvold and Nielsen, 2012). This thesis was devoted to uncovering unknown parasite diversity by using molecular tools. Without these tools, none of the studies in this thesis would have been possible. Many endosymbionts and intracellular parasites do not cause severe pathology, therefore detecting them by eye, even with most standard microscopic approaches, is almost impossible. Next-generation

sequencing (NGS) allows us to obtain a large number of sequences at relatively low cost and less time compared to Sanger sequencing. Although Sanger sequencing has its own merits, using NGS can help uncover a greater diversity of any taxa of interest. For example, many new lineages of microsporidians were discovered by using NGS (Williams et al., 2018). Similarly, by using NGS, more thorough and comprehensive screening of endosymbiont diversity in New Zealand could be achieved in the future.

With some screening efforts, we can reveal and document the diversity of organisms largely unknown from traditional detection methods (for both hosts and parasites). Although the diversity of *Dictyocoela* in amphipod hosts and their ecological impacts on hosts are well known in Europe, the presence of *Dictyocoela* still remains largely unknown in the Southern Hemisphere. The presence or absence of, and the genetic information on *Dictyocoela* from some amphipod groups, should be especially of value. For example, I suggested the ancient origin of *Dictyocoela* and that their intimate association with amphipod hosts may have originated a long time ago. However, we lack data from other regions in the Southern Hemisphere. Specifically, screening endemic amphipods from other areas that used to be part of Gondwana could provide strong evidence for the 'co-vicariance' scenario that I proposed in Chapter 2. In addition, screening cosmopolitan marine species for microsporidian parasites could also provide valuable information regarding their possible vectoring roles in transmitting microsporidian parasites between landmasses. Although marine species comprise ~80% of amphipod diversity, there is no record of *Dictyocoela* from marine amphipods yet. Again, occurrence data and genetic information should provide valuable insights into the transmission and evolutionary history of *Dictyocoela* in amphipod hosts.

The same logic applies to *Rickettsia*. Many lineages of *Rickettsia* show little pathogenic effects on their hosts. Some are even believed to be beneficial (Kikuchi and Fukatsu, 2005). Compared to other groups, Torix *Rickettsia* were discovered relatively recently (Kikuchi et al. 2002). However, a large number of 'mislabelled' sequences in the GenBank and BOLD databases suggest a wide host range (Pilgrim et al. 2020; Chapter 3). Targeted screening is highly likely to uncover a hidden diversity of *Rickettsia* in new host groups. Only then can their interactions with hosts be better understood.

Despite the traditional belief that 'parasite phylogeny mirrors that of its host' (Fahrenholz, 1913), perfectly congruent phylogenies are very rare (de Vienne et al., 2013). It has been highlighted that the link between coevolution and cophylogeny is weak at best (Althoff et al., 2014). There are still many missing connections in our understanding of the processes shaping these patterns. In this regard, microsporidians can be an interesting system to investigate both coevolution and cophylogeny and the elements connecting them. Recently, various scales of research have been explored to better understand the biology and diversity of Microsporidia and their relatives. For example, we now better understand how Microsporidia interact with their hosts at a cellular level. We have learnt that microsporidians and hosts exert selective pressures on each other; that is, they are involved in a coevolutionary arms race. Microsporidians have wide host ranges across almost all animal taxa. A few studies, including Chapter 2, have shown host-parasite cophylogenetic patterns (Andreadis et al., 2012; Baker et al., 1998; Shafer et al., 2009). By using these host-parasite associations, the linkage between microevolutionary processes and macroevolutionary patterns can be more holistically understood.

Unwanted amplification of the nucleotide sequences of untargeted taxa is commonplace in PCRs (Prince and Andrus, 1992). However, instead of discarding these data, compiling them with host and habitat information may provide useful and interesting information on the known and unknown diversity of microorganisms (most of which are likely viruses and bacteria). I detected rickettsial sequences using universal COI primers. The use of blocking primers effectively prevented the amplification of Rickettsia sequences (Chapter 3), but I obtained some sequences of other alphaproteobacteria or oomycetes. This suggests that amphipods (and other hosts) may harbour diverse, but as yet undiscovered, endosymbionts.

Although reports of mitonuclear discordance have rapidly increased in the recent literature, the causes behind these patterns are still poorly understood. Identifying the actual causes of mitonuclear discordance patterns can be very challenging and demanding (Streicher and Day, 2020; Weiss et al., 2018). However, investigating the

possible causes can be highly valuable. Although we provided several possible causes of mitonuclear discordance patterns in *Paracalliope* hosts by focusing on several aspects in Chapter 4 (phylogeographical histories, the presence of endosymbionts, and some host intrinsic factors), other explanations are also possible. Among them, the 'selfish mitochondria' hypothesis is an interesting one (Perlman et al., 2006). In the absence of endosymbionts, highly divergent lineages of mtDNA are often found within a host population; this may be explained by the differences in nature between mtDNA and nuDNA. Mitonuclear discordance patterns in deep time scales are common in other groups of crustaceans such as copepods. Future comparative studies on such groups including copepods and amphipods may provide interesting insights on what could be common underlying causes.

In this thesis, I found interesting 'patterns'. However, many questions remain unanswered. How different individual hosts and parasites interact within the same habitat is still not yet well understood. For example, we still do not know the relative contribution of vertical versus horizontal transmission to the overall success of *Dictyocoela* species in *Paracalliope* hosts. The importance of each mode of transmission could be confirmed in a laboratory setting by quantifying the presence of parasites within host eggs. Also, it has been suggested that microsporidians are not the only feminisers in amphipod hosts (Ironside and Alexander, 2015; Short et al., 2012). Future studies are needed to confirm the occurrence of reproduction manipulation induced by symbionts in *Paracalliope*; subsequently, detailed experiments should be able to differentiate between reproductive manipulation exerted by different endosymbionts.

'Holobionts', which consist of a host and the full set of its microbial symbionts, can function as integrated biological units (Roughgarden et al., 2018; Thompson et al., 2014). In addition to the host itself, holobionts consist of diverse bacteria, archaea, fungi, viruses, and protists (Rohwer et al., 2002). A holobiont is not a static assemblage but continuously changes under the influence of various biotic and abiotic factors (McFall-Ngai et al., 2013; Pita et al., 2018). How each lineage of endosymbionts interacts with their host, and how these lineages interact with each other, not only has

immediate functional implications for the holobiont, but can also lead to various ecological and evolutionary consequences. In this thesis, I investigated several aspects of only two groups of endosymbionts of amphipod hosts. A deeper understanding of specific host-endosymbiont interactions, encompassing a broader range of symbionts, should slowly but surely broaden our understanding of the more complex networks of holobionts.

References

Adrianov, A. V., Rybakov, A. V., 1992. *Kinorhynchospora japonica* gen. n., sp.n. (Microsporidia) from the intestinal epithelium of *Kinorhynchus yushini* (Homalorhagida, Pycnophyidae) from the Sea of Japan. Hydrobiol. J. 28, 52–58.

Alexander, W.G., Wisecaver, J.H., Rokas, A., Hittinger, C.T., 2016. Horizontally acquired genes in early-diverging pathogenic fungi enable the use of host nucleosides and nucleotides. Proc. Natl. Acad. Sci. U. S. A. 113, 4116–4121. https://doi.org/10.1073/pnas.1517242113

Alperi, A., Figueras, M.J., Inza, I., Martínez-Murcia, A.J., 2008. Analysis of 16S rRNA gene mutations in a subset of *Aeromonas* strains and their impact in species delineation. Int. Microbiol. 11, 185–194. https://doi.org/10.2436/20.1501.01.59

Althoff, D.M., Segraves, K.A., Johnson, M.T.J., 2014. Testing for coevolutionary diversification: Linking pattern with process. Trends Ecol. Evol. 29, 82-89. https://doi.org/10.1016/j.tree.2013.11.003

Anderson, R., May, R., 1981. The population dynamics of microparasites and their invertebrate hosts. Philos. Trans. R. Soc. London. B, Biol. Sci. 291, 451–524. https://doi.org/10.1098/rstb.1981.0005

Anderson, T.K., Sukhdeo, M.V.K., 2010. Abiotic versus biotic hierarchies in the assembly of parasite populations. Parasitology 137, 743–754. https://doi.org/10.1017/S0031182009991430

Andersson, S.G.E., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C.M., Podowski, R.M., Näslund, A.K., Eriksson, A.S., Winkler, H.H., Kurland, C.G., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133–140. https://doi.org/10.1038/24094

Andreadis, T.G., Simakova, A. V., Vossbrinck, C.R., Shepard, J.J., Yurchenko, Y.A., 2012. Ultrastructural characterization and comparative phylogenetic analysis of new microsporidia from Siberian mosquitoes: Evidence for coevolution and host switching. J. Invertebr. Pathol. 109, 59–75. https://doi.org/10.1016/j.jip.2011.09.011

Apte, S., Smith, P.J., Wallis, G.P., 2007. Mitochondrial phylogeography of New Zealand freshwater crayfishes, *Paranephrops* spp. Mol. Ecol. 16, 1897–1908. https://doi.org/10.1111/j.1365-294X.2007.03224.x

Araújo, M.B., Rozenfeld, A., 2014. The geographic scaling of biotic interactions. Ecography (Cop.). 37, 406–415. https://doi.org/10.1111/j.1600-0587.2013.00643.x

Archibald, J.M., 2015. Endosymbiosis and eukaryotic cell evolution. Curr. Biol. 25, R911-R921. https://doi.org/10.1016/j.cub.2015.07.055

Ardila-Garcia, A.M., Fast, N.M., 2012. Microsporidian infection in a free-living marine nematode. Eukaryot. Cell 11, 1544–1551. https://doi.org/10.1128/EC.00228-12

Arditi, R., Ginzburg, L.R., 1989. Coupling in predator-prey dynamics: Ratio-dependence. J. Theor. Biol. 139, 311–326. https://doi.org/10.1016/S0022-5193(89)80211-5

Arfianti, T., Wilson, S., Costello, M.J., 2018. Progress in the discovery of amphipod crustaceans. PeerJ 2018, 6, e5187. https://doi.org/10.7717/peerj.5187

Arneberg, P., Skorping, A., Grenfell, B., Read, A.F., 1998. Host densities as determinants of abundance in parasite communities. Proc. R. Soc. B Biol. Sci. 265, 1283–1289. https://doi.org/10.1098/rspb.1998.0431

Arundell, K., Dunn, A., Alexander, J., Shearman, R., Archer, N., Ironside, J.E., 2015. Enemy release and genetic founder effects in invasive killer shrimp populations of Great Britain. Biol. Invasions 17, 1439–1451. https://doi.org/10.1007/s10530-014-0806-y

Asochakov, I.A., 1994. Technique for measuring body length of amphipods. Hydrobiol. J. 30, 107–110.

Azad, A.F., Beard, C.B., 1998. Rickettsial pathogens and their arthropod vectors. Emerg. Infect. Dis. 4, 179–186. https://doi.org/10.3201/eid0402.980205

Bacela-Spychalska, K., Wróblewski, P., Mamos, T., Grabowski, M., Rigaud, T., Wattier, R., Rewicz, T., Konopacka, A., Ovcharenko, M., 2018. Europe-wide reassessment of *Dictyocoela* (Microsporidia) infecting native and invasive amphipods (Crustacea): Molecular versus ultrastructural traits. Sci. Rep. 8, 1–16. https://doi.org/10.1038/s41598-018-26879-3

Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.L., Hahn, M.W., Kitano, J., Mayrose, I., Ming, R., Perrin, N., Ross, L., Valenzuela, N., Vamosi, J.C., Mank, J.E., Peichel, C.L., Ashman, T.L., Blackmon, H., Goldberg, E.E., Hahn, M.W., Kirkpatrick, M., Kitano, J., Mayrose, I., Ming, R., Pennell, M.W., Perrin, N., Valenzuela, N., Vamosi, J.C., 2014. Sex determination: Why so many ways of doing it? PLoS Biol. 12, 1–13. https://doi.org/10.1371/journal.pbio.1001899

Bailly-Bechet, M., Martins-Simões, P., Szöllosi, G.J., Mialdea, G., Sagot, M.F., Charlat, S., 2017. How long does *Wolbachia* remain on board? Mol. Biol. Evol. 34, 1183–1193. https://doi.org/10.1093/molbev/msx073

Baker, M.D., Vossbrinck, C.R., Becnel, J.J., Andreadis, T.G., 1998. Phylogeny of *Amblyospora* (Microsporida: Amblyosporidae) and related genera based on small subunit ribosomal DNA Data: A possible example of host parasite cospeciation. J. Invertebr. Pathol. 71, 199–206. https://doi.org/10.1006/jipa.1997.4725

Baker, M.D., Vossbrinck, C.R., Maddox, J. V., Undeen, A.H., 1994. Phylogenetic relationships among *Vairimorpha* and *Nosema* species (Microspora) based on ribosomal RNA sequence data. J. Invertebr. Pathol. 64, 100–106. https://doi.org/10.1006/jipa.1994.1077

Balbiani, G., 1882. Sur les microsporidies ou psorospermies des articules. Comptes rendus l'Académie des Sci. 95, 1168–1171.

Balbuena, J.A., Míguez-Lozano, R., Blasco-Costa, I., 2013. PACo: a novel procrustes application to cophylogenetic analysis. PLoS One 8, e61048. https://doi.org/10.1371/journal.pone.0061048

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x

Barnard, Jerry Laurens, 1974. Gammaridean Amphipoda of Australia, Part II. Smithson. Contrib. to Zool. 1–148. https://doi.org/10.5479/si.00810282.139

Barnard, J. L., 1974. Gammaridean Amphipoda of Australia, Part II. Smithson. Contrib. to Zool. 1–148. https://doi.org/10.5479/si.00810282.139

Barnard, J.L., Barnard, C., 1983. Freshwater Amphipoda of the world. No. C/595.371 B3. Hayfield Associates, Virginia.

Barthel, A., Gourinat, A.C., Cazorla, C., Joubert, C., Dupont-Rouzeyrol, M., Descloux, E., 2013. Breast milk as a possible route of vertical transmission of dengue virus? Clin. Infect. Dis. 57, 415–417. https://doi.org/10.1093/cid/cit227

Bass, D., Czech, L., Williams, B.A.P., Berney, C., Dunthorn, M., Mahé, F., Torruella,

G., Stentiford, G.D., Williams, T.A., 2018. Clarifying the Relationships between Microsporidia and Cryptomycota. J. Eukaryot. Microbiol. 65, 773–782. https://doi.org/10.1111/jeu.12519
Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67. https://doi.org/10.18637/jss.v067.i01
Becnel, J.J., Andreadis, T.G., 2014. Microsporidia in Insects, in: Microsporidia: Pathogens of opportunity: First Edition. pp. 521–570. https://doi.org/10.1002/9781118395264.ch21
Becnel, J.J., Takvorian, P.M., Cali, A., 2014. Checklist of available generic names for microsporidia with type species and type hosts, in: Weiss, L.M., James J. Becnel (Eds.), Microsporidia: Pathogens of Opportunity. John Wiley & Sons, Inc., pp. 671–686. https://doi.org/10.1002/9781118395264.app1
Bella, M.C., Fish, J.D., 1996. Fecundity and seasonal changes in reproductive output of females of the gravel beach amphipod, *Pectenogammarus Planicrurus*. J. Mar. Biol. Assoc. United Kingdom 76, 37–55.
Bellamy, D.J., 1990. Moa's ark: The voyage of New Zealand. Penguin Books. United Kingdom
Berbee, M.L., James, T.Y., Strullu-Derrien, C., 2017. Early diverging fungi: Diversity and impact at the dawn of terrestrial life. Annu. Rev. Microbiol. 71, 41–60. https://doi.org/10.1146/annurev-micro-030117-020324
Bernardo, P.H., Sánchez-Ramírez, S., Sánchez-Pacheco, S.J., Álvarez-Castañeda, S.T., Aguilera-Miller, E.F., Mendez-de la Cruz, F.R., Murphy, R.W., 2019. Extreme mito-nuclear discordance in a peninsular lizard: the role of drift, selection, and climate. Heredity (Edinb). 123, 359–370. https://doi.org/10.1038/s41437-019-0204-4
Bethel, W.M., Holmes, J.C., 1977. Increased vulnerability of amphipods to predation owing to altered behavior induced by larval acanthocephalans. Can. J. Zool. 55, 110–115. https://doi.org/10.1139/z77-013
Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2006.11.004
Binetruy, F., Buysse, M., Barosi, R., Duron, O., 2020. Novel *Rickettsia* genotypes in ticks in French Guiana, South America. Sci. Rep. 10. https://doi.org/10.1038/s41598-020-59488-0
Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141, 1007–1014.
Boessenkool, S., Epp, L.S., Haile, J., Bellemain, E., Edwards, M., Coissac, E., Willerslev, E., Brochmann, C., 2012. Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. Mol. Ecol. 21, 1806–1815. https://doi.org/10.1111/j.1365-294X.2011.05306.x
Bojko, J., Bącela-Spychalska, K., Stebbing, P.D., Dunn, A.M., Grabowski, M., Rachalewski, M., Stentiford, G.D., 2017. Parasites, pathogens and commensals in the "low-impact" non-native amphipod host *Gammarus roeselii*. Parasites and Vectors 10. https://doi.org/10.1186/s13071-017-2108-6
Bojko, J., Ovcharenko, M., 2019. Pathogens and other symbionts of the Amphipoda: taxonomic diversity and pathological significance. Dis. Aquat. Organ. 136, 3–36. https://doi.org/10.3354/dao03321
Bouchon, D., Rigaud, T., Juchault, P., 1998. Evidence for widespread *Wolbachia*

infection in isopod crustaceans : molecular identification and host feminization Evidence for widespread *Wolbachia* infection in isopod crustaceans : molecular identification and host feminization. https://doi.org/10.1098/rspb.1998.0402
Bourtzis, K., Nirgianaki, A., Markakis, G., Savakis, C., 1996. *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. Genetics 144, 1063–1073.
Bowman, J.C., Petrov, A.S., Frenkel-Pinter, M., Penev, P.I., Williams, L.D., 2020. Root of the Tree: The Significance, Evolution, and Origins of the Ribosome. Chem. Rev. https://doi.org/10.1021/acs.chemrev.9b00742
Bradford, T., Adams, M., Humphreys, W.F., Austin, A.D., Cooper, S.J.B., 2010. DNA barcoding of stygofauna uncovers cryptic amphipod diversity in a calcrete aquifer in Western Australia's arid zone. Mol. Ecol. Resour. 10, 41–50. https://doi.org/10.1111/j.1755-0998.2009.02706.x
Breinholt, J.W., Kawahara, A.Y., 2013. Phylotranscriptomics: Saturated third codon positions radically influence the estimation of trees based on next-gen data. Genome Biol. Evol. 5, 2082–2092. https://doi.org/10.1093/gbe/evt157
Brennan, I.G., Bauer, A.M., Jackman, T.R., 2016. Mitochondrial introgression via ancient hybridization, and systematics of the Australian endemic pygopodid gecko genus *Delma*. Mol. Phylogenet. Evol. 94, 577–590. https://doi.org/10.1016/j.ympev.2015.10.005
Brown, G.C., 1992. Control of respiration and ATP synthesis in mammalian mitochondria and cells. Biochem. J. https://doi.org/10.1042/bj2840001
Brucker, R.M., Bordenstein, S.R., 2012. Speciation by symbiosis. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2012.03.011
Bryant, D., Moulton, V., 2002. Neighbornet: An agglomerative method for the construction of planar phylogenetic networks, in: Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in Bioinformatics). pp. 375–391. https://doi.org/10.1007/3-540-45784-4_28
Buckley, T.R., Simon, C., Chambers, G.K., 2001. Phylogeography of the New Zealand cicada *Maoricicada campbelli* based on mitochondrial DNA sequences: Ancient clades associated with cenozoic environmental change. Evolution (N. Y). 55, 1395–1407. https://doi.org/10.1111/j.0014-3820.2001.tb00661.x
Bucklin, A., Steinke, D., Blanco-Bercial, L., 2011. DNA Barcoding of marine Metazoa. Ann. Rev. Mar. Sci. 3, 471–508. https://doi.org/10.1146/annurev-marine-120308-080950
Buhay, J.E., 2009. "COI-like" Sequences are becoming problematic in molecular systematic and DNA barcoding studies. J. Crustac. Biol. 29, 96–110. https://doi.org/10.1651/08-3020.1
Burbidge, M.L., Colbourne, R.M., Robertson, H.A., Baker, A.J., 2003. Molecular and other biological evidence supports the recognition of at least three species of brown kiwi. Conserv. Genet. 4, 167–177. https://doi.org/10.1023/A:1023386506067
Carrea, C., Anderson, L. V., Craw, D., Waters, J.M., Burridge, C.P., 2013. The significance of past interdrainage connectivity for studies of diversity, distribution and movement of freshwater-limited taxa within a catchment. J. Biogeogr. 41, 536–547. https://doi.org/10.1111/jbi.12207
Casal, G., Matos, E., Garcia, P., Al-Quraishy, S., Azevedo, C., 2012. Ultrastructural and molecular studies of *Microgemma carolinus* n. sp. (Microsporidia), a parasite of the fish *Trachinotus carolinus* (Carangidae) in southern Brazil. Parasitology 139,

1720–1728. https://doi.org/10.1017/S0031182012001011
Castelli, M., Sassera, D., Petroni, G., 2016. Biodiversity of "non-model" Rickettsiales and their association with aquatic organisms, in: Thomas, S. (Ed.), Rickettsiales: Biology, Molecular Biology, Epidemiology, and Vaccine development. Springer International Publishing, pp. 59–91. https://doi.org/10.1007/978-3-319-46859-4_3
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334
Ceccarelli, F.S., Haddad, C.R., Ramírez, M.J., 2016. Endosymbiotic Rickettsiales (Alphaproteobacteria) from the spider genus *Amaurobioides* (Araneae: Anyphaenidae) . J. Arachnol. 44, 251-253. https://doi.org/10.1636/joa-s-15-009
Chaijarasphong, T., Munkongwongsiri, N., Stentiford, G.D., Aldama-Cano, D.J., Thansa, K., Flegel, T.W., Sritunyalucksana, K., Itsathitphaisarn, O., 2020. The shrimp microsporidian *Enterocytozoon hepatopenaei* (EHP): Biology, pathology, diagnostics and control. J. Invertebr. Pathol. In press. https://doi.org/10.1016/j.jip.2020.107458
Chandra, S., Forsyth, M., Lawrence, A.L., Emery, D., Šlapeta, J., 2017. Cat fleas (*Ctenocephalides felis*) from cats and dogs in New Zealand: Molecular characterisation, presence of *Rickettsia felis* and *Bartonella clarridgeiae* and comparison with Australia. Vet. Parasitol. 234, 25–30. https://doi.org/10.1016/j.vetpar.2016.12.017
Chapman, M.A., Lewis, M.H., Winterbourn, M.J., 2011. Guide to the freshwater crustacea of New Zealand., Guide to the freshwater crustacea of New Zealand. New Zealand Freshwater Sciences Society, Wellington.
Charlesworth, B., 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10, 195-205. https://doi.org/10.1038/nrg2526
Chilton, C., 1920. The occurrence in the Philippine Islands of the freshwater amphipod *Paracalliope fluviatilis* (G.M. Thomson). Philipp. J. Sci. 17, 513–514.
Chilton, C., 1906. Note on some Crustacea from the freshwater lakes of New Zealand. Proc. Zool. Soc. London 702–705.
Chilton, C., 1882. On some subterranean Crustacea. Trans. Proc. New Zeal. Inst. 14, 174–180.
Chun, J.Y., Kim, K.J., Hwang, I.T., Kim, Y.J., Lee, D.H., Lee, I.K., Kim, J.K., 2007. Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. Nucleic Acids Res. 35, e40. https://doi.org/10.1093/nar/gkm051
Clausen, C., 2000. Light and ultrastructural observations on a microsporidium in the hydrozoan *Halammohydra intermedia* (Cnidaria). Sarsia 85, 177–180. https://doi.org/10.1080/00364827.2000.10414568
Clayton, D.H., Bush, S.E., Johnson, K.P., 2015. Coevolution of life on hosts: integrating ecology and history. University of Chicago Press. https://doi.org/10.7208/chicago/9780226302300.001.0001
Clayton, D.H., Johnson, K.P., 2003. Linking coevolutionary history to ecological process: doves and lice. Evolution (N. Y). 57, 2335–2341. https://doi.org/10.1111/j.0014-3820.2003.tb00245.x
Conlan, K.E., 1994. Amphipod crustaceans and environmental disturbance: a review. J. Nat. Hist. 28, 519–554. https://doi.org/10.1080/00222939400770241

Conow, C., Fielder, D., Ovadia, Y., Libeskind-Hadas, R., 2010. Jane: A new tool for the cophylogeny reconstruction problem. Algorithms Mol. Biol. 5, 1-10. https://doi.org/10.1186/1748-7188-5-16

Copilaş-Ciocianu, D., Borko, Š., Fišer, C., 2020. The late blooming amphipods: global change promoted post-Jurassic ecological radiation despite Palaeozoic origin. Mol. Phylogenet. Evol. 143, 106664. https://doi.org/https://doi.org/10.1016/j.ympev.2019.106664

Cordaux, R., Bouchon, D., Grève, P., 2011. The impact of endosymbionts on the evolution of host sex-determination mechanisms. Trends Genet. 27, 332-341. https://doi.org/10.1016/j.tig.2011.05.002

Corradi, N., Pombert, J.F., Farinelli, L., Didier, E.S., Keeling, P.J., 2010. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinali*s. Nat. Commun. 1, 1-7. https://doi.org/10.1038/ncomms1082

Corsaro, D., Walochnik, J., Venditti, D., Müller, K.D., Hauröder, B., Michel, R., 2014a. Rediscovery of *Nucleophaga amoebae*, a novel member of the Rozellomycota. Parasitol. Res. 113, 4491–4498. https://doi.org/10.1007/s00436-014-4138-8

Corsaro, D., Walochnik, J., Venditti, D., Steinmann, J., Müller, K.D., Michel, R., 2014b. Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. Parasitol. Res. 113, 1909–1918. https://doi.org/10.1007/s00436-014-3838-4

Corsaro, D., Wylezich, C., Venditti, D., Michel, R., Walochnik, J., Wegensteiner, R., 2019. Filling gaps in the microsporidian tree: rDNA phylogeny of *Chytridiopsis typographi* (Microsporidia: Chytridiopsida). Parasitol. Res. 118, 169–180. https://doi.org/10.1007/s00436-018-6130-1

Craw, D., Upton, P., Waters, J., Wallis, G., 2017. Biological memory of the first pleistocene glaciation in New Zealand. Geology 45, 595–598. https://doi.org/10.1130/G39115.1

Cunha, M.R., Sorbe, J.C., Moreira, M.H., 2000. The amphipod *Corophium multisetosum* (Corophiidae) in Ria de Aveiro (NW Portugal). I. Life history and aspects of reproductive biology. Mar. Biol. 137, 637–650. https://doi.org/10.1007/s002270000384

Czaker, R., 1997. Wittmannia antarctica N. G., n. sp. (Nosematidae), a new hyperparasite in the Antarctic dicyemid mesozoan *Kantharella antarctica*. J. Eukaryot. Microbiol. 44, 438–446. https://doi.org/10.1111/j.1550-7408.1997.tb05721.x

Darriba, D., Taboada, G., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109

de Buron, I., Loubes, C., Maurand, J., 1990. Infection and pathological alterations within the acanthocephalan *Acanthocephaloides propinquus* attributable to the microsporidian hyperparasite *Microsporidium acanthocephali*. Trans. Am. Microsc. Soc. 109, 91. https://doi.org/10.2307/3226598

de Vienne, D.M., 2019. Tanglegrams are misleading for visual evaluation of tree congruence. Mol. Biol. Evol. 36, 174–176. https://doi.org/10.1093/molbev/msy196

de Vienne, D.M., Giraud, T., Shykoff, J.A., 2007. When can host shifts produce congruent host and parasite phylogenies? A simulation approach. J. Evol. Biol. 20, 1428–1438. https://doi.org/10.1111/j.1420-9101.2007.01340.x

de Vienne, D.M., Refrégier, G., López-Villavicencio, M., Tellier, A., Hood, M.E., Giraud, T., 2013. Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. New Phytol. 198, 347–385. https://doi.org/10.1111/nph.12150

Dean, P., Hirt, R.P., Embley, T.M., 2016. Microsporidia: Why make nucleotides if you can steal them? PLoS Pathog. 12, e1005870. https://doi.org/10.1371/journal.ppat.1005870

Després, L., 2019. One, two or more species? Mitonuclear discordance and species delimitation. Mol. Ecol. 28, 3845-3847. https://doi.org/10.1111/mec.15211

Desser, S.S., Koehler, A., Barta, J.R., Kamyab, J., Ringuette, M.J., 2004. *Trichonosema algonquinensis* n. sp. (Phylum Microsporidia) in *Pectinatella magnifica* (Bryozoa: Phylactolaemata) from Algonquin Park, Ontario, Canada. J. Eukaryot. Microbiol. 51, 389-393. https://doi.org/10.1111/j.1550-7408.2004.tb00385.x

Dezfuli, B.S., Giari, L., Poulin, R., 2000. Species associations among larval helminths in an amphipod intermediate host. Int. J. Parasitol. 30, 1143–1146. https://doi.org/10.1016/S0020-7519(00)00093-X

Dick, J.T.A., Faloon, S.E., Elwood, R.W., 1998. Active brood care in an amphipod: influences of embryonic development, temperature and oxygen. Anim. Behav. 56, 663–672. https://doi.org/10.1006/anbe.1998.0797

Dimova, M., Madyarova, E., Gurkov, A., Drozdova, P., Lubyaga, Y., Kondrateva, E., Adelshin, R., Timofeyev, M., 2018. Genetic diversity of Microsporidia in the circulatory system of endemic amphipods from different locations and depths of ancient Lake Baikal. PeerJ 6, e5329. https://doi.org/10.7717/peerj.5329

Doliwa, A., Dunthorn, M., Rassoshanska, E., Mahé, F., Bass, D., Ritter, C.D., 2021. Identifying potential hosts of short-branch Microsporidia. Microb. Ecol. 2021, 1-5. https://doi.org/10.1007/s00248-020-01657-9

Doña, J., Sweet, A.D., Johnson, K.P., Serrano, D., Mironov, S., Jovani, R., 2017. Cophylogenetic analyses reveal extensive host-shift speciation in a highly specialized and host-specific symbiont system. Mol. Phylogenet. Evol. 115, 190–196. https://doi.org/10.1016/j.ympev.2017.08.005

Drozdova, P., Madyarova, E., Dimova, M., Gurkov, A., Vereshchagina, K., Adelshin, R., Timofeyev, M., 2020. The diversity of microsporidian parasites infecting the Holarctic amphipod *Gammarus lacustris* from the Baikal region is dominated by the genus *Dictyocoela*. J. Invertebr. Pathol. 170, 107330. https://doi.org/10.1016/j.jip.2020.107330

Drummond, A.J., Newcomb, R.D., Buckley, T.R., Xie, D., Dopheide, A., Potter, B.C.M., Heled, J., Ross, H.A., Tooman, L., Grosser, S., Park, D., Demetras, N.J., Stevens, M.I., Russell, J.C., Anderson, S.H., Carter, A., Nelson, N., 2015. Evaluating a multigene environmental DNA approach for biodiversity assessment. Gigascience 4, 2015, 4-46. https://doi.org/10.1186/s13742-015-0086-1

Dunn, A.M., Smith, J.E., 2001. Microsporidian life cycles and diversity: The relationship between virulence and transmission. Microbes Infect. 3, 381–388. https://doi.org/10.1016/S1286-4579(01)01394-6

Dunn, A.M., Terry, R.S., Smith, J.E., 2001. Transovarial transmission in the microsporidia. Adv. Parasitol. 48, 57–100. https://doi.org/https://doi.org/10.1016/S0065-308X(01)48005-5

Dupuis, J.R., Sperling, F.A.H., 2020. Phylogenomic test of mitochondrial clues to archaic ancestors in a group of hybridizing swallowtail butterflies. Mol.

Phylogenet. Evol. 152, 106921. https://doi.org/10.1016/j.ympev.2020.106921
Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L., Engelstädter, J., Hurst, G.D., 2008. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. BMC Biol. 6, 27. https://doi.org/10.1186/1741-7007-6-27
Duron, O., Schneppat, U.E., Berthomieu, A., Goodman, S.M., Droz, B., Paupy, C., Obame Nkoghe, J., Rahola, N., Tortosa, P., 2014. Origin, acquisition and diversification of heritable bacterial endosymbionts in louse flies and bat flies. Mol. Ecol. 23, 2105–2117. https://doi.org/10.1111/mec.12704
Dyall, S.D., Brown, M.T., Johnson, P.J., 2004. Ancient invasions: from endosymbionts to organelles. Science 304, 253-257. https://doi.org/10.1126/science.1094884
Dyková, I., Veverková, M., Fiala, I., Macháčková, B., Pecková, H., 2003. *Nuclearia pattersoni* sp. n. (Filosea), a new species of amphizoic amoeba isolated from gills of roach (*Rutilus rutilus*), and its rickettsial endosymbiont. Folia Parasitol. (Praha). 50, 161–170. https://doi.org/10.14411/fp.2003.030
Edlind, T.D., Li, J., Visvesvara, G.S., Vodkin, M.H., McLaughlin, G.L., Katiyar, S.K., 1996. Phylogenetic analysis of β-tubulin sequences from amitochondrial protozoa. Mol. Phylogenet. Evol. 5, 359–367. https://doi.org/10.1006/mpev.1996.0031
Elbrecht, V., Leese, F., 2015. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. PLoS One 10, e0130324. https://doi.org/10.1371/journal.pone.0130324
Elith, J., Leathwick, J.R., 2009. Species distribution models: Ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159
Eme, D., Westfalls, K.M., Kristjánsson, B.K., Pálsson, S., 2017. Genetic variation of the groundwater amphipod *Crangonyx islandicus* Svavarsson et Kristjánsson, 2006, endemic to Iceland. Biodivers. J. 8, 571–573.
Engelbrecht, A., Matthee, S., Du Toit, N., Matthee, C.A., 2016. Limited dispersal in an ectoparasitic mite, *Laelaps giganteus*, contributes to significant phylogeographic congruence with the rodent host, *Rhabdomys*. Mol. Ecol. 25, 1006–1021. https://doi.org/10.1111/mec.13523
Engelstädter, J., Hurst, G.D.D., 2009. The ecology and evolution of microbes that manipulate host reproduction. Annu. Rev. Ecol. Evol. Syst. 40, 127–149. https://doi.org/10.1146/annurev.ecolsys.110308.120206
Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
Fahrenholz, H., 1913. Ectoparasiten und abstammungslehre. Zool. Anz. 41, 371–374.
Fenwick, G., Webber, R., 2008. identification of New Zealand's terrestrial amphipods (crustacea: amphipoda: talitridae). Tuhinga Rec. Museum New Zeal. Te Papa Tongarewa 19, 29–56.
Fenwick, G.D., 2001. the freshwater amphipoda (Crustacea) of New Zealand: a review. J. R. Soc. New Zeal. 31, 341–363. https://doi.org/10.1080/03014223.2001.9517658
Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486
Fišer, C., Zagmajster, M., Ferreira, R.L., 2013. Two new Amphipod families recorded

in South America shed light on an old biogeographical enigma. Syst. Biodivers. 11, 117–139. https://doi.org/10.1080/14772000.2013.788579
Fitzpatrick, B.M., Johnson, J.R., Kump, D.K., Smith, J.J., Voss, S.R., Shaffer, H.B., 2010. Rapid spread of invasive genes into a threatened native species. Proc. Natl. Acad. Sci. U. S. A. 107, 3606–3610. https://doi.org/10.1073/pnas.0911802107
Fleming, C.A., 1979. The geological history of New Zealand and its life. Auckland University Press.
Foissner, I., Foissner, W., 1995. *Ciliatosporidium platyophryae* nov. gen., nov. spec. (Microspora incerta sedis), a parasite of *Platyophrya terricola* (Ciliophora, Colpodea). Eur. J. Protistol. 31, 248–259. https://doi.org/10.1016/S0932-4739(11)80088-X
Fokin, S.I., Di Giuseppe, G., Erra, F., Dini, F., 2008. *Euplotespora binucleata* n. gen., n. sp. (Protozoa: Microsporidia), a parasite infecting the hypotrichous ciliate *Euplotes woodruffi*, with observations on microsporidian infections in Ciliophora. J. Eukaryot. Microbiol. 55, 214–228. https://doi.org/10.1111/j.1550-7408.2008.00322.x
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299. https://doi.org/10.1071/ZO9660275
Forsman, Z.H., Knapp, I.S.S., Tisthammer, K., Eaton, D.A.R., Belcaid, M., Toonen, R.J., 2017. Coral hybridization or phenotypic variation? Genomic data reveal gene flow between *Porites lobata* and *P. Compressa*. Mol. Phylogenet. Evol. 111, 132–148. https://doi.org/10.1016/j.ympev.2017.03.023
Franzen, C., 2008. Microsporidia: A Review of 150 Years of Research. Open Parasitol. J. 2, 1–34. https://doi.org/10.2174/1874421400802010001
Franzen, C., Nassonova, E.S., Schölmerich, J., Issi, I. V., 2006. Transfer of the members of the genus *Brachiola* (Microsporidia) to the genus *Anncaliia* based on ultrastructural and molecular data. J. Eukaryot. Microbiol. 53, 26–35. https://doi.org/10.1111/j.1550-7408.2005.00066.x
Frazer, J.L., Tham, K.M., Reid, M., van Andel, M., McFadden, A.M.J., Forsgren, E., Pettis, J.S., Pharo, H., 2015. Primera detección de *Nosema ceranae* en abejas melíferas de Nueva Zelanda. J. Apic. Res. 54, 358–365. https://doi.org/10.1080/00218839.2016.1161962
Friend, J.A., Richardson, A.M.M., 1986. Biology of terrestrial amphipods. Annu. Rev. Entomol. 31, 25–48. https://doi.org/10.1146/annurev.en.31.010186.000325
Friesen, O.C., Goellner, S., Poulin, R., Lagrue, C., 2019. Parasites shape community structure and dynamics in freshwater crustaceans. Parasitology 147, 1–12. https://doi.org/10.1017/S0031182019001483
Fuertes Aguilar, J., Rosselló, J.A., Nieto Feliner, G., 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Mol. Ecol. 8, 1341–1346. https://doi.org/10.1046/j.1365-294X.1999.00690.x
Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397-423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421
Galindo, L.J., Torruella, G., Moreira, D., Eglit, Y., Simpson, A.G.B., Völcker, E.,

Clauß, S., López-García, P., 2019. Combined cultivation and single-cell approaches to the phylogenomics of nucleariid amoebae, close relatives of fungi. Philos. Trans. R. Soc. B Biol. Sci. 374, 20190094. https://doi.org/10.1098/rstb.2019.0094
Galindo, L.J., Torruella, G., Moreira, D., Timpano, H., Paskerova, G., Smirnov, A., Nassonova, E., López-García, P., 2018. Evolutionary genomics of *Metchnikovella incurvata* (metchnikovellidae): An early branching microsporidium. Genome Biol. Evol. 10, 2736–2748. https://doi.org/10.1093/gbe/evy205
Galtier, N., Nabholz, B., GlÉmin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541-4550. https://doi.org/10.1111/j.1365-294X.2009.04380.x
Garrick, R.C., Sabree, Z.L., Jahnes, B.C., Oliver, J.C., 2017. Strong spatial-genetic congruence between a wood-feeding cockroach and its bacterial endosymbiont, across a topographically complex landscape. J. Biogeogr. 44, 1500–1511. https://doi.org/10.1111/jbi.12992
Gatehouse, H.S., Malone, L.A., 1998. The ribosomal RNA gene region of *Nosema apis* (Microspora): DNA sequence for small and large subunit rRNA genes and evidence of a large tandem repeat unit size. J. Invertebr. Pathol. 71, 97–105. https://doi.org/10.1006/jipa.1997.4737
Geller, J., Meyer, C., Parker, M., Hawk, H., 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol. Ecol. Resour. 13, 851–861. https://doi.org/10.1111/1755-0998.12138
Gerth, M., Wolf, R., Bleidorn, C., Richter, J., Sontowski, R., Unrein, J., Schlegel, M., Gruppe, A., 2017. Green lacewings (Neuroptera: Chrysopidae) are commonly associated with a diversity of rickettsial endosymbionts. Zool. Lett. 3, 1-7. https://doi.org/10.1186/s40851-017-0072-9
Giribet, G., Boyer, S.L., 2010. 'Moa's Ark'or 'Goodbye Gondwana': is the origin of New Zealand's terrestrial invertebrate fauna ancient, recent, or both? Invertebr. Syst. 24, 1–8. https://doi.org/10.1071/IS10009
Giska, I., Sechi, P., Babik, W., 2015. Deeply divergent sympatric mitochondrial lineages of the earthworm *Lumbricus rubellus* are not reproductively isolated. BMC Evol. Biol. 15, 217. https://doi.org/10.1186/s12862-015-0488-9
Gollas-Galván, T., Avila-Villa, L.A., Martínez-Porchas, M., Hernandez-Lopez, J., 2014. *Rickettsia*-like organisms from cultured aquatic organisms, with emphasis on necrotizing hepatopancreatitis bacterium affecting penaeid shrimp: an overview on an emergent concern. Rev. Aquac. 6, 256-269. https://doi.org/10.1111/raq.12043
Gompert, Z., Forister, M.L., Fordyce, J.A., Nice, C.C., 2008. Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. Mol. Ecol. 17, 5231–5244. https://doi.org/10.1111/j.1365-294X.2008.03988.x
Goodacre, S.L., Martin, O.Y., Thomas, C.F.G., Hewitt, G.M., 2006. *Wolbachia* and other endosymbiont infections in spiders. Mol. Ecol. 15, 517–527. https://doi.org/10.1111/j.1365-294X.2005.02802.x
Gorbunov, A.K., Kosova, A.A., 2001. Parasites in rotifers from the Volga delta, in: Hydrobiologia. 446, 51–55. https://doi.org/10.1023/A:1017569004998
Grabner, D.S., 2017. Hidden diversity: parasites of stream arthropods. Freshw. Biol. 62, 52–64. https://doi.org/10.1111/fwb.12848

Grabner, D.S., Weigand, A.M., Leese, F., Winking, C., Hering, D., Tollrian, R., Sures, B., 2015. Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. Parasit. Vectors 8, 419. https://doi.org/10.1186/s13071-015-1036-6

Graf, F., 1984. Presence of bacteria in the posterior caecum in the intestinal lumen of the hypogean Crustacean *Niphargus virei* (Gammaridae: Amphipoda). Can. J. Zool. Can. Zool. 62, 1829–1833.

Grassly, N.C., Fraser, C., 2006. Seasonal infectious disease epidemiology. Proc. R. Soc. B Biol. Sci. https://doi.org/10.1098/rspb.2006.3604

Gray, M.W., 1989. The evolutionary origins of organelles. Trends Genet. 5, 294–299. https://doi.org/10.1016/0168-9525(89)90111-X

Gualtieri, L., Nugnes, F., Nappo, A.G., Gebiola, M., Bernardo, U., 2017. Life inside a gall: closeness does not favour horizontal transmission of *Rickettsia* between a gall wasp and its parasitoid. FEMS Microbiol. Ecol. 93, fix087. https://doi.org/10.1093/femsec/fix087

Guler, Y., Short, S., Kile, P., Ford, A.T., 2012. Integrating field and laboratory evidence for environmental sex determination in the amphipod, *Echinogammarus marinus*. Mar. Biol. 159, 2885–2890. https://doi.org/10.1007/s00227-012-2042-2

Haag, K.L., James, T.Y., Pombert, J.F., Larsson, R., Schaer, T.M.M., Refardt, D., Ebert, D., 2014. Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. Proc. Natl. Acad. Sci. U. S. A. 111, 15480–15485. https://doi.org/10.1073/pnas.1410442111

Hafner, M., Demastes, J., Spradling, T., Reed, D., 2003. Cophylogeny between pocket gophers and chewing lice, in: Page, R.D.M. (Ed.), Tangled trees: phylogeny, cospeciatiion, and coevolution. The University of Chicago Press, Chicago, 195–220.

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, 1087–1090. https://doi.org/10.1126/science.8066445

Hagimori, T., Abe, Y., Date, S., Miura, K., 2006. The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. Curr. Microbiol. 52, 97–101. https://doi.org/10.1007/s00284-005-0092-0

Haine, E.R., Boucansaud, K., Rigaud, T., 2005. Conflict between parasites with different transmission strategies infecting an amphipod host. Proc. R. Soc. B Biol. Sci. 272, 2505–2510. https://doi.org/10.1098/rspb.2005.3244

Hajduskova, E., Literak, I., Papousek, I., Costa, F.B., Novakova, M., Labruna, M.B., Zdrazilova-Dubska, L., 2016. "Candidatus Rickettsia mendelii", a novel basal group rickettsia detected in *Ixodes ricinus* ticks in the Czech Republic. Ticks Tick. Borne. Dis. 7, 482–486. https://doi.org/10.1016/j.ttbdis.2016.02.004

Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N., Hickey, D.A., 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet. 23, 167–172. https://doi.org/10.1016/j.tig.2007.02.001

Haley, A.J., 1954. Microsporidian Parasite, *Glugea Hertwigi*, in American Smelt from the Great Bay Region, New Hampshire. Trans. Am. Fish. Soc. 83, 84–90. https://doi.org/10.1577/1548-8659(1953)83[84:mpghia]2.0.co;2

Havermans, C., Nagy, Z.T., Sonet, G., De Broyer, C., Martin, P., 2011. DNA barcoding reveals new insights into the diversity of Antarctic species of *Orchomene sensu lato* (Crustacea: Amphipoda: Lysianassoidea). Deep. Res. Part II Top. Stud.

Oceanogr. 58, 230–241. https://doi.org/10.1016/j.dsr2.2010.09.028
Heath, T.A., Hedtke, S.M., Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol. 46, 239–257. https://doi.org/10.3724/SP.J.1002.2008.08016
Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. B Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218
Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc. Natl. Acad. Sci. U. S. A. 101, 14812–14817. https://doi.org/10.1073/pnas.0406166101
Heinz, E., Hacker, C., Dean, P., Mifsud, J., Goldberg, A. V., Williams, T.A., Nakjang, S., Gregory, A., Hirt, R.P., Lucocq, J.M., Kunji, E.R.S., Embley, T.M., 2014. Plasma membrane-located purine nucleotide transport proteins are key components for host exploitation by microsporidian intracellular parasites. PLoS Pathog. 10, e1004547. https://doi.org/10.1371/journal.ppat.1004547
Hill, W.G., Robertson, A., 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38, 226–231. https://doi.org/10.1007/BF01245622
Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., Pitteloud, C., Dincă, V., Alvarez, N., Vila, R., 2019. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly *Thymelicus sylvestris*. Mol. Ecol. 28, 3857–3868. https://doi.org/10.1111/mec.15153
Hipsley, C.A., Müller, J., 2014. Beyond fossil calibrations: Realities of molecular clock practices in evolutionary biology. Front. Genet. 5, 138. https://doi.org/10.3389/fgene.2014.00138
Hogg, I.D., Stevens, M.I., Schnabel, K.E., Ann Chapman, M., 2006. Deeply divergent lineages of the widespread New Zealand amphipod *Paracalliope fluviatilis* revealed using allozyme and mitochondrial DNA analyses. Freshw. Biol. 51, 236–248. https://doi.org/10.1111/j.1365-2427.2005.01491.x
Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. Biometrical J. 50, 346-363. https://doi.org/10.1002/bimj.200810425
Hou, Z., Sket, B., 2016. A review of Gammaridae (Crustacea: Amphipoda): the family extent, its evolutionary history, and taxonomic redefinition of genera. Zool. J. Linn. Soc. 176, 323–348. https://doi.org/10.1111/zoj.12318
Hudson, R.R., Turelli, M., 2003. Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. Evolution (N. Y). 57, 182-190. https://doi.org/10.1111/j.0014-3820.2003.tb00229.x
Hurst, G.D.D., Jiggins, F.M., 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc. R. Soc. B Biol. Sci. 272, 1525-1534. https://doi.org/10.1098/rspb.2005.3056
Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254-267. https://doi.org/10.1093/molbev/msj030
Iannilli, V., Ruffo, S., 2007. A new genus and species of Phreatogammaridae (*Caledonietta maryae* n. gen. n. sp) from New Caledonia (Crustacea, Amphipoda). Boll. del Mus. Civ. di Stor. Nat. di Verona 31, 23–29.
Ironside, J.E., Alexander, J., 2015. Microsporidian parasites feminise hosts without

paramyxean co-infection: support for convergent evolution of parasitic feminisation. Int. J. Parasitol. 45, 427–433. https://doi.org/10.1016/j.ijpara.2015.02.001
Ironside, J.E., Wilkinson, T.J., 2018. Accumulation and exchange of parasites during adaptive radiation in an ancient lake. Int. J. Parasitol. 48, 297–307. https://doi.org/10.1016/j.ijpara.2017.10.003
James, A.C., Ballard, J.W.O., 2000. Expression of cytoplasmic incompatibility in *Drosophila simulans* and its impact on infection frequencies and distribution of *Wolbachia pipientis*. Evolution (N. Y). 54, 1661–1672. https://doi.org/10.1111/j.0014-3820.2000.tb00710.x
James, T.Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., Stajich, J.E., 2013. Shared signatures of parasitism and phylogenomics unite cryptomycota and Microsporidia. Curr. Biol. 23, 1548–1553. https://doi.org/10.1016/j.cub.2013.06.057
Janz, N., 2011. Ehrlich and Raven revisited: Mechanisms underlying codiversification of plants and enemies. Annu. Rev. Ecol. Evol. Syst. 42, 71-89. https://doi.org/10.1146/annurev-ecolsys-102710-145024
Janzen, D.H., 1980. When is it coevolution? Evolution (N. Y). 34, 611. https://doi.org/10.2307/2408229
Jasid, S., Simontacchi, M., Bartoli, C.G., Puntarulo, S., 2006. Chloroplasts as a nitric oxide cellular source. Effect of reactive nitrogen species on chloroplastic lipids and proteins. Plant Physiol. 142, 1246–1255. https://doi.org/10.1104/pp.106.086918
Jiggins, F.M., 2003. Male-killing *Wolbachia* and mitochondrial DNA: Selective sweeps, hybrid introgression and parasite population dynamics. Genetics 164, 5–12.
Johnstone, R.A., Hurst, G.D.D., 1996. Maternally inherited male-killing microorganisms may confound interpretation of mitochondrial DNA variability. Biol. J. Linn. Soc. 58, 453–470. https://doi.org/10.1006/bijl.1996.0047
Jones, M.D.M., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., Richards, T.A., 2011. Discovery of novel intermediate forms redefines the fungal tree of life. Nature 474, 200–204. https://doi.org/10.1038/nature09984
Jones, S.R.M., Prosperi-Porta, G., Kim, E., 2012. The diversity of microsporidia in parasitic copepods (Caligidae: Siphonostomatoida) in the northeast pacific ocean with description of *Facilispora margolisi* n. g., n. sp. and a new family Facilisporidae n. fam. J. Eukaryot. Microbiol. 59, 206–217. https://doi.org/10.1111/j.1550-7408.2011.00614.x
Kamp, P.J.J., 1986. Late Cretaceous-Cenozoic tectonic development of the southwest Pacific region. Tectonophysics 121, 225–251. https://doi.org/10.1016/0040-1951(86)90045-4
Karpov, S.A., Mamkaeva, M.A., Aleoshin, V. V., Nassonova, E., Lilje, O., Gleason, F.H., 2014. Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. Front. Microbiol. 5, 112. https://doi.org/10.3389/fmicb.2014.00112
Karvonen, A., Jokela, J., Laine, A.L., 2019. Importance of sequence and timing in parasite coinfections. Trends Parasitol. 35, 109-118. https://doi.org/10.1016/j.pt.2018.11.007
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

Keeling, P.J., Doolittle, W.F., 1996. Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol. Biol. Evol. 13, 1297–1305. https://doi.org/10.1093/oxfordjournals.molbev.a025576
Kikuchi, Y., Fukatsu, T., 2005. *Rickettsia* infection in natural leech populations. Microb. Ecol. 49, 265–271. https://doi.org/10.1007/s00248-004-0140-5
Kikuchi, Y., Sameshima, S., Kitade, O., Kojima, J., Fukatsu, T., 2002. Novel clade of *Rickettsia* spp. from leeches. Appl. Environ. Microbiol. 68, 999–1004. https://doi.org/10.1128/AEM.68.2.999-1004.2002
Kim, Y., Neilsen, R., 2004. Linkage disequilibrium as a signature of selective sweeps. Genetics 167, 1513–1524. https://doi.org/10.1534/genetics.103.025387
Kimura, M., 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature. 267, 275-276. https://doi.org/10.1038/267275a0
King, R.A., Bradford, T., Austin, A.D., Humphreys, W.F., Cooper, S.J.B., 2012. Divergent molecular lineages and not-so-cryptic species: The first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. J. Crustac. Biol. 32, 465–488. https://doi.org/10.1163/193724012X626566
Knott, B., 1975. A new species of freshwater amphipod, *Paracalliope larai* (Family Eusiridae) from Tasmania, in: Papers and Proceedings of the Royal Society of Tasmania. 39–52.
Knowlton, N., Weigt, L.A., 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. R. Soc. B Biol. Sci. 265, 2257–2263. https://doi.org/10.1098/rspb.1998.0568
Knox, M.A., Hogg, I.D., Pilditch, C.A., 2011. The role of vicariance and dispersal on New Zealand's estuarine biodiversity: The case of *Paracorophium* (Crustacea: Amphipoda). Biol. J. Linn. Soc. 103, 863–874. https://doi.org/10.1111/j.1095-8312.2011.01675.x
Kornobis, E., Pálsson, S., 2011. Discordance in variation of the ITS Region and the Mitochondrial COI Gene in the Subterranean Amphipod *Crangonyx islandicus*. J. Mol. Evol. 73, 34–44. https://doi.org/10.1007/s00239-011-9455-2
Kornobis, E., PÁlsson, S., KristjÁnsson, B.K., Svavarsson, J., 2010. Molecular evidence of the survival of subterranean amphipods (Arthropoda) during Ice Age underneath glaciers in Iceland. Mol. Ecol. 19, 2516–2530. https://doi.org/10.1111/j.1365-294X.2010.04663.x
Krebes, L., Blank, M., Frankowski, J., Bastrop, R., 2010. Molecular characterisation of the Microsporidia of the amphipod *Gammarus duebeni* across its natural range revealed hidden diversity, wide-ranging prevalence and potential for co-evolution. Infect. Genet. Evol. 10, 1027–1038. https://doi.org/10.1016/j.meegid.2010.06.011
Kristjánsson, B.K., Svavarsson, J., 2007. Subglacial refugia in Iceland enabled groundwater amphipods to survive glaciations. Am. Nat. 170, 292–296. https://doi.org/10.1086/518951
Küchler, S.M., Kehl, S., Dettner, K., 2009. Characterization and localization of *Rickettsia* sp. in water beetles of genus *Deronectes* (Coleoptera: Dytiscidae). FEMS Microbiol. Ecol. 68, 201–211. https://doi.org/10.1111/j.1574-6941.2009.00665.x
Kuzmenkova, Z. V, Sherbakov, D.U., Smith, J.E., 2008. Diversity of microsporidia, that parasitize on Baikal amphipods *Gmelinoides fasciatus* from different populations. Bull. Irkutsk State Univ. Ser. Biol. Ecol. 1, 56–61.

Kvie, K.S., Hogner, S., Aarvik, L., Lifjeld, J.T., Johnsen, A., 2013. Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*). Ecol. Evol. 3, 126–144. https://doi.org/10.1002/ece3.434

Lacerda, A.L. d. F., Proietti, M.C., Secchi, E.R., Taylor, J.D., 2020. Diverse groups of Fungi are associated with plastics in the surface waters of the Western South Atlantic and the Antarctic Peninsula. Mol. Ecol. 29, 1903-1918. https://doi.org/10.1111/mec.15444

Lagrue, C., Joannes, A., Poulin, R., Blasco-Costa, I., 2016. Genetic structure and host-parasite co-divergence: evidence for trait-specific local adaptation. Biol. J. Linn. Soc. 118, 344–358. https://doi.org/10.1111/bij.12722

Landis, C.A., Campbell, H.J., Begg, J.G., Mildenhall, D.C., Paterson, A.M., Trewick, S.A., 2008. The Waipounamu Erosion Surface: questioning the antiquity of the New Zealand land surface and terrestrial fauna and flora. Geol. Mag. 145, 173–197. https://doi.org/10.1017/S0016756807004268

Lara, E., Moreira, D., López-García, P., 2010. The environmental clade LKM11 and *Rozella* form the deepest branching clade of Fungi. Protist 161, 116–121. https://doi.org/10.1016/j.protis.2009.06.005

Larose, C., Schwander, T., 2016. Nematode endoparasites do not codiversify with their stick insect hosts. Ecol. Evol. 6, 5446–5458. https://doi.org/10.1002/ece3.2264

Larsson, J.I.R., 2014. The primitive Microsporidia, in: Microsporidia: pathogens of opportunity: First Edition. 605–634. https://doi.org/10.1002/9781118395264.ch24

Larsson, J.I.R., 1992. The ultrastructural cytology of *Bacillidium strictum* (Léger and Hesse, 1916) Jírovec, 1936 (Microspora, Bacillidiidae). Eur. J. Protistol. 28, 175–183. https://doi.org/10.1016/S0932-4739(11)80047-7

Larsson, R., 1982. A rickettsial pathogen of the amphipod *Rivulogammarus pulex*. J. Invertebr. Pathol. 40, 28–35. https://doi.org/10.1016/0022-2011(82)90033-7

Lartillot, N., Brinkmann, H., Philippe, H., 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model, in: BMC Evolutionary Biology. 7, 1-14. https://doi.org/10.1186/1471-2148-7-S1-S4

Lartillot, N., Lartillot, N., Bayesian, P., Using, P., Scornavacca, S.M., Lartillot, N., 2020. PhyloBayes: Bayesian Phylogenetics using site-heterogeneous models, in: Phylogenetics in the Genomic Era. No commercial publisher.

Lass, S., Ebert, D., 2006. Apparent seasonality of parasite dynamics: analysis of cyclic prevalence patterns. Proc. R. Soc. B Biol. Sci. 273, 199–206. https://doi.org/10.1098/rspb.2005.3310

Lee, C.E., Frost, B.W., 2002. Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae), in: Hydrobiologia. 111–128. https://doi.org/10.1023/A:1021293203512

Legendre, P., Desdevises, Y., Bazin, E., 2002. A statistical test for host-parasite coevolution. Syst. Biol. 51, 217–234. https://doi.org/10.1080/10635150252899734

Leigh, J.W., Bryant, D., 2015. POPART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front. Zool. 10, 1-14. https://doi.org/10.1186/1742-9994-10-34

Leung, T.L.F., Poulin, R., 2008. Parasitism, commensalism, and mutualism: Exploring the many shades of symbioses. Vie Milieu 58, 107–115.
Levron, C., Ternengo, S., Toguebaye, B.S., Marchand, B., 2005. Ultrastructural description of the life cycle of *Nosema monorchis* n. sp. (Microspora, Nosematidae), hyperparasite of *Monorchis parvus* (Digenea, Monorchiidae), intestinal parasite of *Diplodus annularis* (Pisces, Teleostei). Eur. J. Protistol. 41, 251–256. https://doi.org/10.1016/j.ejop.2005.05.001
Lewis, K.B., Carter, L., Davey, F.J., 1994. The opening of Cook Strait: interglacial tidal scour and aligning basins at a subduction to transform plate edge. Mar. Geol. 116, 293–312. https://doi.org/10.1016/0025-3227(94)90047-7
Lievens, E.J.P., Rode, N.O., Landes, J., Segard, A., Jabbour-Zahab, R., Michalakis, Y., Lenormand, T., 2019. Long-term prevalence data reveals spillover dynamics in a multi-host (*Artemia*), multi-parasite (Microsporidia) community. Int. J. Parasitol. 49, 471–480.
Linnen, C.R., Farrell, B.D., 2008. Phylogenetic analysis of nuclear and mitochondrial genes reveals evolutionary relationships and mitochondrial introgression in the sertifer species group of the genus *Neodiprion* (Hymenoptera: Diprionidae). Mol. Phylogenet. Evol. 48, 240–257. https://doi.org/10.1016/j.ympev.2008.03.021
Lipsitch, M., Nowak, M.A., Ebert, D., May, R.M., 1995. The population dynamics of vertically and horizontally transmitted parasites. Proc. R. Soc. London. Ser. B Biol. Sci. 260, 321–327. https://doi.org/10.1098/rspb.1995.0099
López, J.Á.R., Husemann, M., Schmitt, T., Kramp, K., Habel, J.C., 2018. Mountain barriers and trans-Saharan connections shape the genetic structure of *Pimelia* darkling beetles (Coleoptera: Tenebrionidae). Biol. J. Linn. Soc. 124, 547-556. https://doi.org/10.1093/biolinnean/bly053
Lowry, J.K., Myers, A.A., 2017. A phylogeny and classification of the Amphipoda with the establishment of the new order Ingolfiellida (Crustacea: Peracarida). Zootaxa 4265, 1–89. https://doi.org/http://dx.doi.org/10.11646/zootaxa.4265.1.1
Lowry, J.K., Myers, A.A., 2013. A phylogeny and classification of the Senticaudata subord. nov.(Crustacea: Amphipoda). Zootaxa 3610, 1–80. https://doi.org/10.11646/zootaxa.3610.1.1
Łukasik, P., van Asch, M., Guo, H., Ferrari, J., Charles, H., 2013. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. Ecol. Lett. 16, 214–218. https://doi.org/10.1111/ele.12031
Macdonald, K.S., Yampolsky, L., Duffy, J.E., 2005. Molecular and morphological evolution of the amphipod radiation of Lake Baikal. Mol. Phylogenet. Evol. 35, 323–343. https://doi.org/10.1016/j.ympev.2005.01.013
Machtelinckx, T., Van Leeuwen, T., Van De Wiele, T., Boon, N., De Vos, W.H., Sanchez, J.A., Nannini, M., Gheysen, G., De Clercq, P., 2012. Microbial community of predatory bugs of the genus *Macrolophus* (Hemiptera: Miridae). BMC Microbiol. 12, S9. https://doi.org/10.1186/1471-2180-12-S1-S9
Malone, L.A., Wigley, P.J., Dhana, S.D., 1987. Identity of a microsporidium from three New Zealand pasture insects: *Costelytra zealandica* (Coleoptera, scarabeidae), *Wiseana* spp. (Lepidoptera, Hepialidae), and *Listronotus bonariensis* (Coleoptera, Curculionidae). J. Invertebr. Pathol. 49, 135–144. https://doi.org/10.1016/0022-2011(87)90154-6
Manylov, O., 1999. First finding of a microsporidian parasite in the gastrotrich, *Turbanella lutheri* (Gastrotricha: Macrodasyida). Protistology 1.

Martins, A.S., Haimovici, M., 1997. Distribution, abundance and biological interactions of the cutlassfish *Trichiurus lepturu*s in the southern Brazil subtropical convergence ecosystem. Fish. Res. 30, 217–227.
Mayr, E., 2000. Species concepts and phylogenetic theory: a debate, in: Q.D., W., Meier, R. (Eds.), species concepts and phylogenetic theory: a debate. Columbia University Press, 17–29.
McCabe, J., Dunn, A.M., 1997. Adaptive significance of environmental sex determination in an amphipod. J. Evol. Biol. 10, 515–527. https://doi.org/10.1046/j.1420-9101.1997.10040515.x
McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H. V., Domazet-Lošo, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A.H., Kremer, N., Mazmanian, S.K., Metcalf, J.L., Nealson, K., Pierce, N.E., Rawls, J.F., Reid, A., Ruby, E.G., Rumpho, M., Sanders, J.G., Tautz, D., Wernegreen, J.J., 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. U. S. A. 110, 3229-3236. https://doi.org/10.1073/pnas.1218525110
McGill, B.J., 2010. Matters of scale. Science 328, 575-576. https://doi.org/10.1126/science.1188528
McGlone, M.S., 2005. Goodbye gondwana. J. Biogeogr. 32, 739–740. https://doi.org/10.1111/j.1365-2699.2005.01278.x
McGuire, J.A., Linkem, C.W., Koo, M.S., Hutchison, D.W., Lappin, A.K., Orange, D.I., Lemos-Espinal, J., Riddle, B.R., Jaeger, J.R., 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution (N. Y). 61, 2879–2897. https://doi.org/10.1111/j.1558-5646.2007.00239.x
Merkle, D., Middendorf, M., 2005. Reconstruction of the cophylogenetic history of related phylogenetic trees with divergence timing information, in: Theory in Biosciences. 277–299. https://doi.org/10.1016/j.thbio.2005.01.003
Messick, G.A., Overstreet, R.M., Nalepa, T.F., Tyler, S., 2004. Prevalence of parasites in amphipods *Diporeia* spp. from Lakes Michigan and Huron, USA. Dis. Aquat. Organ. 59, 159–170. https://doi.org/10.3354/dao059159
Michel, R., Schmid, E.N., Böker, T., Hager, D.G., Müller, K.D., Hoffmann, R., Seitz, H.M., 2000. *Vannella* sp. harboring Microsporidia-like organisms isolated from the contact lens and inflamed eye of a female keratitis patient. Parasitol. Res. 86, 514–520. https://doi.org/10.1007/s004360050704
Mikhailov, K. V., Simdyanov, T.G., Aleoshin, V. V., 2017. Genomic survey of a hyperparasitic microsporidian *Amphiamblys* sp. (Metchnikovellidae). Genome Biol. Evol. 9, 454–467. https://doi.org/10.1093/gbe/evw235
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in: 2010 Gateway Computing Environments Workshop (GCE). Ieee, 1–8. https://doi.org/10.1109/GCE.2010.5676129
Minard, G., Mavingui, P., Moro, C.V., 2013. Diversity and function of bacterial microbiota in the mosquito holobiont. Parasites and Vectors 6, 1-12. https://doi.org/10.1186/1756-3305-6-146
Mioduchowska, M., Czyz, M.J., Gołdyn, B., Kur, J., Sell, J., 2018. Instances of erroneous DNA barcoding of metazoan invertebrates: Are universal cox1 gene primers too “universal”? PLoS One 13, e0199609.

https://doi.org/10.1371/journal.pone.0199609
Moon, K.L., Chown, S.L., Fraser, C.I., 2019. Tandem host-parasite dispersal inferred from similarities in phylogeographical patterns among Little Penguins and their 'terrestrial' ectoparasites. J. Biogeogr. 46, 2520–2531. https://doi.org/10.1111/jbi.13714
Morales, H.E., Pavlova, A., Joseph, L., Sunnucks, P., 2015. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. Mol. Ecol. 24, 2820–2837. https://doi.org/10.1111/mec.13203
Murlas Cosmides, L., Tooby, J., 1981. Cytoplasmic inheritance and intragenomic conflict. J. Theor. Biol. 89, 83–129. https://doi.org/10.1016/0022-5193(81)90181-8
Murphy, N.P., Adams, M., Austin, A.D., 2009. Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. Mol. Ecol. 18, 109–122. https://doi.org/10.1111/j.1365-294X.2008.04007.x
Myers, A.A., 1993. Dispersal and endemicity in gammaridean Amphipoda. J. Nat. Hist. 27, 901–908. https://doi.org/10.1080/00222939300770551
Naegeli, C., 1957. Über die neue Krankheit der Seidenraupe und verwandte Organismen. Bot. Zeit. 15, 760–761.
Nakanishi, K., Hoshino, M., Nakai, M., Kunimi, Y., 2008. Novel RNA sequences associated with late male killing in *Homona magnanima*. Proc. R. Soc. B Biol. Sci. 275, 1249–1254. https://doi.org/10.1098/rspb.2008.0013
Nathan, R., 2001. The challenges of studying dispersal, in: Trends in Ecology and Evolution. pp. 481–483. https://doi.org/10.1016/S0169-5347(01)02272-8
Nei, M., Tajima, F., 1981. Genetic drift and estimation of effective population size. Genetics 98, 625–640.
Nieberding, C., Jousselin, E., Desdevises, Y., 2010. The use of co-phylogeographic patterns to predict the nature of host–parasite interactions, and vice versa, in: The Biogeography of Host-Parasite Interactions. 59–69.
Noda, H., Watanabe, K., Kawai, S., Yukuhiro, F., Miyoshi, T., Tomizawa, M., Koizumi, Y., Nikoh, N., Fukatsu, T., 2012. Bacteriome-associated endosymbionts of the green rice leafhopper *Nephotettix cincticeps* (Hemiptera: Cicadellidae). Appl. Entomol. Zool. 47, 217–225. https://doi.org/10.1007/s13355-012-0110-1
Nomura, T., 2002. Effective size of populations with unequal sex ratio and variation in mating success. J. Anim. Breed. Genet. 119, 297–310. https://doi.org/10.1046/j.1439-0388.2002.00347.x
Norte, A.C., Margos, G., Becker, N.S., Albino Ramos, J., Núncio, M.S., Fingerle, V., Araújo, P.M., Adamík, P., Alivizatos, H., Barba, E., Barrientos, R., Cauchard, L., Csörgő, T., Diakou, A., Dingemanse, N.J., Doligez, B., Dubiec, A., Eeva, T., Flaisz, B., Grim, T., Hau, M., Heylen, D., Hornok, S., Kazantzidis, S., Kováts, D., Krause, F., Literak, I., Mänd, R., Mentesana, L., Morinay, J., Mutanen, M., Neto, J.M., Nováková, M., Sanz, J.J., Pascoal da Silva, L., Sprong, H., Tirri, I.S., Török, J., Trilar, T., Tyller, Z., Visser, M.E., Lopes de Carvalho, I., 2020. Host dispersal shapes the population structure of a tick-borne bacterial pathogen. Mol. Ecol. 29, 485–501. https://doi.org/10.1111/mec.15336
Ovcharenko, M.O., Bacela, K., Wilkinson, T., Ironside, J.E., Rigaud, T., Wattier, R.A., 2010. *Cucumispora dikerogammari* n. gen.(Fungi: Microsporidia) infecting the invasive amphipod *Dikerogammarus villosus*: a potential emerging disease in European rivers. Parasitology 137, 191–204.

https://doi.org/10.1017/S0031182009991119
Page, R.D.M., 2003. Tangled trees: phylogeny, cospeciation, and coevolution. University of Chicago Press.
Paradis, E., 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26, 419–420. https://doi.org/10.1093/bioinformatics/btp696
Park, E., Jorge, F., Poulin, R., 2020. Shared geographic histories and dispersal contribute to congruent phylogenies between amphipods and their microsporidian parasites at regional and global scales. Mol. Ecol. 29, 3330–3345. https://doi.org/10.1111/mec.15562
Parola, P., Paddock, C.D., Raoult, D., 2005. Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. Clin. Microbiol. Rev. 18, 719–756. https://doi.org/10.1128/CMR.18.4.719-756.2005
Pasteur, L., 1870. Études sur la maladie des vers à soie. Gauthier-Villars, Paris, France 1870. Gauthier-Villars, Paris, France.
Pazhenkova, E.A., Lukhtanov, V.A., 2019. Nuclear genes (but not mitochondrial DNA barcodes) reveal real species: Evidence from the *Brenthis fritillary* butterflies (Lepidoptera, Nymphalidae). J. Zool. Syst. Evol. Res. 57, 298–313. https://doi.org/10.1111/jzs.12252
Perlman, S.J., Hunter, M.S., Zchori-Fein, E., 2006. The emerging diversity of *Rickettsia*. Proc. R. Soc. B Biol. Sci. 273, 2097–2106. https://doi.org/10.1098/rspb.2006.3541
Peyretaillade, E., Biderre, C., Peyret, P., Duffieux, F., Méténier, G., Gouy, M., Michot, B., Vivarès, C.P., 1998. Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. Nucleic Acids Res. 26, 3513–3520. https://doi.org/10.1093/nar/26.15.3513
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 Biol. 9, e1000602. https://doi.org/10.1371/journal.pbio.1000602
Pickup, J., Ironside, J.E., 2018. Multiple origins of parasitic feminization: thelygeny and intersexuality in beach-hoppers are caused by paramyxid parasites, not microsporidia. Parasitology 145, 408–415. https://doi.org/10.1017/S0031182017001597
Pilgrim, J., Ander, M., Garros, C., Baylis, M., Hurst, G.D.D., Siozios, S., 2017. Torix group *Rickettsia* are widespread in *Culicoides* biting midges (Diptera: Ceratopogonidae), reach high frequency and carry unique genomic features. Environ. Microbiol. 19, 4238–4255. https://doi.org/10.1111/1462-2920.13887
Pilgrim, J., Thongprem, P., Davison, H., Siozios, S., Baylis, M., Zakharov, E., Ratnasingham, S., DeWaard, J., Smith, A., Hurst, G., 2020. Torix *Rickettsia* are widespread in arthropods and reflect a neglected symbiosis. Authorea. https://doi.org/10.22541/au.159534851.19125003
Pita, L., Rix, L., Slaby, B.M., Franke, A., Hentschel, U., 2018. The sponge holobiont in a changing ocean: from microbes to ecosystems. Microbiome. https://doi.org/10.1186/s40168-018-0428-1
Pollard, W., Omelon, C., Andersen, D., McKay, C., 1999. Perennial spring occurrence in the Expedition Fiord area of western Axel Heiberg Island, Canadian High

Arctic. Can. J. Earth Sci. 36, 105–120. https://doi.org/10.1139/e98-097
Porter, T.M., Hajibabaei, M., 2018. Over 2.5 million COI sequences in GenBank and growing. PLoS One 13, e0200177. https://doi.org/10.1371/journal.pone.0200177
Poulin, R., 2020. Meta-analysis of seasonal dynamics of parasite infections in aquatic ecosystems. Int. J. Parasitol. 50, 501-510. https://doi.org/10.1016/j.ijpara.2020.03.006
Poulin, R., 2011. Evolutionary ecology of parasites (Second Edition). Princeton University Press. https://doi.org/10.2307/3285737
Poulin, R., 1999. The intra- and interspecific relationships between abundance and distribution in helminth parasites of birds. J. Anim. Ecol. 68, 719–725. https://doi.org/10.1046/j.1365-2656.1999.00319.x
Poulin, R., Latham, A.D.M., 2002. Parasitism and the burrowing depth of the beach hopper *Talorchestia quoyana* (Amphipoda: Talitridae). Anim. Behav. 63, 269–275. https://doi.org/10.1006/anbe.2001.1938
Pratlong, M., Rancurel, C., Pontarotti, P., Aurelle, D., 2017. Monophyly of Anthozoa (Cnidaria): why do nuclear and mitochondrial phylogenies disagree? Zool. Scr. 46, 363–371. https://doi.org/10.1111/zsc.12208
Prince, A.M., Andrus, L., 1992. PCR: How to kill unwanted DNA. Biotechniques 12, 358-360.
Qu, X.J., Jin, J.J., Chaw, S.M., Li, D.Z., Yi, T.S., 2017. Multiple measures could alleviate long-branch attraction in phylogenomic reconstruction of Cupressoideae (Cupressaceae). Sci. Rep. 7, 1-11. https://doi.org/10.1038/srep41005
Quandt, C.A., Beaudet, D., Corsaro, D., Walochnik, J., Michel, R., Corradi, N., James, T.Y., 2017. The genome of an intranuclear parasite, *Paramicrosporidium saccamoebae*, reveals alternative adaptations to obligate intracellular parasitism. Elife 6, e29594. https://doi.org/10.7554/eLife.29594
Quicke, K.M., Bowen, J.R., Johnson, E.L., McDonald, C.E., Ma, H., O'Neal, J.T., Rajakumar, A., Wrammert, J., Rimawi, B.H., Pulendran, B., Schinazi, R.F., Chakraborty, R., Suthar, M.S., 2016. Zika Virus infects human placental macrophages. Cell Host Microbe 20, 83–90. https://doi.org/10.1016/j.chom.2016.05.015
Quiles, A., Bacela-Spychalska, K., Teixeira, M., Lambin, N., Grabowski, M., Rigaud, T., Wattier, R.A., 2019. Microsporidian infections in the species complex *Gammarus roeselii* (Amphipoda) over its geographical range: evidence for both host–parasite co-diversification and recent host shifts. Parasit. Vectors 12, 327. https://doi.org/10.1186/s13071-019-3571-z
Quilter, C.G., 1976. Microsporidan parasite *Thelohania contejeani* Henneguy from New Zealand freshwater crayfish. New Zeal. J. Mar. Freshw. Res. 10, 225–231. https://doi.org/10.1080/00288330.1976.9515609
R Development Core Team, R., 2013. R: A language and environment for statistical computing, R foundation for statistical computing. https://doi.org/10.1007/978-3-540-74686-7
Ranwez, V., Chantret, N., 2020. Strengths and limits of multiple sequence alignment and filtering methods, in: Phylogenetics in the Genomic Era. 1-36.
Raoult, D., La Scola, B., Enea, M., Fournier, P.E., Roux, V., Fenollar, F., Galvao, M.A.M., De Lamballerie, X., 2001. A flea-associated *Rickettsia* pathogenic for humans. Emerg. Infect. Dis. 7, 73–81. https://doi.org/10.3201/eid0701.010112
Ratnasingham, S., Hebert, P.D.N., 2013. A DNA-based registry for all animal species:

The Barcode Index Number (BIN) System. PLoS One 8, e66213. https://doi.org/10.1371/journal.pone.0066213
Raupach, M.J., Hannig, K., Morinière, J., Hendrich, L., 2016. A DNA barcode library for ground beetles (Insecta, Coleoptera, Carabidae) of Germany: The genus *Bembidion Latreille*, 1802 and allied taxa. Zookeys 2016, 121–141. https://doi.org/10.3897/zookeys.592.8316
Raupach, M.J., Radulovici, A.E., 2015. Looking back on a decade of barcoding crustaceans. Zookeys 2015, 53–81. https://doi.org/10.3897/zookeys.539.6530
Reddy, A.K., Balne, P.K., Garg, P., Krishnaiah, S., 2011. Is microsporidial keratitis a seasonal infection in India? Clin. Microbiol. Infect. 17, 1114–1116. https://doi.org/10.1111/j.1469-0691.2010.03084.x
Reeves, W.K., Kato, C.Y., Gilchriest, T., 2008. Pathogen screening and bionomics of *Lutzomyia apache* (Diptera: Psychodidae) in wyoming, USA. J. Am. Mosq. Control Assoc. 24, 444–447. https://doi.org/10.2987/5745.1
Refrégier, G., Le Gac, M., Jabbour, F., Widmer, A., Shykoff, J.A., Yockteng, R., Hood, M.E., Giraud, T., 2008. Cophylogeny of the anther smut fungi and their caryophyllaceous hosts: Prevalence of host shifts and importance of delimiting parasite species for inferring cospeciation. BMC Evol. Biol. 8, 1-17. https://doi.org/10.1186/1471-2148-8-100
Reichgelt, T., Kennedy, E.M., Conran, J.G., Mildenhall, D.C., Lee, D.E., 2015. The early Miocene paleolake Manuherikia: vegetation heterogeneity and warm-temperate to subtropical climate in southern New Zealand. J. Paleolimnol. 53, 349–365. https://doi.org/10.1007/s10933-015-9827-5
Řezáč, M., Gasparo, F., Král, J., Heneberg, P., 2014. Integrative taxonomy and evolutionary history of a newly revealed spider *Dysdera ninnii* complex (Araneae: Dysderidae). Zool. J. Linn. Soc. 172, 415-474. https://doi.org/10.1111/zoj.12177
Rigaud, T., Perrot-Minnot, M., Brown, M., 2010. Parasite and host assemblages: Embracing the reality will improve our knowledge of parasite transmission and virulence. Proc. R. Soc. B Biol. Sci. 277, 3696-3702.
Robertson, D.R., 1996. Interspecific competition controls abundance and habitat use of territorial caribbean damselfishes. Ecology 77, 885–899. https://doi.org/10.2307/2265509
Roger, A.J., Muñoz-Gómez, S.A., Kamikawa, R., 2017. The origin and diversification of Mitochondria. Curr. Biol. 27, 1177–1192. https://doi.org/10.1016/j.cub.2017.09.015
Rohwer, F., Seguritan, V., Azam, F., Knowlton, N., 2002. Diversity and distribution of coral-associated bacteria. Mar. Ecol. Prog. Ser. 243, 1–10. https://doi.org/10.3354/meps243001
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M., Huelsenbeck, J., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029
Roughgarden, J., Gilbert, S.F., Rosenberg, E., Zilber-Rosenberg, I., Lloyd, E.A., 2018. Holobionts as units of selection and a model of their population dynamics and evolution. Biol. Theory 13, 44–65. https://doi.org/10.1007/s13752-017-0287-1
Rousset, F., Bouchon, D., Pintureau, B., Juchault, P., Solignac, M., 1992. *Wolbachia* endosymbionts responsible for various alterations of sexuality in arthropods. Proc. R. Soc. B Biol. Sci. 250, 91–98. https://doi.org/10.1098/rspb.1992.0135

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248

Ryman, N., Laikre, L., Hössjer, O., 2019. Do estimates of contemporary effective population size tell us what we want to know? Mol. Ecol. 28, 1904–1918. https://doi.org/10.1111/mec.15027

Sabaneyeva, E., Castelli, M., Szokoli, F., Benken, K., Lebedeva, N., Salvetti, A., Schweikert, M., Fokin, S., Petroni, G., 2018. Host and symbiont intraspecific variability: The case of *Paramecium calkins*i and "Candidatus Trichorickettsia mobilis." Eur. J. Protistol. 62, 79–94. https://doi.org/10.1016/j.ejop.2017.12.002

Sagristà, E., Bozzo, M.G., Bigas, M., Poquet, M., Durfort, M., 1998. Developmental cycle and ultrastructure of *Steinhausia mytilovum*, a microsporidian parasite of oocytes of the mussel, *Mytilus galloprovincialis* (Mollusca, Bivalvia). Eur. J. Protistol. 34, 58–68. https://doi.org/10.1016/S0932-4739(98)80040-0

Sahni, S.K., Rydkina, E., 2009. Host-cell interactions with pathogenic *Rickettsia* species. Future Microbiol. 2009, 323-339. https://doi.org/10.2217/fmb.09.6

Sanders, J.L., Watral, V., Kent, M.L., 2012. Microsporidiosis in zebrafish research facilities. ILAR J. 53, 106–113. https://doi.org/10.1093/ilar.53.2.106

Sanderson, M.J., 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302. https://doi.org/10.1093/bioinformatics/19.2.301

Sandholzer, L., Nostrand, T., Young, L., 1945. Studies on an ichthyosporidian-like parasite of ocean pout (*Zoareces anguillaris*). United States Dep. Inter. Fish Wildl. Serv.

Sapir, A., Dillman, A.R., Connon, S.A., Grupe, B.M., Ingels, J., Mundo-Ocampo, M., Levin, L.A., Baldwin, J.G., Orphan, V.J., Sternberg, P.W., 2014. Microsporidia-nematode associations in methane seeps reveal basal fungal parasitism in the deep sea. Front. Microbiol. 5, 43. https://doi.org/10.3389/fmicb.2014.00043

Scheid, P., 2007. Mechanism of intrusion of a microspordian-like organism into the nucleus of host amoebae (*Vannella* sp.) isolated from a keratitis patient. Parasitol. Res. 101, 1097–1102. https://doi.org/10.1007/s00436-007-0590-z

Scholz, F., Fringuelli, E., Bolton-Warberg, M., Marcos-López, M., Mitchell, S., Prodhol, P., Moffet, D., Savage, P., Murphy O'Sullivan, S., O'Connor, I., McCarthy, E., Rodger, H.D., 2017. First record of *Tetramicra brevifilum* in lumpfish (*Cyclopterus lumpus*, L.). J. Fish Dis. 40, 757–771. https://doi.org/10.1111/jfd.12554

Schrallhammer, M., Ferrantini, F., Vannini, C., Galati, S., Schweikert, M., Görtz, H.D., Verni, F., Petroni, G., 2013. "Candidatus Megaira polyxenophila" gen. nov., sp. nov.: Considerations on evolutionary history, host range and shift of early divergent Rickettsiae. PLoS One 8, e72581. https://doi.org/10.1371/journal.pone.0072581

Schulenburg, J.H.G.V.D., Hurst, G.D.D., Tetzlaff, D., Booth, G.E., Zakharov, I.A., Majerus, M.E.N., 2002. History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA sequence analysis. Genetics 160, 1075–1086.

Scott, M.E., Dobson, A., 1989. The role of parasites in regulating host abundance. Parasitol. Today. 5, 176-183. https://doi.org/10.1016/0169-4758(89)90140-3

Scrimgeour, G.J., Winterbourn, M.J., 1989. Effects of floods on epilithon and benthic macroinvertebrate populations in an unstable New Zealand river. Hydrobiologia 171, 33–44. https://doi.org/10.1007/BF00005722

Shafer, A.B.A., Williams, G.R., Shutler, D., Rogers, R.E.L., Stewart, D.T., 2009. Cophylogeny of *Nosema* (Microsporidia: Nosematidae) and bees (Hymenoptera: Apidae) suggests both cospeciation and a host-switch. J. Parasitol. 95, 198–203. https://doi.org/10.1645/GE-1724.1

Shoemaker, D.D., Katju, V., Jaenike, J., 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. Evolution (N. Y). 53, 1157–1164. https://doi.org/10.1111/j.1558-5646.1999.tb04529.x

Short, S., Guler, Y., Yang, G., Kille, P., Ford, A.T., 2012. Paramyxean-microsporidian co-infection in amphipods: Is the consensus that Microsporidia can feminise their hosts presumptive? Int. J. Parasitol. 42, 683–691. https://doi.org/10.1016/j.ijpara.2012.04.014

Shylla, J.A., Ghatani, S., Tandon, V., 2013. Utility of divergent domains of 28S ribosomal RNA in species discrimination of paramphistomes (Trematoda: Digenea: Paramphistomoidea). Parasitol. Res. 112, 4239–4253. https://doi.org/10.1007/s00436-013-3616-8

Sibley, L.D., 2004. Intracellular parasite invasion strategies. Science 304, 248-253. https://doi.org/10.1126/science.1094717

Siddall, M.E., Fontanella, F.M., Watson, S.C., Kvist, S., Erséus, C., 2009. Barcoding bamboozled by bacteria: convergence to metazoan mitochondrial primer targets by marine microbes. Syst. Biol. 58, 445-451. https://doi.org/10.1093/sysbio/syp033

Slamovits, C.H., Williams, B.A.P., Keeling, P.J., 2004. Transfer of *Nosema locustae* (Microsporidia) to *Antonospora locustae* n. comb. based on molecular and ultrastructural data. J. Eukaryot. Microbiol. 51, 207–213. https://doi.org/10.1111/j.1550-7408.2004.tb00547.x

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science 236, 787–792. https://doi.org/10.1126/science.3576198

Slothouber Galbreath, J.G., Smith, J.E., Terry, R.S., Becnel, J.J., Dunn, A.M., 2004. Invasion success of *Fibrillanosema crangonycis*, n. sp., ng: a novel vertically transmitted microsporidian parasite from the invasive amphipod host *Crangonyx pseudogracilis*. Int. J. Parasitol. 34, 235–244. https://doi.org/10.1016/j.ijpara.2003.10.009

Smith, M.A., Bertrand, C., Crosby, K., Eveleigh, E.S., Fernandez-Triana, J., Fisher, B.L., Gibbs, J., Hajibabaei, M., Hallwachs, W., Hind, K., Hrcek, J., Huang, D.W., Janda, M., Janzen, D.H., Li, Y., Miller, S.E., Packer, L., Quicke, D., Ratnasingham, S., Rodriguez, J., Rougerie, R., Shaw, M.R., Sheffield, C., Stahlhut, J.K., Steinke, D., Whitfield, J., Wood, M., Zhou, X., 2012. *Wolbachia* and DNA barcoding insects: patterns, potential, and problems. PLoS One 7, e36514. https://doi.org/10.1371/journal.pone.0036514

Snowden, K.F., 2014. Microsporidia in higher vertebrates, in: Microsporidia: Pathogens of Opportunity: First Edition. 469–491. John Wiley & Sons. https://doi.org/10.1002/9781118395264.ch19

Sokolova, Y.Y., Senderskiy, I. V., Tokarev, Y.S., 2016. Microsporidia *Alfvenia sibirica* sp. n. and *Agglomerata cladocera* (Pfeiffer) 1895, from Siberian microcrustaceans and phylogenetic relationships within the “Aquatic outgroup” lineage of fresh

water microsporidia. J. Invertebr. Pathol. 136, 81–91. https://doi.org/10.1016/j.jip.2016.03.009
Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. U. S. A. 105, 13486–13491. https://doi.org/10.1073/pnas.0803076105
Stamatakis, A., 2014. The RAxML v8 . 1 . X Manual (Updated). Tutoriais e Manuais 1–61.
Stensvold, C.R., Nielsen, H.V., 2012. Comparison of microscopy and PCR for detection of intestinal parasites in Danish patients supports an incentive for molecular screening platforms. J. Clin. Microbiol. 50, 540-541. https://doi.org/10.1128/JCM.06012-11
Stentiford, G.D., Bass, D., Williams, B.A.P., 2019. Ultimate opportunists—the emergent enterocytozoon group microsporidia. PLoS Pathog. 15, e1007668. https://doi.org/10.1371/journal.ppat.1007668
Stentiford, G.D., Bateman, K.S., Feist, S.W., Chambers, E., Stone, D.M., 2013a. Plastic parasites: extreme dimorphism creates a taxonomic conundrum in the phylum Microsporidia. Int. J. Parasitol. 43, 339–352. https://doi.org/10.1016/j.ijpara.2012.11.010
Stentiford, G.D., Bateman, K.S., Small, H.J., Moss, J., Shields, J.D., Reece, K.S., Tuck, I., 2010. *Myospora metanephrops* (n. g., n. sp.) from marine lobsters and a proposal for erection of a new order and family (Crustaceacida; Myosporidae) in the class Marinosporidia (Phylum Microsporidia). Int. J. Parasitol. 40, 1433–1446. https://doi.org/10.1016/j.ijpara.2010.04.017
Stentiford, G.D., Becnel, J.J., Weiss, L.M., Keeling, P.J., Didier, E.S., Bjornson, S., Freeman, M.A., Brown, M.J.F., Roesel, K., Sokolova, Y., 2016. Microsporidia–emergent pathogens in the global food chain. Trends Parasitol. 32, 336–348. https://doi.org/10.1016/j.pt.2015.12.004
Stentiford, G.D., Feist, S.W., Stone, D.M., Bateman, K.S., Dunn, A.M., 2013b. Microsporidia: diverse, dynamic, and emergent pathogens in aquatic systems. Trends Parasitol. 29, 567–578. https://doi.org/10.1016/j.pt.2013.08.005
Stentiford, G.D., Ramilo, A., Abollo, E., Kerr, R., Bateman, K.S., Feist, S.W., Bass, D., Villalba, A., 2017. *Hyperspora aquatica* n.gn., n.sp. (Microsporidia), hyperparasitic in *Marteilia cochillia* (Paramyxida), is closely related to crustacean-infecting microspordian taxa. Parasitology 144, 186–199. https://doi.org/10.1017/S0031182016001633
Stevens, G.R., Stevens, D.L., 1980. New Zealand adrift: the theory of continental drift in a New Zealand setting. Reed Wellington.
Stevens, M.I., Hogg, I.D., Pilditch, C.A., 2006. Evidence for female-biased juvenile dispersal in corophiid amphipods from a New Zealand estuary. J. Exp. Mar. Bio. Ecol. 331, 9–20. https://doi.org/10.1016/j.jembe.2005.09.018
Stewart, A.D., Logsdon, J.M., Kelley, S.E., 2005. An empirical study of the evolution of virulence under both horizontal and vertical transmission. Evolution (N. Y). 59, 730–739. https://doi.org/10.1111/j.0014-3820.2005.tb01749.x
Stoch, F., Christian, E., Flot, J.F., 2020. Molecular taxonomy, phylogeny and biogeography of the *Niphargus tatrensis* species complex (Amphipoda, Niphargidae) in Austria. Org. Divers. Evol. 20, 701–722. https://doi.org/10.1007/s13127-020-00462-z

Stouthamer, R., 1997. *Wolbachia*-induced parthenogenesis, in: Influential Passengers: Inherited Microorganisms and Arthropod Reproduction. Oxford University Press. 102-124.
Streicher, J.W., Day, J.J., 2020. The toad's warts: Discordance creates bumpy expectations of mitochondrial-nuclear evolution between species. Mol. Ecol. 29, 3400–3402. https://doi.org/10.1111/mec.15568
Strogen, D.P., Bland, K.J., Nicol, A., King, P.R., 2014. Paleogeography of the Taranaki Basin region during the latest Eocene–early Miocene and implications for the 'total drowning'of Zealandia. New Zeal. J. Geol. Geophys. 57, 110–127. https://doi.org/10.1080/00288306.2014.901231
Sturmbauer, C., Meyer, A., 1992. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. Nature 358, 578–581. https://doi.org/10.1038/358578a0
Sutherland, D.L., Hogg, I.D., Waas, J.R., 2009. Phylogeography and species discrimination in the *Paracalliope fluviatilis* species complex (Crustacea: Amphipoda): can morphologically similar heterospecifics identify compatible mates? Biol. J. Linn. Soc. 99, 196–205. https://doi.org/10.1111/j.1095-8312.2009.01344.x
Sweet, A.D., Johnson, K.P., 2018. The role of parasite dispersal in shaping a host–parasite system at multiple evolutionary scales. Mol. Ecol. 27, 5104–5119. https://doi.org/10.1111/mec.14937
Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x
Takezaki, N., Rzhetsky, A., Nei, M., 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12, 823–833. https://doi.org/10.1093/oxfordjournals.molbev.a040259
Tamim El Jarkass, H., Reinke, A.W., 2020. The ins and outs of host-microsporidia interactions during invasion, proliferation and exit. Cell. Microbiol. 2020, e13247. https://doi.org/10.1111/cmi.13247
Temereva, E.N., Sokolova, Y.Y., 2018. A Microsporidian Infection in phoronids (Phylum Phoronida): *Microsporidium phoronidi* n. sp. from a *Phoronis embryolabi*. J. Eukaryot. Microbiol. 65, 427–431. https://doi.org/10.1111/jeu.12490
Terry, R.S., Dunn, A.M., Smith, J.E., 1997. Cellular distribution of a feminizing microsporidian parasite: a strategy for transovarial transmission. Parasitology 115, 157–163. https://doi.org/10.1017/S0031182097001236
Terry, R.S., Smith, J.E., Dunn, A.M., 1998. Impact of a novel, feminising microsporidium on its crustacean host. J. Eukaryot. Microbiol. 45, 497–501. https://doi.org/10.1111/j.1550-7408.1998.tb05106.x
Terry, R.S., Smith, J.E., Sharpe, R.G., Rigaud, T., Timothy, D., Littlewood, J., Ironside, J.E., Rollinson, D., Bouchon, D., Macneil, C., Dick, J.T.A., Dunn, A.M., 2004. Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum Microspora. Proc. R. Soc. B Biol. Sci. 271, 1783–1789. https://doi.org/10.1098/rspb.2004.2793
Tham, A.C., Sanjay, S., 2012. Clinical spectrum of microsporidial keratoconjunctivitis. Clin. Exp. Ophthalmol. 40, 512-518. https://doi.org/10.1111/j.1442-9071.2011.02712.x

Thiel, M., 1999. Duration of extended parental care in marine amphipods. J. Crustac. Biol. 19, 60–71. https://doi.org/10.2307/1549547
Thomas, F., Poulin, R., 1998. Manipulation of a mollusc by a trophically transmitted parasite: convergent evolution or phylogenetic inheritance? Parasitology 116, 431–436.
Thompson, J.R., Rivera, H.E., Closek, C.J., Medina, M., 2014. Microbes in the coral holobiont: partners through evolution, development, and ecological interactions. Front. Cell. Infect. Microbiol. 4, 176. https://doi.org/10.3389/fcimb.2014.00176
Thomson, 1987. New Zealand Crustacea, with descriptions of new species. Trans. Proc. New Zeal. Inst. 11, 230–248.
Thongprem, P., Davison, H.R., Thompson, D.J., Lorenzo-Carballa, M.O., Hurst, G.D.D., 2020. Incidence and diversity of Torix *Rickettsia*–Odonata Symbioses. Microb. Ecol. 81, 203-212. https://doi.org/10.1007/s00248-020-01568-9
Thum, R.A., Harrison, R.G., 2009. Deep genetic divergences among morphologically similar and parapatric *Skistodiaptomus* (Copepoda: Calanoida: Diaptomidae) challenge the hypothesis of Pleistocene speciation. Biol. J. Linn. Soc. 96, 150–165. https://doi.org/10.1111/j.1095-8312.2008.01105.x
Timofeev, S., Tokarev, Y., Dolgikh, V., 2020. Energy metabolism and its evolution in Microsporidia and allied taxa. Parasitol. Res. 119, 1433-1441. https://doi.org/10.1007/s00436-020-06657-9
Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907-3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x
Tokarev, Y.S., Huang, W.F., Solter, L.F., Malysh, J.M., Becnel, J.J., Vossbrinck, C.R., 2020. A formal redefinition of the genera *Nosema* and *Vairimorpha* (Microsporidia: Nosematidae) and reassignment of species based on molecular phylogenetics. J. Invertebr. Pathol. 169, 107279. https://doi.org/10.1016/j.jip.2019.107279
Tóth, J.P., Varga, Z., Verovnik, R., Wahlberg, N., Váradi, A., Bereczki, J., 2017. Mito-nuclear discordance helps to reveal the phylogeographic patterns of *Melitaea ornata* (Lepidoptera: Nymphalidae). Biol. J. Linn. Soc. 121, 267–281. https://doi.org/10.1093/BIOLINNEAN/BLW037
Tourtip, S., Wongtripop, S., Stentiford, G.D., Bateman, K.S., Sriurairatana, S., Chavadej, J., Sritunyalucksana, K., Withyachumnarnkul, B., 2009. *Enterocytozoon hepatopenaei* sp. nov. (Microsporida: Enterocytozoonidae), a parasite of the black tiger shrimp *Penaeus monodon* (Decapoda: Penaeidae): Fine structure and phylogenetic relationships. J. Invertebr. Pathol. 102, 21–29. https://doi.org/10.1016/j.jip.2009.06.004
Trewick, S.A., Wallis, G.P., Morgan-Richards, M., 2011. The invertebrate life of New Zealand: a phylogeographic approach. Insects 2, 297–325. https://doi.org/10.3390/insects2030297
Troemel, E.R., Félix, M.A., Whiteman, N.K., Barrière, A., Ausubel, F.M., 2008. Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. PLoS Biol. 6, e309. https://doi.org/10.1371/journal.pbio.0060309
Tsaousis, A.D., Kunji, E.R.S., Goldberg, A. V., Lucocq, J.M., Hirt, R.P., Embley, T.M., 2008. A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. Nature 453, 553–556. https://doi.org/10.1038/nature06903

Turelli, M., Hoffmann, A.A., McKechnie, S.W., 1992. Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics 132, 713–723.
Väinölä, R., Witt, J.D.S., Grabowski, M., Bradbury, J.H., Jazdzewski, K., Sket, B., 2008. Global diversity of amphipods (Amphipoda; Crustacea) in freshwater. Hydrobiologia 595, 241–255. https://doi.org/10.1007/s10750-007-9020-6
Vávra, J., Fiala, I., Krylová, P., Petrusek, A., Hyliš, M., 2018. Molecular and structural assessment of microsporidia infecting daphnids: The "obtusa-like" microsporidia, a branch of the monophyletic Agglomeratidae clade, with the establishment of a new genus *Conglomerata*. J. Invertebr. Pathol. 159, 95–104. https://doi.org/10.1016/j.jip.2018.10.003
Vavra, J., Hylis, M., Vossbrinck, C.R., Pilarska, D.K., Linde, A., Weiser, J., McManus, M.L., Hoch, G., Solter, L.F., 2006. *Vairimorpha disparis* n. comb. (Microsporidia: Burenellidae): A redescription and taxonomic revision of *Thelohania disparis* Timofejeva 1956, a microsporidian parasite of the gypsy moth *Lymantria dispar* (L.) (Lepidoptera: Lymantriidae). J. Eukaryot. Microbiol. 53, 292–304. https://doi.org/10.1111/j.1550-7408.2006.00108.x
Vávra, J., Larsson, J.I.R., 2014. Structure of Microsporidia, in: Microsporidia: Pathogens of Opportunity: First Edition. John Wiley & Sons. 1–70. https://doi.org/10.1002/9781118395264.ch1
Verheye, M.L., Martin, P., Backeljau, T., D'Acoz, C., 2016. DNA analyses reveal abundant homoplasy in taxonomically important morphological characters of Eusiroidea (Crustacea, Amphipoda). Zool. Scr. 45, 300–321. https://doi.org/10.1111/zsc.12153
Vestheim, H., Deagle, B.E., Jarman, S.N., 2011. Application of blocking oligonucleotides to improve signal-to-noise ratio in a PCR. Methods Mol. Biol. 687, 265–274. https://doi.org/10.1007/978-1-60761-944-4_19
Vestheim, H., Jarman, S.N., 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples - A case study on prey DNA in Antarctic krill stomachs. Front. Zool. 5, 1-11. https://doi.org/10.1186/1742-9994-5-12
Vossbrinck, C.R., Debrunner-Vossbrinck, B.A., 2005. Molecular phylogeny of the Microsporidia: ecological, ultrastructural and taxonomic considerations. Folia Parasitol. (Praha). 52, 131. https://doi.org/10.14411/fp.2005.017
Vossbrinck, C.R., Debrunner-Vossbrinck, B.A., Weiss, L.M., 2014. Phylogeny of the Microsporidia. Microsporidia Pathog. Oppor. First Ed. John Wiley & Sons. 203–220. https://doi.org/10.1002/9781118395264.ch6
Vossbrinck, C.R., Woese, C.R., 1986. Eukaryotic ribosomes that lack a 5.8S RNA. Nature 320, 287–288. https://doi.org/10.1038/320287a0
Vucetich, J.A., Waite, T.A., Nunney, L., 1997. Fluctuating population size and the ratio of effective to census population size. Evolution (N. Y). 51, 2017–2021. https://doi.org/10.1111/j.1558-5646.1997.tb05123.x
Wadi, L., Reinke, A.W., 2020. Evolution of microsporidia: An extremely successful group of eukaryotic intracellular parasites. PLoS Pathog. 16, e1008276. https://doi.org/10.1371/journal.ppat.1008276
Walker D.H., 1995. Rocky Mountain spotted fever: a seasonal alert. Clin. Infect. Dis. 20, 1111–1117.
Wallis, G.P., Jorge, F., 2018. Going under down under? Lineage ages argue for extensive survival of the Oligocene marine transgression on Zealandia. Mol. Ecol.

27, 4368–4396. https://doi.org/10.1111/mec.14875
Wang, H.L., Lei, T., Wang, X.W., Maruthi, M.N., Zhu, D.T., Cameron, S.L., Rao, Q., Shan, H.W., Colvin, J., Liu, Y.Q., Liu, S.S., 2020. A newly recorded *Rickettsia* of the Torix group is a recent intruder and an endosymbiont in the whitefly *Bemisia tabaci*. Environ. Microbiol. 22, 1207-1221. https://doi.org/10.1111/1462-2920.14927
Wang, Z., Wu, M., 2015. An integrated phylogenomic approach toward pinpointing the origin of mitochondria. Sci. Rep. 5, 1-12. https://doi.org/10.1038/srep07949
Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D.N., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. B Biol. Sci. 360, 1847–1857. https://doi.org/10.1098/rstb.2005.1716
Watanabe, K., Yukuhiro, F., Matsuura, Y., Fukatsu, T., Noda, H., 2014. Intrasperm vertical symbiont transmission. Proc. Natl. Acad. Sci. U. S. A. 111, 7433–7437. https://doi.org/10.1073/pnas.1402476111
Waters, J.M., Wallis, G.P., 2001. Mitochondrial DNA phylogenetics of the *Galaxias vulgaris* complex from South Island, New Zealand: rapid radiation of a species flock. J. Fish Biol. 58, 1166–1180. https://doi.org/10.1006/jfbi.2000.1520
Weckstein, J.D., 2004. Biogeography explains cophylogenetic patterns in toucan chewing lice. Systematic Biology. 53, 154–164. https://doi.org/10.1080/10635150490265085
Weinert, L.A., 2015. The diversity and phylogeny of Rickettsia, in: Morand, S., Krasnov, B.R., Littlewood, D.T.J. (Eds.), Parasite diversity and diversification: evolutionary ecology meets phylogenetics. Cambridge University Press, 150–181. https://doi.org/10.1017/CBO9781139794749.010
Weinert, L.A., Werren, J.H., Aebi, A., Stone, G.N., Jiggins, F.M., 2009. Evolution and diversity of *Rickettsia* bacteria. BMC Biol. 7, 1–15. https://doi.org/10.1186/1741-7007-7-6
Weiss, L.M., Vossbrinck, C.R., 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia, in: The Microsporidia and Microsporidiosis. American Society of Microbiology, 129–171. https://doi.org/10.1128/9781555818227.ch4
Weiss, L.M., Zhu, X., Cali, A., Tanowitz, H.B., Wittner, M., 1994. Utility of microsporidian rRNA in diagnosis and phylogeny. Folia Parasitol. (Praha). 41, 81–90.
Weiss, M., Weigand, H., Weigand, A.M., Leese, F., 2018. Genome-wide single-nucleotide polymorphism data reveal cryptic species within cryptic freshwater snail species—The case of the *Ancylus fluviatilis* species complex. Ecol. Evol. 8, 1063–1072. https://doi.org/10.1002/ece3.3706
Welton, J.S., Clarke, R.T., 1980. Laboratory studies on the reproduction and growth of the amphipod, *Gammarus pulex* (L.). J. Anim. Ecol. 49, 581. https://doi.org/10.2307/4265
Werren, J.H., 1998. *Wolbachia* and speciation, in: Howard, D., Berlocher, S. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, 245–260.
Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: Master manipulators of invertebrate biology. Nat. Rev. 6, 741-751. Microbiol. https://doi.org/10.1038/nrmicro1969
White, J.A., Giorgini, M., Strand, M.R., Pennacchio, F., 2013. Arthropod endosymbiosis and evolution, in: Minelli, A., Boxshall, G., Fusco, G. (Eds.),

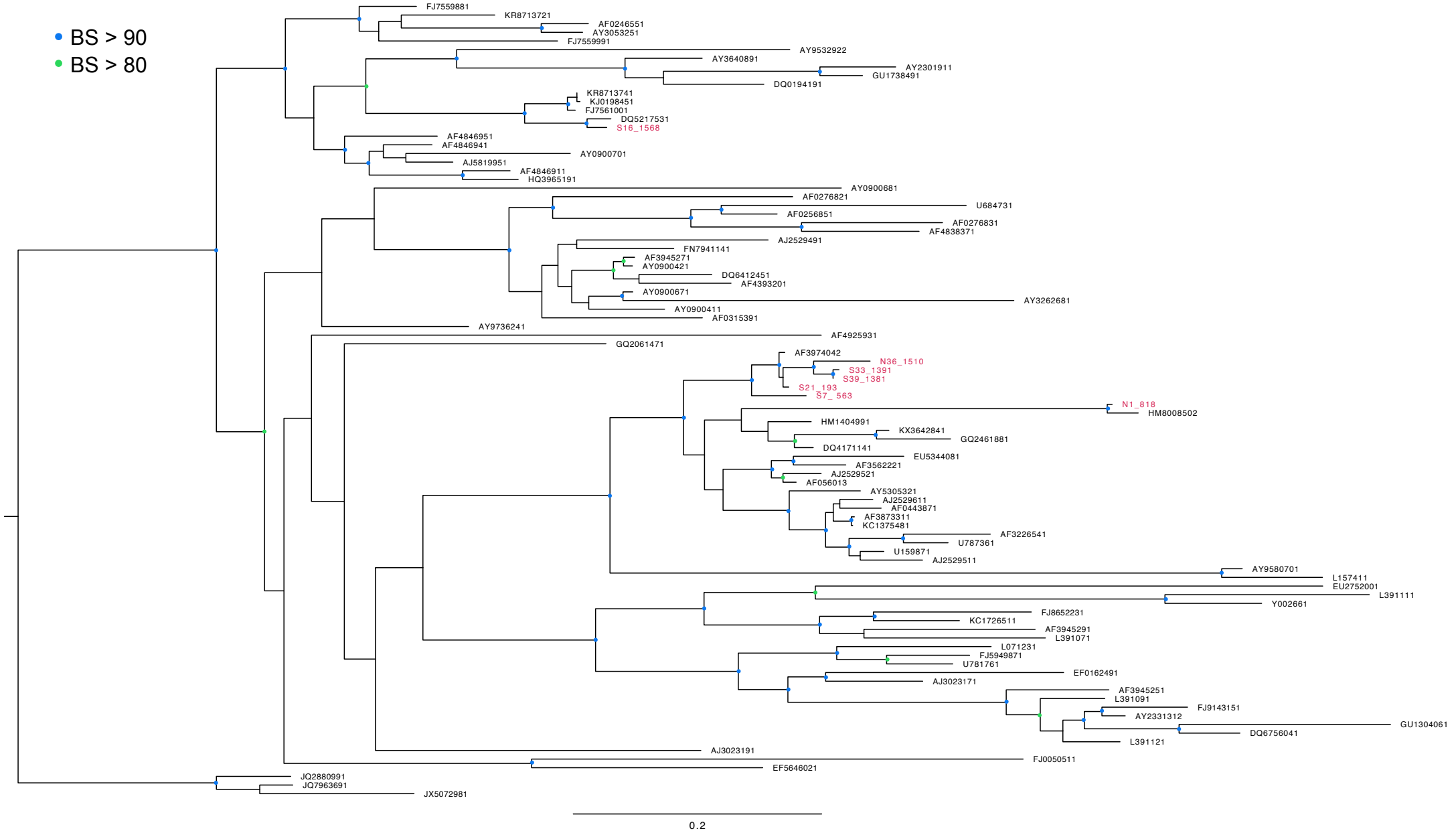
Arthropod biology and evolution: molecules, development, morphology. Springer, Berlin, Heidelberg, 441–477. https://doi.org/10.1007/978-3-642-36160-9_17
White, T.A., Campbell, B.D., Kemp, P.D., Hunt, C.L., 2000. Sensitivity of three grassland communities to simulated extreme temperature and rainfall events. Glob. Chang. Biol. 6, 671–684. https://doi.org/10.1046/j.1365-2486.2000.00344.x
Wijayawardene, N.N., Dai, D.Q., Hyde, K.D., Tang, L.Z., 2020. Outline of Fungi and fungus-like taxa. Mycosphere 11, 1060–1456. https://doi.org/10.5943/mycosphere/11/1/8
Wildish, D.J., 2012. Long distance dispersal and evolution of talitrids (Crustacea: Amphipoda: Talitridae) in the northeast Atlantic islands. J. Nat. Hist. 46, 2329–2348. https://doi.org/10.1080/00222933.2012.708458
Williams, B.A.P., Hamilton, K.M., Jones, M.D., Bass, D., 2018. Group-specific environmental sequencing reveals high levels of ecological heterogeneity across the microsporidian radiation. Environ. Microbiol. Rep. 10, 328–336. https://doi.org/10.1111/1758-2229.12642
Winters, A.D., Marsh, T.L., Brenden, T.O., Faisal, M., 2015. Analysis of bacterial communities associated with the benthic amphipod *Diporeia* in the Laurentian great lakes basin. Can. J. Microbiol. 61, 72–81. https://doi.org/10.1139/cjm-2014-0434
Witt, J.D.S., Threloff, D.L., Hebert, P.D.N., 2006. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: Implications for desert spring conservation. Mol. Ecol. 15, 3073–3082. https://doi.org/10.1111/j.1365-294X.2006.02999.x
Wolff, J.N., Ladoukakis, E.D., Enríquez, J.A., Dowling, D.K., 2014. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Philos. Trans. R. Soc. B Biol. Sci. 369, 20130443. https://doi.org/10.1098/rstb.2013.0443
Wolinska, J., Seda, J., Koerner, H., Smilauer, P., Petrusek, A., 2011. Spatial variation of *Daphnia* parasite load within individual water bodies. J. Plankton Res. 33, 1284–1294. https://doi.org/10.1093/plankt/fbr016
Woodin, S.A., 1974. Polychaete abundance patterns in a marine soft-sediment environment: the importance of biological interactions. Ecol. Monogr. 44, 171–187.
Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7. https://doi.org/10.1016/S1055-7903(02)00326-3
Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, Di., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., Stott, M.B., Nunoura, T., Banfield, J.F., Schramm, A., Baker, B.J., Spang, A., Ettema, T.J.G., 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature 541, 353–358. https://doi.org/10.1038/nature21031
Zhang, D., Tang, L., Cheng, Y., Hao, Y., Xiong, Y., Song, G., Qu, Y., Rheindt, F.E., Alström, P., Jia, C., Lei, F., 2019. “ghost Introgression” as a cause of deep mitochondrial divergence in a bird species complex. Mol. Biol. Evol. 36, 2375–2386. https://doi.org/10.1093/molbev/msz170
Zhao, Y., Zhang, W.Y., Wang, R.L., Niu, D.L., 2020. Divergent domains of 28S ribosomal RNA gene: DNA barcodes for molecular classification and identification of mites. Parasites and Vectors 13, 1-12. https://doi.org/10.1186/s13071-020-04124-z

Zhu, X., Wittner, M., Tanowitz, H.B., Kotler, D., Cali, A., Weiss, L.M., 1993. Small subunit rRNA sequence of *Enterocytozoon bieneusi* and its potential diagnostic role with use of the polymerase chain reaction. J. Infect. Dis. 168, 1570–1575.

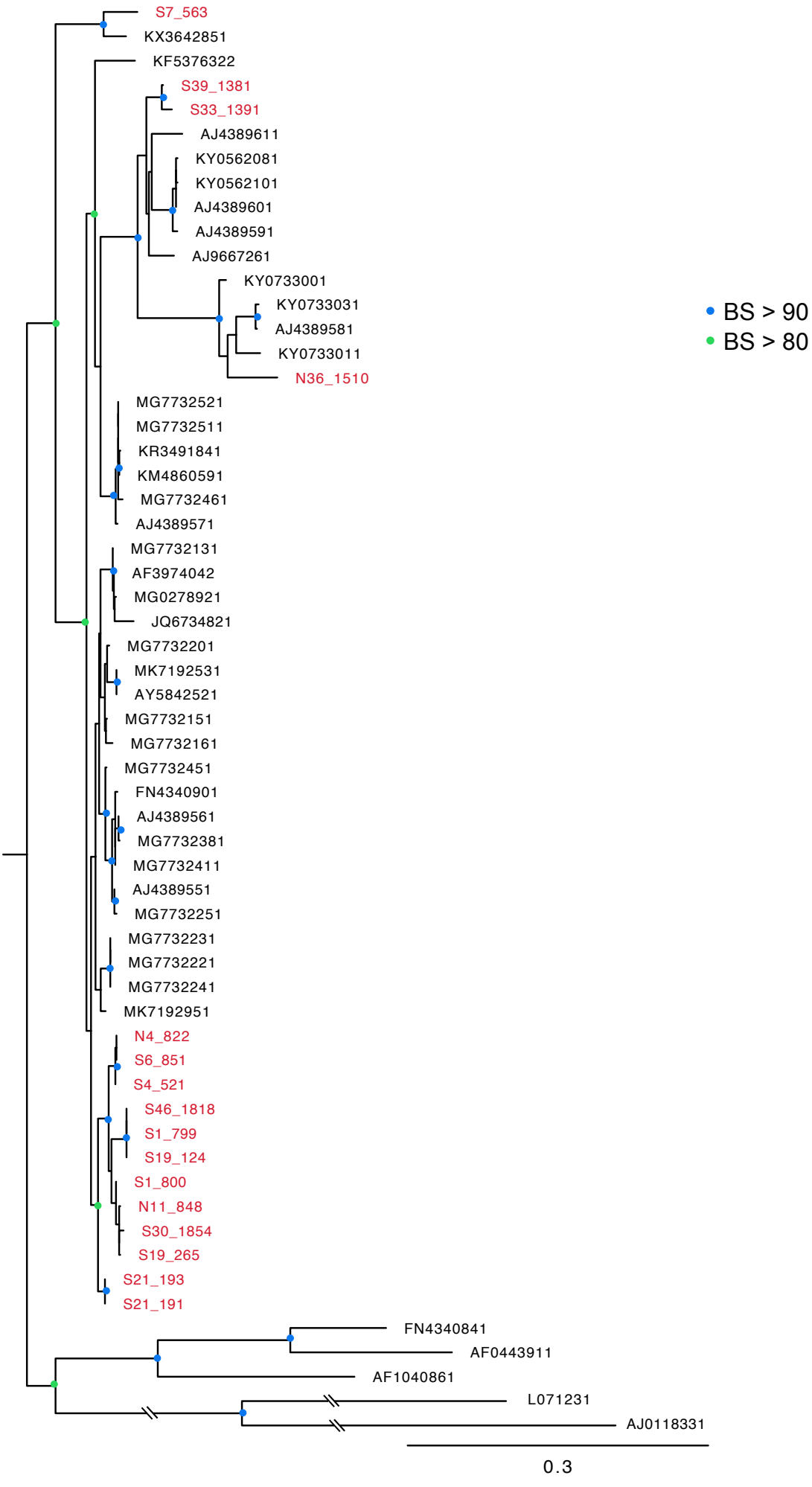
Supplementary Figures for Chapter 2

Supplementary Figure 2.1 A Maximum likelihood tree of the phylum Microsporidia
Supplementary Figure 2.2 A Maximum likelihood tree of the genus *Dictyocoela*
Supplementary Figure 2.3 Figure 2.4C with full detail
Supplementary Figure 2.4 Figure 2.5C with full detail

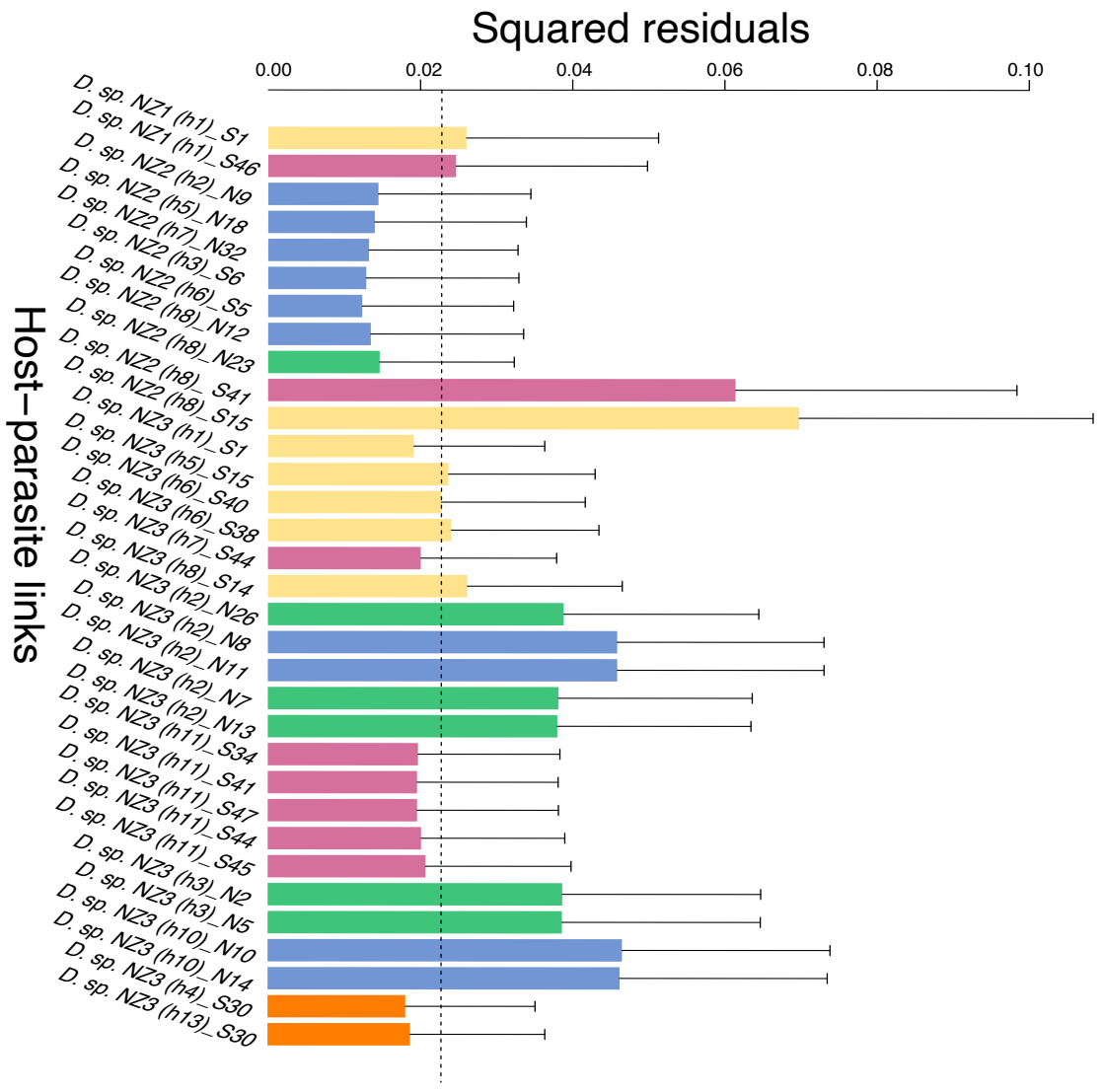
Supplementary Figure 2.1



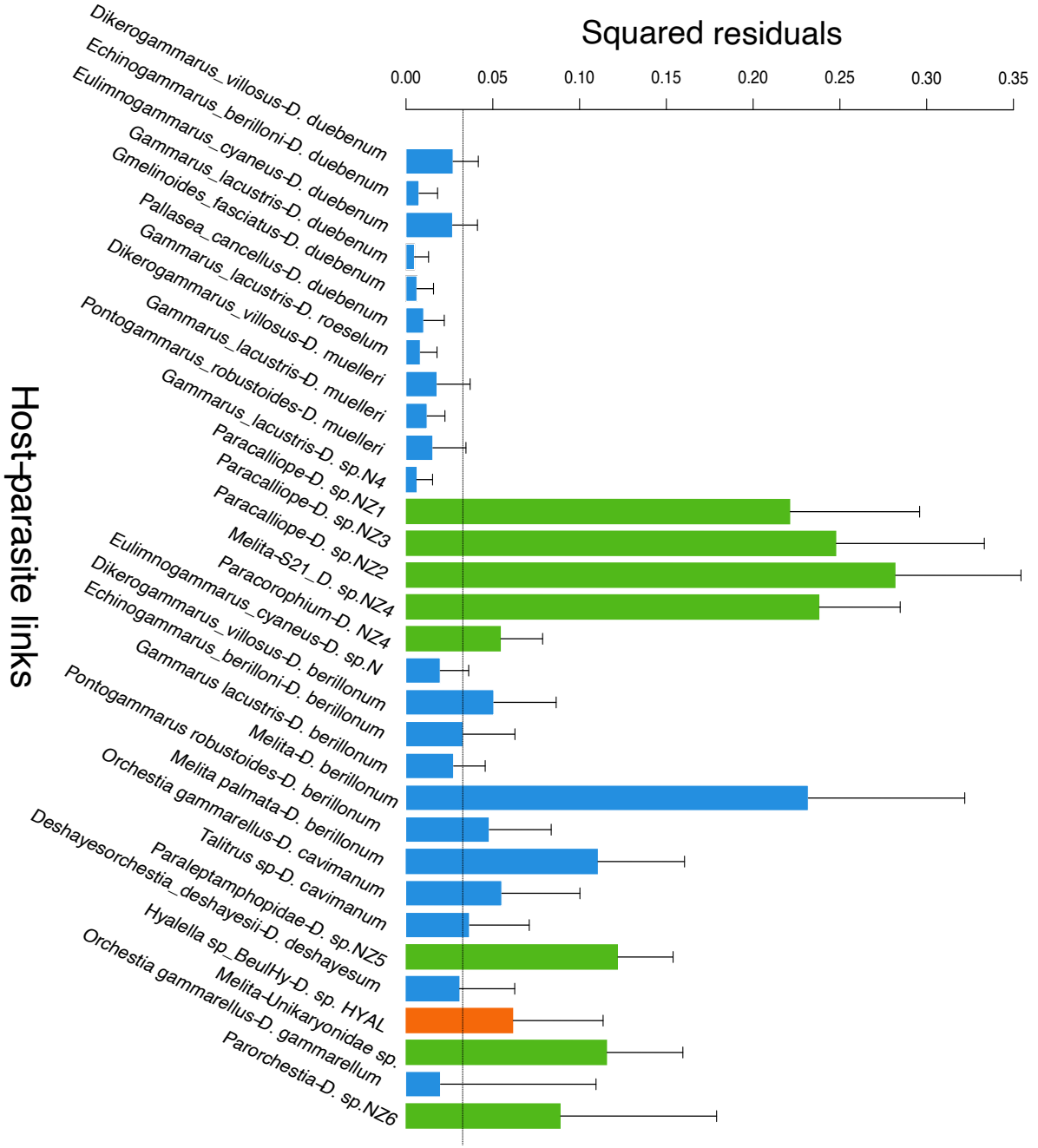
Supplementary Figure 2.2



Supplementary Figure 2.3



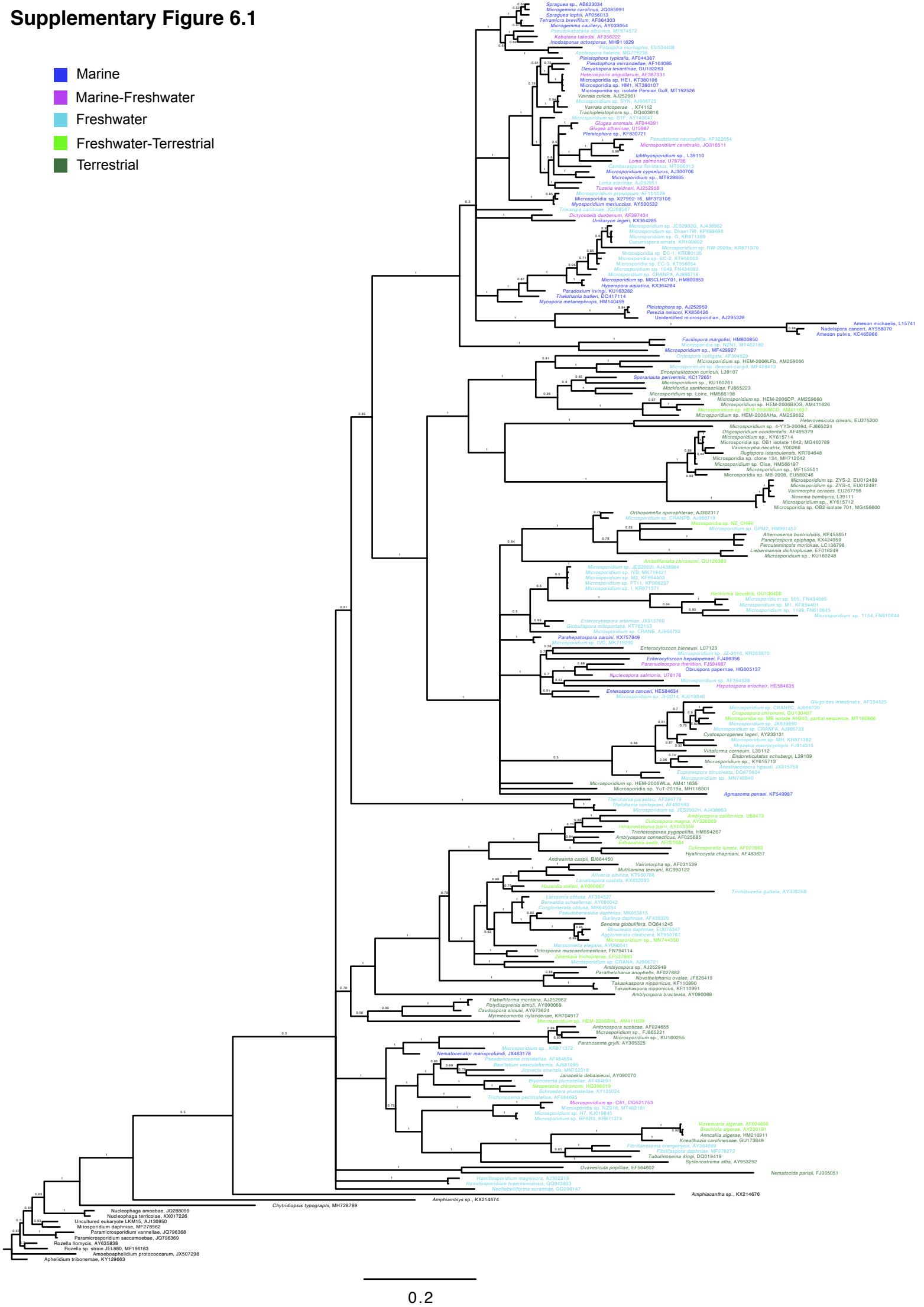
Supplementary Figure 2.4



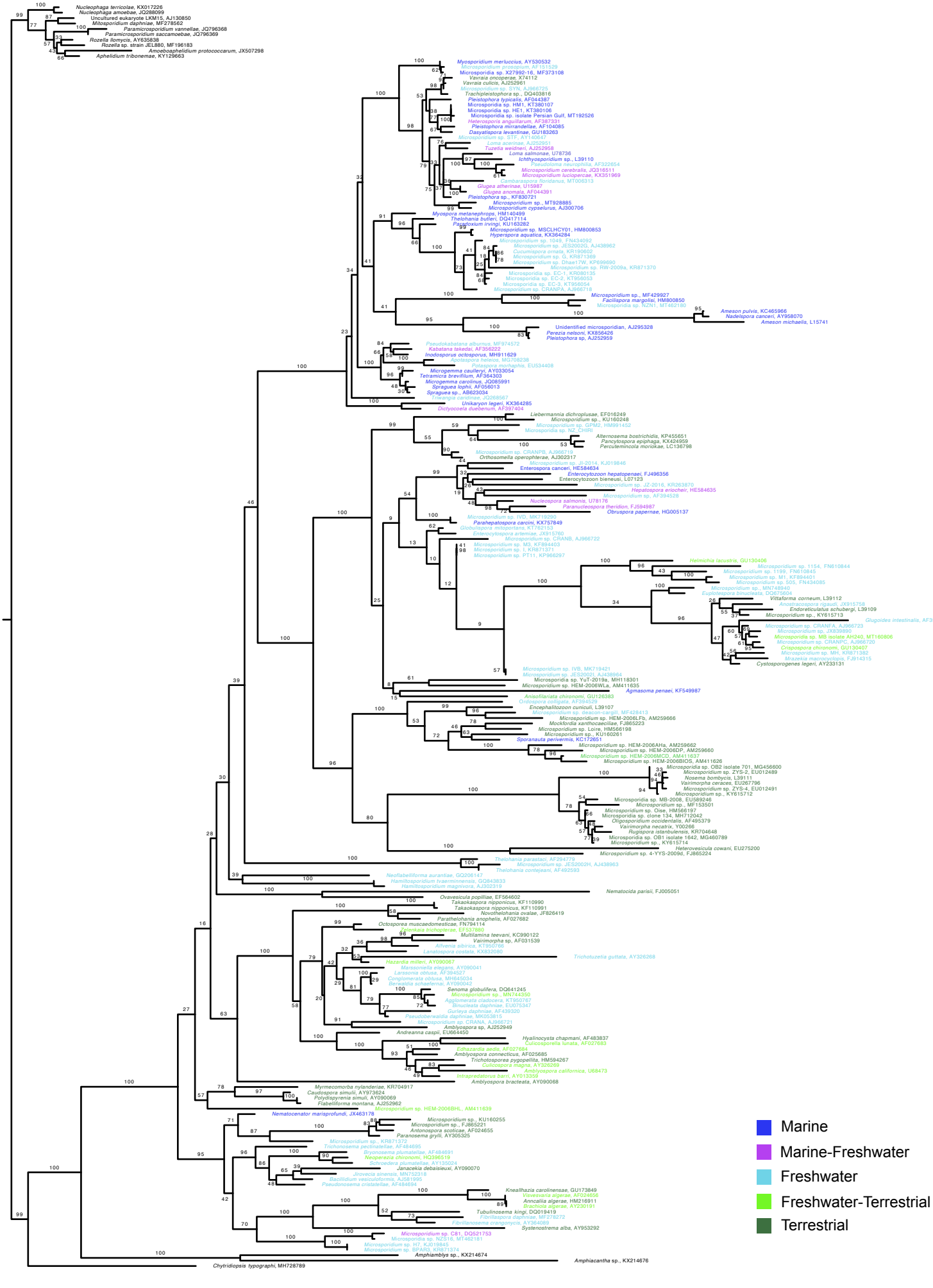
Supplementary Figures for Chapter 6

Supplementary Figure 6.1 A full Bayesian tree of canonical microsporidians
Supplementary Figure 6.2 A raxml tree of canonical microsporidians
Supplementary Figure 6.3 A rooted Bayesian tree of microsporididians and their relatives
Supplementary Figure 6.4 A rooted ML tree of microsporididians and their relatives
Supplementary Figure 6.5 An unrooted ML tree of microsporididians and their relatives

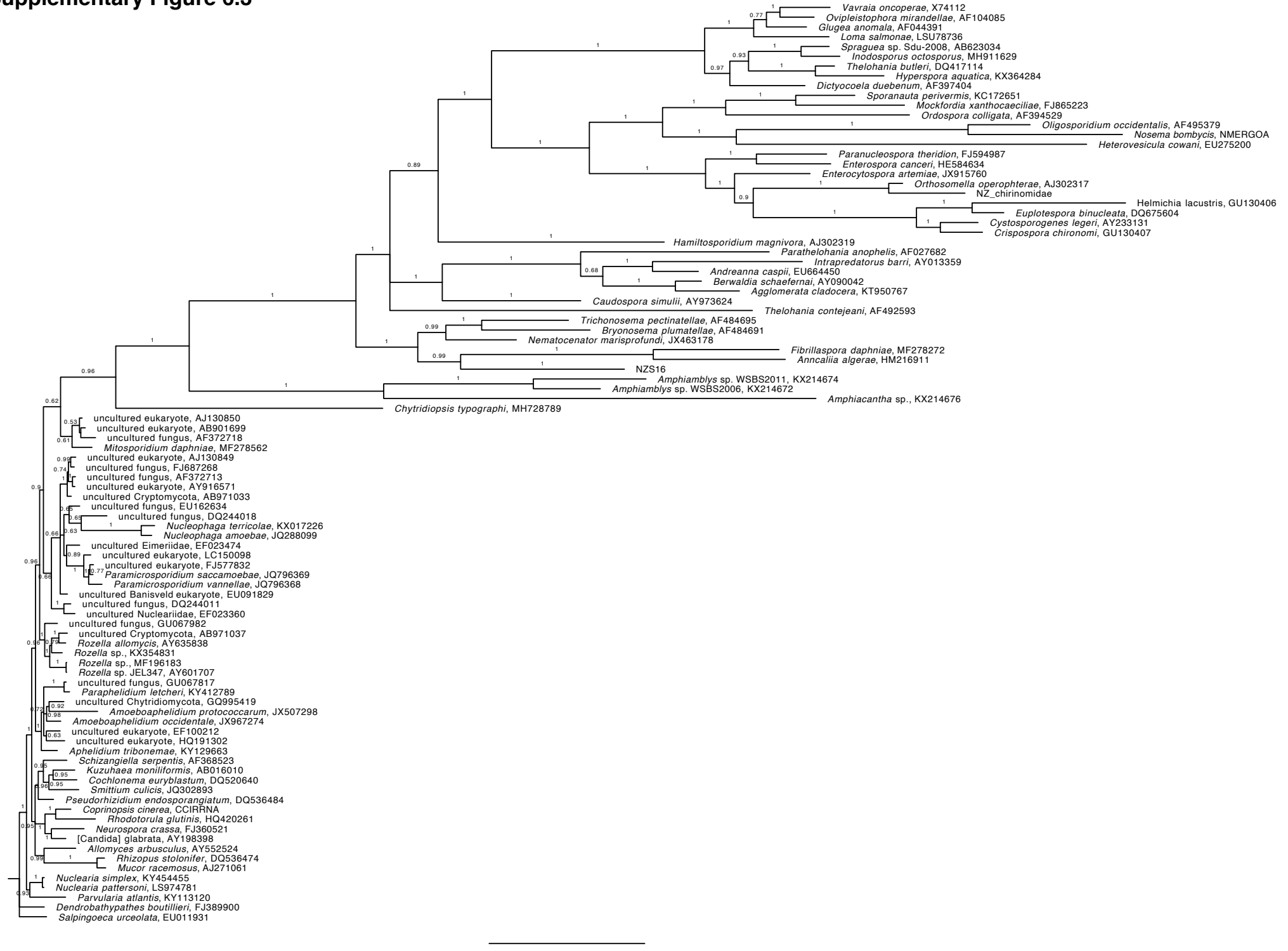
Supplementary Figure 6.1



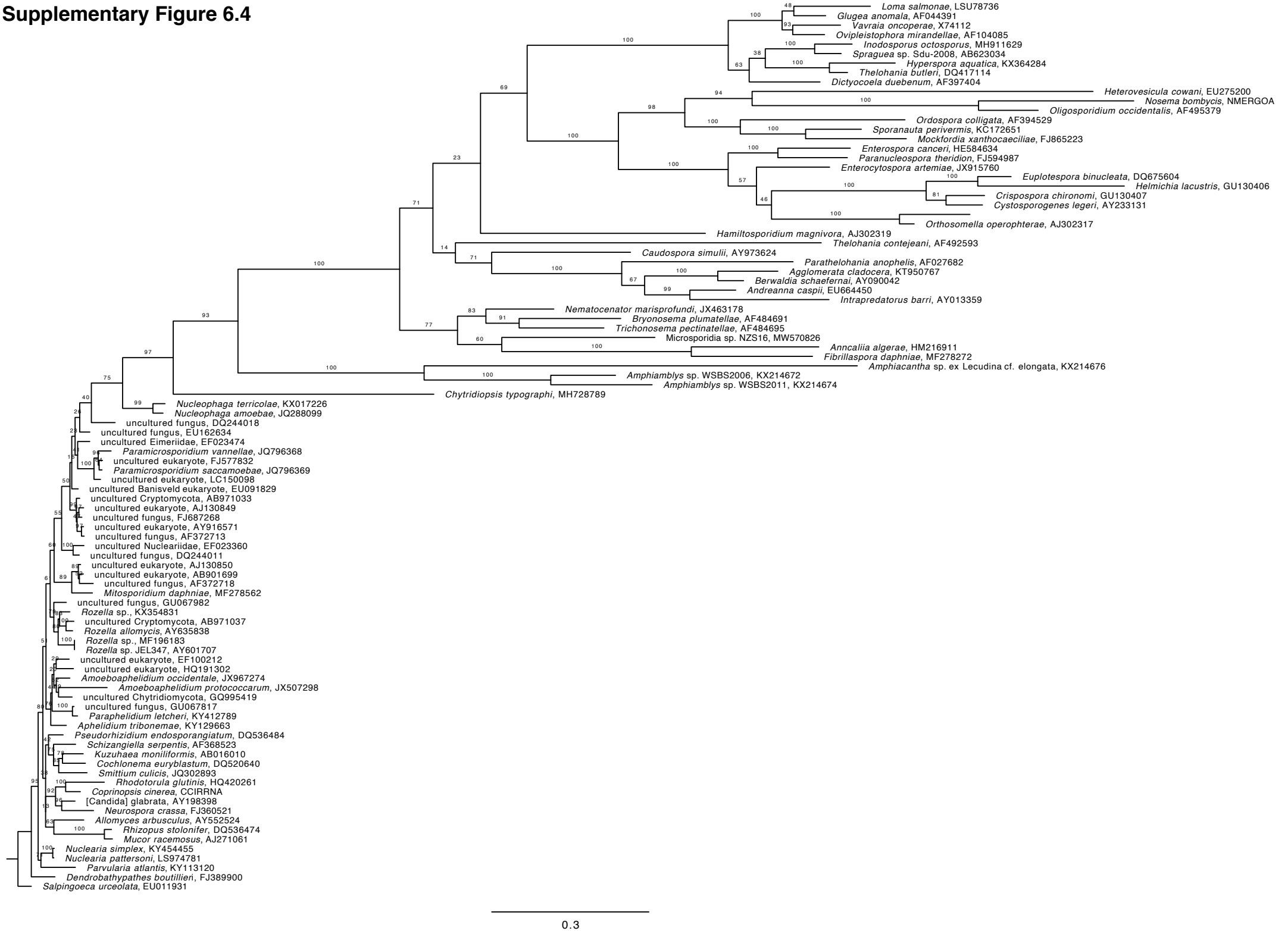
Supplementary Figure 6.2



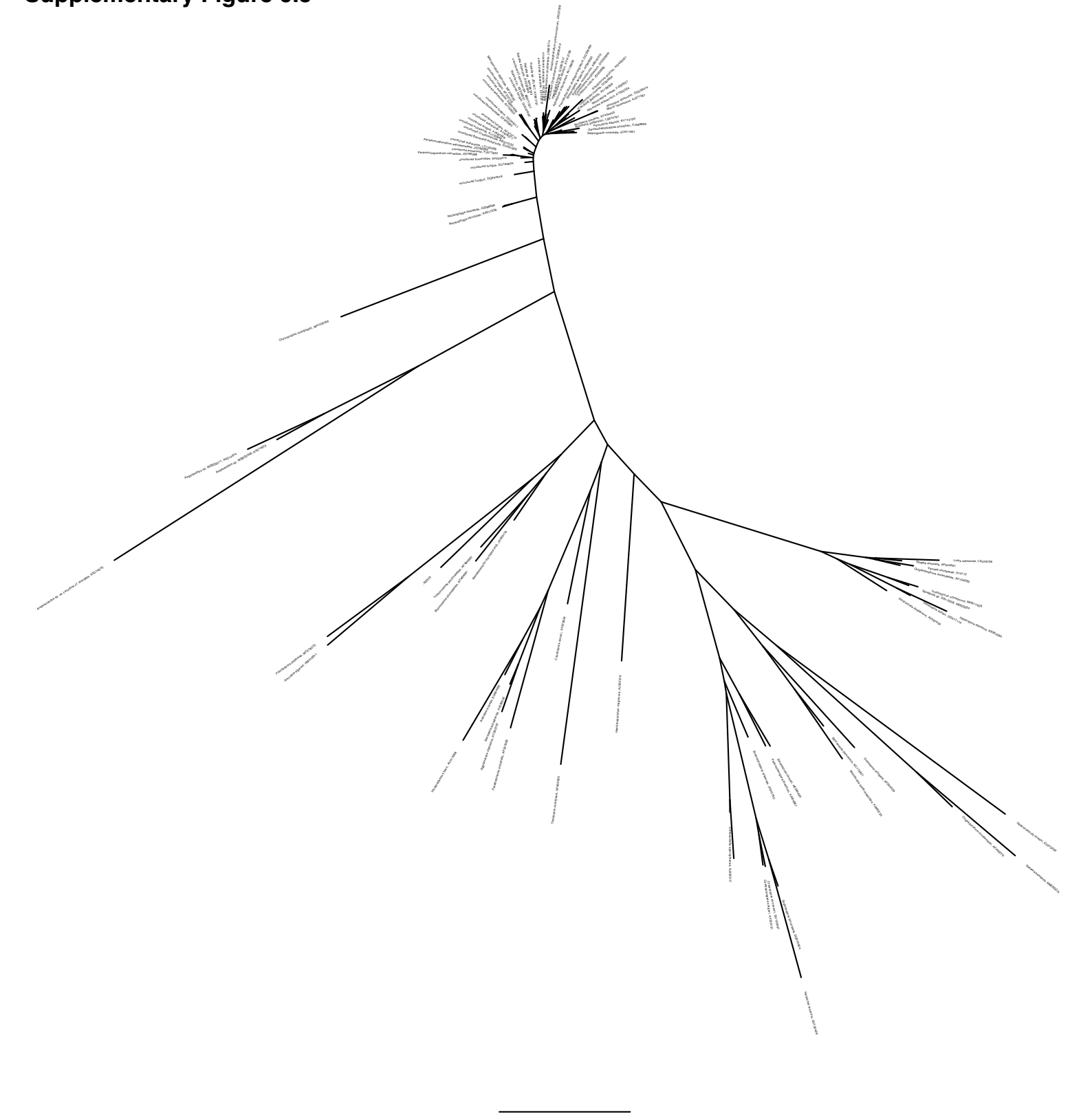
Supplementary Figure 6.3



Supplementary Figure 6.4



Supplementary Figure 6.5



Supplementary table 2.1 Known host range and geographic distribution of each Dictyocoela species.

Species	Host species	References	Host taxa	Region	GenBank ID used for Figure5
D. berillonum	*Dikerogammarus villosus*, *D. haemobaphes*, *Echinogammarus berilloni, E. ischnus*, *E. marinus, E. trichiatus*, *Gammarus duebeni duebeni*, *G. pulex*, *G. tigrinus*, *Melita palmata*, *Pontogammarus robustoides*	Terry et al. 2004; Krebes et al. 2010; Etxabe et al. 2015; Grabner et al. 2015; Wilkinson et al. 2011	Gammaroidea, Melitidae	Europe	AJ438957
D. duebenum	*Acanthogammarus lappaceus*, *Ac. victorii*, *Dikerogammarus villosus, Dorogostaiskia parasitica*, *Echinogammarus berilloni*, *E. marinus*, *Eulimnogammarus cyaneus*, *Eu. verrucosus*, *Eu. marituji Bazikalova*, *Gammarus duebeni duebeni*, *G. pulex*, *G. tigrinus*, *G. roeselii*, *G. fossarum*, *G. lacustris*, *G. locusta*, *G. pseudolimneus*, *Gmelinoides fasciatus*, *Pallasea cancellous*	Terry et al. 2004; Hogg et al. 2002; Grabner et al. 2015; Krebes et al. 2010; Wilkinson et al. 2011; Kuzmenkova et al. 2008	Gammaroidea	Europe, Asia (Lake Baikal)	AF397404
D. muelleri	*Dikerogammarus bispinosus*, *Dikerogammarus haemobaphes*, *Dikerogammarus villosus, Gammarus duebeni celticus*, *G. duebeni duebeni*, *G. aeqicauda*, *G. varsoviensis*, *G. pulex*, *G. roeselii*, *Pontogammarus robustoides*	Terry et al. 2004; Grabner et al. 2015; Krebes et al. 2010; Wilkinson et al. 2011	Gammaroidea	Europe	AJ438955
D. roeselum	*Gammarus balcanicus*, *G. duebeni celticus*, *G. fossarum*, *G. lacustris*, *G. pullex*, *G. roeseli*, *G. varsoviensis,*	Haine et al. 2004; Terry et al. 2004, Quiles et al. 2019	Gammaridae	Europe	AY584252
D. sp. N	*Echinogammarus ischnus*	Bacela-Spychalska et al. 2018	Gammaridae	Europe	MG773222
D. sp. N4	*Gammarus roeselii*	Quiles et al. 2019	Gammaridae	Europe	MK719295
D. cavimanum	*Talitrus* sp., *Orchestia aestuarensis*, *O. cavimana*,	Terry et al. 2004; Pickup and Ironside 2017	Talitridae	Europe	AJ438959
D. deshayesum	*Deshayesorchestia deshayesii*	Terry et al. 2004	Talitridae	Europe	AJ438961
D. gammarellum	*Orchestia gammarellus*, *O. mediterranea*	Terry et al. 2004; Pickup and Ironside 2017	Talitridae	Europe	AJ438958
D. sp. HYAL	*Hyalella* sp. BeulHy	Slothouber Galbreath et al. 2010	Hyalellidae	USA (Lake Beulah)	AJ966726
D. diporeiae	*Diporeia* sp.	Winters and Faisal 2014	Pontoporeiidae	USA (Lake Superior)	KF537632
D. sp NZ1	*Paracalliope* spp.	This study	Paracalliopidae	New Zealand	MT462166
D. sp NZ2	*Paracalliope* spp.	This study	Paracalliopidae	New Zealand	MT462168
D. sp NZ3	*Paracalliope* spp.	This study	Paracalliopidae	New Zealand	MT462172
D. sp NZ4	*Paracorophium excavatum*, *Melita* sp.	This study	Paracorophidae, Melitidae	New Zealand	MT462176
D. sp NZ5	Paraorchestia sp.	This study	Talitridae	New Zealand	MT462178
D. sp NZ6	Paraleptamphopidae sp.	This study	Paraleptamphopidae	New Zealand	MT462179

Supplementary table 2.2 Sampling sites and screening results. A total of 2,670 individuals from 69 sites were used for molecular screening. Microsporidians were detected from 51 sites. The sites that SSU rDNA sequences were obtained from pooled samples are marked with †.

Species	Location ID	Region	Location	Type	Microsporidia	number of microsporidian species obtained	number of tubes positive for Microsporidia	number of tubes	number of individuals per tube	total number of individuals used for screening
Paracalliope	N1	Northland	Mangatete River	River	Y	3	9	12	4	48
	N2		Hukerenui	Stream	Y	1	7	12	4	48
	N23		Ruakaka	Stream	Y	1	3	12	4	48
	N26		Kaiwaka	Stream	Y	1†	4	12	4	48
	N22	Auckland	Makarau	Stream	Y	1	1	2	4	8
	N3		Kaukapakapa	Stream	Y	1†	1	12	4	48
	N5		Wairere Reserve	Pond	Y	1†	1	12	4	48
	N27	Waikato	Mangatarata	Stream	Y	2†	2	3	4	12
	N28		Tahuna	Stream			0	12	4	48
	N4		Waerenga	Stream	Y	1	8	12	4	48
	N6		Kaihere	Stream			0	12	4	48
	N7		Tauwhare	Stream	Y	1	3	8	4	32
	N12	Taranaki	Manaia	Stream	Y	1	2	12	4	48
	N13		Tirorangi	Stream	Y	2†	3	10	4	40
	N17		Mokoia	Stream	Y	1	4	4	4	16
	N18		Rahotu	Stream	Y	1†	2	12	4	48
	N8	Hawke's Bay	Haumoana	Stream	Y	1	8	12	4	48
	N30		Clive	Stream	Y	1†	5	12	4	48
	N31		Tutaekuri River	River			0	4	4	16
	N32		Pakipaki	Stream	Y	1†	2	3	4	12
	N34		Waipawa River	River	Y	1†	1	2	4	8
	N11	Manawatu-Wanganui	Okoia	Stream	Y	2	10	12	4	48
	N9		Woodville	Stream	Y	1	5	12	4	48

N10	Wellington	Waikanae River	River	Y	1†	2	12	4	48
N14		Porirua stream	Stream	Y	1†	5	12	4	48
N35		Waipoua River	River	Y	1	4	12	4	48
N36		South Featherston	Stream			0	4	4	16
N37		Hutt River	River			0	3	4	12
S27	Tasman	Pohara	Stream			0	12	4	48
S3		Motueka River	River			0	12	4	48
S4		Riwaka	Stream	Y	1†	3	12	4	48
S5		Waimea River	River	Y	1†	1	12	4	48
S12	Nelson	Queen's Gardens	Pond	Y	1	1	12	4	48
S2		Matai River	River			0	12	4	48
S11	Malborough	Opawa River	River	Y	1	5	12	4	48
S6		Grovetown	Stream	Y	1	4	12	4	48
S1	Canterbury	Waitaki River	River	Y	1	3	12	4	48
S28		Otaio River	River	Y	1†	1	2	4	8
S29		Opihi River	River	Y	1†	1	1	4	4
S30		Winchester	Stream	Y	3	6	12	4	48
S32		Lake Benmore	Stream			0	3	4	12
S46		Amberley	River	Y	1	5	12	4	48
S47		Waikuku	Stream	Y	1	12	12	4	48
S16	Otago	Lake Waihola	Lake	Y	1	4	12	4	48
S19		Tomahawk Lagoon	Lagoon	Y	3	4	12	4	48
S34		Tokomairiro	Stream	Y	1	4	12	4	48
S35		Clutha River	River	Y	1†	1	12	4	48
S36		Waiwera	Stream			0	12	4	48
S37		Clinton	Stream	Y	1†	1	12	4	48
S45		Owaka	Stream	Y	1†	1	12	4	48
S48		Waitati	Stream			0	12	4	48
S14	Southland	Waiau river	River	Y	1†	1	6	4	24

	S15		Tokanui	Stream	Y	2†	3	16	4	64
	S38		Pukerau	Stream	Y	1	7	12	4	48
	S39		Mataura River	River			0	12	4	48
	S40		Waimumu	Stream	Y	2	5	12	4	48
	S41		Waimahaka	Stream	Y	2	12	12	4	48
	S43		Haldane	Southland			0	12	4	48
	S44		Niagara Falls	Southland	Y	3	6	12	4	48
Paraleptamphopidae (black)	S33	Otago	Crown Range	Stream	Y	1†	1	6	8	48
Paraleptamphopidae (black_small)	S39	Southland	Mataura River	River	Y	1†	1	12	4	48
Paraleptamphopidae (eyeless)	S50	Otago	Herbert	Stream			0	24	1	24
Paraleptamphopidae (eyeless)	N17	Taranaki	Mokoia	Stream			0	2	1	2
Parorchestia tenuis	S26	Tasman	Takaka River	River			0	2	1	2
	N17	Taranaki	Mokoia	Stream			0	1	1	1
	N36	Wellington	South Featherston	Stream	Y	1	1	1	1	1
	S8	Canterbury	Anzac drive Reserve	Water reservoir			0	24	1	24
	N15	Manawatu-Wanganui	Manawatu River	River			0	8	1	8
	N24	Northland	Kaeo	Stream			0	8	1	8
Talitridae sp.	S16	Otago	Lake Waihola	Lake			0	4	1	4
Senticaudata sp.	S33	Otago	Crown Range	Stream			0	4	7	28
Melita awa	S7	Canterbury	Avon Heathcote estuary	Estuary	Y	2†	2	12	4	48
	S21	Otago	Kaikorai estuary	Estuary	Y	1†	1	12	4	48
Paracorophium	S16	Otago	Lake Waihola	Lake	Y	1	1	7	1	7
	S21	Otago	Kaikorai estuary	estuary	Y	2†	2	12	4	48
	S10	Tasman	Aorere River	Estuary			0	19	3	57
Phreatogammarus sp.	S13	Tasman	Waimea Inlet	Inlet			0	15	4	60

Supplementary table 2.3 Genbank accession IDs for microsporidian sequences obtained from this study

Species	Haplotype	Genbank ID	Location ID_Sample ID
D. sp. NZ1	Haplotype 1	MT462166	S46_1818
			S1_799
	Haplotype 2	MT462167	S19_124
D. sp. NZ2	Haplotype 1	MT462168	S4_521
	Haplotype 2	MT462169	N9_846
	Haplotype 3	MT462170	S6_851
	Haplotype 4	MT462171	N4_822
	Haplotype 5	MT462183	N18_1450
	Haplotype 6	MT462184	S5_538
	Haplotype 7	MT462185	N30_1472
			N32_1502
	Haplotype 8	MT462186	S15_327
			S37_1262
			N35_1625
			S41_1549
			N12_766
			N23_1415
D. sp. NZ3	Haplotype 1	MT462172	S1_800
	Haplotype 2	MT462173	S12_638
			N7_843
			N8_827
			N11_848
			N11_849
			N13_782
			N13_785
			N22_1641
			N26_1427
			N27_1627
			N27_1628
	Haplotype 3	MT462174	S19_265
			S19_121
			N2_820
			N3_683
			N5_703
	Haplotype 4	MT462175	S30_1854
			S30_1862
	Haplotype 5	MT462187	S15_340
			S29_1366
	Haplotype 6	MT462188	S40_1761
			S40_1773
			S38_1813
			S35_1240
	Haplotype 7	MT462189	S44_1916
			S44_1920
	Haplotype 8	MT462190	S14_344
	Haplotype 9	MT462191	S4_527
			S11_853
	Haplotype 10	MT462192	N10_747

			N17_1617
			N14_1456
	Haplotype 11	MT462193	S28_1364
			S34_1176
			S41_1548
			S44_1913
			S45_1231
			S47_1564
	Haplotype 12	MT462194	N34_1505
	Haplotype 13	MT462195	S30_1859
D. sp. NZ4	Haplotype 1	MT462176	S21_191
			S21_203
	Haplotype 2	MT462196	S21_193
D. sp. NZ5	Haplotype 1	MT462177	S33_1391
	Haplotype 2	MT462178	S39_1381
D. sp. NZ6	Haplotype 1	MT462179	N36_1510
Unikaryonidae sp.	Haplotype 1	MT462182	S7_561
			S7_563
Microsporidium sp. NZN1	Haplotype 1	MT462180	S16_7
			S16_1568
Microsporidium sp. NZS16	Haplotype 1	MT462181	N1_818
			N1_819
			N1_836

Supplementary table 2.4 Genbank accession IDs for microsporidian sequences obtained from this study

Species	Genbank ID
Vairimorpha necatrix	Y00266.1
Enterocytozoon bieneusi	L07123.1
Ameson michaelis	L15741.1
Glugea atherinae	U15987.1
Endoreticulatus schubergi	L39109.1
Encephalitozoon cuniculi	L39107.1
Nosema bombycis	L39111.1
Vittaforma corneum	L39112.1
Nucleospora salmonis	U78176.1
Loma salmonae	U78736.1
Amblyospora californica	U68473.1
Parathelohania anophelis	AF027682.1
Culicosporella lunata	AF027683.1
Spraguea lophii	AF056013.1
Antonospora scoticae	AF024655.1
Vairimorpha sp.	AF031539.1
Amblyospora connecticus	AF025685.1
Pleistophora typicalis	AF044387.1
Amblyospora sp.	AJ252949.1
Loma acerinae	AJ252951.1
Microgemma sp.	AJ252952.1
Vavraia culicis	AJ252961.1
Pseudoloma neurophilia	AF322654.1
Kabatana takedai	AF356222.1
Orthosomella operophterae	AJ302317.1
Hamiltosporidium magnivora	AJ302319.1
Glugoides intestinalis	AF394525.1
Larssonia obtusa	AF394527.1
Ordospora colligata	AF394529.1
Gurleya daphniae	AF439320.1
Dictyocoela duebenum	AF397404.2
Thelohania contejeani	AF492593.1
Hyalinocysta chapmani	AF483837.1
Bryonosema plumatellae	AF484691.1
Pseudonosema cristatellae	AF484694.1
Trichonosema pectinatellae	AF484695.1
Heterosporis anguillarum	AF387331.1
Marssoniella elegans	AY090041.1
Berwaldia schaefei	AY090042.1
Hazardia milleri	AY090067.1
Amblyospora bracteata	AY090068.1
Janacekia debaisieuxi	AY090070.1
Paranosema grylli	AY305325.1
Trichotuzetia guttata	AY326268.1
Brachiola algerae	AY230191.1
Fibrillanosema crangonycis	AY364089.1
Myosporidium merluccius	AY530532.1
Bacillidium vesiculoformis	AJ581995.1
Cystosporo legeri	AY233131.2

Systenostrema alba	AY953292.2
Nadelspora canceri	AY958070.1
Caudospora simulii	AY973624.1
Tubulinosema kingi	DQ019419.1
Thelohania butleri	DQ417114.1
Microsporidium sp.	DQ521753.1
Senoma globulifera	DQ641245.1
Euplotespora binucleata	DQ675604.1
Liebermannia dichroplusae	EF016249.1
Ovavesicula popilliae	EF564602.1
Heterovesicula cowani	EU275200.1
Potaspora morhaphis	EU534408.1
Nematocida parisii	FJ005051.1
Mockfordia xanthocaeciliae	FJ865223.1
Cucumispora dikerogammari	GQ246188.1
Paranucleospora theridion	FJ594987.1
Mrazekia macrocyclopis	FJ914315.1
Kneallhazia carolinensae	GU173849.1
Microsporidium sp.	FJ755988.1
Microsporidium sp.	FJ755999.1
Microsporidium sp.	FJ756100.1
Neoflabelliforma aurantiae	GQ206147.1
Octosporea muscaedomesticae	FN794114.1
Helmichia lacustris	GU130406.1
Myospora metanephrops	HM140499.1
Neoperezia chironomi	HQ396519.1
Nucleophaga amoebae	JQ288099.1
Facilispora margolisi	HM800850.2
Amoeboaphelidium protococcarum	JX507298.1
Paramicrosporidium saccamoebae	JQ796369.1
Sporanauta perivermis	KC172651.1
Heterosporis sp.	KC137548.1
Microsporidium sp. *H7*	KJ019845.1
Microsporidium sp. *RR1 haplotype 1*	KR871372.1
Microsporidium sp. *BPAR3 haplotype 1*	KR871374.1
Hyperspora aquatica	KX364284.1

Supplementary table 2.5 Genbank accession IDs used for inferring phylogenetic tree of the amphipods

Species	18S	28S	COI	Family	Superfamily	Parvorder
Dorogostaiskia parasitica	FJ756227	KF586548	FJ756322	Acanthogammaridae	Gammaroidea	Gammaridira
Eulimnogammarus cyaneus		MN005057	MK887676	Eulimnogammaridae	Gammaroidea	Gammaridira
Dikerogammarus villosus	MK160013	MK159870	KM208868	Gammaridae	Gammaroidea	Gammaridira
Echinogammarus berilloni	MK160056	MK159922	KT075261	Gammaridae	Gammaroidea	Gammaridira
Gammarus lacustris	JF966145	JF965733	HM425346	Gammaridae	Gammaroidea	Gammaridira
Gmelinoides fasciatus	MH001610		FJ715919	Micruropodidae	Gammaroidea	Gammaridira
Pallasea cancellus	AY926865		MG936153	Pallaseidae	Gammaroidea	Gammaridira
Pontogammarus robustoides	MK160018	MK159875	MK159947	Pontogammaridae	Gammaroidea	Gammaridira
Hyalella sp BeulHy	AY743944	DQ464742	AJ968915	Hyalellidae	Hyaloidea	Talitridira
Deshayesorchestia deshayesii	KP010834	MG655759	MG655900	Talitridae	Talitroidea	Talitridira
Orchestia gammarellus	KP010845		MG320264	Talitridae	Talitroidea	Talitridira
Talitrus sp	AY826955		KT209182	Talitridae	Talitroidea	Talitridira
Melita palmata	MN089550		KX224070	Melitidae	Hadzioidea	Hadziidira
Melita awa		MT465143	MT466579	Melitidae	Hadzioidea	Hadziidira
Parorchestia tenuis	MT465139	MT465139	MT466576	Talitridae	Talitroidea	Talitridira
Paraleptamphopidae sp.		MT465134	MT466574	Paraleptamphopidae	Gammaroidea	Gammaridira
Paracalliope fluviatilis		MT465170	KR336880	Paracalliopidae	Oedicerotoidea	Oedicerotidira
Paracorophium excavatum		MT465144	MT466578	Corophiidae	Corophioidea	Corophiidira
Homarus gammarus	DQ079749	DQ079789	KT209166			

Supplementary table 3.1 Sampling sites and screening results.

Location ID	Region	Type	Location	Species	number of tubes	number of individuals per tube	total number of individuals used for screening	Rickettsia	Sequence	number of tubes positive for Microsporidia	Location ID
N1	Northland	River	Mangatete River	*Paracalliope*	12	4	48				N1
N2	Northland	Stream	Hukerenui	*Paracalliope*	12	4	48				N2
N23	Northland	Stream	Ruakaka	*Paracalliope*	12	4	48				N23
N26	Northland	Stream	Kaiwaka	*Paracalliope*	12	4	48				N26
N22	Auckland	Stream	Makarau	*Paracalliope*	2	4	8				N22
N3	Auckland	Stream	Kaukapakapa	*Paracalliope*	12	4	48				N3
N5	Auckland	Pond	Wairere Reserve	*Paracalliope*	12	4	48				N5
N27	Waikato	Stream	Mangatarata	*Paracalliope*	3	4	12				N27
N28	Waikato	Stream	Tahuna	*Paracalliope*	12	4	48				N28
N4	Waikato	Stream	Waerenga	*Paracalliope*	12	4	48				N4
N6	Waikato	Stream	Kaihere	*Paracalliope*	12	4	48				N6
N7	Waikato	Stream	Tauwhare	*Paracalliope*	8	4	32				N7
N8	Hawke's Bay	Stream	Haumoana	*Paracalliope*	12	4	48				N8
N30	Hawke's Bay	Stream	Clive	*Paracalliope*	12	4	48				N30
N31	Hawke's Bay	River	Tutaekuri River	*Paracalliope*	4	4	16				N31
N32	Hawke's Bay	Stream	Pakipaki	*Paracalliope*	3	4	12				N32
N34	Hawke's Bay	River	Waipawa River	*Paracalliope*	2	4	8	Y	1*	2	N34

N12	Taranaki	Stream	Manaia	*Paracalliope*	12	4	48				N12
N13	Taranaki	Stream	Tirorangi	*Paracalliope*	10	4	40				N13
N17	Taranaki	Stream	Mokoia	*Paracalliope*	4	4	16	Y	1*	2	N17
N18	Taranaki	Stream	Rahotu	*Paracalliope*	12	4	48				N18
N11	Manawatu-Wanganui	Stream	Okoia	*Paracalliope*	12	4	48	Y	1*	5	N11
N9	Manawatu-Wanganui	Stream	Woodville	*Paracalliope*	12	4	48				N9
N10	Wellington	River	Waikanae River	*Paracalliope*	12	4	48				N10
N14	Wellington	Stream	Porirua stream	*Paracalliope*	12	4	48	Y	1	1	N14
N35	Wellington	River	Waipoua River	*Paracalliope*	12	4	48	Y	1*	1	N35
N36	Wellington	Stream	South Featherston	*Paracalliope*	4	4	16	Y	1*	4	N36
N37	Wellington	River	Hutt River	*paracalliope*	3	4	12	Y	1*	1	N37
S27	Tasman	Stream	Pohara	*Paracalliope*	12	4	48				S27
S3	Tasman	River	Motueka River	*Paracalliope*	12	4	48				S3
S4	Tasman	Stream	Riwaka	*Paracalliope*	12	4	48				S4
S5	Tasman	River	Waimea River	*Paracalliope*	12	4	48				S5
S12	Nelson	Pond	Queen's Gardens	*Paracalliope*	12	4	48				S12
S2	Nelson	River	Matai River	*Paracalliope*	12	4	48				S2
S11	Malborough	River	Opawa River	*Paracalliope*	12	4	48				S11
S6	Malborough	Stream	Grovetown	*Paracalliope*	12	4	48				S6
S1	Canterbury	River	Waitaki River	*Paracalliope*	12	4	48	Y	1	9	S1
S28	Canterbury	River	Otaio River	*Paracalliope*	2	4	8	Y	1*	1	S28

S29	Canterbury	River	Opihi River	*Paracalliope*	1	4	4	Y	1*	1	S29
S30	Canterbury	Stream	Winchester	*Paracalliope*	12	4	48	Y	1	12	S30
S32	Canterbury	Stream	Lake Benmore	*Paracalliope*	3	4	12	Y	1	3	S32
S46	Canterbury	River	Amberley	*Paracalliope*	12	4	48				S46
S47	Canterbury	Stream	Waikuku	*Paracalliope*	12	4	48				S47
S16	Otago	Lake	Lake Waihola	*Paracalliope*	12	4	48	Y	1	6	S16
S19	Otago	Lagoon	Tomahawk Lagoon	*paracalliope*	12	4	48				S19
S34	Otago	Stream	Tokomairiro	*Paracalliope*	12	4	48	Y	1*	2	S34
S35	Otago	River	Clutha River	*Paracalliope*	12	4	48				S35
S36	Otago	Stream	Waiwera	*Paracalliope*	12	4	48				S36
S37	Otago	Stream	Clinton	*Paracalliope*	12	4	48	Y	1	3	S37
S45	Otago	Stream	Owaka	*Paracalliope*	12	4	48	Y	1	4	S45
S48	Otago	Stream	Waitati	*Paracalliope*	12	4	48	Y	1	10	S48
S14	Southland	River	Waiau river	*Paracalliope*	6	4	24				S14
S15	Southland	Stream	Tokanui	*Paracalliope*	16	4	64	Y	1	11	S15
S38	Southland	Stream	Pukerau	*Paracalliope*	12	4	48	Y	1	1	S38
S39	Southland	River	Mataura River	*Paracalliope*	12	4	48	Y	1	12	S39
S40	Southland	Stream	Waimumu	*Paracalliope*	12	4	48	Y	1	11	S40
S41	Southland	Stream	Waimahaka	*Paracalliope*	12	4	48	Y	1	11	S41
S43	Southland	Southland	Haldane	*Paracalliope*	12	4	48	Y	1	5	S43
S44	Southland	Southland	Niagara Falls	*Paracalliope*	12	4	48	Y	1	8	S44

S33	Otago	Stream	Crown Range	*Paraleptamphopus* sp.	6	8	48				S33
S50	Otago	Stream	Herbert	*Paraleptamphopus* sp.	24	1	24	Y	1*	9	S50
N17	Taranaki	Stream	Mokoia	*Paraleptamphopus* sp.	2	1	2	Y	1	1	N17
S33	Otago	Stream	Crown Range	*Paraleptamphopus* sp.	8	1	8				S33
N17	Taranaki	Stream	Mokoia	Parorchestia tenuis	1	1	1				N17
S26	Tasman	River	Takaka River	Parorchestia tenuis	2	1	2				S26
N36	Wellington	Stream	South Featherston	Parorchestia tenuis	1	1	1				N36
S8	Canterbury	Water reservoir	Anzac drive Reserve	Parorchestia tenuis	24	1	24				S8
N15	Manawatu-Wanganui	River	Manawatu River	Parorchestia tenuis	8	1	8				N15
N24	Northland	Stream	Kaeo	Parorchestia tenuis	8	1	8				N24
S33	Otago	Stream	Crown Range	undescribed sp.	4	7	28	Y	1	4	S33
S39	Southland	River	Mataura River	*Paraleptamphopus* sp.	12	4	48	Y	1*	12	S39
S16	Otago	Lake	Lake Waihola	Talitridae sp.	4	1	4				S16
S7	Canterbury	Estuary	Avon Heathcote estuary	*Melita*	12	4	48				S7
S21	Otago	Estuary	Kaikorai estuary	*Melita*	12	4	48				S21
S16	Otago	Lake	Lake Waihola	*Paracorophium*	7	1	7				S16
S21	Otago	estuary	Kaikorai estuary	*Paracorophium*	12	4	48				S21
S10	Tasman	Estuary	Aorere River	*Paracorophium*	19	3	57				S10
S13	Tasman	Inlet	Waimea Inlet	*Phreatogammarus* sp.	15	4	60				S13

Supplementary table 3.2 GenBank accession IDs for the sequences obtained in this study

	16S	glta	atpa	COI
N11_830	MT507651			
N14_1462	MT507652			
N17_1479	MT507653			
N17_eyeless_1621	MT507654	MT524989	MT515460	
N34_1505	MT507655	MT524990		MT515479
N35_1493	MT507656	MT524991		
N36_1509	MT507657		MT515461	
N37_1513	MT507658		MT515462	
S1_1148	MT507659		MT515463	
S15_1167	MT507660		MT515464	MT515480
S16_1568	MT507661	MT524992	MT515465	
S28_1364	MT507662			
S29_1366	MT507663	MT524993	MT515466	
S30_1526	MT507664	MT524994	MT515467	MT515481
S32_1565	MT507665	MT524995	MT515468	
S33_1582			MT515469	MT515482
S34_1177			MT515470	
S37_1535		MT524996	MT515471	MT515483
S38_1540	MT507666	MT524997	MT515472	
S39_1542	MT507667	MT524998		
S39_black_1374	MT507668			
S40_1546			MT515473	
S41_1549	MT507669	MT524999	MT515474	MT515484
S43_1551	MT507670		MT515475	
S44_1556	MT507671	MT525000	MT515476	
S45_1560	MT507672			MT515485
S48_1575	MT507673	MT525001	MT515477	
S50_1965	MT507674	MT525002	MT515478	MT515486

Supplementary table 3.3 16S rRNA sequences of Torix Rickettsia that were obtained from GenBank

Host Order	Host	Accession ID
Amoeba	*Nuclearia pattersoni*	AY364636.1
Amoeba	*Pompholyxophrys punicea*	MK616427.1
Annelida	*Hemiclepsis marginata*	AB113215.1
Annelida	*Torix tsukubana*	AB113214.1
Annelida	*Hemiclepsis marginata*	AB066352.1
Annelida	*Torix tagoi*	AB066351.1
Arachnida	*Mermessus fradeorum*	MN028388.1
Coleoptera	*Deronectes*	FM955312.1
Coleoptera	*Deronectes*	FM955310.1
Coleoptera	*Deronectes*	FM955311.1
Coleoptera	*Deronectes*	FM177877.1
Coleoptera	*Deronectes*	FM177876.1
Coleoptera	*Deronectes*	FM177869.1
Coleoptera	*Deronectes*	FM177873.1
Coleoptera	*Deronectes*	FM177872.1
Coleoptera	*Deronectes*	FM177871.1
Coleoptera	*Deronectes*	FM177868.1
Coleoptera	*Deronectes*	FM177874.1
Coleoptera	*Deronectes*	FM177875.1
Coleoptera	*Deronectes*	FM177870.1
Diptera	*Phlebotomus chinensis*	KX363668.1
Diptera	*Phlebotomus chinensis*	KX363666.1
Diptera	*Limonia chorea*	AF322442.1
Diptera	*Culicoides*	KY777733.1
Diptera	*Culicoides*	KY777726.1
Diptera	*Culicoides*	KY777724.1
Diptera	*Culicoides*	KY777731.1
Diptera	*Culicoides*	KY777730.1
Diptera	*Culicoides*	KY777729.1
Diptera	*Culicoides*	KY777727.1
Diptera	*Culicoides*	KY777723.1
Diptera	*Culicoides*	KY777725.1
Diptera	*Culicoides*	KY777728.1
Diptera	*Culicoides*	KY777722.1
Diptera	*Lutzomyia apache*	EU223247.1
Hemiptera	*Centrotus cornutus*	MN082138.1
Hemiptera	*Gargara genistae*	MN082141.1
Hemiptera	*Macrolophus*	HE583203.1
Hemiptera	*Platypleura kaempferi*	KR911839.1
Hemiptera	*Nephotettix cincticeps*	AB702995.1
Hemiptera	*Nephotettix cincticeps*	KU586121.1
Hemiptera	*Nephotettix cincticeps*	KU586122.1
Hemiptera	*Bemisia tabaci*	MG063879.1
Hymenoptera	*Atta colombica*	LN570502.1
Hymenoptera	*Asobara tabida*	FJ603467
Megaloptera	*Sialis lutaria*	MF156636
Megaloptera	*Sialis lutaria*	MF156635
NA	*environmental sample*	AM697554.1
Oligochaeta	*Mesenchytraeus sp.*	AB991365.1
Psocoptera	*Cerobasis guestfalica*	DQ652595.1
Psocoptera	*Cerobasis guestfalica*	DQ652596.1

Supplementary table 3.4 GltA sequences of Torix Rickettsia that were obtained from GenBank

Host Order	Sequence name	Accession ID
Arachnida	Rickettsia endosymbiont of Oedothorax gibbosus isolate W035 citrate synthase gene, partial cds	HQ286289.1
Arachnida	Rickettsia endosymbiont of Pityohyphantes phrygianus citrate synthase gene, complete sequence	DQ231491.1
Arachnida	Rickettsia endosymbiont of Araneus diadematus citrate synthase gene, complete sequence	DQ231490.1
Arachnida	Rickettsia endosymbiont of Oedothorax retusus clone RickD315 citrate synthase gene, partial cds	JN889707.1
Arachnida	Rickettsia endosymbiont of Theridiidae sp. citrate synthase gene, complete sequence	DQ231486.1
Arachnida	Rickettsia endosymbiont of Erigone dentipalpis citrate synthase gene, complete sequence	DQ231492.1
Arachnida	Rickettsia endosymbiont of Hylaphantes graminicola citrate synthase gene, complete sequence	DQ231487.1
Arachnida	Rickettsia endosymbiont of Lepthyphantes zimmermani citrate synthase gene, complete sequence	DQ231488.1
Arachnida	Rickettsia endosymbiont of Microneta viaria citrate synthase gene, complete sequence	DQ231493.1
Arachnida	Rickettsia endosymbiont of Walckenaeria cuspidata citrate synthase gene, complete sequence	DQ231489.1
Arachnida	Rickettsia endosymbiont of Meta mengei citrate synthase gene, complete sequence	DQ231482.1
Arachnida	Rickettsia endosymbiont of Troxochrus scabriculus citrate synthase gene, complete sequence	DQ231485.1
Arachnida	Rickettsia endosymbiont of Gnathonarium dentatum citrate synthase gene, complete sequence	DQ231484.1
Arachnida	Rickettsia endosymbiont of Meta mengei citrate synthase gene, complete sequence	DQ231483.1
Coleoptera	Rickettsia endosymbiont of Deronectes delarouzei partial gltA gene for citrate synthase	FM955313.1
Coleoptera	Rickettsia endosymbiont of Deronectes aubei partial gltA gene for citrate synthase	FM955315.1
Coleoptera	Rickettsia endosymbiont of Deronectes platynotus partial gltA gene for citrate synthase	FM177878.1
Coleoptera	Rickettsia endosymbiont of Deronectes semirufus partial gltA gene for citrate synthase	FM955314.1
Diptera	Rickettsia endosymbiont of Achalcus cinereus strain 86 citrate synthase (gltA) gene, partial cds	JQ925562.1
Diptera	Rickettsia endosymbiont of Rhaphium appendiculatum strain 8 citrate synthase (gltA) gene, partial cds	JQ925544.1
Diptera	Rickettsia endosymbiont of Lutzomyia apache GltA (gltA) gene, partial cds	EU368001.1
Diptera	Rickettsia endosymbiont of Medetera muralis strain 187 citrate synthase (gltA) gene, partial cds	JQ925597.1
Diptera	Rickettsia endosymbiont of Chrysotimus molliculus strain 159 citrate synthase (gltA) gene, partial cds	JQ925586.1
Diptera	Rickettsia endosymbiont of Dolichopus wahlbergi strain 76B citrate synthase (gltA) gene, partial cds	JQ925557.1
Diptera	Rickettsia endosymbiont of Dolichopus wahlbergi strain 76A citrate synthase (gltA) gene, partial cds	JQ925556.1
Diptera	Rickettsia endosymbiont of Clinocera sp. strain 249 citrate synthase (gltA) gene, partial cds	JQ925615.1
Diptera	Rickettsia endosymbiont of Phyllodromia melanocephala strain 238 citrate synthase (gltA) gene, partial cds	JQ925609.1
Diptera	Rickettsia endosymbiont of Chrysotimus flaviventris strain 144 citrate synthase (gltA) gene, partial cds	JQ925578.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-imp14 citrate synthase (gltA) gene, partial cds	KY765379.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-rie1 citrate synthase (gltA) gene, partial cds	KY765383.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-N1 citrate synthase (gltA) gene, partial cds	KY765376.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-st3 citrate synthase (gltA) gene, partial cds	KY765385.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-N28 citrate synthase (gltA) gene, partial cds	KY765374.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-pul23uk citrate synthase (gltA) gene, partial cds	KY765381.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-sal1 citrate synthase (gltA) gene, partial cds	KY765384.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-pul3swe citrate synthase (gltA) gene, partial cds	KY765380.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-N27 citrate synthase (gltA) gene, partial cds	KY765375.1
Diptera	Uncultured Rickettsia sp. clone Rs.IIB-dudd5 citrate synthase (gltA) gene, partial cds	KY765378.1
Diptera	Rickettsia endosymbiont of Gymnopternus metallicus strain 30 citrate synthase (gltA) gene, partial cds	JQ925549.1
Diptera	Rickettsia endosymbiont of Sphyrotarsus argyrostomus strain 325 citrate synthase (gltA) gene, partial cds	JQ925623.1
Diptera	Rickettsia endosymbiont of Microphor holosericeus strain 252 citrate synthase (gltA) gene, partial cds	JQ925617.1
Diptera	Rickettsia endosymbiont of Hilara interstincta strain 248 citrate synthase (gltA) gene, partial cds	JQ925614.1
Diptera	Rickettsia endosymbiont of Chrysotus blepharosceles strain 77 citrate synthase (gltA) gene, partial cds	JQ925558.1
Diptera	Rickettsia endosymbiont of Argyra atriceps strain 19 citrate synthase (gltA) gene, partial cds	JQ925548.1
Diptera	Rickettsia endosymbiont of Empis bicuspidata strain 250 citrate synthase (gltA) gene, partial cds	JQ925616.1
Diptera	Rickettsia endosymbiont of Campsicnemus picticornis strain 56 citrate synthase (gltA) gene, partial cds	JQ925555.1
Diptera	Rickettsia endosymbiont of Argyra vestita strain 160 citrate synthase (gltA) gene, partial cds	JQ925587.1
Diptera	Rickettsia endosymbiont of Gymnopternus brevicornis strain 190 citrate synthase (gltA) gene, partial cds	JQ925600.1

Diptera	Rickettsia endosymbiont of Argyra atriceps strain 122 citrate synthase (gltA) gene, partial cds	JQ925569.1
Diptera	Rickettsia endosymbiont of Neurigona lineata strain 183 citrate synthase (gltA) gene, partial cds	JQ925596.1
Diptera	Rickettsia endosymbiont of Medetera dendrobaena strain 129 citrate synthase (gltA) gene, partial cds	JQ925571.1
Diptera	Rickettsia endosymbiont of Medetera parenti strain 226 citrate synthase (gltA) gene, partial cds	JQ925607.1
Diptera	Rickettsia endosymbiont of Empis nigripes strain 254 citrate synthase (gltA) gene, partial cds	JQ925618.1
Diptera	Rickettsia endosymbiont of Dolichopus claviger strain 15 citrate synthase (gltA) gene, partial cds	JQ925546.1
Diptera	Rickettsia endosymbiont of Medetera saxatilis strain 189 citrate synthase (gltA) gene, partial cds	JQ925599.1
Diptera	Rickettsia endosymbiont of Chrysotus laesus strain D2 citrate synthase (gltA) gene, partial cds	JQ925625.1
Diptera	Rickettsia endosymbiont of Hydrophorus borealis strain 107 citrate synthase (gltA) gene, partial cds	JQ925566.1
Diptera	Rickettsia endosymbiont of Hercostomus praeceps strain 141 citrate synthase (gltA) gene, partial cds	JQ925577.1
Hemiptera	Rickettsia endosymbiont of Bemisia tabaci GltA (gltA) gene, partial cds	MG063880.1
Hemiptera	Rickettsia symbiont of Nephotettix cincticeps isolate WHCUTA-130 citrate (Si)-synthase (gltA) gene, partial cds	KU586334.1
Hemiptera	Rickettsia symbiont of Nephotettix cincticeps isolate WHCUTA-121 citrate (Si)-synthase (gltA) gene, partial cds	KU586333.1
Hemiptera	Rickettsia endosymbiont of Cimex lectularius isolate SouthDakota1 citrate synthase (gltA) gene, partial cds	MN788122.1
Hymenoptera	Rickettsia endosymbiont of Quadrastichus mendeli clone RiQm_13 citrate synthase (gltA) gene, partial cds	KX673391.1
Hymenoptera	Rickettsia endosymbiont of Quadrastichus mendeli clone RiQm_10 citrate synthase (gltA) gene, partial cds	KX673390.1
Hymenoptera	Rickettsia endosymbiont of Quadrastichus mendeli clone RiQm_7 citrate synthase (gltA) gene, partial cds	KX673389.1
Hymenoptera	Rickettsia endosymbiont of Quadrastichus mendeli clone RiQm_4 citrate synthase (gltA) gene, partial cds	KX673388.1
Hymenoptera	Rickettsia endosymbiont of Quadrastichus mendeli clone RiQm_1 citrate synthase (gltA) gene, partial cds	KX673387.1
Siphonaptera	Rickettsia endosymbiont of Nosopsyllus laeviceps laeviceps GltA gene, partial cds	KX457954.1

Supplementary table 3.5 COI sequences of Torix Rickettsia that were obtained from GenBank

Order	Similarity to *Rickettsia* from *Paracalliope*	Accession ID	Date of release	Country	Aware	Voucher?
Amphipoda	100.00%	KR336922.1	23-May-16	New Zealand	N	N
Amphipoda	99.84%	KR336946.1	23-May-16	New Zealand	N	N
Amphipoda	99.84%	KR336928.1	23-May-16	New Zealand	N	N
Amphipoda	99.84%	KR336924.1	23-May-16	New Zealand	N	N
Amphipoda	99.69%	KR336948.1	23-May-16	New Zealand	N	N
Amphipoda	99.69%	KR336947.1	23-May-16	New Zealand	N	N
Amphipoda	99.69%	KR336945.1	23-May-16	New Zealand	N	N
Amphipoda	99.69%	KR336943.1	23-May-16	New Zealand	N	N
Amphipoda	99.69%	KR336925.1	23-May-16	New Zealand	N	N
Amphipoda	99.53%	KR336944.1	23-May-16	New Zealand	N	N
Amphipoda	99.53%	KR336929.1	23-May-16	New Zealand	N	N
Amphipoda	99.38%	KR336923.1	23-May-16	New Zealand	N	N
Arachnida	91.28%	KU600821.1	7-Aug-16	South Africa	Y	NA
Arachnida	91.28%	KU600820.1	7-Aug-16	South Africa	Y	NA
Arachnida	91.26%	KU600824.1	7-Aug-16	South Africa	Y	NA
Arachnida	91.26%	KU600823.1	7-Aug-16	South Africa	Y	NA
Arachnida	91.25%	KF005604.1	28-Oct-14	Czech Republic	Y	NA
Arachnida	91.21%	KU600822.1	7-Aug-16	South Africa	Y	NA
Arachnida	90.68%	KU600819.1	7-Aug-16	South Africa	Y	NA
Coeloptera	96.08%	MH158029.1	28-Jan-19	Germany	N	N
Coleoptera	97.94%	MH158027.1	28-Jan-19	Germany	N	N
Coleoptera	97.81%	MH158030.1	28-Jan-19	Germany	N	N
Diptera	92.12%	KY765403.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.83%	KY765405.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.83%	KY765401.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.71%	KY765399.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.41%	KY765400.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.39%	KY765404.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.39%	KY765398.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.35%	KY765408.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.35%	KY765406.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	91.10%	KM679400.1	22-Jan-15	Estonia?	N	Y
Diptera	91.07%	KY765402.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	89.04%	KY765407.1	17-Aug-17	Sweden, UK, Corsica	Y	NA
Diptera	82.47%	KM966683.1	11-Jul-18	Canada	N	Y
Endopterygota	98.91%	KP422277.1	3-Feb-17	New Zealand	N	Y
Endopterygota	97.97%	KP421136.1	3-Feb-17	New Zealand	N	Y
Hemiptera	91.34%	HM452248.1	28-Aug-10	China	Y	NA
Hemiptera	91.26%	MN609260.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609259.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609258.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609257.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609256.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609255.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609254.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609253.1	24-Nov-19	Korea	N	Y

Hemiptera	91.26%	MN609252.1	24-Nov-19	Korea	N	Y
Hemiptera	91.26%	MN609251.1	24-Nov-19	Korea	N	Y
Hemiptera	91.16%	HE583223.1	2-Oct-11	Laboratory strain, originating from Koppert BV	Y	NA
Hymenoptera	95.01%	KC182291.1	19-Nov-13	Canada	N	Y
Hymenoptera	94.85%	KC182318.1	19-Nov-13	Canada	N	Y
Hymenoptera	92.89%	KM562170.1	9-May-19	Canada	N	Y
Hymenoptera	92.18%	KP421783.1	3-Feb-17	New Zealand	N	Y
Hymenoptera	92.03%	KR931610.1	1-Jun-18	Canada	N	Y
Hymenoptera	91.86%	KM996857.1	30-May-18	Canada	N	Y
Hymenoptera	91.73%	KM996859.1	30-May-18	Canada	N	Y
Hymenoptera	91.61%	KM996197.1	30-May-18	Canada	N	Y
Hymenoptera	91.39%	KM995931.1	30-May-18	Canada	N	Y
Hymenoptera	81.71%	KM568038.1	9-May-19	Canada	N	Y
Hymenoptera	80.42%	KM565576.1	9-May-19	Canada	N	Y
Mandibulata	96.74%	KP421611.1	3-Feb-17	New Zealand	N	Y
Megaloptera	90.95%	MF156672.1	29-Aug-17	UK	Y	NA
Megaloptera	90.95%	MF156671.1	29-Aug-17	UK	Y	NA
Odonata	92.98%	KM383849.1	15-Jun-16	USA	N	N

Supplementary table 3.6 GenBank accession IDs for the sequences of non-Torix Rickettsia that were used for tree inference in this study

16S	glta	COI
AM494475	FJ666764	KM561203
AP008981	FJ666768	CP015010
AJ867656	CP003308	CP000849
AM159487	CP040325	CP000087
DQ223223	CP003393	CP016305
AF523878	KJ882311	CP003393
AJ630204	KJ882309	CP003304
MG699452	MH458574	CP000409
FR822999	MN450397	CP003319
EF520410	KX457952	CP047359
EF074039	FJ666753	CP032049
EF667899	LC388780	AP017602
EF667896	FJ666754	AP017601
DQ395479	NC009883	AP017600
AJ319724	MF156688	AP017599
HE648945	MF156689	AP017598
HG315613	KY678093	AP017597
KJ494868	KT834984	AP017596
KJ494867	GU559856	AP017595
AB746406	KX137900	AP017594
AB604674	FJ666758	AP017593
AB604675	KX457953	AP017592
AB746411	FJ666757	AP017591
AB746409	KF646706	AP017590
AB604676	HE583221	CP018914
KJ995968	U597141	CP018913
AB604677	JN315968	CP006010
AB021128	JF966774	CP006009
NR_074485	AF516331	CP000766
MF496166	MG893576	CP003318
MF496165	KF666472	CP003309
EU036982	KF963607	CP003308
MF002588	KU499847	CP003307
MF002587	KT257873	CP003306
KJ410262	AF497585	CP000848
L36224	MG022117	CP003311
NR_0744802	KX963394	CP003305
MF002584	MF511244	CP001612
KY069267	KX963391	CP040325
KF318168	MN450399	CP000847
JQ339355	MF511252	CP003338
JQ412124	MG545017	AM494475
LC127421	AF120027	
AB604672	MG022119	
CP000849	KX963393	
NR_0744842	KY678106	
NR_0367741	AF207827	
CP0000871	AF176091	

L361031_12
U110141_14
KT3406171
KT3406191
KT3406151
EU8815031
EU8815051
EU8815011
EU8814991
EU8814971
EU8814961
HF9350681
KY7990721
MF0397441
HG3156121
HG3156111
HG3156101
HG3156181
HG3156191
HG3156171
HG3156141
HG3156091
FJ609401
FJ609404
CP003308
CP040325
CP003393
KJ882318
KJ882316
KY348769
MN446745
KX457947
FJ609388
LC388765
FJ609389

Supplementary table 6.1 The list of rRNA sequences used to infer the phylogenetic tree of canonical microsporidia

Group	Species	Accession No.
Canononical microsporidia	Cambaraspora floridanus	MT006313
Canononical microsporidia	Cucumispora ornata	KR190602
Canononical microsporidia	Pseudonosema cristatellae	AF484694
Canononical microsporidia	Bryonosema plumatellae	AF484691
Canononical microsporidia	Trichonosema pectinatellae	AF484695
Canononical microsporidia	Potaspora morhaphis	EU534408
Canononical microsporidia	Loma acerinae	AJ252951
Canononical microsporidia	Euplotespora binucleata	DQ675604
Canononical microsporidia	Microsporidium sp.	KR871372
Canononical microsporidia	Hamiltosporidium magnivora	AJ302319
Canononical microsporidia	Hamiltosporidium tvaerminnensis	GQ843833
Canononical microsporidia	Mrazekia macrocyclopis	FJ914315
Canononical microsporidia	Microsporidium prosopium	AF151529
Canononical microsporidia	Pseudokabatana alburnus	MF974572
Canononical microsporidia	Jirovecia sinensis	MN752318
Canononical microsporidia	Thelohania contejeani	AF492593
Canononical microsporidia	Pseudoloma neurophilia	AF322654
Canononical microsporidia	Thelohania parastaci	AF294779
Canononical microsporidia	Schroedera plumatellae	AY135024
Canononical microsporidia	Neoflabelliforma aurantiae	GQ206147
Canononical microsporidia	Bacillidium vesiculoformis	AJ581995
Canononical microsporidia	Glugoides intestinalis	AF394525
Canononical microsporidia	Gurleya daphniae	AF439320
Canononical microsporidia	Larssonia obtusa	AF394527
Canononical microsporidia	Microsporidium sp	AF394528
Canononical microsporidia	Ordospora colligata	AF394529
Canononical microsporidia	Binucleata daphniae	EU075347
Canononical microsporidia	Anostracospora rigaudi	JX915758
Canononical microsporidia	Enterocytospora artemiae	JX915760
Canononical microsporidia	Microsporidium sp	JX839890
Canononical microsporidia	Fibrillaspora daphniae	MF278272
Canononical microsporidia	Fibrillanosema crangonycis	AY364089
Canononical microsporidia	Agglomerata cladocera	KT950767
Canononical microsporidia	Alfvenia sibirica	KT950766
Canononical microsporidia	Apotaspora heleios	MG708238
Canononical microsporidia	Globulispora mitoportans	KT762153
Canononical microsporidia	Lanatospora costata	KX832080
Canononical microsporidia	Conglomerata obtusa	MH645034
Canononical microsporidia	Pseudoberwaldia daphniae	MK053815
Canononical microsporidia	Berwaldia schaefernai	AY090042
Canononical microsporidia	Marssoniella elegans	AY090041
Canononical microsporidia	Trichotuzetia guttata	AY326268
Canononical microsporidia	Triwangia caridinae	JQ268567
Canononical microsporidia	Edhazardia aedis	AF027684
Canononical microsporidia	Amblyospora californica	U68473
Canononical microsporidia	Culicosporella lunata	AF027683

Canononical microsporidia	Brachiola algerae	AY230191
Canononical microsporidia	Zelenkaia trichopterae	EF537880
Canononical microsporidia	Neoperezia chironomi	HQ396519
Canononical microsporidia	Intrapredatorus barri	AY013359
Canononical microsporidia	Anisofilariata chironomi	GU126383
Canononical microsporidia	Crispospora chironomi	GU130407
Canononical microsporidia	Helmichia lacustris	GU130406
Canononical microsporidia	Visvesvaria algerae	AF024656
Canononical microsporidia	Culicospora magna	AY326269
Canononical microsporidia	Hazardia milleri	AY090067
Canononical microsporidia	Sporanauta perivermis	KC172651
Canononical microsporidia	Ichthyosporidium sp.	L39110
Canononical microsporidia	Myosporidium merluccius	AY530532
Canononical microsporidia	Unidentified microsporidian	AJ295328
Canononical microsporidia	Parahepatospora carcini	KX757849
Canononical microsporidia	Thelohania butleri	DQ417114
Canononical microsporidia	Microgemma carolinus	JQ085991
Canononical microsporidia	Pleistophora sp	AJ252959
Canononical microsporidia	Obruspora papernae	HG005137
Canononical microsporidia	Dasyatispora levantinae	GU183263
Canononical microsporidia	Facilispora margolisi	HM800850
Canononical microsporidia	Microgemma caulleryi	AY033054
Canononical microsporidia	Tetramicra brevifilum	AF364303
Canononical microsporidia	Spraguea sp.	AB623034
Canononical microsporidia	Pleistophora typicalis	AF044387
Canononical microsporidia	Pleistophora mirrandellae	AF104085
Canononical microsporidia	Nematocenator marisprofundi	JX463178
Canononical microsporidia	Perezia nelsoni	KX856426
Canononical microsporidia	Agmasoma penaei	KF549987
Canononical microsporidia	Myospora metanephrops	HM140499
Canononical microsporidia	Paradoxium irvingi	KU163282
Canononical microsporidia	Hyperspora aquatica	KX364284
Canononical microsporidia	Unikaryon legeri	KX364285
Canononical microsporidia	Inodosporus octosporus	MH911629
Canononical microsporidia	Enterospora canceri	HE584634
Canononical microsporidia	Ameson pulvis	KC465966
Canononical microsporidia	Enterocytozoon hepatopenaei	FJ496356
Canononical microsporidia	Nadelspora canceri	AY958070
Canononical microsporidia	Pleistophora sp.	KF830721
Canononical microsporidia	Spraguea lophii	AF056013
Canononical microsporidia	Microsporidium cypselurus	AJ300706
Canononical microsporidia	Ameson michaelis	L15741
Canononical microsporidia	Tuzetia weidneri	AJ252958
Canononical microsporidia	Loma salmonae	U78736
Canononical microsporidia	Nucleospora salmonis	U78176
Canononical microsporidia	Dictyocoela duebenum	AF397404
Canononical microsporidia	Glugea atherinae	U15987
Canononical microsporidia	Kabatana takedai	AF356222
Canononical microsporidia	Glugea anomala	AF044391
Canononical microsporidia	Paranucleospora theridion	FJ594987

Canononical microsporidia	Hepatospora eriocheir	HE584635
Canononical microsporidia	Heterosporis anguillarum	AF387331
Canononical microsporidia	Microsporidium cerebralis	JQ316511
Canononical microsporidia	Novothelohania ovalae	JF826419
Canononical microsporidia	Takaokaspora nipponicus	KF110991
Canononical microsporidia	Takaokaspora nipponicus	KF110990
Canononical microsporidia	Kneallhazia carolinensae	GU173849
Canononical microsporidia	Encephalitozoon cuniculi	L39107
Canononical microsporidia	Endoreticulatus schubergi	L39109
Canononical microsporidia	Nosema bombycis	L39111
Canononical microsporidia	Vittaforma corneum	L39112
Canononical microsporidia	Parathelohania anophelis	AF027682
Canononical microsporidia	Oligosporidium occidentalis	AF495379
Canononical microsporidia	Multilamina teevani	KC990122
Canononical microsporidia	Rugispora istanbulensis	KR704648
Canononical microsporidia	Anncaliia algerae	HM216911
Canononical microsporidia	Orthosomella operophterae	AJ302317
Canononical microsporidia	Amblyospora sp	AJ252949
Canononical microsporidia	Flabelliforma montana	AJ252962
Canononical microsporidia	Vavraia culicis	AJ252961
Canononical microsporidia	Antonospora scoticae	AF024655
Canononical microsporidia	Tubulinosema kingi	DQ019419
Canononical microsporidia	Cystosporogenes legeri	AY233131
Canononical microsporidia	Alternosema bostrichidis	KP455651
Canononical microsporidia	Vairimorpha ceraces	EU267796
Canononical microsporidia	Vavraia oncoperae	X74112
Canononical microsporidia	Vairimorpha sp	AF031539
Canononical microsporidia	Percutemincola moriokae	LC136798
Canononical microsporidia	Myrmecomorba nylanderiae	KR704917
Canononical microsporidia	Senoma globulifera	DQ641245
Canononical microsporidia	Andreanna caspii	EU664450
Canononical microsporidia	Trichotosporea pygopellita	HM594267
Canononical microsporidia	Paranosema grylli	AY305325
Canononical microsporidia	Heterovesicula cowani	EU275200
Canononical microsporidia	Microsporidium sp.	FJ865221
Canononical microsporidia	Mockfordia xanthocaeciliae	FJ865223
Canononical microsporidia	Systenostrema alba	AY953292
Canononical microsporidia	Liebermannia dichroplusae	EF016249
Canononical microsporidia	Nematocida parisii	FJ005051
Canononical microsporidia	Trachipleistophora sp.	DQ403816
Canononical microsporidia	Ovavesicula popilliae	EF564602
Canononical microsporidia	Caudospora simulii	AY973624
Canononical microsporidia	Vairimorpha necatrix	Y00266
Canononical microsporidia	Amblyospora connecticus	AF025685
Canononical microsporidia	Amblyospora bracteata	AY090068
Canononical microsporidia	Hyalinocysta chapmani	AF483837
Canononical microsporidia	Janacekia debaisieuxi	AY090070
Canononical microsporidia	Polydispyrenia simuli	AY090069
Canononical microsporidia	Octosporea muscaedomesticae	FN794114
Canononical microsporidia	Pancytospora epiphaga	KX424959

Canononical microsporidia	Enterocytozoon bieneusi	L07123
Canononical microsporidia	Microsporidia sp. clone 134	MH712042
Canononical microsporidia	Microsporidia sp. EC-1	KR080135
Canononical microsporidia	Microsporidia sp. EC-2	KT956053
Canononical microsporidia	Microsporidia sp. EC-3	KT956054
Canononical microsporidia	Microsporidia sp. HE1	KT380106
Canononical microsporidia	Microsporidia sp. HM1	KT380107
Canononical microsporidia	Microsporidia sp. isolate Persian Gulf	MT192526
Canononical microsporidia	Microsporidia sp. isolate X27992-16	MF373108
Canononical microsporidia	Microsporidia sp. MB	MT160806
Canononical microsporidia	Microsporidia sp. MB-2008	EU589246
Canononical microsporidia	Microsporidia sp. NZ_CHIRI	
Canononical microsporidia	Microsporidia sp. NZN1	MT462180
Canononical microsporidia	Microsporidia sp. NZS16	MT462181
Canononical microsporidia	Microsporidia sp. OB1	MG460789
Canononical microsporidia	Microsporidia sp. OB2	MG456600
Canononical microsporidia	Microsporidia sp. YuT-2019a	MH118301
Canononical microsporidia	Microsporidium luciopercae	KX351969
Canononical microsporidia	Microsporidium sp.	MT928885
Canononical microsporidia	Microsporidium sp.	MF429927
Canononical microsporidia	Microsporidium sp.	KU160261
Canononical microsporidia	Microsporidium sp.	KY615714
Canononical microsporidia	Microsporidium sp.	MF153501
Canononical microsporidia	Microsporidium sp.	KY615712
Canononical microsporidia	Microsporidium sp.	KU160248
Canononical microsporidia	Microsporidium sp.	KY615713
Canononical microsporidia	Microsporidium sp.	MN748940
Canononical microsporidia	Microsporidium sp.	MN744350
Canononical microsporidia	Microsporidium sp.	KU160255
Canononical microsporidia	Microsporidium sp. 1049	FN434092
Canononical microsporidia	Microsporidium sp. 1154	FN610844
Canononical microsporidia	Microsporidium sp. 1199	FN610845
Canononical microsporidia	Microsporidium sp. 4-YYS-2009d	FJ865224
Canononical microsporidia	Microsporidium sp. 505	FN434085
Canononical microsporidia	Microsporidium sp. BPAR3	KR871374
Canononical microsporidia	Microsporidium sp. C81	DQ521753
Canononical microsporidia	Microsporidium sp. CRANA	AJ966721
Canononical microsporidia	Microsporidium sp. CRANB	AJ966722
Canononical microsporidia	Microsporidium sp. CRANFA	AJ966723
Canononical microsporidia	Microsporidium sp. CRANPA	AJ966718
Canononical microsporidia	Microsporidium sp. CRANPB	AJ966719
Canononical microsporidia	Microsporidium sp. CRANPC	AJ966720
Canononical microsporidia	Microsporidium sp. deacon-cargill	MF428413
Canononical microsporidia	Microsporidium sp. Dhae17W	KP699690
Canononical microsporidia	Microsporidium sp. G	KR871369
Canononical microsporidia	Microsporidium sp. GPM2	HM991452
Canononical microsporidia	Microsporidium sp. H7	KJ019845
Canononical microsporidia	Microsporidium sp. HEM-2006AHa	AM259662
Canononical microsporidia	Microsporidium sp. HEM-2006BHL	AM411639
Canononical microsporidia	Microsporidium sp. HEM-2006BIOS	AM411626

Canononical microsporidia	Microsporidium sp. HEM-2006DP	AM259660
Canononical microsporidia	Microsporidium sp. HEM-2006LFb	AM259666
Canononical microsporidia	Microsporidium sp. HEM-2006MCD	AM411637
Canononical microsporidia	Microsporidium sp. HEM-2006WLa	AM411635
Canononical microsporidia	Microsporidium sp. I	KR871371
Canononical microsporidia	Microsporidium sp. IVB	MK719421
Canononical microsporidia	Microsporidium sp. IVD	MK719290
Canononical microsporidia	Microsporidium sp. JES2002G	AJ438962
Canononical microsporidia	Microsporidium sp. JES2002H	AJ438963
Canononical microsporidia	Microsporidium sp. JES2002I	AJ438964
Canononical microsporidia	Microsporidium sp. JI-2014	KJ019846
Canononical microsporidia	Microsporidium sp. JZ-2016	KR263870
Canononical microsporidia	Microsporidium sp. Loire	HM566198
Canononical microsporidia	Microsporidium sp. M1	KF894401
Canononical microsporidia	Microsporidium sp. M3	KF894403
Canononical microsporidia	Microsporidium sp. MH	KR871382
Canononical microsporidia	Microsporidium sp. MSCLHCY01	HM800853
Canononical microsporidia	Microsporidium sp. Oise	HM566197
Canononical microsporidia	Microsporidium sp. PT11	KP966297
Canononical microsporidia	Microsporidium sp. RW-2009a	KR871370
Canononical microsporidia	Microsporidium sp. STF	AY140647
Canononical microsporidia	Microsporidium sp. SYN	AJ966725
Canononical microsporidia	Microsporidium sp. ZYS-2	EU012489
Canononical microsporidia	Microsporidium sp. ZYS-4	EU012491
Outgroups	Nucleophaga amoebae	JQ288099
Outgroups	Paramicrosporidium saccamoebae	JQ796369
Outgroups	Paramicrosporidium vannellae	JQ796368
Outgroups	Nucleophaga terricolae	KX017226
Outgroups	Mitosporidium daphniae	MF278562
Outgroups	Uncultured eukaryote LKM15	AJ130850
Outgroups	Rozella llomycis	AY635838
Outgroups	Amoeboaphelidium protococcarum	JX507298
Outgroups	Aphelidium tribonemae	KY129663
Outgroups	Rozella sp. strain JEL880	MF196183
Outgroups	Amphiamblys sp.	KX214674
Outgroups	Chytridiopsis typographi	MH728789
Outgroups	Amphiacantha sp.	KX214676

Supplementary table 6.2 The list of rRNA sequences used to infer the phylogenetic tree of Opisthosporidia

Group	Organism	Accession No.
Amphiacantha	Amphiacantha sp. ex Lecudina cf. elongata	KX214676
Amphiamblys	Amphiamblys sp. WSBS2006	KX214672
Amphiamblys	Amphiamblys sp. WSBS2011	KX214674
Anthozoa	Dendrobathypathes boutillieri	FJ389900
Aphelida	uncultured eukaryote	EF100212
Aphelida	uncultured eukaryote	HQ191302
Aphelida	Amoeboaphelidium protococcarum	JX507298
Aphelida	Amoeboaphelidium occidentale	JX967274
Aphelida	Aphelidium tribonemae	KY129663
Aphelidea	uncultured Chytridiomycota	GQ995419
Aphelidea	uncultured fungus	GU067817
Aphelidea	Paraphelidium letcheri	KY412789
Choanoflagellates	Salpingoeca urceolata	EU011931
Chytridiopsis	Chytridiopsis typographi	MH728789
Clade C	uncultured fungus	GU067982
Fungi	Kuzuhaea moniliformis	AB016010
Fungi	Schizangiella serpentis	AF368523
Fungi	Mucor racemosus	AJ271061
Fungi	[Candida] glabrata	AY198398
Fungi	Allomyces arbusculus	AY552524
Fungi	Coprinopsis cinerea	CCIRRNA
Fungi	Cochlonema euryblastum	DQ520640
Fungi	Rhizopus stolonifer	DQ536474
Fungi	Pseudorhizidium endosporangiatum	DQ536484
Fungi	Neurospora crassa	FJ360521
Fungi	Rhodotorula glutinis	HQ420261
Fungi	Smittium culicis	JQ302893
Laz IX	uncultured Nucleariidae	EF023360
Laz V	uncultured Eimeriidae	EF023474
Laz VII	uncultured Banisveld eukaryote	EU091829
Laz VIII	uncultured fungus	EU162634
LKM11	uncultured Cryptomycota	AB971033
LKM11	uncultured fungus	AF372713
LKM11	uncultured eukaryote	AJ130849
LKM11	uncultured eukaryote	AY916571
LKM11	uncultured fungus	FJ687268
LKM46	uncultured fungus	DQ244018
Microsporidia	Spraguea sp. Sdu-2008	AB623034
Microsporidia	Parathelohania anophelis	AF027682
Microsporidia	Glugea anomala	AF044391
Microsporidia	Ovipleistophora mirandellae	AF104085
Microsporidia	Ordospora colligata	AF394529
Microsporidia	Dictyocoela duebenum	AF397404
Microsporidia	Bryonosema plumatellae	AF484691
Microsporidia	Trichonosema pectinatellae	AF484695
Microsporidia	Thelohania contejeani	AF492593
Microsporidia	Oligosporidium occidentalis	AF495379

Microsporidia	Orthosomella operophterae	AJ302317
Microsporidia	Hamiltosporidium magnivora	AJ302319
Microsporidia	Intrapredatorus barri	AY013359
Microsporidia	Berwaldia schaefernai	AY090042
Microsporidia	Cystosporogenes legeri	AY233131
Microsporidia	Caudospora simulii	AY973624
Microsporidia	Thelohania butleri	DQ417114
Microsporidia	Euplotespora binucleata	DQ675604
Microsporidia	Heterovesicula cowani	EU275200
Microsporidia	Andreanna caspii	EU664450
Microsporidia	Paranucleospora theridion	FJ594987
Microsporidia	Mockfordia xanthocaeciliae	FJ865223
Microsporidia	Helmichia lacustris	GU130406
Microsporidia	Crispospora chironomi	GU130407
Microsporidia	Enterospora canceri	HE584634
Microsporidia	Anncaliia algerae	HM216911
Microsporidia	Nematocenator marisprofundi	JX463178
Microsporidia	Enterocytospora artemiae	JX915760
Microsporidia	Sporanauta perivermis	KC172651
Microsporidia	Agglomerata cladocera	KT950767
Microsporidia	Hyperspora aquatica	KX364284
Microsporidia	Loma salmonae	LSU78736
Microsporidia	Fibrillaspora daphniae	MF278272
Microsporidia	Inodosporus octosporus	MH911629
Microsporidia	Nosema bombycis	NMERGOA
Microsporidia	Microsporidia sp. NZ_CHIRI	NZ_chirinomidae
Microsporidia	Microsporidia sp. NZ16	MT462181
Microsporidia	Vavraia oncoperae	X74112
Mitosporidium	Mitosporidium daphniae	MF278562
Morellospora	uncultured eukaryote	AB901699
Morellospora	uncultured fungus	AF372718
Morellospora	uncultured eukaryote	AJ130850
Nuclearia	Nuclearia simplex	KY454455
Nuclearia	Nuclearia pattersoni	LS974781
Nucleophaga	Nucleophaga amoebae	JQ288099
Nucleophaga	Nucleophaga terricolae	KX017226
Paramicrosporidia	uncultured eukaryote	FJ577832
Paramicrosporidia	Paramicrosporidium vannellae	JQ796368
Paramicrosporidia	Paramicrosporidium saccamoebae	JQ796369
Paramicrosporidia	uncultured eukaryote	LC150098
Parvularia	Parvularia atlantis	KY113120
Rozella	uncultured Cryptomycota	AB971037
Rozella	Rozella sp. JEL347	AY601707
Rozella	Rozella allomycis	AY635838
Rozella	Rozella sp.	KX354831
Rozella	Rozella sp.	MF196183
WIM27	uncultured fungus	DQ244011