UNIVERSITÉ DE BOURGOGNE FRANCHE-COMTÉ, France - UMR CNRS 6282 Biogéosciences, Equipe Ecologie Evolutive.
UNIVERSITY OF LODZ, Poland - Department of Invertebrate Zoology and Hydrobiology.

Philosophiæ doctor in Life Sciences, Ecology and Evolution

Adrien QUILES

Evolutionary histories of symbioses between microsporidia and their amphipod hosts : contribution of studying two hosts over their geographic ranges.

Ph.D defense will be held the 12/13/2019 amphitheatre Monge, 6 boulevard Gabriel, 21000 Dijon, 13/12/2019

Jury :

Didier BOUCHON, Professor	Poitiers University	Reviewer
Łukasz KAJTOCH, Doctor habilitated	Polish Academy of Sciences (Kraków)	Reviewer
Isabelle FLORENT,Professor	National Museum of Natural History (Paris)	Examiner
Jerzy BANBURA, Professor	Lodz University	Examiner
Thierry RIGAUD, Research director	Scientific Research National Center	Supervisor
Rémi WATTIER, Doctor	Burgundy University	Co-supervisor
Karolina BACELA, Doctor habilitated	Lodz University	Co-Supervisor

PROJET FINANCÉ PAR L'ANR - PROJECT FUNDED BY THE ANR

"Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution."

Albert Einstein

As quoted in "What Life Means to Einstein: An Interview by George Sylvester Viereck"
in The Saturday Evening Post (26 October 1929)

ACKNOWLEDGMENT

Special thanks to; All the people who helped me during the development of my thesis and especially my directors, Thierry RIGAUD, Rémi WATTIER and Karolina BACELA-SPYCHALSKA how trusted me to carry this project, their great availability and all their advices during those 3 years of team work.

I would like, to thank the members of my jury to accept the evaluation of this thesis work: Didier BOUCHON, Łukasz KAJTOCH, Isabelle FLORENT, Jerzy BANBURA. I know your time is precious and I am grateful for the comments, criticisms and advice you will give me and I will make the best use of it to improve my work.

This work would not have been possible without the support of the Agence Nationale de la Recherche (ANR-15-CE32-0006), and the Polish ministry of science and education (N N303 579439 and N N304 350139), which allowed me, thanks to those research allocation and various financial support, to devote myself serenely to the development of my thesis.

Michal GRABOWSKI who help me improving articles and part of the thesis. Tomasz REWICZ, Tomasz MAMOS which provided me inestimable valuable materials of Gammarids samples from all across Europe, to advance this research.

Sébastien DUPERRON and Richard CORDEAU for their help and many advise during thesis supervisory committee.

This work could not have been carried out without the availability and friendship of Maria TEIXEIRA which showed me the molecular biological lab and help me doing experiences. Manon CHOGNE, Camille SABARLY, Sebastien MOTREUIL and Alexandre BAUER for their friendship in the laboratory and outside.

Aleksandra JABLONSKA for her friendship when I was in cold winter time in Poland.

The help of tree hard worker third degree student, Nicolas LAMBIN, Vincent GARDET, Laurène CHEVALIER for dissecting, extracting DNA and screening thousands of individuals for microsporidia with me.

At the end of this journey, I finally thank those who are dear to me and that I have somewhat neglected during those tree year to complete this thesis. Their attentions and encouragements have accompanied me throughout these years. To my parents, Muriel and Jean-Marc QUILES my brother Robin QUILES. I have a particular thought for my grandmothers, Daisy CAMIER and Nadine QUILES.

I would like to thanks all my friends for their support and trust.

CONTENTS

NOTIFICATIONS

This thesis is composed of a general background, three chapters and a general conclusion.

Each chapter is constructed like a ‘’journal paper’’. Therefore, chapters can be taken separately. As a consequence, no general material and methods or references list section is provided.

Chapter I was published in Parasites & Vectors in 2019:
Adrien Quiles, Karolina Bacela‐Spychalska , Maria Teixeira, Nicolas Lambin, Michal Grabowski,Thierry Rigaud and Rémi André Wattier. 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. Parasites & Vectors, 12:327.

Chapers II and III are unpublished manuscripts.

All suplementary material are available at:
https://cloud.u-bourgogne.fr:443/index.php/s/XP9ZW8wnr87kP7z

or scanning the following QR code:

GENERAL BACKGROUND

SYMBIOSES AND THEIR EVOLUTIONARY CONSEQUENCES

Co-evolutionary patterns and processes

Species do not live in isolation but interact with others and the environments. Interaction between species is so powerful that it profoundly influences all organisms. Paradoxically however, symbiosis is a concept which appears very recently in the history of sciences. The first definition of symbiosis was given in 1878 by Heinrich Anton de Bary who defined symbiosis as organisms belonging to different species living together (English translation of Heinrich Anton de Bary's 1878 speech, 'Die Erscheinung der Symbiose' (Oulhen *et al.* 2016)). De Bary was the first to suggest that lichen could be an association between algae and a fungi, and not a plant as they were considered at that time. The vocabulary used to describe the different natures of the symbiotic relationships really developed ten years later with the work of the Zoologist Pierre-Joseph van Beneden. He formalized the concepts of parasitism, mutualism and commensalism in his book entitled “Les commensaux et les parasites” (Beneden 1875). He defined mutualism as a reciprocal benefits relationship without any cost for both partners. Commensalism is described as relationship that benefits from its association to one partner without damaging the other. Parasitism or Predation are described as a symbiotic relationship where one partner/species gains benefits from the association (at the level of survival or fitness) to the detriment of the other. Amensalism is an association in which one is inhibited or destroyed by the other which is still unaffected. Finally, competition described as neither benefits for both partners. This cots-benefits categorization helped to understand and study the symbioses, but most evolutionary biologists now agree on the fact that symbiotic relationships may always be described as relationships situated somewhere in the continuum between mutualism and parasitism (Leigh 2010; Lin & Koskella 2015; Matthews *et al.* 2019).

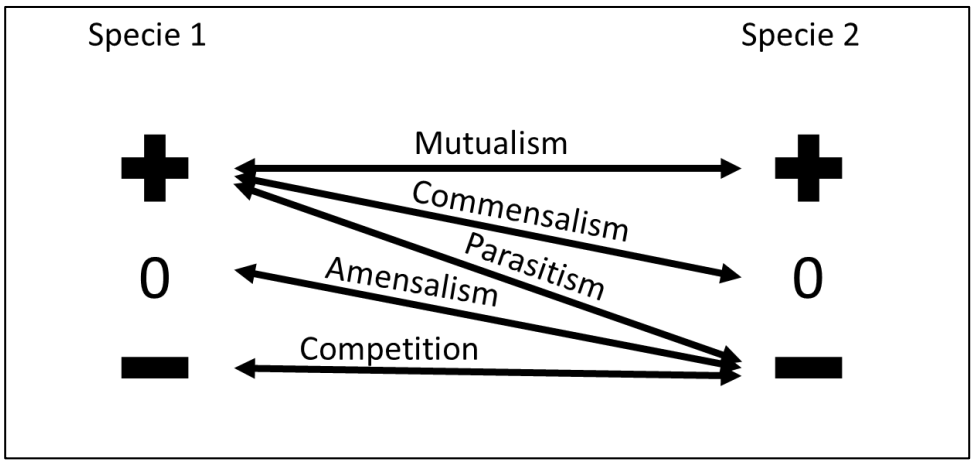


Fig. 1. Symbiosis, inter-species relationships according to their impact on each of the two species. Reprinted from (Duperron 2016).

Many of these symbiotic interaction can lead to evolutionary effects on both interacting species, in particular co-adaptations or co-evolutionary processes. We will define here co-evolution in its broader sense, *i.e.* as the phenomenon in which the evolution of one organism influences that of one or several others. The term of co-adaptation was used by Darwin for two organisms adapting to one another (Darwin 1859). This constant conflict between host and parasite is not fixed, and each species must continuously adapt just to maintain the existing relationship. The description of this process was formalized by Van Valen (1973), and named the Red Queen hypothesis in homage to Lewis Carroll's "Alice in Wonderland" (quoted : "Now , here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"). Coevolution between host and parasite is following this principle: any changes in one partner will have an impact and unbalance the relationship with the other, which will be forced to adapt. The narrow sense of the Red Queen Hypothesis is now mostly used by many evolutionary biologists (more precisely population geneticists) studying parasitic interactions, as the process by which genetic variation in host populations is tracking the genetic variation of parasites in a runaway process of ever-evolving defences (in the hosts) and counter defences (in parasites) (see Carius *et al.* 2001; Dybdahl & Lively 1998, etc.). However, in its original sense, the Red Queen Hypothesis was englobing many other broader processes, and was not restricted to parasitic relationships. These processes led to diverse co-evolutionary patterns such as impacts on host genome evolution (*e.g.* Bourgeois *et al.* 2017), co-diversification of the partners at large evolutionary time scales (*e.g.* Harmon *et al.* 2019) and the evolution of specificity (*e.g.* Boots *et al.* 2014). Mutualisms can also be included in co-evolutionary processes (*e.g.* Rubin & Moreau 2016).

An extreme pattern of coevolution is when interacting species have a speciation event in parallel with eachother, leading to co-speciation (Fig. 2a). Co-speciation can appear in very close associations (*e.g.* symbiosis) but can also be due to external causes (*e.g.* geographic barrier, sexual selection, etc.). Co-phylogeny studies are trying to unravel congruence between such interacting groups. Strict co-phylogeny occurs if every speciation event in one group (*e.g.* host) is accompanied by speciation in the other group (*e.g.* parasites) (Fig. 2a). However, many observed patterns may obscure coevolutionary processes and phylogenetic studies per se may not help to understand the processes. For example, co-phylogenetic pattern will be lost if associations are changed (*e.g.* following a host switch, Fig. 2.b or if a symbiont go extinct, Fig. 2g, the two phenomena being possibly additive, Fig. 2 c,d,e). In addition, co-speciation should not be taken, *sensu stricto*, as a pattern of co-evolutionary congruence. For example, a co-phylogenetic congruence can happen between two groups which are not interacting, but that experienced a

common isolation event (*e.g.* after the rise of a geographic barrier)(Althoff *et al.* 2014; Smith *et al.* 2008). This pattern can also happen when parasites are colonizing a succession of hosts. In that case, parasite phylogenies may be congruent, even if they are not interacting but 'just' share the same host phylogeny (Hafner & Page 1995).

In fact, clear co-phylogenies are very rare, or at least uncommon (De Vienne *et al.* 2007). It is more common two show some evidence of co-speciation mixed with another event that scrambles strict congruence.

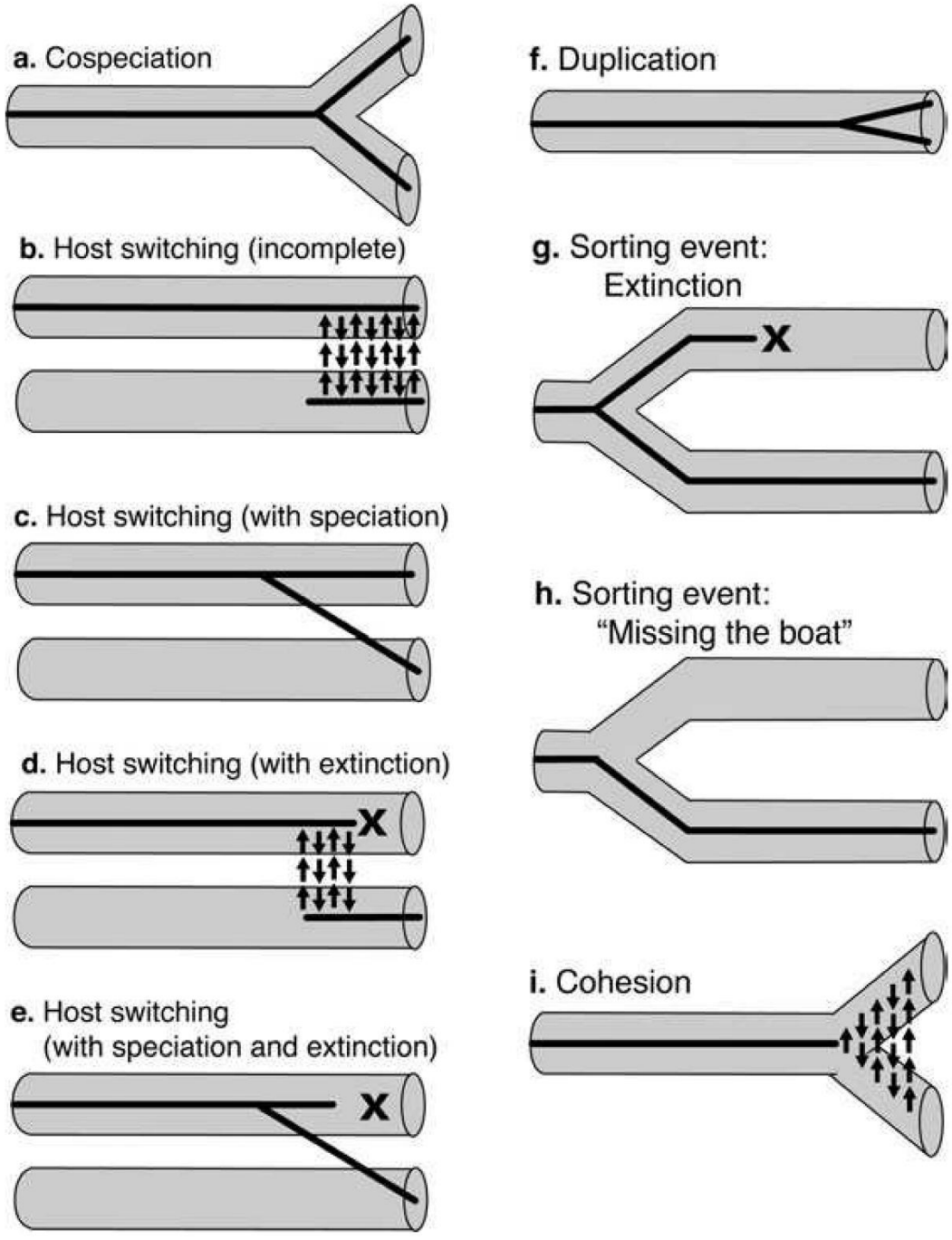


Fig. 2. Different processes leading to different evolutionary patterns of coevolving hosts and symbionts. Grey pipes represent hosts, and black lines represent parasites. Arrows represent gene flow and X's represent extinction. (a) Co-speciation, in which the host and parasite speciate at the same time, yielding congruent phylogenies; (b) Incomplete host switching, in which a parasite colonizes a novel host, but maintains gene flow with parasites on the original host; (c) Host switching with speciation; (d) Host switching in which the parasite colonizes a new host, then goes extinct on the original host; (e) Host switching in which the parasite colonizes a new host, speciates, then goes extinct on the original host; (f) Duplication of parasites on a single host; (g) Cospeciation followed by extinction of one parasite; (h) "Missing the boat," in which the parasite fails to colonize one of two diverging host lineages; (i) Parasite cohesion, in which a parasite maintains gene flow between diverging host populations (Reprinted from (Clayton *et al.* 2015)).

Parasitism

Among interspecific interactions, parasitism is the most common way of living on earth (Lafferty 2010; Price 1980; Windsor 1998). Parasites are representing one-third of all eukaryotic animal (De Meeûs & Renaud 2002). A parasite is described as an organism living in (or on) another living organism, obtaining from its host part, or all, of its nutriments, usually showing some degree of adaptive structural modification, and causing damages to its host (Price 1980). Parasites are very diverse, including microbial pathogens, plants rust, endoparasitic organisms, herbivorous insect, ectoparasite, etc. (De Meeûs & Renaud 2002). Interactions could be seen at the community levels as almost any organism is in interaction with one or more other species. Obligate parasites cannot complete their life-cycle without exploiting a suitable host, making such host-parasite associations an ideal framework to study co-evolution and possible co-phylogenies. This is particularly true for obligate intracellular parasites.

A key characteristics of the parasite life-cycle may also be important when aiming to sudy host-parasite coevolution: the way the parasite is transmitted to its next host. Indeed, parasites may be transmitted in two different ways (not always exclusive). The most represented association involved parasitic re-establishment in a new host at each generation (Bright & Bulgheresi 2010). In that case, the host can live without its parasite during all or part of its life cycle, and parasites are not inherited from host' parents. This type of parasite acquisition is called horizontal transmission (Fig. 3). This lateral transmission comes from the environment where the parasite can live briefly. Interaction between partners begins before contact with detection, recognition and acquisition.

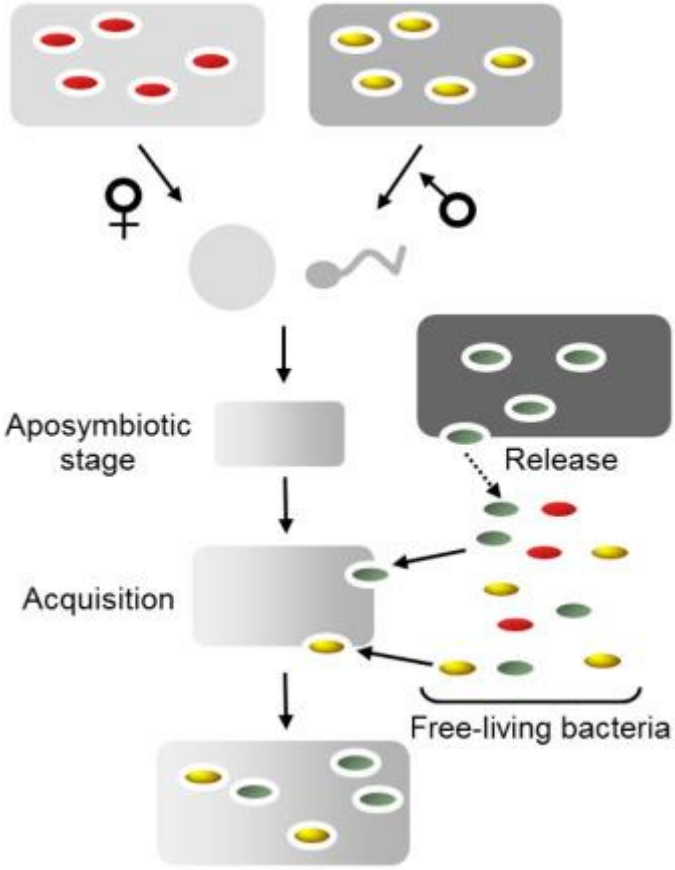


Fig. 3. Horizontal transmission of symbionts (here bacteria) from one host generation to the next. Host gametes do not contain any parasites. Parasites are acquired from another hosts (lateral transmission) or from the environment, in which free-living forms of the symbionts exist (environmental acquisition). These free-living forms may be rare or abundant, active or dormant. Acquisition may be limited to a particular phase of the cycle (competence window) or continue throughout the life cycle. The parasite lineages present in a host individual may, therefore, differ from those present in its parents. Reprinted from (Duperron 2016).

The other parasite transmission way is vertical transmission. In that case, most of the time, the infected host does not have any phase of its life cycle without the parasites. Parental infections through males are sometimes possible, for example in case of paternal parental care (Moran &

Dunbar 2006), but vertically-transmitted parasites generally pass from one host generation to another by the ability to infect the eggs (either because they are intracytoplasmic parasites, or because they infect the egg surface). In case of intracytoplasmic parasites, this transmission is strictly maternal due to the more substantial intracellular content of the female gamete. Most of the parasite predominantly develop near the ovaries of the host, but the transmission to the egg is poorly understood. In some well-studied biological models, it nevertheless seems that intracytoplasmic symbionts infect oocytes during their maturation process, suggesting that symbionts are taking advantage of the intense cytological processes occurring during vitellogenesis (*e.g.* intense phagocytosis) to entre the gametes (see for example Genty *et al.* 2014 for *Wolbachia* bacteria, and Dubuffet *et al.* 2013 for microsporidia).

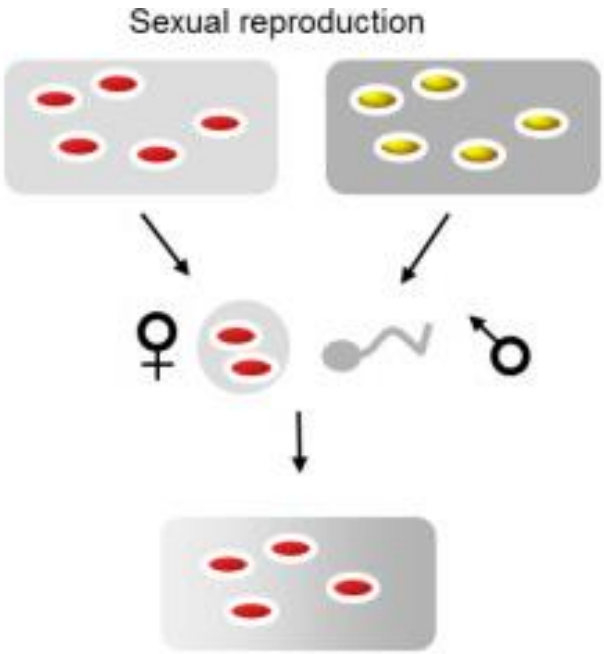


Fig. 4. Strictly maternal vertical transmission from one generation to the next one. The parents produce gametes, here a large female gamete containing a few parasites from the mother (left) and a small mobile male gamete containing no parasites. The zygote, the individual resulting from the fusion of the gametes, will have mixed genetics input from the two parents, but all of its symbionts will come from the mother. This is a strictly maternal transmission. Cases of co-transmission by both parents, in which both gametes contain symbionts, and the descendants inherit symbionts from both parents, are rare.. Reprinted from (Duperron 2016).

Most of the time, the transmission strategy is not strictly vertical or strictly horizontal (Ebert 2013). Using strictly one mode of transmission can have significant evolutionary consequences enabling parasite-host shuffling leading to possible extinction. In addition, it is to be noted that in the case of vertical transovarial transmission, reproduction transmits only a tiny fraction of the symbionts, inducing a potentially strong bottleneck into their life cycle at each generation (Kaltenpoth *et al.* 2010).

MICROSPORIDIA: the ultimate parasites?

General characteristics and cycle

Microsporidia are obligatory intracellular eukaryote parasites that can infect both animals and some protists in diversified habitats (Keeling & Fast 2002; Vávra & Lukeš 2013). These eukaryotes are obligate unicellular endoparasites composed of an extremely ancient and phylogenetically diverse phylum, with more than 1300 to 1500 species in 187 genera (Wittner 2014). The first microsporidia formally taxonomically described was *Nosema bombycis* (Naegeli 1857). This parasite was slightly earlier presented by Guérin-Méneville (1849-1850) as the agent causing the "pepper disease" ("pébrine" in French) ravaging silkworms (*Bombyx mori*) industry in southern Europe. Then Pasteur published a paper "*Etude sur la maladie des vers à soie*" in which he described how to prevent and control pébrine disease in the industry (Pasteur, 1870). However, at that time, microsporidia were considered as yeast. Later on, another microsporidia was described: *Nosema apis* causing similar and frequent infection in honeybees, inducing loss of many bees and collapse of the colonies. These microsporidia were recognised as a distinct group when Balbani (1882) proposed to name this division (Stentiford & Dunn 2014): Microsporidia. Then Sprague creates the phylum Microspora, which was updated several time afterwards (Sprague *et al.* 1992; Vossbrinck *et al.* 2014). Microsporidia are infecting many vertebrate, including humans, but they are also responsible for many diseases of insect, fish and crustaceans (Becnel & Andreadis 2014; Becnel & Takvorian 2014; Bulnheim 1975; Stentiford *et al.* 2016; Weber *et al.* 1994).

Microsporidia life cycle consists of two main phases, merogony and sporulation (Fig. 5) The germinating spore, injects throw an 'injection tube' the spore contents 'sporoplasm', into the cytoplasm of a host cell (Delbac & Polonais 2008; Franzen 2004; Franzen *et al.* 2005; Weidner 1972). The sporoplasm grows into cells called meronts, which divide by 'merogony' into daughter meronts. The meronts progressively fill the cytoplasm of the host cell. Then, sporulation starts after an unknown signal, notably associated with synthesis of proteins progressively deposited on the plasma membrane that will constitute part of the spore wall (Fig. 5) (Bohne *et al.* 2000; Brosson *et al.* 2005; Hayman *et al.* 2001; Li *et al.* 2009; Peuvel-Fanget *et al.* 2006; Southern *et al.* 2007; Wu *et al.* 2008, 2009; Xu *et al.* 2006). The spore is the product of internal differentiation of a single cell (Vávra & Larsson 2014). Spore is the only stage in which the microsporidia can survive in the environment and be disseminated. The spores are often released in the environment after the death of the host cell, and may infect the next host *via* an horizontal way. However, a number of spores (sometimes spores differentiated in a different way) also infect other cells of the very same

individual host by an "auto-infectious" process (Fig. 5) (Dunn *et al.* 2001). These spores allow to infect other organs of the host body. They may also infect gonads and germ cells, therefore opening the way for a vertical transmission pathway.

Microsporidia exhibit different transmission strategies and impose different virulence on their hosts. Horizontal transmission (HT) is occurring between conspecifics or between host species thanks to environmental spores. HT is linked mostly to high virulence. Vertical transmission (VT) may also be achieved, from mother to eggs and offspring thanks to auto-infectious spores. VT is usually associated with low or supposedly no virulence. Combination of both VT and HT are also found in a number of host species (Dunn & Smith 2001; Ebert 2013).

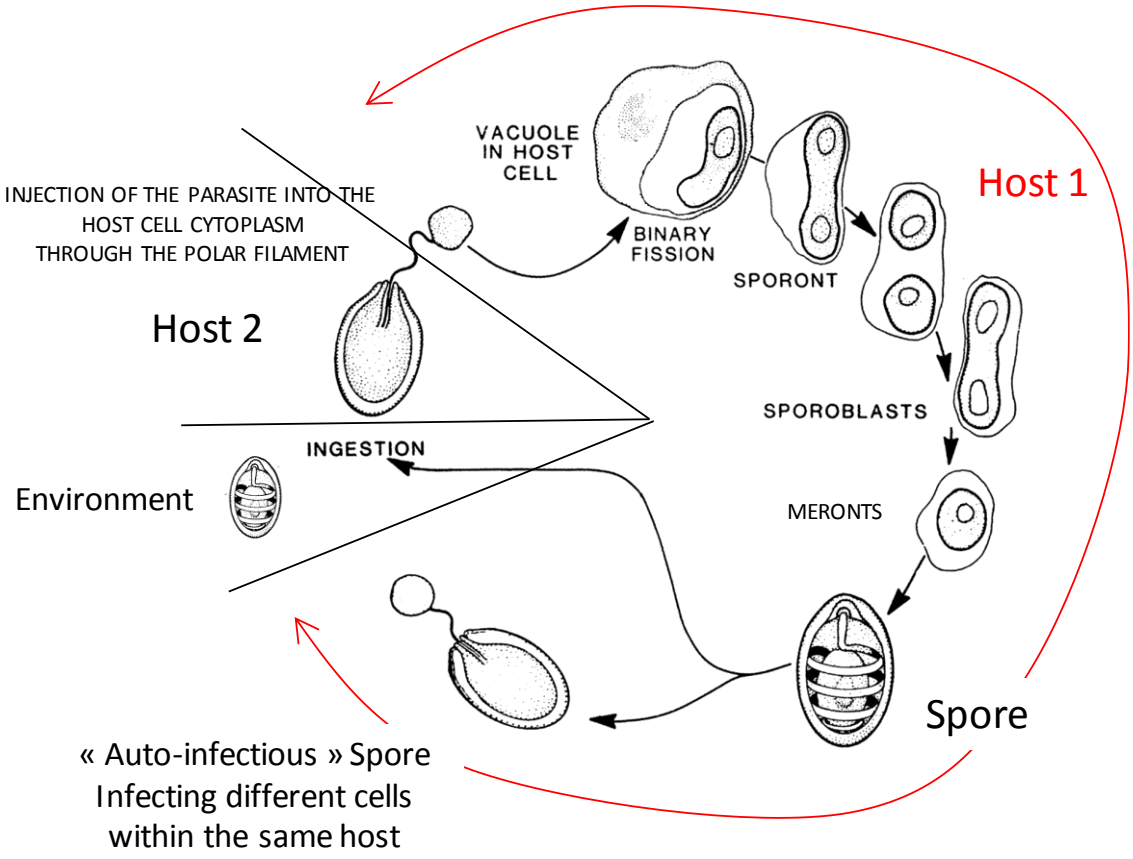


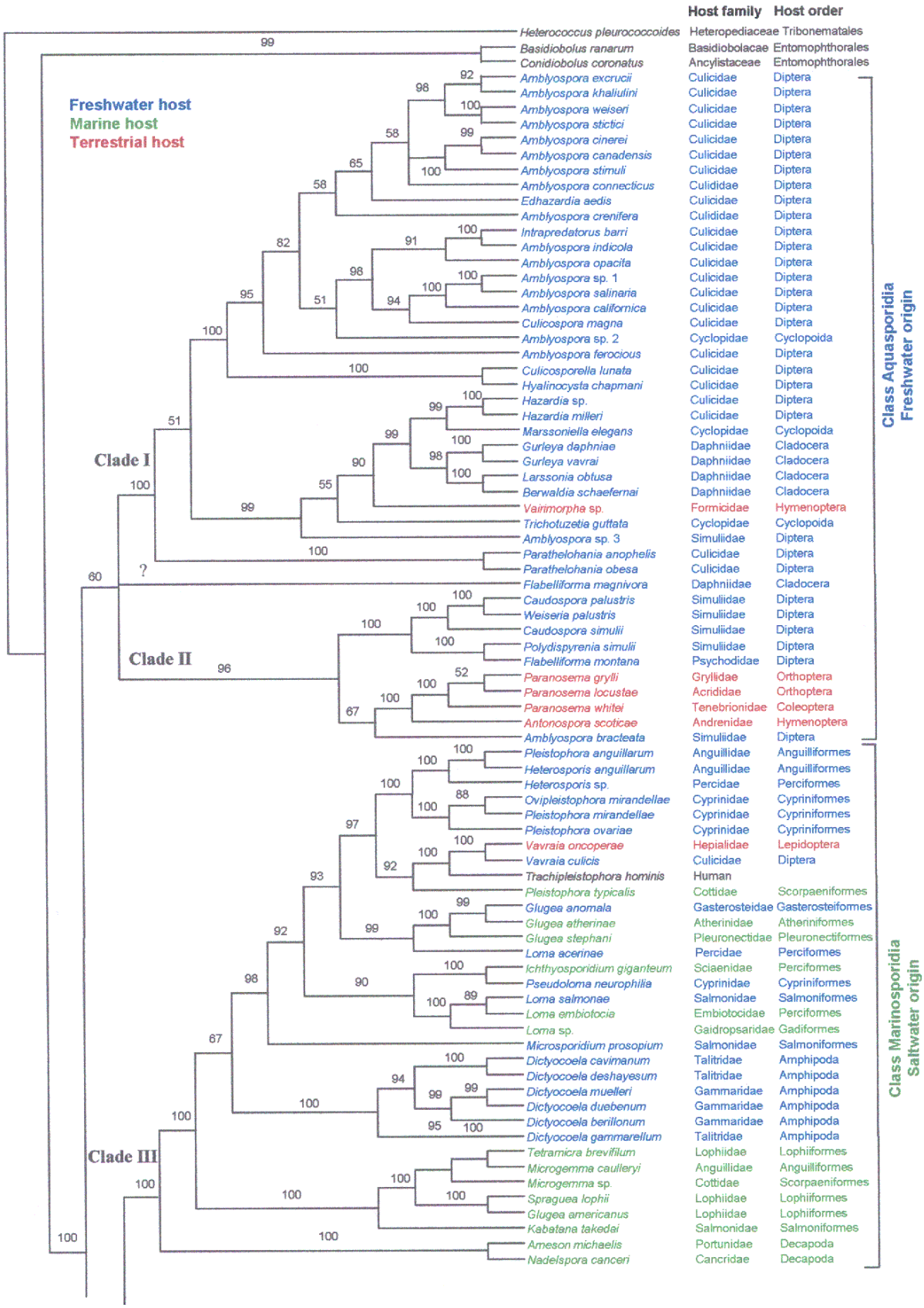
Fig. 5. An example of life cycle of microsporidia (fictive species). Redrawn from (Gardiner 1988).

Phylogeny

The microsporidia possess several characters call apomorphies that distinguish them as a taxon: (i) the presence, in the spore stage, of a coiled polar filament, (ii) the posterior vacuole, (iii) the anterior polaroplast, and (iv) the diplokaryon (not present in all species). Most recently, molecular analyses has revolutionised the microsporidia taxonomy (Vossbrinck *et al.* 2014). Microsporidia are now recognized to belong to fungi (Corradi & Selman 2013; Hibbett *et al.* 2007),

or are at least a sister group to fungi (Capella-Gutiérrez *et al.* 2012; Karpov *et al.* 2014). Modern microsporidian species descriptions are based on detailed microscopic and ultrastructural images as fare as the use of molecular markers (most of the time sequence of small subunit ribosomal DNA (SSU rDNA)) (*e.g.* Bacela-Spychalska *et al.* 2018). Literature continually increases recording of new microsporidia genera and species especialy in amphipods (Bacela-Spychalska *et al.* 2018; Bojko *et al.* 2015, 2017b; Dimova *et al.* 2018; Grabner 2017; Grabner *et al.* 2015; Haine *et al.* 2004; Ironside *et al.* 2003b, 2008; Ironside 2007; Ironside & Alexander 2015; Ironside & Wilkinson 2018; Krebes *et al.* 2010, 2014; Ovcharenko *et al.* 2010; Slothouber Galbreath *et al.* 2009; Terry *et al.* 1999; Weigand *et al.* 2016; Wilkinson *et al.* 2011; Winters & Faisal 2014; Yang *et al.* 2011).

Vossbrinck & Debrunner-Vossbrinck (2005) defined five major clades within the microsporidia phylogeny (Fig. 6). These parasite clades are mostly, but not exclusively, fitting the distributions of their hosts ecology, also allowing classifying microsporidia according to their host habitat. Aquasporidia (Clade I, II and V), are designated as primarily parasites infecting freshwater organisms. Marinosporidia (Clade III), represent parasites of marine hosts. Exceptions to this rule include *Dictyocoela* and *Cucumispora* parasites infecting some freshwater amphipods. It was argued by authors that *Dictyocoela* were initially parasites of marine organisms whose hosts have adapted freshwater habitats, bringing their parasites with them (Vossbrinck & Debrunner-Vossbrinck 2005). Finally, Terresporidia (clade IV), are terrestrial microsporidia mainly infecting insects. However, even if mircrosporidia of this group parasitize many Lepidoptera, Hymenoptera, Coleoptera, some also infect freshwater hosts (Moodie *et al.* 2003; Terry *et al.* 1999). This clade also includes infections of vertebrates (*e.g.* mammals, birds and reptiles). These clade numbering have been used since in différents studies (*e.g.* Stentiford *et al.* 2010). Nevertheless, since formal taxonomy rules, such as a holistic approach using morphological characters and molecular ones, haven't been used to produce this classification, this subdivision have faced many criticisms (Larsson 2005).


Freshwater host
Marine host
Terrestrial host
Host family
Host order
Heterococcus pleurococcoides Heteropediaceae Tribonematales
Basidiobolus ranarum Basidiobolacae Entomophthorales
Conidiobolus coronatus Ancylistaceae Entomophthorales
Amblyospora excrucii Culicidae Diptera
Amblyospora khaliulini Culicidae Diptera
Amblyospora weiseri Culicidae Diptera
Amblyospora stictici Culicidae Diptera
Amblyospora cinerei Culicidae Diptera
Amblyospora canadensis Culicidae Diptera
Amblyospora stimuli Culicidae Diptera
Amblyospora connecticus Culididae Diptera
Edhazardia aedis Culicidae Diptera
Amblyospora crenifera Culididae Diptera
Intrapredatorus barri Culicidae Diptera
Amblyospora indicola Culicidae Diptera
Amblyospora opacita Culicidae Diptera
Amblyospora sp. 1 Culicidae Diptera
Amblyospora salinaria Culicidae Diptera
Amblyospora californica Culicidae Diptera
Culicospora magna Culicidae Diptera
Amblyospora sp. 2 Cyclopidae Cyclopoida
Amblyospora ferocious Culicidae Diptera
Culicosporella lunata Culicidae Diptera
Hyalinocysta chapmani Culicidae Diptera
Hazardia sp. Culicidae Diptera
Hazardia milleri Culicidae Diptera
Marssoniella elegans Cyclopidae Cyclopoida
Gurleya daphniae Daphniidae Cladocera
Gurleya vavrai Daphniidae Cladocera
Larssonia obtusa Daphniidae Cladocera
Berwaldia schaefernai Daphniidae Cladocera
Vairimorpha sp. Formicidae Hymenoptera
Trichotuzetia guttata Cyclopidae Cyclopoida
Amblyospora sp. 3 Simuliidae Diptera
Parathelohania anophelis Culicidae Diptera
Parathelohania obesa Culicidae Diptera
Flabelliforma magnivora Daphniidae Cladocera
Caudospora palustris Simuliidae Diptera
Weiseria palustris Simuliidae Diptera
Caudospora simulii Simuliidae Diptera
Polydispyrenia simulii Simuliidae Diptera
Flabelliforma montana Psychodidae Diptera
Paranosema grylli Gryllidae Orthoptera
Paranosema locustae Acrididae Orthoptera
Paranosema whitei Tenebrionidae Coleoptera
Antonospora scoticae Andrenidae Hymenoptera
Amblyospora bracteata Simuliidae Diptera
Pleistophora anguillarum Anguillidae Anguilliformes
Heterosporis anguillarum Anguillidae Anguilliformes
Heterosporis sp. Percidae Perciformes
Ovipleistophora mirandellae Cyprinidae Cypriniformes
Pleistophora mirandellae Cyprinidae Cypriniformes
Pleistophora ovariae Cyprinidae Cypriniformes
Vavraia oncoperae Hepialidae Lepidoptera
Vavraia culicis Culicidae Diptera
Trachipleistophora hominis Human
Pleistophora typicalis Cottidae Scorpaeniformes
Glugea anomala Gasterosteidae Gasterosteiformes
Glugea atherinae Atherinidae Atheriniformes
Glugea stephani Pleuronectidae Pleuronectiformes
Loma acerinae Percidae Perciformes
Ichthyosporidium giganteum Sciaenidae Perciformes
Pseudoloma neurophilia Cyprinidae Cypriniformes
Loma salmonae Salmonidae Salmoniformes
Loma embiotocia Embiotocidae Perciformes
Loma sp. Gaidropsaridae Gadiformes
Microsporidium prosopium Salmonidae Salmoniformes
Dictyocoela cavimanum Talitridae Amphipoda
Dictyocoela deshayesum Talitridae Amphipoda
Dictyocoela muelleri Gammaridae Amphipoda
Dictyocoela duebenum Gammaridae Amphipoda
Dictyocoela berillonum Gammaridae Amphipoda
Dictyocoela gammarellum Talitridae Amphipoda
Tetramicra brevifilum Lophiidae Lophiiformes
Microgemma caulleryi Anguillidae Anguilliformes
Microgemma sp. Cottidae Scorpaeniformes
Spraguea lophii Lophiidae Lophiiformes
Glugea americanus Lophiidae Lophiiformes
Kabatana takedai Salmonidae Salmoniformes
Ameson michaelis Portunidae Decapoda
Nadelspora canceri Cancridae Decapoda
Clade I
Clade II
Clade III
?
Class Aquasporidia
Freshwater origin
Class Marinosporidia
Saltwater origin

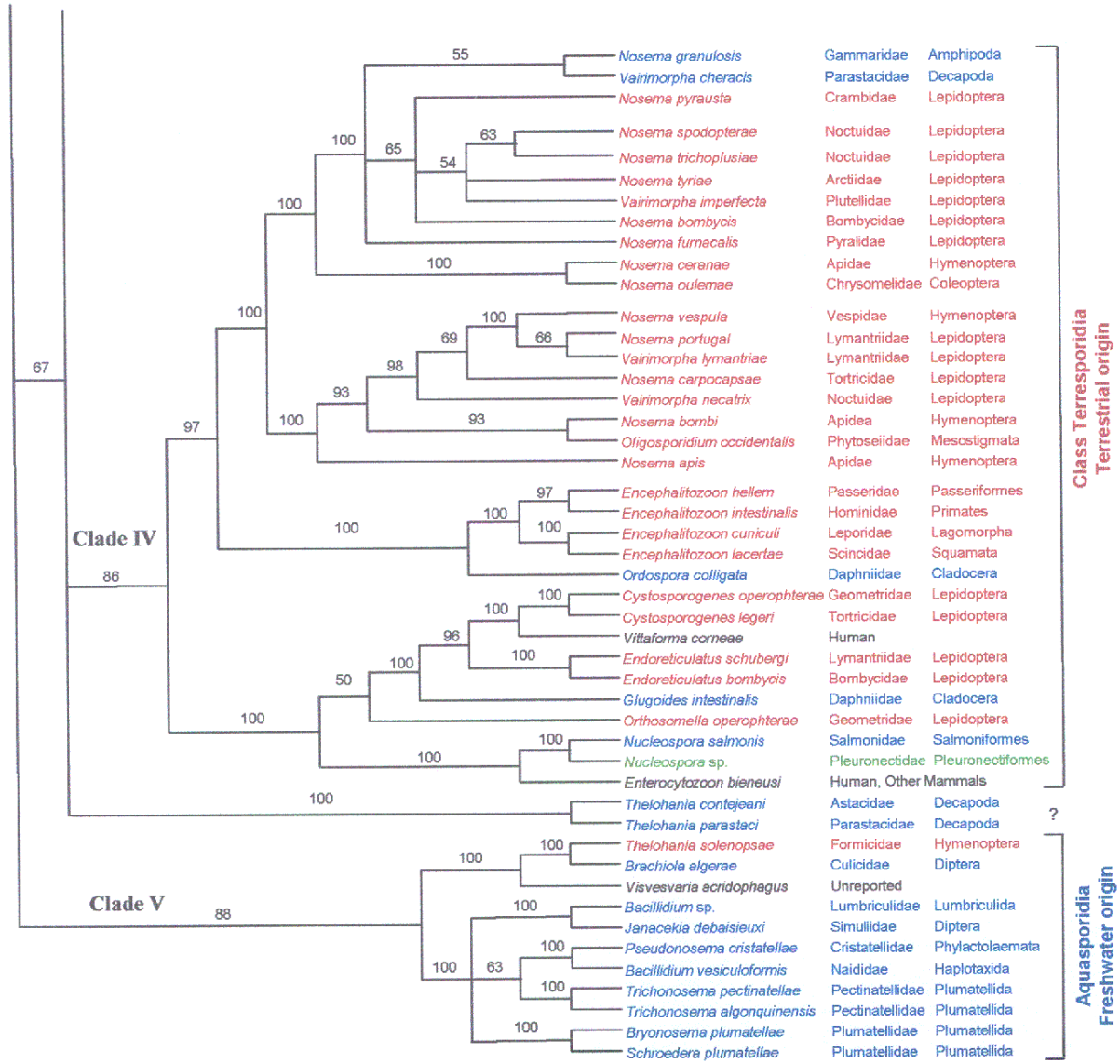


Fig. 6. Microsporidia phylogeny. Clades I–V are indicated. Microsporidia from freshwater hosts are indicated in blue, from marine hosts in green and from terrestrial hosts in red. Outgroups and microsporidians from human or unreported hosts are in black. Reprinted from (Vossbrinck & Debrunner-Vossbrinck 2005).

As noted before, microsporidia can infect many host taxa. There is however no real consensus on the host range a single microsporidian species can infect. Some Microsporidia are reported to infect many, even sometimes unrelated, host species (Weigand *et al.* 2016). For example, *Octosporea muscaedomesticae* was reported to infect seven species of Calyptratae flies (Vossbrinck *et al.* 2010). *Cytosporoenes* sp. is also known to infect many hosts with 19 lepidopteran species (Solter *et al.* 2000). *Pleistophora muelleri,* which is known to infect the

crustacean amphipod *Gammarus duebeni,* is suspected to used fish as intermediate or reservoir hosts (Ironside *et al.* 2008). Despite this apparent wide range of host possibility, some authors suggested high host specificity for many microsporidia (Shaw *et al.* 2000). These two opposite views are probably due to different ways of approaching this problem of specificity. First, a single microsporidia species can be found infecting several host species under experimental conditions, but this does not inevitably reflect natural condictions where ecological barriers exist between potential hosts and the parasites (Solter & Maddox 1998). Host-parasite ecology plays a role in the maintenance of host-parasite specificity and many hosts do not live in sympatry with other related hosts, preventing possible parasite exchanges. Second, 'specificity' may depends on the phylogenetic level under study. Microsporidia "species" is not a clear concept (this is probably true for most living entities, but especially true for microsporidia). Indeed, a few studies showed that similar morphological microsporidia may hide a large genetic variability (*e.g.* Bacela-Spychalska *et al.* 2018). Therefore, at what level should we have to consider that "a microsporidia is specific to its host"? At the morphological level? At the haplotype level? It has nevertheless been showed that microsporidia could evolve and co-speciated with their host's species or a group of related species (Baker *et al.* 1998; Shafer *et al.* 2009; Smith 2009). For example, *Loma* spp. microsporidia are showing some degree of co-speciation with their gadid fish host (Brown *et al.* 2010) as well as *Amblyospora* microsporidia infecting siberian mosquitoes (with nonetheless certain amount of host switching, Andreadis *et al.* 2012) (Fig. 7).

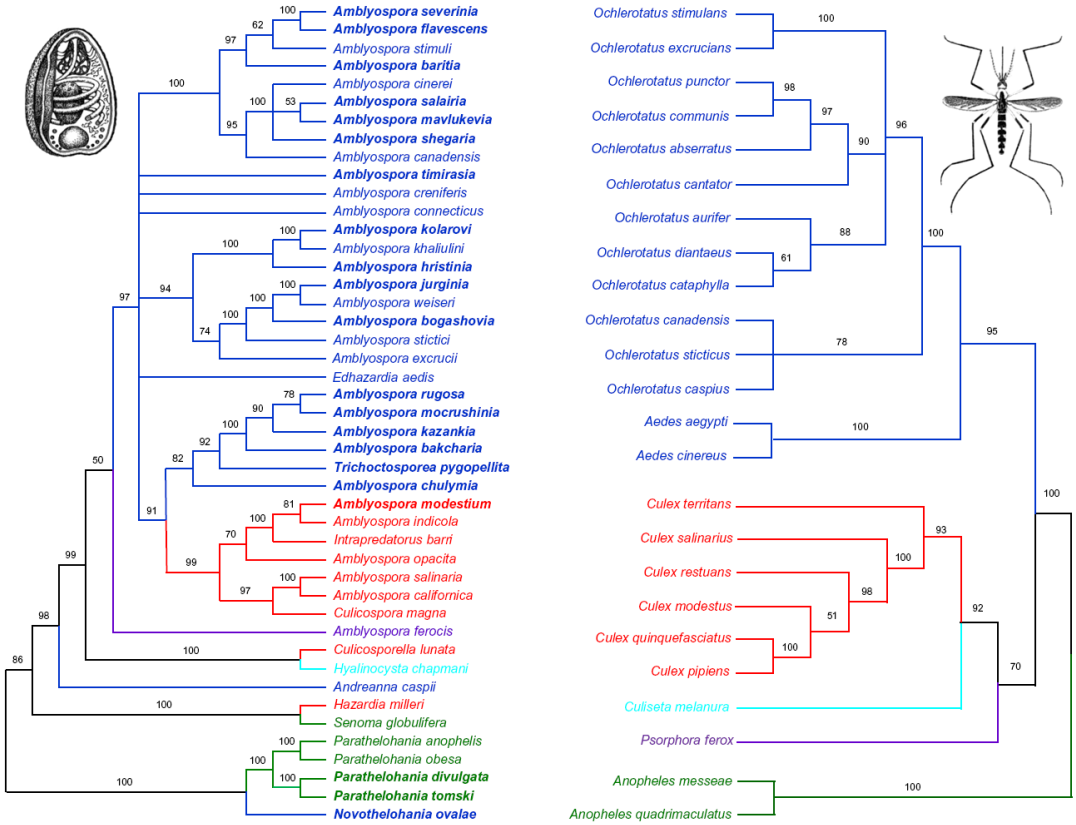


Fig. 7. Comparison of microsporidia and mosquito phtlogentic trees. Bootsraps analysis based on 1000 replicates using Neighbor Joining analysis. Microsporidia and corresponding mosquito host genera are color-coded. Reprinted from Andreadis *et al.* 2012.

Microsporidia are particularly found at high prevalence in aquatic organisms, causing several diseases (Stentiford *et al.* 2013). The impact of microsporidian diseases can impact the population and/or ecosystem levels and some of them have the potential to be emergent diseases (Stentiford *et al.* 2013). Those infecting the crustacean amphipods received inscreasing attention, both because of their unusual role in sex determination, as agents of modulation of competition between host species, and as potential emerging diseases (Bacela-Spychalska *et al.* 2012; Hatcher *et al.* 2012; MacNeil *et al.* 2003b; Terry *et al.* 2004; Tofts *et al.* 1995).

GAMMARIDS

Distribution and Ecological role

Within Crustacean Amphipoda, the sub-order Gammaridea is the most diverse with at least 19 families (*e.g.* Gammaridae, Niphargidae, etc.) (Vainola *et al.* 2008, Hou *et al.* 2014, 2016). The family Gammaridae itself is diverse with at least 15 genera described (*e.g. Gammarus*,

Pontogammarus, *Echinogammarus*, *Dikerogammarus*) which phylogenetic affinities have been tentatively reassesed recently by Hou *et al.* (2014, 2016) (Fig. 8).

The genus *Gammarus* Fabricius, 1775 (Amphipoda, Crustacea, Fig. 7) is distributed all across the Northern hemisphere in marine, brackish, freshwater and subterranean aquatic habitats up to 4500 m in altitude (Hou & Li 2004). Based solely on morphological definition of species, in North America *Gammarus* species are estimated *c.* 50 species according to the Smithsonian collection; Europe has approximately *c.* 100 species (Karaman, Gordan S; Pinkster 1987; Karaman & Pinkster 1977b; Väinölä *et al.* 2008) and Asia has *c.* 54 species (Hou & Li 2002, 2004).

Gammarids are recognised to play key roles in the structure and function of worldwide aquatic ecosystems. *Gammarus* mostly occur in large numbers and have an essential impact on the carbon transfer in the food chain as detritivores, shredders, grazers or predators of smaller animals (eggs and larvae) (Dick 1995; Graça *et al.* 2001; Kelly *et al.* 2002; Macneil *et al.* 1997). They also constitute the primary food source of many animals, such as fish (Macneil *et al.* 1997). This key role is very fragile, and a rising number of studies are describing invasive gammarid species and their impact on the aquatic communities (Diffie & Landau 2008; Grabowski *et al.* 2007; MacNeil *et al.* 2003a; Piscart *et al.* 2007; Rewicz *et al.* 2014, 2017). *Gammarus* are also used in ecotoxicological research as sentinels sensitive to environmental pollution (Clason & Zauke 2000; Fialkowski & Rainbow 2006; Larsson 2007; Neuparth *et al.* 2005). These organisms are also used to study host-parasite interactions such as *Gammarus*-microsporidia associations (*e.g.* Grabner *et al.* 2015; Krebes *et al.* 2010; Weigand *et al.* 2016) or *Gammarus*-Acanthocephala associations (*e.g.* Dianne *et al.* 2011; Franceschi *et al.* 2008).

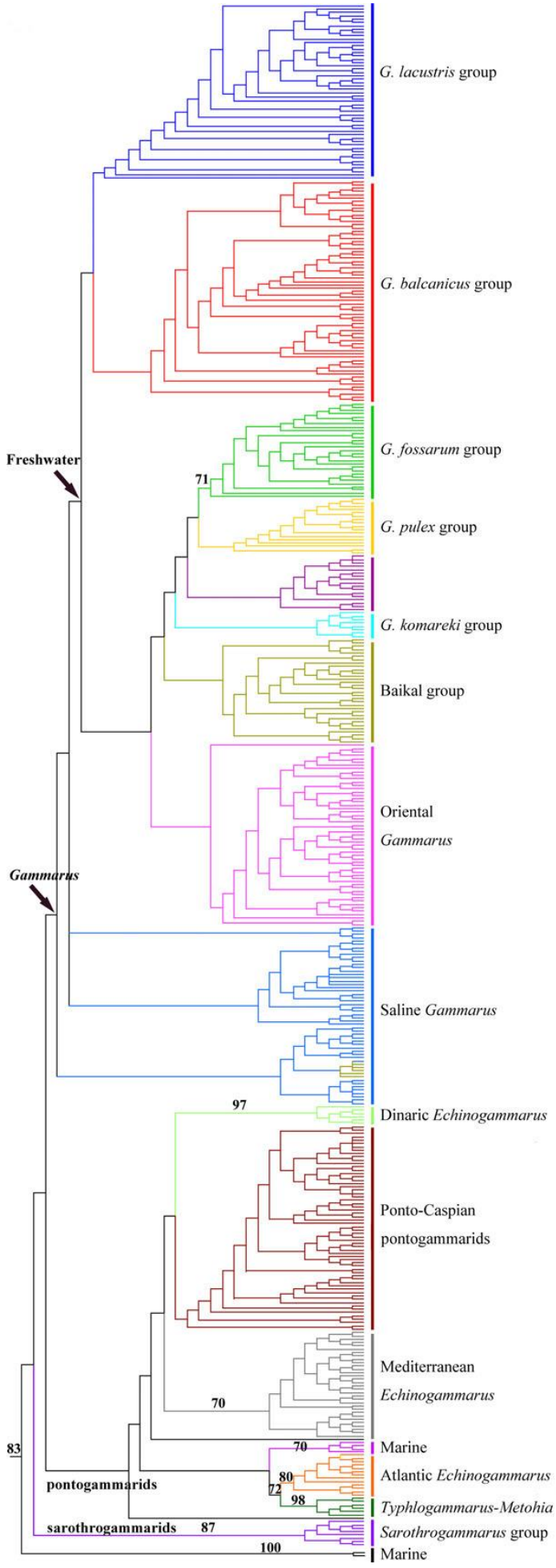


Fig. 8. Phylogenetic trees of Gammaridae derived from maximum parsimony analysis on a combined dataset with 28S, COI, 18S, and EF1a genes for 483 taxa. Support values > 70 are shown above branches for major lineage separations. Reprinted from Hou *et al.* 2014

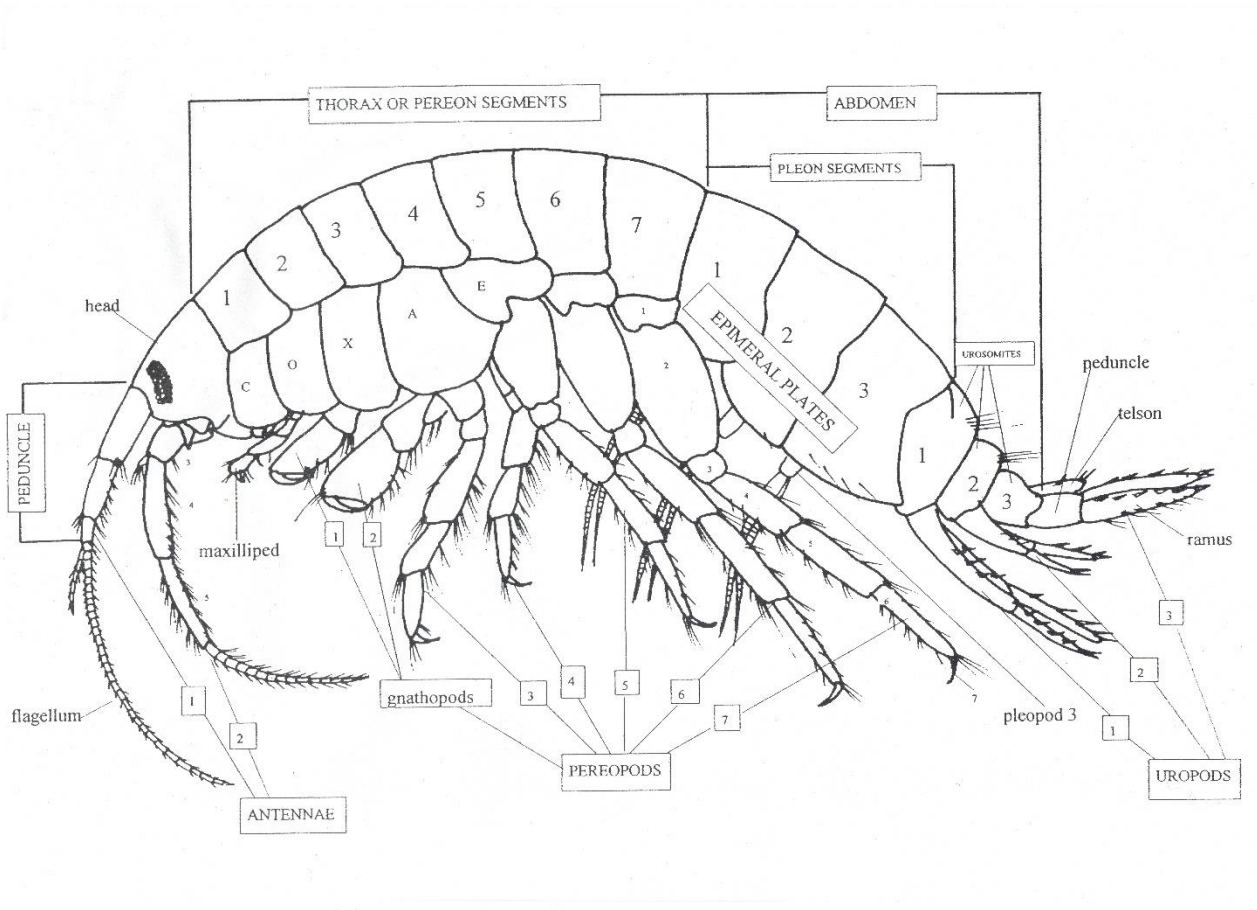


Fig. 9. Lateral view of a gammarid. This displays the morphological features which are used to narrow it down to a species level by the use of dichotomous keys. Antennas, Flagellum, Coxal plate, Maxilliped, 1-7: pereiopods (or thoracic legs) 1 to 7 (the first and second pereiopods are respectively the first and second gnathopods), pleopods1-3 : uropods, telson.

Phylogeny and cryptic diversity

Based on morphological characters 117 species were distinguished within *Gammarus* genus include (Barnard, J.L. & Barnard 1983; Hou & Li 2002, 2004; Karaman, Gordan S; Pinkster 1987; Karaman & Pinkster 1977b). Stock (1967) first erected *Gammarus locusta*-group, which primarily occurs in brackish water. Karaman and Pinkster (1977a, b, 1987) used morphological characters to classify the freshwater *Gammarus* species into three groups: the *Gammarus pulex* group (without dorsal carina and with dense setation on the third and fourth pereopods and the third uropod, Fig. 7), the *Gammarus balcanicus* group (without dorsal carina and poorly setose on the third and fourth pereopods and the third uropod), and the *Gammarus roeselii* group (with dorsal carina, Fig. 7). Nevertheless, this classification remains controversial because of the large number of species in the genus, the relatively few characters defining the groups, and the amount of morphological and ecological diversity present among species (Pinkster, 1983).

New phylogeny based on molecular data such as the study made by Hou *et al.* (2011) is giving a more definitive evolution history of the genus. This analysis showed that *Gammarus* genus early diverged in Paleocene from saline ancestry in the Tethyan region, and then later colonized freshwater habitat in the Middle Eocene (Fig. 10). *Gammarus* genus is composed of five groups: One group still saline lineage and dispersed to both sides of the Atlantic due to few barriers between the Tethys and the Atlantic (*c.* 55 Ma), and diversified throughout its evolutionary history. Three other groups are now freshwater *Gammarus*, underwent a rapid diversification phase until the Middle Miocene, and lineages successively diversified from west to east across Southern Europe and Central Asia (Fig. 10). The Oriental *Gammarus* group diverged from European freshwater *Gammarus* (37 Ma), due to the separation of the Eurasian lineages. Groups *G. lacustris* and *G. balcanicus* clade diverged at 36 Ma (Fig. 10). finally, the European freshwater *Gammarus* groups are composed of many endemic species complexes *G. fossarum*, *G. pulex*, *G. roeselii*, and *G. komareki* which diversified later at 33 Ma (Fig. 10) (Hou *et al.* 2011). The subsequent inclusion of gammarids from the Lake Baikal into this phylogeny did not change this topology, but showed that this group belong to the Oriental/European large clade (Fig. 8 above).

Studies made in the recent years on the within-species genetic diversity of gammarids all evidenced a large amount of cryptic species diversity. Cryptic diversity is defined as the occurrence of distinct evolutionary lineages that are otherwise morphologically indistinguishable within a nominal species (Bickford *et al.* 2007; Struck *et al.* 2018). Despite the morphological similarities, cryptic lineages may carry not only unique evolutionary trajectories but also the potential of differing responses to ongoing and future global change (Feckler *et al.* 2014; Paaby & Rockman 2014). Therefore, biodiversity assessments that ignore cryptic lineages, including them in a single species or a single conservation unit, may severely undervalue biodiversity patterns (Bálint *et al.* 2011; Fišer *et al.* 2018; Riddle *et al.* 2011). Since the pionnier studies of Müller (1998) and Hogg *et al.* (1998), the amount of cryptic diversity appears to be very high in amphipods and particularly in gammarids (Grabowski *et al.* 2017; Hupało *et al.* 2019; Katouzian *et al.* 2016; Lagrue *et al.* 2014; Lefébure *et al.* 2006; Mamos *et al.* 2016; Müller 2000; Sutherland *et al.* 2010; Weiss *et al.* 2014; Westram *et al.* 2011; Witt & Hebert 2011; Witt *et al.* 2006).

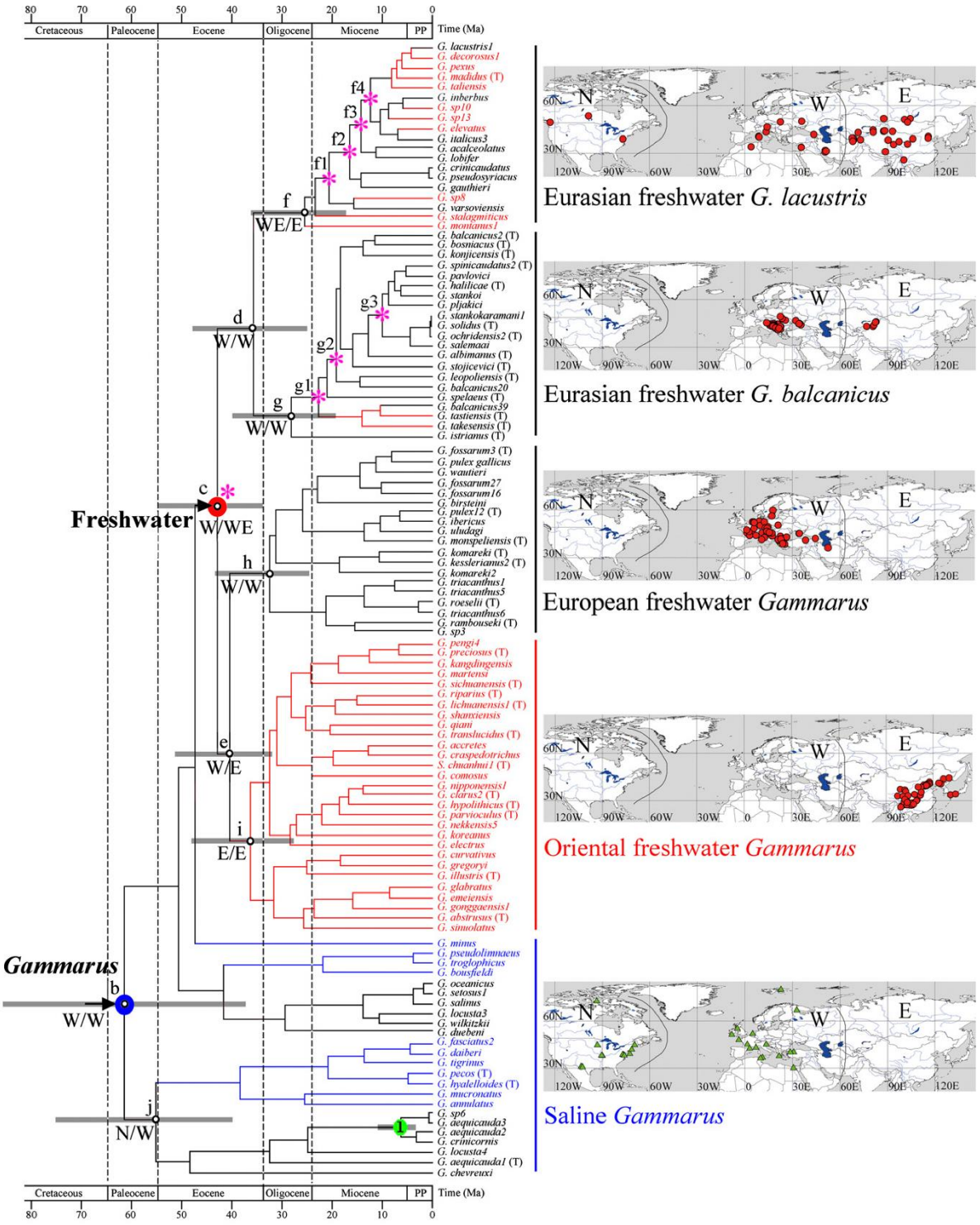


Fig. 10. Maximum clade credibility chronogram inferred from a BEAST dating analysis. Horizontal bars indicating the 95% highest posterior density interval. Major ancestral range reconstructions are shown below the branches (N, Nearctic; W, West Palearctic; E, East Palearctic), and arrows show main dispersal events. The blue circle (node b) indicates the saline origin of the genus *Gammarus* and the red circle (node c) indicates the origin of freshwater species. Outgroups and Baikal amphipods are not shown. (Right) Maps correspond to the sampling sites for the adjacent clade. Reprinted from (Hou *et al.* 2011).

In particular, the *Gammarus* species that we are going to focus on in this thesis show high cryptic diversity with numerous divergent molecular taxonomic units representing probable cryptic and many morphologically undescribed species (Grabowski *et al.* 2017; Mamos *et al.* 2016). *Gammarus balcanicus* is a morphospecies widely distributed in the Balkanids *sensus lato*, from the Eastern Carpathians to the Balkan Peninsula and to the eastern Alps. The biogeographic history of this species has been resolved recently (Copilaş-Ciocianu & Petrusek 2017; Mamos *et al.* 2016). *Gammarus roeselii* is presenting the same pattern with a high cryptic diversity found primarily in the early diversification region of the host in the Balkans, and his evolutionary history has also been elucidated recently (Grabowski *et al.* 2017). The precise evolution history of these two morphospecies will be addressed in more details in the following chapter.

MICROSPORIDIA – AMPHIPODA ASSOCIATIONS

Among the aquatic arthropods, freshwater amphipod crustaceans, and especially Gammaridae, are common hosts for microsporidia (for overviews see Grabner 2017; Stentiford *et al.* 2013; Weiss & Becnel 2015).

All microsporidia found in amphipods, excluding those evidenced during this thesis, are summarised in Table 1. As far as we know, 21, 8 and 17 species of amphipods have been found infected by microsporidia in Europe, America, and in Lake Baikal, respectively (Table 1). We limited the list provided in Table 1 to microsporidia for which at least a SSU rDNA sequence was available. Indeed, descriptions only based on morphological evidences, especially light microscopy of spores, often proved to be misleading (*e.g.* Ovcharenko *et al.* 2010). Since none of the investigated host species has been found free of microsporidia, it is probable that this list is only limited by the lack of prospect in other host species. It is to be noted that many taxa are only described using Microsporidium sp X, and not a full binomial descriptor including genus and species name. How many of these taxa are in fact synonymous is a good question. Although a consensus is emerging within the community working on Microsporidia that a binomial name should be given to a ‘’fully’’ described taxa (using both molecular and ultrastructural traits; *e.g.* (Bacela-Spychalska *et al.* 2018), it is also to be noted that many record in Table 1 with a full name are only based on SSU rDNA data. In other words, it is worth noting that some microsporidia identification can be misleading or imprecise, depending on the year of description.

Table 1. A review of microsporidia infected amphipods, either based on literature or Genbank record (when unpublished). European Amphipods In white: American amphipods in grey: Amphipods from the Lake Baikal in blue:, in bold: Microsporidia reported after the beginning of this thesis. [a]Species or not fully named species (*e.g.* *Microsporidia* spp.) for which an SSU rDNA sequence is available. Clades labels correspond to those proposed by the authors. Different haplotypes/strains and different locations may be present under a single name provided here. See references for details. [b]Microsporidian clades refer to those of Vossbrinck and Debrunner-Vossbrinck (2005).

Amphipod host species	Microsporidia species[a]	Closest identified microsporidia species (and/or microsporidian clade)[b]	Reference
Chaetogammarus ischnus	***Dictyocoela berillonum***	Clade III	Bacela-Spychalska *et al.* 2018
	***Dictyocoela sp* N1-N3**	*Dictyocoela muelleri* (Clade III)	
	Dictyocoela muelleri	Clade III	
Dikerogammarus haemobaphes	*Dictyocoela berillonum*	Clade III	Wilkinson *et al.* 2011
	Dictyocoela muelleri	Clade III	Green Etxabe *et al.* 2015
	Cucumispora ornata	Clade III	Bojko *et al.* 2015
	Microsporidium RW-2009a	*Cucumispora dikerogammari* (Clade III)	Grabner *et al.* 2015
	Microsporidium sp G	*Cucumispora ornata* (Clade III)	Bacela-Spychalska *et al.* 2018
	Dictyocoela sp 30-33	*Dictyocoela duebenum* (Clade III)	Wroblewski & Ovcharenko., unpublished
Dikerogammarus villosus	*Dictyocoela berillonum*	Clade III	Wattier *et al.* 2007
	Dictyocoela muelleri	Clade III	Ovcharenko *et al.* 2010
	Cucumispora dikerogammari	Clade III	Bacela-Spychalska *et al.* 2018
	Dictyocoela berillonum	Clade III	Arundell *et al.* 2015
	Dictyocoela roeselum	Clade III	
Echinogammarus berilloni	*Dictyocoela duebenum*	Clade III	Terry *et al.* 2004
	Dictyocoela duebenum/ muelleri	Clade III	Wilkinson *et al.* 2011
	Dictyocoela berillonum	Clade III	Grabner *et al.* 2015
	Dictyocoela sp 42 JEI-2011	*Dictyocoela berillonum* (Clade III)	Bacela-Spychalska *et al.* 2018
Echinogammarus marinus	*Dictyocoela berillonum*	Clade III	Yang *et al.* 2011
	Dictyocoela duebenum	Clade III	
Echinogammarus trichiatus	*Microsporidium* sp H7	*Trichonosema pectinatellae* (Clade V)	Arundell *et al.* 2015
Gammarus aequicauda	***Dictyocoela muelleri***	Clade III	Bacela-Spychalska *et al.* 2018
Gammarus balcanicus	***Dictyocoela roeselum***	Clade III	Bacela-Spychalska *et al.* 2018
Gammarus duebeni	*Dictyocoela duebenum*	Clade III	Terry *et al.* 1999
(G.d. duebeni or *G.d. celticus)*	*Dictyocoela muelleri*	Clade III	Hogg *et al.* 2002
	Dictyocoela berillonum	Clade III	MacNeil *et al.* 2003
	Pleistophora mulleri	Clade III	Terry *et al.* 2004
	Microsporidium sp.1049	*Cucumispora roeselii* (Clade III)	Krebes *et al.* 2010
	Microsporidium sp. 505	*Helmichia lacustris* (Clade IV)	Wilkinson *et al.* 2011
	Microsporidium sp. 515	*Helmichia lacustris* (Clade IV)	Ironside 2013
	Microsporidium sp. 1199	*Helmichia lacustris* (Clade IV)	
	Microsporidium sp. 1154	*Helmichia lacustris* (Clade IV)	
	Microsporidium sp. 711	*Cystoporogenes operophterae* (Clade IV)	
	Nosema granulosis	Clade IV	
	Microsporidium sp. 616	*Nosema furnacalis* (Clade IV)	
Gammarus fossarum	*Dictyocoela roeselum*	Clade III	Grabner *et al.* 2015

	Dictyocoela duebenum	Clade III	Weigand *et al.* 2016
	Dictyocoela duebenum/ muelleri	Clade III	Grabner 2017
	Microsporidium RR2	*Orthosomella operophterae* (Clade IV)	Bacela-Spychalska *et al.* 2018
	Microsporidium 505	*Helmichia lacustris* (Clade IV)	
	Microsporidium 515	*Helmichia lacustris* (Clade IV)	
	Nosema granulosis	Clade IV	
	Orthosomella DG-2015	*Orthosomella operophterae* (Clade IV)	
	Dictyocoela roeselum	Clade III	Bacela-Spychalska *et al.* 2018
Gammarus lacustris	*Dictyocoela* sp.	*Dictyocoela duebenum* (Clade III)	Wilkinson *et al.* 2011
	Dictyocoela GL	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
Gammarus pulex	*Dictyocoela duebenum*	Clade III	Terry *et al.* 2004
	Dictyocoela duebenum/ muelleri	Clade III	Wilkinson *et al.* 2011
	Dictyocoela sp.	*Dictyocoela duebenum* (Clade III)	Grabner *et al.* 2015
	Dictyocoela 25	*Dictyocoela duebenum* (Clade III)	Ironside & Alexander 2015
	***Microsporidium* sp G**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
	Microsporidium 505	*Helmichia lacustris* (Clade IV)	
	Microsporidium 515	*Helmichia lacustris* (Clade IV)	
	Microsporidium RR2	*Orthosomella operophterae* (Clade IV)	
	Microsporidium sp. I	*Glugoides intestinalis* (Clade IV)	
	Nosema granulosis	Clade IV	
	Microsporidium RR1	*Paranosema locustae* (Clade V)	
Gammarus roeselii	*Dictyocoela muelleri*	Clade III	Haine *et al.* 2004
	Dictyocoela roeselum	Clade III	Terry *et al.* 2004
	Dictyocoela duebenum	Clade III	Grabner *et al.* 2015
	Dictyocoela duebenum/ muelleri	Clade III	Grabner 2017
	***Microsporidium* sp G**	*Cucumispora ornata* (Clade III)	Bacela-Spychalska *et al.* 2018
	Microsporidium sp. E	*Cucumispora ornata* (Clade III)	
	Microsporidium 505	*Helmichia lacustris* (Clade IV)	
	Microsporidium sp I	*Glugoides intestinalis* (Clade IV)	
	Microsporidium 515	*Helmichia lacustris* (Clade IV)	
	Nosema granulosis	Clade IV	
	Microsporidium RR1	*Paranosema locustae* (Clade V)	
Gammarus varsoviensis	***Dictyocoela muelleri***	Clade III	Bacela-Spychalska *et al.* 2018
	Dictyocoela roeselum	Clade III	Bacela-Spychalska, unpublished
	Cucumispora ornata	Clade III	
Niphargogammarus intermedius	***Anncaliia azovica***	Clade III	Tokarev *et al.* 2018
Niphargus schellenbergi	*Microsporidium* sp I	*Glugoides intestinalis* (Clade IV)	Weigand *et al.* 2016
	Nosema granulosis	Clade IV	
Orchestia cavimana	*Dictyocoela cavimanum*	Clade III	Terry *et al.* 2004

Pontogammarus robustoides	***Dictyocoela berillonum***	**Clade III**	Wilkinson *et al.* 2011
	Dictyocoela muelleri	**Clade III**	Bacela-Spychalska *et al.* 2018
	Dictyocoela sp 40-41 JEI-2011	*Dictyocoela berillonum* (Clade III)	
Talorchestia deshayesei	*Microsporidium* JES2002H	*Thelohania contejeani* (Clade uncertain)	Terry *et al.* 2004
Corophium volutator	*Microsporidium* sp. C81	*Anncaliia algerae* (Clade V)	Mautner *et al.* 2007
Crangonyx floridanus	*Microsporidium* sp. CRANFA	*Cystosporogenes operoptherae* (Clade IV)	Slothouber Galbreath *et al.* 2009
	Microsporidium sp. CRANFB	*Spraguea lophii* (Clade III)	
Crangonyx pseudogracilis	*Microsporidium* sp. CRANPA	*Spraguea lophii* (Clade III)	Slothouber Galbreath *et al.* 2009
	Fibrillanosema crangonycis	Clade IV	
	Microsporidium sp. CRANPB	*Orthosomella operoptherae* (Clade IV)	
	Microsporidium sp. CRANPC	*Cystosporogenes operoptherae* (Clade IV)	
Diporeia sp.	*Dictyocoela diporeiae*	Clade III	Winters & Faisal 2014
Gammarus pseudolimnaeus	*Dictyocoela duebenum*	Clade III	Ryan & Kohler 2010
	Microsporidium sp.	*Orthosomella operoptherae* (Clade IV)	
Gammarus tigrinus	*Microsporidium* BPAR3	Clade V	Grabner *et al.* 2015
Hyalella sp.	*Dictyocoela* sp. HYAL	*Dictyocoela cavimanum* (Clade III)	Slothouber Galbreath *et al.* 2009
Synurella sp.	*Microsporidium* sp. SYN	*Vavraia culicis* (Clade III)	Slothouber Galbreath *et al.* 2009
Acanthogammarus lapaceus	***Microsporidia sp.***	***Hazardia milleri (Clade I)***	***Ironside & Wilkinson 2018***
	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	
	***Helmichia* sp. Clade 2**	*Helmichia lacustris* (Clade IV)	
Acanthogammarus victorii	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
Brachyuropus grewingkii	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
Brandtia latissima	***Microsporidia* sp.**	*Hazardia milleri* (Clade I)	Wilkinson *et al.* 2011
	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
Dorogostaiskia parasitica	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
	***Helmichia* sp. Clade 1**	*Helmichia lacustris* (Clade IV)	
	***Helmichia* sp. Clade 2**	*Helmichia lacustris* (Clade IV)	
	***Microsporidium* sp.**	*Anncaliia algerae* (Clade V)	
Eulimnogammarus cyaneus	***Dictyocoela* sp. Clade 4**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
Eulimnogammarus marituji	***Dictyocoela* sp. Clade D**	*Dictyocoela duebenum* (Clade III)	Dimova *et al.* 2018
Eulimnogammarus verrucosus	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Dictyocoela* sp. Clade 4**	*Dictyocoela duebenum* (Clade III)	Dimova *et al.* 2018
	***Helmichia* sp. Clade 1**	*Helmichia lacustris* (Clade IV)	
	***Microsporidium* sp. Clade E**	*Enterocytospora artemiae* (Clade IV)	

	***Microsporidium* sp. Clade NV**	*Nosema bombycis* (Clade IV)	
Garjajewia cabanisii	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
Gmelinoides fasciatus	***Dictyocoela* sp. Clade 2**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
	***Helmichia* sp. Clade 1**	*Helmichia lacustris* (Clade IV)	
Linevichella vortex	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
	***Helmichia* sp. Clade 1**	*Helmichia lacustris* (Clade IV)	
Micruropus platycercus	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
Micruropus wahlii	***Microsporidium* sp.**	*Hazardia milleri* (Clade I)	Ironside & Wilkinson 2018
	***Helmichia* sp. Clade 1**	*Helmichia lacustris* (Clade IV)	
	***Helmichia* sp. Clade 3**	*Helmichia lacustris* (Clade IV)	
Ommatogammarus albinus	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	Ironside & Wilkinson 2018
Pallasea cancellus	***Dictyocoela* sp. Clade 1**	*Dictyocoela duebenum* (Clade III)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	
	***Dictyocoela* sp. Clade D**	*Dictyocoela berillonum* (Clade III)	Dimova *et al.* 2018
Pallaseopsis kessleri	***Microsporidia* sp.**	*Hazardia milleri* (Clade I)	Ironside & Wilkinson 2018
	***Cucumispora* sp.**	*Cucumispora ornata* (Clade III)	

The three main microsporidian genera found commonly infecting this freshwater or brackish water amphipods were: *Nosema* (Naegeli 1857), *Cucumispora* (Ovcharenko *et al.* 2010) and *Dictyocoela* (Terry *et al*. 2004).

The most recurrent microsporidian genus found infecting amphipods is *Dictyocoela*. Discovered by Terry *et al.* (2004), *Dictyocoela* spp. were not found, at our very best knowledge, in hosts other than amphipods and could therefore be specific to this host clade. They infect numerous taxa of European and Lake Baikal gammarids, and are also present in American hosts (albeit apparently rarer there) (Table 1). In total, eight species of *Dictycolea* with Linnean binomial names have been proposed: *Dictyocoela berillonum*, *D. cavimanum*, *D. deshayesum*, *D. duebenum*, *D. gammarellum*, *D. muelleri*, *D. roeselum* (Terry *et al.* 2004; Haine *et al.* 2004; Ironside & Alexander 2015; Krebes *et al.* 2010) and *D. diporeiae* (Winters *et al.* 2014). Most of these denominated species were based on SSU rDNA sequences only. A recent study nevertheless re-defined the *Dictyocoela* genus infecting amphipods in Europe, and confirmed most of these species using both molecular and ultrastructural traits (Bacela-Spychalska *et al.* 2018). Before this study, many imprecisions were observed for denominating *Dictyocoela* spp. Some authors did not took the risk to provide any name using "*Dictyocoela* sp." for describing the variation they found within the genus (*e.g.* Ironside & Alexander 2015), and others referred to the "*D. duebenum/muelleri* species complex" for their newly obtain sequences (Grabner *et al.* 2015).

The *Cucumispora* genus, which is also very abundant in European and Baikal amphipods, is an example of microsporidian poor identification in databases. Species of this genus have been fully described (using both molecular and ultrastructural traits) very recently: *Cucumispora dikerogammari* infecting *Dikerogammarus villosus* was described less than 10 years ago (Ovcharenko *et al.* 2010). Even more recently, *Cucumispora ornata* was described in *Dikerogammarus haemobaphes* (Bojko *et al.* 2015) and *Cucumispora roeselii* found in *Gammarus roeselii* (Bojko *et al.* 2017a). Many sequences found prior these formal descriptions were attributed to *Microsporidium* sp. with various acronyms, and their status have not been actualized in the databases, generating a lot of confusion (see Table 1 and Bojko *et al.* 2017).

The *Nosema* genus is infecting around half of European amphipods studies, but only a few American or Baikal amphipods (Table 1). *Nosema granulosis* infecting *Gammarus duebeni* were described only twenty years ago (Terry *et al.* 1999). It has been well identify using SSU rDNA and morphology.

Numerous studies showed that *Nosema granulosis*, but also *Dictyocoela roeselum*, *D. duebenum* and *D. muelleri* are vertically transmitted parasites (Dubuffet *et al.* 2013; Haine *et al.*

2004; Terry *et al.* 1999). They induce very low virulence and, above all, they are remarkable because they induce sex-ratio distortion in their host populations, by reversing male host into functional females (Dunn & Smith 2001; Ironside & Alexander 2015; Kelly *et al.* 2002). This trait is well known in some vertically-transmitted bacteria, such as the typical example of *Wolbachia* infecting woodlice (Bouchon *et al.* 1998) or lepidoptera (Kageyama *et al.* 2017), and the microsporidia-gammarid association was the first to evidence such a phenomenon in eukaryotic symbionts. Feminization allow the vertically transmitted symbionts to maintain and spread in host populations, even in the absence of a beneficial effect on their host fitness (Hurst & Majerus 1993). This characteristic also induces a trait very convenient to detect a putative feminizing symbiots: a bias in prevalence according to gender. Indeed, high prevalences were found in females, but none or very rare infection are found in males (Haine *et al.* 2004; Terry *et al.* 1999, 2004). However, it is to be noticed that some strains of *Dictyocoela duebenum* are not feminizing their host (Ironside & Alexander 2015), indicating possible variation on this trait, as well on the vertical transmission trait within a single microsporidia species. The maintenance of horizontal transmission in vertically-transmitted sex ratio distorters has been shown to provide an advantage (Ironside *et al.* 2011). Nevertheless, for *Nosema granulosis* infecting *G. duebeni*, vertical transmission appears to be strict (Ironside *et al.* 2003a).

Cucumispora genus (as well as *Dictyocoela diporae*, Winters *et al.* 2014) is described as horizontally-transmitted microsporidia. The infection occur after ingestion of infected tissues or environmental spores (Bacela-Spychalska *et al.* 2012; Ovcharenko *et al.* 2010). These horizontally transmitted microsporidia induce severe pathology, by filling and replacing muscle tissue of host (Bojko *et al.* 2015, 2017a; Ovcharenko *et al.* 2010). These species are able to switch hosts occasionally, from one species of gammarid to another (Bacela-Spychalska *et al.* 2012; Bojko *et al.* 2017c).

In addition to these three main genera infecting amphipods, a quantity of additional microsporidian lineages has been found based on molecular surveys. Most of them were observed occasionally, thanks to various host population screenings. Those rare microsporidia are not yet fully described since no anatomical and ultrastructural descriptions are available (Grabner 2017; Grabner *et al.* 2015; Ironside *et al.* 2008; Krebes *et al.* 2010; Quiles *et al.* 2019; Terry *et al.* 2004; Wilkinson *et al.* 2011). Some of these undescribed parasites may nevertheless be found in relatively high prevalence in some host species. For example, three parasite 'species' are found in prevalence as high as *Dictyocoela* and *Nosema* in *Gammarus duebeni* : *Pleistophora mulleri, Microsporidium* sp. 505 and *Microsporidium* sp. 515 (Krebes *et al.* 2010). The two later microsporidia were then found, but at lower rates, in *Gammarus pulex* (Grabner *et al.* 2015).

Even if Table 1 provides a broad picture of the pattern of microsporidian infections in amphipod, it is worth noting that it is a coarse picture and that the global evolutionary history of these infections remain confusing. It is mainly due to the very diverse types of studies that have been conducted in investigating these host-parasite relationships. Some of these studies explored in depth the relationships between the host *Gammarus duebeni* and the two vertically-transmitted parasites *Nosema granulosis* and *Dictyocoela duebenum*. While they were often conducted in a few populations (Hatcher *et al.* 2005), they allowed to understand the details of the evolutionary impact and mechanisms of the peculiar trait of feminization (*e.g.* (Dunn *et al.* 2006a, 2006b; Ironside *et al.* 2003b; Rodgers-Gray *et al.* 2004; Terry *et al.* 1999). In the same host, a wide biogeographic study allowed to show that these two parasite species were not the only microsporidia infecting *G. duebeni* (Krebes *et al.* 2010), making this system the most studied and the most understood among microsporidia-amphipod relationships. However, a survey conducted on several other amphipod species through Europe indicated that *G. duebeni* was far to be the only gammarid species infected by microsporidia (Terry *et al.* 2004). From this study, it became obvious that *G. duebeni* was only the tip of the iceberg for understanding the whole story. Other studies conducted on single host species (in a restricted number of populations) began to deepens our understanding by describing new species of parasites and/or expanding the host range of some already-detected parasites (*e.g.* Bojko *et al.* 2015, 2017b; Haine *et al.* 2004, 2007; Ironside *et al.* 2008; Ovcharenko *et al.* 2010; Winters & Faisal 2014). In addition, ecological studies, either focusing on single invasive species (Wattier *et al.* 2007) or gammarid assemblages on restricted areas (Grabner 2017; Grabner *et al.* 2015), began to show that several species of gammarids may share the same microsporidian groups, but always with some doubts on the precise level of parasite identification. It was therefore impossible to affirm that a single parasite strain may (or not) share different host species. At the opposite, a parasite-centered survey showed that *Dictyocoela* parasites may infect a wide host range (Ironside & Alexander 2015). From this study, it appears that each host species seems to host a peculiar parasite strain; however, the limited number of host populations tested here has precluded extending this assumption to the whole species.

From the accumulation of all these studies, it also appeared that the SSU rDNA could be an efficient genetic marker, both to assign an individual to a parasite species level taxa but also to reveal variation within this taxonomic level. Within single parasite species, several SSU rDNA variants were most of the time detected, some of them appearing only once in a single host species, some other shared by different host species (*e.g.* see the variation detected within each *Dictyocoela* species by Bacela-Spychalska *et al.* 2018).

Therefore, in absence of a synthesis of all these studies, it remains unclear if a single microsporidian species is specific to one host species, or if different variants of a single parasite infect different host species. On the opposite, since some microsporidia species can be acquired from potential host communities through horizontal transfers (Bacela-Spychalska *et al.* 2012; Bojko *et al.* 2017b), there may be no specificity at all, and our inability to evidence such generalist parasites would simply be limited by our limited sample sizes. On the opposite, it is to be pointed out that Table 1 is purely qualitive in essence and hides extremely heretogenous sampling efforts. If a given parasite appears to be associated with two hosts sepecies, one being represented by only one tested individial and the other by hundreds of individuals from tens of populations over large geographic range, one could conclude that the single record is possibly sporrious, representing a best a recent horizontal inter-specific transfert, or even simply a transient association associated with fortious foraging or scavenging.

In addition to these uncertainties, the recent discovering that most gammarid morpho-species were consisting in fact of several cryptic species (see above) further obscured our understanding of gammarid-microsporidia evolutionary history. Indeed, the cryptic diversity of the host seems not cryptic at all for some other parasites (*e.g.* acanthocephalan parasites: *e.g.* Galipaud *et al.* 2017; Westram *et al.* 2011). This lack of information comes mostly from the profound ignorance of the evolutionary and biogeographic history of most of the gammarid host studied so far. In addition, some species show recent geographical expansions of both parasite and associated parasite in their native range (*e.g.* Wattier *et al.* 2007), favouring a mixture of fauna and thus potentially parasitic exchanges.

Ph.D AIMS

The main aim of this Ph.D was to characterise extensively the diversity of microsporidia infecting two gammarid morpho-species: *Gammarus roeselii* and *Gammarus balcanicus* over their entire geographical range, and across the highly divergent host MOTUs recently evidenced (Grabowski *et al.* 2017; Mamos *et al.* 2016). We took advantage of several sampling campaigns made for these two species between 2004 and 2016 by members of the Department of Invertebrate Zoology and Hydrobiology, University of Lodz, Poland and the laboratory Biogeosciences, University of Burgundy, France, initially dedicated to the study of gammarid evolutionary histories. The samples, either under the form of extracted DNA kept from these campaigns, or as individuals kept in ethanol for which the DNA has been extracted (see Chapter I and chapter II for methods), allowed to study the very same populations (sometimes individuals) than those studies by Grabowski *et al.* (2017) and Mamos *et al.* (2016).

The central question will be to test if the biogeographic history of the hosts had affected the diversity of their microsporidia. Using parasite genetic markers, mostly SSU rDNA sequences, several analyses were made to address this general question.

First, we try to clarify the infection pattern and specificity, by tentatively exploring the phylogenetic relationships of the detected microsporidia within the different cryptic diversity of *Gammarus roeselii* and *G. balcanicus*, and comparing these parasites with those of other Gammaridae. Special attention was given to compare parasites of *G. balcanicus* with those of *G. roeselii*, as the two species present range overlap in the Balkans.

Second, we hypothesis ecological or evolutionary scenarios which can lead to the observed infection patterns. The following questions will be explored. (i) Can parasite clades be explained by the diversification pattern of the hosts (suggesting possible host-parasite co-diversifications)? We predict that the vertically-transmitted *Nosema* spp. and *Dictyocoela* spp., tightly linked to their hosts life-cycle, would more correspond to such a pattern than the horizontally-transmitted parasites (*e.g. Cucumispora*), (ii) For *G. roeselii*, which is a recent colonizer of the western part of Europe, are some parasites restricted to the recently-colonised zone, suggesting host shifts form local indigenous fauna?

Third, when possible, we explored the hypothetical bias of microsporidian prevalence between male and female hosts. Following Terry *et al.* (2004), finding a high prevalence in females and almost no infection in males could be a first step toward the evidence of vertical transmission associated with a feminizing effect. Unfortunately, this was only possible for *G. balcanicus* hosts since *G. roeselii* individuals hasn't been identified for gender prior DNA extractions. We

nevertheless have the information that, in some *G. roeselii* French populations, *Nosema granulosis* and two *Dictyocoela* parasites are vertically-transmitted, and have high probability to induce sex ratio distortion (Haine *et al.* 2004, 2007)

Finally, because the SSU rDNA marker appeared to be limited to understand phylogenetic relationships for *Nosema* parasites, we used the RPB1 gene to (i) try to detect more genetic variation within parasite variants detected with the SSU rDNA, (ii) determine with more precisions the phylogenetic relationships between the different parasites variants and (iii) propose host-parasite evolutionary history scenarios to explain the diversification of *N. granulosis* among host species.

REFERENCES

Althoff, D. M., Segraves, K. A., & Johnson, M. T. J. (2014). Testing for coevolutionary diversification: Linking pattern with process. *Trends in Ecology and Evolution*, **29**(2), 82–89.

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. *Journal of Invertebrate Pathology*, **109**(1), 59–75.

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. *Biological Invasions*, **17**(5), 1439–1451.

Bacela-Spychalska, K., Wattier, R. A., Genton, C., & Rigaud, T. (2012). Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. *Biological Invasions*, **14**(9), 1831–1842.

Bacela-Spychalska, K., Wróblewski, P., Mamos, T., … Ovcharenko, M. (2018). Europe-wide reassessment of Dictyocoela (Microsporidia) infecting native and invasive amphipods (Crustacea): Molecular versus ultrastructural traits. *Scientific Reports*, **8**(1), 1–16.

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. *Journal of Invertebrate Pathology*, **71**(3), 199–206.

Bálint, M., Domisch, S., Engelhardt, C. H. M., … Nowak, C. (2011). Cryptic biodiversity loss linked to global climate change. *Nature Climate Change*, **1**(6), 313–318.

Barnard, J.L. & Barnard, C. M. (1983). (1983). *Freshwater Amphipoda of the World*. *Arctic*, Mt. Vernon Va.: Hayfield Associates. Retrieved from https://www.worldcat.org/title/freshwater-amphipoda-of-the-world/oclc/10019274

Becnel, J. J., & Andreadis, T. G. (2014). Microsporidia in Insects. *Microsporidia: Pathogens of Opportunity: First Edition*, 521–570.

Becnel, J. J., & Takvorian, P. M. (2014). Checklist of Available Generic Names for Microsporidia with Type Species and Type Hosts. *Microsporidia: Pathogens of Opportunity: First Edition*, 671–686.

Beneden. (1875). *Les commensaux et les parasites dans le règne animal. Les commensaux et les parasites dans le règne animal.*, 2. éd., Paris,: G. Baillière,. doi:10.5962/bhl.title.1823

Bickford, D., Lohman, D. J., Sodhi, N. S., … Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22**(3), 148–155.

Bohne, W., Ferguson, D. J. P., Kohler, K., & Gross, U. (2000). Developmental expression of a tandemly repeated, glycine- and serine- rich spore wall protein in the microsporidian pathogen Encephalitozoon cuniculi. *Infection and Immunity*, **68**(4), 2268–2275.

Bojko, J., Bącela-Spychalska, K., Stebbing, P. D., … Stentiford, G. D. (2017a). Parasites, pathogens and commensals in the "low-impact" non-native amphipod host *Gammarus roeselii*. *Parasites & Vectors*, **10**(1), 193.

Bojko, J., Clark, F., Bass, D., … Stentiford, G. D. (2017b). *Parahepatospora carcini* n. gen., n. sp., a parasite of invasive *Carcinus maenas* with intermediate features of sporogony between the Enterocytozoon clade and other microsporidia. *Journal of Invertebrate Pathology*, **143**, 124–134.

Bojko, J., Dunn, A. M., Stebbing, P. D., Ross, S. H., Kerr, R. C., & Stentiford, G. D. (2015). *Cucumispora ornata* n. sp. (Fungi: Microsporidia) infecting invasive "demon shrimp" (*Dikerogammarus haemobaphes*) in the United Kingdom. *Journal of Invertebrate Pathology*, **128**, 22–30.

Bojko, J., Stentiford, G., Stebbing, P., … Dunn, A. (2017c). Pathogens of Dikerogammarus haemobaphes regulate host activity and survival, but also threaten native amphipod populations in the UK. *Diseases of Aquatic Organisms*, **View**(January), 1–16.

Boots, M., White, A., Best, A., & Bowers, R. (2014). How specificity and epidemiology drive the coevolution of static trait diversity in hosts and parasites. *Evolution*. doi:10.1111/evo.12393

Bouchon, D., Rigaud, T., & Juchault, P. (1998). Evidence for widespread Wolbachia infection in isopod crustaceans: Molecular identification and host feminization. *Proceedings of the Royal Society B: Biological Sciences*. doi:10.1098/rspb.1998.0402

Bourgeois, Y., Roulin, A. C., Müller, K., & Ebert, D. (2017). Parasitism drives host genome evolution: Insights from the Pasteuria ramosa–Daphnia magna system. *Evolution*. doi:10.1111/evo.13209

Bright, M., & Bulgheresi, S. (2010). A complex journey: Transmission of microbial symbionts. *Nature Reviews Microbiology*, **8**(3), 218–230.

Brosson, D., Kuhn, L., Prensier, G., Vivarès, C. P., & Texier, C. (2005). The putative chitin deacetylase of *Encephalitozoon cuniculi*: A surface protein implicated in microsporidian spore-wall formation. *FEMS Microbiology Letters*, **247**(1), 81–90.

Brown, A. M. V., Kent, M. L., & Adamson, M. L. (2010). Description of five new Loma (microsporidia) species in pacific fishes with redesignation of the type species *Loma morhua* Morrison & Sprague, 1981, based on morphological and molecular species-boundaries tests. *Journal of Eukaryotic Microbiology*, **57**(6), 529–553.

Bulnheim, P. (1975). Microsporidian Infections of Amphipods with Special Reference to Host-Parasite Relationships : A Review. *Marine Fisheries Review*, **37**, 39–45.

Capella-Gutiérrez, S., Marcet-Houben, M., & Gabaldón, T. (2012). Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. *BMC Biology*, **10**(1), 47.

Carius, H. J., Little, T. J., & Ebert, D. (2001). Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. *Evolution*. doi:10.1111/j.0014-3820.2001.tb00633.x

Clason, B., & Zauke, G. P. (2000). Bioaccumulation of trace metals in marine and estuarine amphipods: Evaluation and verification of toxicokinetic models. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**(7), 1410–1422.

Clayton, D. H., Bush, S. E., & Johnson, K. P. (2015). Coevolution of Life on Hosts. *Coevolution of Life on Hosts*. doi:10.7208/chicago/9780226302300.001.0001

Copilaş-Ciocianu, D., & Petrusek, A. (2017). Phylogeography of a freshwater crustacean species complex reflects a long-gone archipelago. *Journal of Biogeography*, **44**(2), 421–432.

Corradi, N., & Selman, M. (2013). Latest progress in microsporidian genome research. *Journal of Eukaryotic Microbiology*, **60**(3), 309–312.

Darwin, C. (1859). On the origin of species, 1859.

De Meeûs, T., & Renaud, F. (2002). Parasites within the new phylogeny of eukaryotes. *Trends in Parasitology*, **18**(6), 247–250.

De Vienne, D. M., Giraud, T., & Shykoff, J. A. (2007). When can host shifts produce congruent host and parasite phylogenies? A simulation approach. *Journal of Evolutionary Biology*, **20**(4), 1428–1438.

Delbac, F., & Polonais, V. (2008). The microsporidian polar tube and its role in invasion. *Sub-Cellular Biochemistry*, **47**, 208–220.

Dianne, L., Perrot-Minnot, M. J., Bauer, A., Gaillard, M., Léger, E., & Rigaud, T. (2011). Protection first then facilitation: A manipulative parasite modulates the vulnerability to predation of its intermediate host according to its own developmental stage. *Evolution*. doi:10.1111/j.1558-5646.2011.01330.x

Dick, J. T. A. (1995). The cannibalistic behaviour of two *Gammarus* species (Crustacea: Amphipoda). *Journal of Zoology*, **236**(4), 697–706.

Diffie, W., & Landau, S. (2008). Enhanced fecundity and parasite release in the first amphipod invader on the Iberian Peninsula. *ACM SIGACT News*, **39**(4), 30.

Dimova, M., Madyarova, E., Gurkov, A., … 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.

Dubuffet, A., Smith, J. E., Solter, L., Alejandra Perotti, M., Braig, H. R., & Dunn, A. M. (2013). Specific detection and localization of microsporidian parasites in invertebrate hosts by using in situ hybridization. *Applied and Environmental Microbiology*, **79**(1), 385–388.

Dunn, A. M., Andrews, T., Ingrey, H., Riley, J., & Wedell, N. (2006a). Strategic sperm allocation under parasitic sex-ratio distortion. *Biol. Lett*, **2**, 78–80.

Dunn, A. M., Hogg, J. C., & Hatcher, M. J. (2006b). Transmission and burden and the impact of temperature on two species of vertically transmitted microsporidia. *International Journal for Parasitology*, **36**(4), 409–414.

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

Duperron, S. (2016). *Microbial Symbioses*. (I. P.- Elsevier, Ed.), 1st edn.

Dybdahl, M. F., & Lively, C. M. (1998). Host-Parasite Coevolution: Evidence for Rare Advantage and Time-Lagged Selection in a Natural Population. *Evolution*. doi:10.2307/2411236

Ebert, D. (2013). The epidemiology and evolution of symbionts with mixed-mode transmission. *Annual Review of Ecology, Evolution, and Systematics*, **44**(1), 623–643.

Feckler, A., Zubrod, J. P., Thielsch, A., Schwenk, K., Schulz, R., & Bundschuh, M. (2014). Cryptic species diversity: An overlooked factor in environmental management? *Journal of Applied Ecology*, **51**(4), 958–967.

Fialkowski, W., & Rainbow, P. S. (2006). The discriminatory power of two biomonitors of trace metal bioavailabilities in freshwater streams. *Water Research*, **40**(9), 1805–1810.

Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, **27**(3), 613–635.

Franceschi, N., Bauer, A., Bollache, L., & Rigaud, T. (2008). The effects of parasite age and intensity on variability in acanthocephalan-induced behavioural manipulation. *International Journal for Parasitology*. doi:10.1016/j.ijpara.2008.01.003

Franzen, C. (2004). Microsporidia: How can they invade other cells? *Trends in Parasitology*, **20**(6), 275–279.

Franzen, C., Hartmann, P., & Salzberger, B. (2005). Cytokine and nitric oxide responses of monocyte-derived human macrophages to microsporidian spores. *Experimental Parasitology*, **109**(1), 1–6.

Galipaud, M., Bollache, L., & Lagrue, C. (2017). Variations in infection levels and parasite-induced mortality among sympatric cryptic lineages of native amphipods and a congeneric invasive species: Are native hosts always losing? *International Journal for Parasitology: Parasites and Wildlife*, **6**(3), 439–447.

Gardiner, C. (1988). An atlas of protozoan parasites in animal tissues.

Genty, L. M., Bouchon, D., Raimond, M., & Bertaux, J. (2014). Wolbachia infect ovaries in the course of their maturation: Last minute passengers and priority travellers? *PLoS ONE*. doi:10.1371/journal.pone.0094577

Grabner, D. S. (2017). Hidden diversity: parasites of stream arthropods. *Freshwater Biology*, **62**(1), 52–64.

Grabner, D. S., Weigand, A. M., Leese, F., … Sures, B. (2015). Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. *Parasites and Vectors*, **8**(1), 1–4.

Grabowski, M., Jazdzewski, K., & Konopacka, A. (2007). Alien crustacea in polish waters - Amphipoda. *Aquatic Invasions*, **2**(1), 25–38.

Grabowski, M., Mamos, T., Bącela-Spychalska, K., Rewicz, T., & Wattier, R. A. (2017). Neogene paleogeography provides context for understanding the origin and spatial distribution of cryptic diversity in a widespread Balkan freshwater amphipod. *PeerJ*, **5**, e3016.

Graça, Cressa, Gessner, Feio, Callies, & Barrios. (2001). Food quality, feeding preferences, survival and growth of shredders from temperate and tropical streams. *Freshwater Biology*, **46**(7), 947–957.

Green Etxabe, A., Short, S., Flood, T., Johns, T., & Ford, A. T. (2015). Pronounced and prevalent intersexuality does not impede the 'Demon Shrimp' invasion. *PeerJ*, **3**, e757.

Hafner, M. S., & Page, R. D. M. (1995). Molecular phylogenies and host-parasite cospeciation: gophers and lice as a model system. *Philosophical Transactions - Royal Society of London, B*, **349**(1327), 77–83.

Haine, E. R., Brondani, E., Hume, K. D., Perrot-Minnot, M. J., Gaillard, M., & Rigaud, T. (2004). Coexistence of three microsporidia parasites in populations of the freshwater amphipod *Gammarus roeseli*: Evidence for vertical transmission and positive effect on reproduction. *International Journal for Parasitology*, **34**(10), 1137–1146.

Haine, E. R., Motreuil, S., & Rigaud, T. (2007). Infection by a vertically-transmitted microsporidian parasite is associated with a female-biased sex ratio and survival advantage in the amphipod *Gammarus roeseli*. *Parasitology*, **134**(Pt 10), 1363–1367.

Harmon, L. J., Andreazzi, C. S., Débarre, F., … Matthews, B. (2019). Detecting the macroevolutionary signal of species interactions. *Journal of Evolutionary Biology*. doi:10.1111/jeb.13477

Hatcher, M. J., Dick, J. T. A., & Dunn, A. M. (2012). Diverse effects of parasites in ecosystems: Linking interdependent processes. *Frontiers in Ecology and the Environment*. doi:10.1890/110016

Hatcher, M. J., Hogg, J. C., & Dunn, A. M. (2005). Local adaptation and enhanced virulence of *Nosema granulosis* artificially introduced into novel populations of its crustacean host, *Gammarus duebeni*. *International Journal for Parasitology*, **35**(3), 265–274.

Hayman, J. R., Hayes, S. F., Amon, J., & Nash, T. E. (2001). Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. *Infection and Immunity*, **69**(11), 7057–7066.

Hibbett, D. S., Binder, M., Bischoff, J. F., … Zhang, N. (2007). A higher-level phylogenetic classification of the Fungi. *Mycological Research*, **111**(5), 509–547.

Hogg, I. D., Larose, C., Lafontaine, Y. de, & Doe, K. G. (1998). Genetic evidence for a *Hyalella* species complex within the Great Lakes - St. Lawrence River drainage basin: implications for ecotoxicology and conservation biology. *Canadian Journal of Zoology*, **76**(6), 1134–1152.

Hogg, J. C., Ironside, J. E., Sharpe, R. G., Hatcher, M. J., Smith, J. E., & Dunn, a M. (2002). Infection of *Gammarus duebeni* populations by two vertically transmitted microsporidia; parasite detection and discrimination by PCR-RFLP. *Parasitology*, **125**(Pt 1), 59–63.

Hou, Z. E., & Li, S. (2002). Two new species of troglobitic amphipod crustaceans (Gammaridae) from Hubei Province, China. *Raffles Bulletin of Zoology*, **50**(1), 27–36.

Hou, Z. E., & Li, S. (2004). *Gammarus* species from Tibet Plateau, China (Crustacea: Amphipoda: Gammaridae). *Raffles Bulletin of Zoology*, **52**(1), 147–170.

Hou, Z., Sket, B., Fiser, C., & Li, S. (2011). Eocene habitat shift from saline to freshwater promoted Tethyan amphipod diversification. *Proceedings of the National Academy of Sciences*, **108**(35), 14533–14538.

Hupało, K., Teixeira, M. A. L., Rewicz, T., … Costa, F. O. (2019). Persistence of phylogeographic footprints helps to understand cryptic diversity detected in two marine amphipods widespread in the Mediterranean basin. *Molecular Phylogenetics and Evolution*. doi:10.1016/j.ympev.2018.11.013

Hurst, G. D. D., & Majerus, M. E. N. (1993). Why do maternally inherited microorganisms kill males? *Heredity*, **71**(1), 81–95.

Ironside, J. E. (2007). Multiple losses of sex within a single genus of Microsporidia. *BMC Evolutionary Biology*, **7**(1), 48.

Ironside, J. E. (2013). Diversity and recombination of dispersed ribosomal DNA and protein coding genes in Microsporidia. *PLoS ONE*, **8**(2). doi:10.1371/journal.pone.0055878

Ironside, J. E., & Alexander, J. (2015). Microsporidian parasites feminise hosts without paramyxean co-infection: Support for convergent evolution of parasitic feminisation. *International Journal for Parasitology*, **45**(6), 427–433.

Ironside, J. E., Dunn, A. M., Rollinson, D., & Smith, J. E. (2003a). Association with host mitochondrial haplotypes suggests that feminizing microsporidia lack horizontal transmission. *Journal of Evolutionary Biology*, **16**(6), 1077–1083.

Ironside, J. E., Smith, J. E., Hatcher, M. J., & Dunn, A. M. (2011). Should sex-ratio distorting parasites abandon horizontal transmission? *BMC Evolutionary Biology*, **11**(1), 370.

Ironside, J. E., Smith, J. E., Hatcher, M. J., Sharpe, R. G., Rollinson, D., & Dunn, A. M. (2003b). Two species of feminizing microsporidian parasite coexist in populations of *Gammarus duebeni*. *Journal of Evolutionary Biology*, **16**(3), 467–473.

Ironside, J. E., & Wilkinson, T. J. (2018). Accumulation and exchange of parasites during adaptive radiation in an ancient lake. *International Journal for Parasitology*, **48**(3–4), 297–307.

Ironside, J. E., Wilkinson, T. J., & Rock, J. (2008). Distribution and host range of the microsporidian *Pleistophora mulleri*. *Journal of Eukaryotic Microbiology*, **55**(4), 355–362.

Kageyama, D., Ohno, M., Sasaki, T., … Sahara, K. (2017). Feminizing Wolbachia endosymbiont disrupts maternal sex chromosome inheritance in a butterfly species . *Evolution Letters*. doi:10.1002/evl3.28

Kaltenpoth, M., Goettler, W., Koehler, S., & Strohm, E. (2010). Life cycle and population dynamics of a protective insect symbiont reveal severe bottlenecks during vertical transmission. *Evolutionary Ecology*, **24**(2), 463–477.

Karaman, Gordan S; Pinkster, S. (1987). Freshwater *Gammarus* species from Europe , North Africa and adjacent regions of Asia (Crustacea-Amphipoda) Part III. *Gammarus balcanicus*-group and related species. *Repository.Naturalis.Nl*. Retrieved from http://repository.naturalis.nl/document/547910

Karaman, G. S., & Pinkster, S. (1977a). Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea-Amphipoda). Part II. *Gammarus roeseli*-group and related species. *Bijdragen Tot de Dierkunde*, **47**(1), 165–196.

Karaman, G. S., & Pinkster, S. (1977b). Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea - Amphipoda) Part I. *Gammarus pulex*-group and related species. *Bijdragen Tot De Dierkunde*, **47**(1), 1–86.

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. *Frontiers in Microbiology*. doi:10.3389/fmicb.2014.00112

Katouzian, A. R., Sari, A., Macher, J. N., … Weigand, A. M. (2016). Drastic underestimation of amphipod biodiversity in the endangered Irano-Anatolian and Caucasus biodiversity hotspots. *Scientific Reports*. doi:10.1038/srep22507

Keeling, P. J., & Fast, N. M. (2002). Microsporidia: Biology and Evolution of Highly Reduced Intracellular Parasites. *Annual Review of Microbiology*, **56**(1), 93–116.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2002). Incomplete feminisation by the microsporidian sex ratio distorter, *Nosema granulosis*, and reduced transmission and feminisation efficiency at low temperatures. *International Journal for Parasitology*, **32**(7), 825–831.

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. *Infection, Genetics and Evolution*, **10**(7), 1027–1038.

Krebes, L., Zeidler, L., Frankowski, J., & Bastrop, R. (2014). (Cryptic) sex in the microsporidian *Nosema granulosis* - Evidence from parasite rDNA and host mitochondrial DNA. *Infection, Genetics and Evolution*, **21**, 259–268.

Lafferty, K. D. (2010). Interacting parasites. *Science*, **330**(6001), 187–188.

Lagrue, C., Wattier, R., Galipaud, M., … Bollache, L. (2014). Confrontation of cryptic diversity and mate discrimination within *Gammarus pulex* and *Gammarus fossarum* species complexes. *Freshwater Biology*, **59**(12), 2555–2570.

Larsson, J. I. R. (2005). Molecular versus morphological approach to microsporidian classification. *Folia Parasitologica*, **52**(1–2), 143–144.

Larsson, J. I. R. (2007). Cytological variation and pathogenicity of the bumble bee parasite *Nosema bombi* (Microspora, Nosematidae). *Journal of Invertebrate Pathology*, **94**(1), 1–11.

Lefébure, T., Douady, C. J., Gouy, M., & Gibert, J. (2006). Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*, **40**(2), 435–447.

Leigh, E. G. (2010). The evolution of mutualism. *Journal of Evolutionary Biology*. doi:10.1111/j.1420-9101.2010.02114.x

Li, Y., Wu, Z., Pan, G., … Zhou, Z. (2009). Identification of a novel spore wall protein (SWP26) from microsporidia *Nosema bombycis*. *International Journal for Parasitology*, **39**(4), 391–398.

Lin, D., & Koskella, B. (2015). Friend and foe: Factors influencing the movement of the bacterium *Helicobacter pylori* along the parasitism-mutualism continuum. *Evolutionary Applications*. doi:10.1111/eva.12231

Macneil, C., A Dick, J. T., & Elwood, R. W. (1997). Amphipoda: Problems and perspectives concerning the functional feeding group concept. *Biol. Rev*, **72**, 349–364.

MacNeil, C., A Dick, J. T., Hatcher, M. J., … Dunn, A. M. (2003a). Differential drift and parasitism in invading and native *Gammarus* spp. (Crustacea: Amphipoda). *Ecography*, **26**, 467–473.

MacNeil, C., Dick, J. T. A., Hatcher, M. J., Fielding, N. J., Hume, K. D., & Dunn, A. M. (2003b). Parasite transmission and cannibalism in an amphipod (Crustacea). *International Journal for Parasitology*, **33**(8), 795–798.

MacNeil, C., Dick, J. T. A., Hatcher, M. J., Terry, R. S., Smith, J. E., & Dunn, A. M. (2003c). Parasite-mediated predation between native and invasive amphipods. *Proceedings of the Royal Society B: Biological Sciences*, **270**(1521), 1309–1314.

Mamos, T., Wattier, R., Burzyński, A., & Grabowski, M. (2016). The legacy of a vanished sea: A high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. *Molecular Ecology*, **25**(3), 795–810.

Matthews, A. C., Mikonranta, L., & Raymond, B. (2019). Shifts along the parasite-mutualist continuum are opposed by fundamental trade-offs. *Proceedings of the Royal Society B: Biological Sciences*. doi:10.1098/rspb.2019.0236

Mautner, S. I., Cook, K. A., Forbes, M. R., McCurdy, D. G., & Dunn, A. M. (2007). Evidence for sex ratio distortion by a new microsporidian parasite of a Corophiid amphipod. *Parasitology*, **134**(Pt 11), 1567–1573.

Moodie, E. G., Le Jambre, L. F., & Katz, M. E. (2003). Thelohania montirivulorum sp. nov. (Microspora: Thelohaniidae), a parasite of the Australian freshwater crayfish, *Cherax destructor*

(Decapoda: Parastacidae): Fine ultrastructure, molecular characteristics and phylogenetic relationships. *Parasitology Research*, **91**(3), 215–228.

Moran, N. A., & Dunbar, H. E. (2006). Sexual acquisition of beneficial symbionts in aphids. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(34), 12803–12806.

Müller, J. (1998). Genetic population structure of two cryptic *Gammarus fossarum* types across a contact zone. *Journal of Evolutionary Biology*. doi:10.1007/s000360050067

Müller, J. (2000). Mitochondrial DNA variation and the evolutionary history of cryptic *Gammarus fossarum* types. *Molecular Phylogenetics and Evolution*. doi:10.1006/mpev.1999.0740

Naegeli, K. (1857). Über die neue Krankheit der Seidenraupe und verwandte Organismen. *Bot Ztg*, **15**, 760–761.

Neuparth, T., Correia, A. D., Costa, F. O., Lima, G., & Costa, M. H. (2005). Multi-level assessment of chronic toxicity of estuarine sediments with the amphipod *Gammarus locusta*: I. Biochemical endpoints. *Marine Environmental Research*, **60**(1), 69–91.

Oulhen, N., Schulz, B. J., & Carrier, T. J. (2016, July 2). English translation of Heinrich Anton de Bary's 1878 speech, 'Die Erscheinung der Symbiose' ('De la symbiose'). *Symbiosis*, Springer Netherlands, pp. 131–139.

Ovcharenko, M., 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**(2), 191–204.

Paaby, A. B., & Rockman, M. V. (2014). Cryptic genetic variation: Evolution's hidden substrate. *Nature Reviews Genetics*, **15**(4), 247–258.

Peuvel-Fanget, I., Polonais, V., Brosson, D., … Delbac, F. (2006). EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall. *International Journal for Parasitology*, **36**(3), 309–318.

Piscart, C., Manach, A., Copp, G. H., & Marmonier, P. (2007). Distribution and microhabitats of native and non-native gammarids (Amphipoda, Crustacea) in Brittany, with particular reference to the endangered endemic sub-species *Gammarus duebeni celticus*. *Journal of Biogeography*, **34**(3), 524–533.

Price, P. W. (1980). Evolutionary biology of parasites. *Monographs in Population Biology*, pp. 1–237.

Quiles, A., Bacela-Spychalska, K., Teixeira, M., … 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. *Parasites and Vectors*, **12**(1), 1–20.

Rewicz, T., Grabowski, M., MacNeil, C., & Bącela-Spychalska, K. (2014). The profile of a 'perfect' invader – the case of killer shrimp, *Dikerogammarus villosus*. *Aquatic Invasions*, **9**(3), 267–288.

Rewicz, T., Wattier, R., Rigaud, T., Grabowski, M., Mamos, T., & Bącela-Spychalska, K. (2017). The killer shrimp, *Dikerogammarus villosus*, invading European Alpine Lakes: A single main source but independent founder events with an overall loss of genetic diversity. *Freshwater Biology*, **62**(6), 1036–1051.

Riddle, B. R., Ladle, R. J., Lourie, S. A., & Whittaker, R. J. (2011). Basic Biogeography: Estimating Biodiversity and Mapping Nature. In *Conservation Biogeography*, Chichester, UK: John Wiley & Sons, Ltd, pp. 45–92.

Rodgers-Gray, T. P., Smith, J. E., Ashcroft, A. E., Isaac, R. E., & Dunn, A. M. (2004). Mechanisms of parasite-induced sex reversal in *Gammarus duebeni*. *International Journal for Parasitology*, **34**(6), 747–753.

Rubin, B. E. R., & Moreau, C. S. (2016). Comparative genomics reveals convergent rates of evolution in ant-plant mutualisms. *Nature Communications*. doi:10.1038/ncomms12679

Ryan, J. A., & Kohler, S. L. (2010). Virulence is context-dependent in a vertically transmitted aquatic host-microparasite system. *International Journal for Parasitology*, **40**(14), 1665–1673.

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. *The Journal of Parasitology*, **95**(1), 198–203.

Shaw, R. W., Kent, M. L., Brown, A. M. V., Whipps, C. M., & Adamson, M. L. (2000). Experimental and natural host specificity of *Loma salmonae* (Microsporidia). *Diseases of Aquatic Organisms*, **40**(2), 131–136.

Slothouber Galbreath, J. G. M., Smith, J. E., Becnel, J. J., Butlin, R. K., & Dunn, A. M. (2009). Reduction in post-invasion genetic diversity in *Crangonyx pseudogracilis* (Amphipoda: Crustacea): A genetic bottleneck or the work of hitchhiking vertically transmitted microparasites? *Biological Invasions*, **12**(1), 191–209.

Smith, C. I., Godsoe, W. K. W., Tank, S., Yoder, J. B., & Pellmyr, O. (2008). Distinguishing coevolution from covicariance in an obligate pollination mutualism: Asynchronous divergence in Joshua tree and its pollinators. *Evolution*, **62**(10), 2676–2687.

Smith, J. E. (2009). The ecology and evolution of microsporidian parasites. *Parasitology*, **136**(14), 1901–1914.

Solter, F., L., Pilarska, K., D., Vossbrinck, & F., C. (2000). Host specificity of microsporidia pathogenic to forest lepidoptera. *Biological Control*, **19**(1), 48–56.

Solter, L. F., & Maddox, J. V. (1998). Physiological host specificity of microsporidia as an indicator of ecological host specificity. *Journal of Invertebrate Pathology*, **71**(3), 207–216.

Southern, T. R., Jolly, C. E., Lester, M. E., & Hayman, J. R. (2007). EnP1, a microsporidian spore wall protein that enables spores to adhere to and infect host cells in vitro. *Eukaryotic Cell*, **6**(8), 1354–1362.

Sprague, V., Becnel, J. J., & Hazard, E. I. (1992). Taxonomy of phylum microspora. *Critical Reviews in Microbiology*, **18**(5–6), 285–395.

Stentiford, G. D., Bateman, K. S., Small, H. J., … 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). *International Journal for Parasitology*, **40**(12), 1433–1446.

Stentiford, G. D., Becnel, J. J., Weiss, L. M., … Solter, L. (2016). Microsporidia – Emergent Pathogens in the Global Food Chain. *Trends in Parasitology*, **32**(4), 336–348.

Stentiford, G. D., & Dunn, A. M. (2014). Microsporidia in Aquatic Invertebrates. In *Microsporidia: Pathogens of Opportunity: First Edition*, Chichester, UK: John Wiley & Sons, Inc., pp. 579–604.

Stentiford, G. D., Feist, S. W., Stone, D. M., Bateman, K. S., & Dunn, A. M. (2013). Microsporidia: Diverse, dynamic, and emergent pathogens in aquatic systems. *Trends in Parasitology*, **29**(11), 567–578.

Struck, T. H., Feder, J. L., Bendiksby, M., … Dimitrov, D. (2018). Finding Evolutionary Processes Hidden in Cryptic Species. *Trends in Ecology and Evolution*, **33**(3), 153–163.

Sutherland, D. L., Hogg, I. D., & Waas, J. R. (2010). Phylogeography and species discrimination in the *Paracalliope fluviatilis* species complex (Crustacea: Amphipoda): Can morphologically similar heterospecifics identify compatible mates? *Biological Journal of the Linnean Society*, **99**(1), 196–205.

Terry, R. S., Smith, J. E., Bouchon, D., … Dunn, A. M. (1999). Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF) microsporidium. *Journal of Eukaryotic Microbiology*, **46**(5), 492–499.

Terry, R. S., Smith, J. E., Sharpe, R. G., … Dunn, A. M. (2004). Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum Microspora. *Proceedings of the Royal Society B: Biological Sciences*, **271**(1550), 1783–1789.

Tofts, C., Hatcher, M. J., Terry, R. S., … Tofts, C. (1995). Evolutionary ecology of vertically transmitted parasites: transovarial transmission of a microsporidian sex ratio distorter in *Gammarus duebeni* . *Parasitology*, **111**(S1), S91.

Tokarev, Y. S., Sokolova, Y. Y., Vasilieva, A. A., & Issi, I. V. (2018). Molecular and Morphological Characterization of *Anncaliia azovica* sp. n. (Microsporidia) Infecting *Niphargogammarus intermedius* (Crustacea, Amphipoda) from the Azov Sea. *Journal of Eukaryotic Microbiology*, **65**(3), 296–307.

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**(1), 241–255.

Van Valen, L. (1973). A new evolutionary theory. *Evolutionary Theory*. doi:Evolutinary Theory

Vávra, J., & Larsson, J. I. R. (2014). Structure of Microsporidia. In *Microsporidia: Pathogens of Opportunity: First Edition*, pp. 1–70.

Vávra, J., & Lukeš, J. (2013). Microsporidia and 'The Art of Living Together.' In *Advances in Parasitology*, Vol. 82, pp. 253–319.

Vossbrinck, C. R., Baker, M. D., & Andreadis, T. G. (2010). Phylogenetic position of Octosporea muscaedomesticae (Microsporidia) and its relationship to Octosporea bayeri based on small subunit rDNA analysis. *Journal of Invertebrate Pathology*, **105**(3), 366–370.

Vossbrinck, C. R., & Debrunner-Vossbrinck, B. A. (2005). Molecular phylogeny of the Microsporidia: Ecological, ultrastructural and taxonomic considerations. *Folia Parasitologica*, **52**(1–2), 131–142.

Vossbrinck, C. R., Debrunner-Vossbrinck, B. A., & Weiss, L. M. (2014). Phylogeny of the Microsporidia. In *Microsporidia: Pathogens of Opportunity: First Edition*, Chichester, UK: John Wiley & Sons, Inc., pp. 203–220.

Wattier, R. A., Haine, E. R., Beguet, J., … Rigaud, T. (2007). No genetic bottleneck or associated microparasite loss in invasive populations of a freshwater amphipod. *Oikos*, **116**(11), 1941–1953.

Weber, R., Bryan, R. T., Schwartz, D. A., & Owen, R. L. (1994). Human microsporidial infections. *Clinical Microbiology Reviews*, **7**(4), 426–461.

Weidner, E. (1972). Ultrastructural study of microsporidian invasion into cells. *Zeitschrift Fur Parasitenkunde*, **40**(3), 227–242.

Weigand, A. M., Kremers, J., Grabner, D. S., … S., G. (2016). Shared microsporidian profiles between an obligate (*Niphargus*) and facultative subterranean amphipod population (*Gammarus*) at sympatry provide indications for underground transmission pathways. *Limnologica*, **58**, 7–10.

Weiss, L. M., & Becnel, J. J. (2015). Microsporidia: pathogens ofopportunity. *Ames: John Wiley & Sons*, (JANUARY 2013). doi:DOI 10.1002/9781118395264.fmatter

Weiss, M., Macher, J. N., Seefeldt, M. A., & Leese, F. (2014). Molecular evidence for further overlooked species within the *Gammarus fossarum* complex (Crustacea: Amphipoda). *Hydrobiologia*. doi:10.1007/s10750-013-1658-7

Westram, A. M. M., Baumgartner, C., Keller, I., & Jokela, J. (2011). Are cryptic host species also cryptic to parasites? Host specificity and geographical distribution of acanthocephalan parasites infecting freshwater *Gammarus*. *Infection, Genetics and Evolution*, **11**(5), 1083–1090.

Wilkinson, T. J., Rock, J., Whiteley, N. M., Ovcharenko, M. O., & Ironside, J. E. (2011). Genetic diversity of the feminising microsporidian parasite *Dictyocoela*: New insights into host-specificity, sex and phylogeography. *International Journal for Parasitology*, **41**(9), 959–966.

Windsor, D. A. (1998). Most of the species on Earth are parasites. *International Journal for Parasitology*, **28**(12), 1939–41.

Winters, A. D., & Faisal, M. (2014). Molecular and ultrastructural characterization of *Haplosporidium diporeiae* n. sp., a parasite of *Diporeia* sp. (Amphipoda, Gammaridea) in the Laurentian Great Lakes (USA). *Parasite*, **7**(1), 343.

Winters, A. D., Fitzgerald, S., Brenden, T. O., Nalepa, T., & Faisal, M. (2014). Spatio-temporal dynamics of parasites infecting *Diporeia* spp. (Amphipoda, Gammaridae) in southern Lake Michigan (USA). *Journal of Invertebrate Pathology*, **121**, 37–45.

Witt, J. D., & Hebert, P. D. (2011). Cryptic species diversity and evolution in the amphipod genus *Hyalella* within central glaciated North America: a molecular phylogenetic approach . *Canadian Journal of Fisheries and Aquatic Sciences*, **57**(4), 687–698.

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. *Molecular Ecology*. doi:10.1111/j.1365-294X.2006.02999.x

Wittner, M. (2014). Historic Perspective on the Microsporidia: Expanding Horizons. In M. Wittner & L. M. Weiss, eds., *The Microsporidia and Microsporidiosis*, Washington: ASM Press, pp. 1–6.

Wu, Z., Li, Y., Pan, G., … Xiang, Z. (2008). Proteomic analysis of spore wall proteins and identification of two spore wall proteins from *Nosema bombycis* (Microsporidia). *Proteomics*, **8**(12), 2447–2461.

Wu, Z., Li, Y., Pan, G., Zhou, Z., & Xiang, Z. (2009). SWP25, a novel protein associated with the *Nosema bombycis* endospore. *Journal of Eukaryotic Microbiology*, **56**(2), 113–118.

Xu, J., Pan, G., Fang, L., … Xiang, Z. (2006). The varying microsporidian genome: Existence of long-terminal repeat retrotransposon in domesticated silkworm parasite *Nosema bombycis*. *International Journal for Parasitology*, **36**(9), 1049–1056.

Yang, G., Short, S., Kille, P., & Ford, A. T. (2011). Microsporidia infections in the amphipod, *Echinogammarus marinus* (Leach): Suggestions of varying causal mechanisms to intersexuality. *Marine Biology*, **158**(2), 461–470.

CHAPTER I. MICROSPORIDIAN INFECTIONS IN THE SPECIES COMPLEX *GAMMARUS ROESELII* (AMPHIPODA) OVER ITS GEOGRAPHICAL RANGE: EVIDENCE FOR BOTH HOST–PARASITE CO-DIVERSIFICATION AND RECENT HOST SHIFTS

Parasites & Vectors

RESEARCH Open Access

Microsporidian infections in the species complex *Gammarus roeselii* (Amphipoda) over its geographical range: evidence for both host–parasite co-diversification and recent host shifts

Adrien Quiles[1,2*], Karolina Bacela-Spychalska[2], Maria Teixeira[1], Nicolas Lambin[1], Michal Grabowski[2], Thierry Rigaud[1] and Rémi André Wattier[1]

Abstract

Background: Microsporidians are obligate endoparasites infecting taxonomically diverse hosts. Both vertical (from mother to eggs) and horizontal (between conspecifics or between species) transmission routes are known. While the former may promote co-speciation and host-specificity, the latter may promote shifts between host species. Among aquatic arthropods, freshwater amphipod crustaceans are hosts for many microsporidian species. However, despite numerous studies, no general pattern emerged about host specificity and co-diversification. In south-eastern Europe, the gammarid *Gammarus roeselii* is composed of 13 cryptic lineages of Miocene to Pleistocene age but few genotypes of one lineage have spread postglacially throughout north-western Europe. Based on nearly 100 sampling sites covering its entire range, we aim to: (i) explore the microsporidian diversity present in *G. roeselii* and their phylogenetic relationships, especially in relation to the parasites infecting other Gammaridae; (ii) test if the host phylogeographical history might have impacted host–parasite association (e.g. co-diversifications or recent host shifts from local fauna).

Methods: We used part of the small subunit rRNA gene as source of sequences to identify and determine the phylogenetic position of the microsporidian taxa infecting *G. roeselii*.

Results: Microsporidian diversity was high in *G. roeselii* with 24 detected haplogroups, clustered into 18 species-level taxa. Ten microsporidian species were rare, infecting a few individual hosts in a few populations, and were mostly phylogenetically related to parasites from other amphipods or various crustaceans. Other microsporidians were represented by widespread genera with high prevalence: *Nosema*, *Cucumispora* and *Dictyocoela*. Two contrasting host association patterns could be observed. First, two vertically transmitted microsporidian species, *Nosema granulosis* and *Dictyocoela roeselum*, share the pattern of infecting *G. roeselii* over most of its range and are specific to this host suggesting the co-diversification scenario. This pattern contrasted with that of *Dictyocoela muelleri*, the three species of *Cucumispora*, and the rare parasites, present only in the recently colonised region by the host. These patterns suggest recent acquisitions from local host species, after the recent spread of *G. roeselii*.

Conclusions: Microsporidians infecting *G. roeselii* revealed two scenarios of host–parasite associations: (i) ancient associations with vertically transmitted parasites that probably co-diversified with their hosts, and (ii) host shifts from local host species, after the postglacial spread of *G. roeselii*.

*Correspondence: quiles.adrien@gmail.com
[1] Université Bourgogne Franche-Comté, Laboratoire Biogéosciences, UMR CNRS 6282, 6 Boulevard Gabriel, 21000 Dijon, France
Full list of author information is available at the end of the article

Keywords: Host–parasite interactions, Biogeography, Phylogeography, Microsporidia, Amphipods

Background

Microsporidians are obligate endoparasites infecting taxonomically diverse hosts inhabiting various environments [1]. These unicellular eukaryotes form a very old and phylogenetically highly diverse phylum, with more than 1300 species in 160 genera [2]. They exhibit different transmission and host exploitation strategies, such as horizontal transmission (HT) often linked to high virulence, or vertical transmission (VT), often associated with low or presumably no virulence, or a combination of both VT and HT [3, 4]. Microsporidians are particularly frequent in aquatic ecosystems [5]. Among the aquatic arthropods, freshwater amphipod crustaceans, and especially species of the Gammaridae, are regular hosts for microsporidians (for overviews see [6, 7]). Since the early descriptions of microsporidians in amphipods (e.g. [6, 8]) there is a constantly increasing number of full descriptions of new genera and species, combining ultrastructural and molecular phylogenetic support [9–11]. Species of three major microsporidian genera (*Nosema* [12], *Cucumispora* [9] and *Dictyocoela* [6]) are infecting freshwater amphipods. They infect many host species across Europe and are highly prevalent. *Nosema granulosis* is so far the only species of its genus known to infect gammarids. This parasite is transovarially-transmitted, feminises host offspring and in consequence induces excess of females in the infected populations [6, 13–15]. This parasite causes limited pathology and shows little evidence of horizontal transmission [3, 8, 15, 16]. *Nosema granulosis* was found to infect other gammarid species such as *Dikerogammarus villosus* [17], *Gammarus fossarum* [16, 18], *Gammarus roeselii* [19] and *Niphargus schellenbergi* [16]. The second frequent genus, *Cucumispora* [9], was found to date only in gammarids. Three species have been described: *C. dikerogammari* infecting *Dikerogammarus villosus* [9], *C. ornata* infecting *Dikerogammarus haemobaphes* [10] and *C. roeselii* infecting *Gammarus roeselii* [20]. These parasites are infecting mostly muscles and show high rates of horizontal trophic transmission [6, 9, 21]. *Cucumispora dikerogammari* was also observed to manipulate predatory behaviour of their hosts [22, 23]. Finally, *Dictyocoela* was found to be the dominant microsporidian genus infecting freshwater amphipods [6]. Its phylogeny was recently reassessed using an integrative approach (molecular and ultrastructural traits, see [24]). On one hand, four species were fully described: *Dictyocoela duebenum*, *D. muelleri*, *D. berillonum* and *D. roeselum.* On the other hand, at least four other species belonging to this genus await formal description, as only ribosomal sequences are available to date [24]. *Dictyocoela* spp. are usually described as vertically-transmitted [6, 8], inducing sex ratio distortion by feminising males, similar to that induced by *N. granulosis*. They have been found, often at high prevalence, in numerous *Gammarus* spp. (*Gammarus aequicauda* [24], *G. balcanicus* [24], *G. duebeni* [25], *G. fossarum* [25, 26], *G. lacustris* [25], *G. pulex* [20, 25, 27], *G. roeselii* [6, 19, 24], *G. setosus* [24, 25], *G. varsoviensis* [24]) and in 15 other amphipod species [6, 17, 20, 24, 25, 28, 29]. In addition to these three major genera, a dozen of additional microsporidian lineages have been identified based on molecular divergence. They were not fully described (i.e. lacking anatomical and ultrastructural descriptions) and, in most cases, they were only observed sporadically [5, 6, 18, 25, 30, 31].

The nature of the studies investigating microsporidian infections in European gammarids is diverse. For example, *G. duebeni* and *D. villosus* host populations were investigated at both small [9, 32] and large geographical scales [17, 18]. However, other studies focusing on single hosts were more limited geographically, restricted to one or a few populations. This was the case for *Gammarus roeselii* [19, 33] or *D. haemobaphes* [21]. Other studies were targeting one parasite in several host species such as *Dictyocoela* spp. [24, 25] or *Pleistophora muelleri* [30]. Finally, recent studies explored host and microsporidian assemblages at local geographical scales, such as part of the Ruhr drainage [5] or in Lake Baikal [28, 34, 35]. From all these studies, no clear specificity pattern emerged. In restrained geographical areas, a single host species may be infected by numerous parasite species (e.g. [5, 34]). Conversely, parasite species of a single genus may infect several host species (e.g. [25]). Furthermore, invasive amphipod species may introduce novel pathogens to a colonised area [21], which may promote parasite adaptation to novel hosts and, therefore, emergence of a new disease [36]. These patterns are sometimes obscured by poor resolution in parasite identification: it is often difficult to assert if microsporidians infecting several hosts belong to single or several "species" [5]. An additional level of complexity comes from many recent studies pointing out that most widespread gammarid species are in fact species complexes characterised by high cryptic diversity. Indeed, individuals ascribed to a single 'species' based on shared diagnostic morphological features may belong to highly divergent phylogenetic lineages (e.g. [37–44]). In most studies upon microsporidia-amphipod relationships, this cryptic diversity was not taken into account. For example, it remains unknown if a single

microsporidian species is specific to one of the cryptic host species or can infect the whole species complex. Indeed, in gammarids, the cryptic diversity of the host seems not cryptic at all for their parasites, e.g. acanthocephalans [45, 46].

The present study aimed to investigate the microsporidian infection patterns in a host with high cryptic diversity. We used *Gammarus roeselii* as a biological model because its biogeography and diversification patterns have recently been investigated [41, 47]. *Gammarus roeselii* populations are widely distributed across European freshwater ecosystems, but this morphospecies is characterised by extensive cryptic diversity with at least 13 highly divergent phylogenetic lineages (molecular operational taxonomic units, MOTUs, named A-M) [41]. These MOTUs diversified mostly over Miocene (starting *c.*18 Mya) in south-east Europe, predominantly in the Balkan Peninsula [41]. We define this area as primary diversification region (Region 1 in Fig. 1). However, one of these MOTUs (MOTU C) diversified further during Pleistocene in the Pannonian basin, north from the Balkans. We define it as secondary diversification area (Region 2a, Fig. 1). One of the lineages of MOTU C expanded postglacially its geographical range in northern and western Europe (Region 2b, Fig. 1), probably by a natural range expansion facilitated occasionally by human activities [47]. Some *G. roeselii* populations were shown to harbour microsporidian infections belonging to species of the three main genera infecting amphipods: *Nosema*, *Dictyocoela* and *Cucumispora* [5, 19, 20, 33]. These studies were conducted only in the area recently colonised by *G. roeselii* (France, Germany and Poland), thus providing a limited overview of the microsporidian assemblage associated with this host. In addition, the extent of infections in the host cryptic lineages remains unknown, with only *G. roeselii* MOTU C being studied

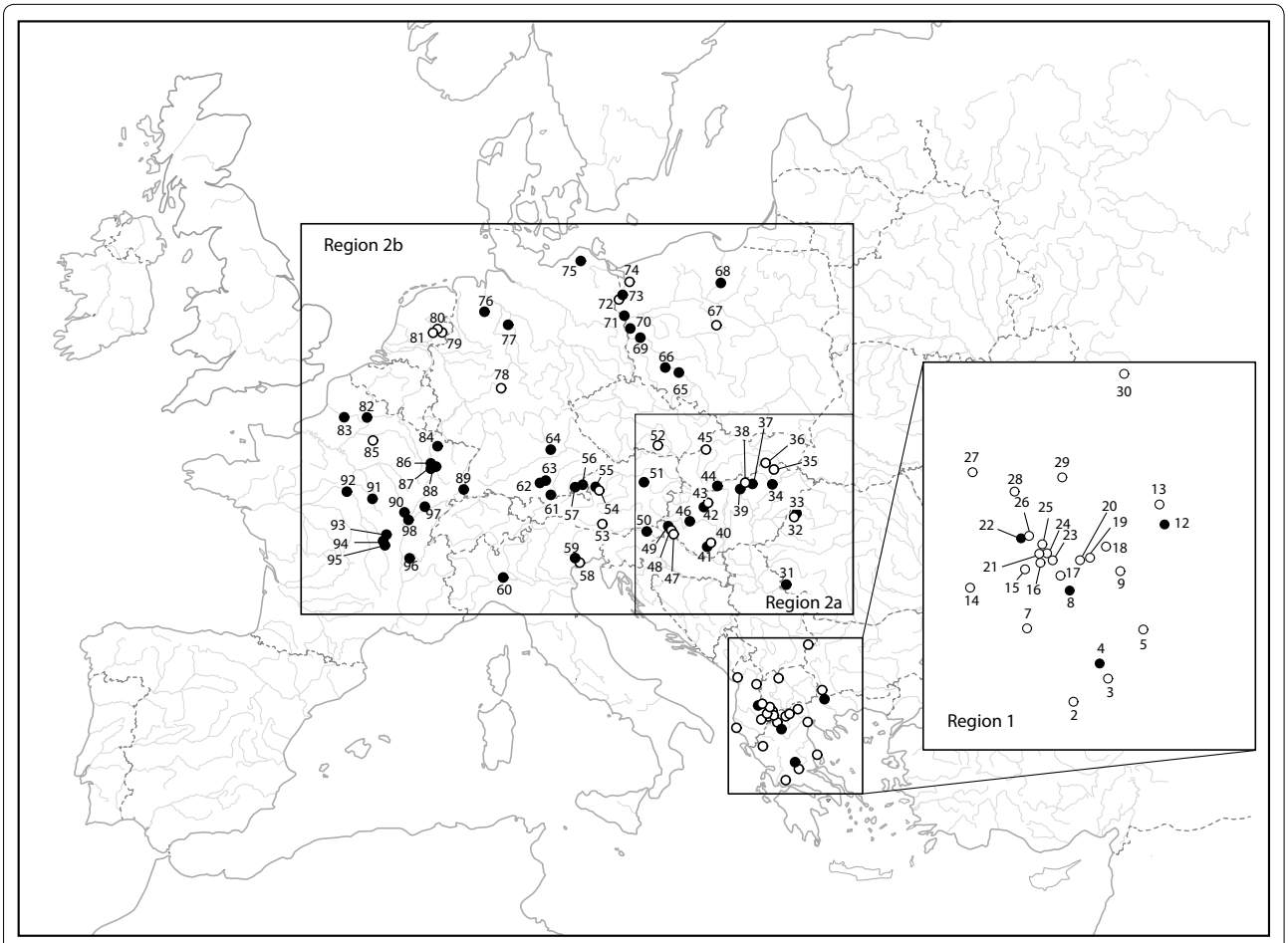


Fig. 1 *Gammarus roeselii* sampling sites in this study. Empty dots refer to sites with no microsporidian infections. Black dots refer to sites where at least one infection was found in a *G. roeselii* individual and for which a sequence was obtained. Site labels refer to Additional file 1: Table S1. The enlarged geographical zone corresponds to area of ancient diversification of *G. roeselii* (thereafter named Region 1). Other sites are located either in the geographical zone of secondary diversification of *G. roeselii* (Region 2a) or in the area of recent, post-glacial, expansion (Region 2b)

so far. In these studies, *Nosema granulosis*, *Dictyocoela roeselum* and *D. muelleri* showed vertical transmission, and induced sex-ratio bias in host populations [19, 48], while the muscle-infecting *Cucumispora roeselum* was shown to be pathogenic, suggesting horizontal transmission [20].

The main aim of our study was to characterise extensively the diversity of microsporidians infecting *G. roeselii* over its entire geographical range and across the highly divergent host MOTUs. This will allow addressing the following issues: (i) clarify the infection pattern and specificity, by exploring the phylogenetic relationships of the detected microsporidia, and comparing these parasites with those of other Gammaridae; and (ii) disentangle the ecological or evolutionary scenarios leading to the observed infection pattern. The following questions were explored: (i) Are all parasite clades present in the host diversification zone in south-east Europe and do these clades show a diversification pattern matching that of the host (suggesting possibilities of host–parasite co-diversifications)? We predict that vertically-transmitted *Nosema* spp. and *Dictyocoela* spp., intimately linked to their hosts, correspond to such a pattern; (ii) Are some parasites restricted to the recently colonised region, and are they taxonomically identical to microsprodian infections found in local gammarids (suggesting possibilities for recent host shifts from local fauna in a given zone)? We predict that most horizontally-transmitted parasites may show such a pattern.

Methods

Sampling

Gammarus roeselii individuals were collected using hand nets and kick-sampling method, at 94 sites in 19 countries, during several sampling campaigns between 2004 and 2016, covering the area of western, central and south-eastern Europe; *c.*4 million km^2 (Additional file 1: Table S1). Sites were plotted on a map (Fig. 1) using Qgis 2.18.4 [49]. The Balkans, hereafter referred to Region 1 (Fig. 1, 26 sites sampled), is known as an area of ancient (mostly Miocene) diversification of the host [41]. *Gammarus roeselii* secondarily (in Pleistocene) diversified in the Pannonian Plain, north from the Balkans [47], hereafter referred to as Region 2a (Fig. 1, 22 sites sampled). The rest of the distribution of *G. roeselii*, hereafter referred to as Region 2b (Fig. 1, 46 sites sampled), is the area colonised postglacially as a result of natural and anthropogenic processes [47]. Individuals were fixed in 96% ethanol directly in the field, and stored at room temperature after returning to the laboratory. Amphipods were identified to the species level using morphological characters described in available keys (e.g. [50, 51]). All the material was stored at the Department of Invertebrate Zoology and Hydrobiology, University of Lodz, Poland, and the laboratory of Biogeosciences, University Bourgogne Franche-Comté, Dijon, France.

Host dissection and total DNA extraction

Dissection of each gammarid was performed under stereomicroscope taking *c.*2 mm^3 of animal tissue (including muscles and gonads), from the thoracic segments 6 and 7. As microsporidia are intracellular parasites, their DNA was co-extracted with host DNA. Among the 1904 individuals used in the present study, DNA of 1108 individuals was already extracted as within the study by Grabowski et al. [41, 47] and 796 were newly extracted. While the sex of host was not recorded at the time of dissection for already extracted samples, it was determined for the newly extracted individuals for 10 sites, mainly from Region 1, we newly extracted DNA for up to 24 males and 24 females in addition to the initial set of DNA samples (Additional file 1: Table S1). Finally, *G. roeselii* individuals from additional sites (relative to Grabowski et al. [41, 47]) were dissected and newly extracted, for Region 1 and 2b. Altogether, Region 1, 2a and 2b accounted, respectively, for 931, 327 and 646 individuals. DNA extraction was performed using either (i) standard phenol-chloroform protocol or (ii) Biobasic EZ-10 96 Well Plate Animal Genomic DNA Isolation Kit and eluted in 100 µl of TE (pH 8). The DNA samples were kept at 4 °C until amplification and subsequently at −20 °C for long-term storage.

Molecular screening for microsporidians

All 1904 individuals were screened for the presence of microsporidians using a short (*c.*350 bp long) diagnostic fragment of the small ribosomal subunit (*SSU* rDNA) marker. The microsporidia-specific primer V1f (forward) (5′-CAC CAG GTT GAT TCT GCC TGA C-3′) [52] paired with newly designed UNIr (reverse) (5′-TCA GGC TCC CTC TCC GGA AT-3′) was used. The use of this short fragment maximised the ability to detect the presence of microsporidians even in case of low infection intensity. As negative and positive controls in PCR reactions, we used, respectively, water and microsporidian DNA (*Dictyocoela roeselum*). The PCR program consisted of an initial denaturing phase at 95 °C for 2 min, followed by 35 cycles of 95 °C for 20 s, 57 °C for 20 s and 72 °C for 20 s, and a final extension at 72 °C for 5 min. The PCR products were visualised by electrophoresis after 20 min migration under 100V in 2% agarose gel.

Sequencing of the microsporidian *SSU* rDNA

For all individuals positively diagnosed for microsporidian infection by PCR screening (see above) our target was to sequence the *c*.800 bp long fragment matching the 5′ part of the *SSU* rRNA gene. This target was tentatively achieved following two strategies: one based on a single amplicon using V1f as the forward primer and HG4r (5′-GCG GCT TAA TTT GAC TCAA C-3′) as the reverse primer (amplicon of *c*.850 bp) and one based on two amplicons with overlapping region i.e. V1F as the forward primer and 530r (5′-CCG CGG CTG CTG GCA C-3′) as the reverse primer; in addition to MC2F as the forward primer (5′-TCC GGA GAG GGA GCC TGA GAG A-3′) and 964r as the reverse primer (5′-CGC GTT GAG TCA AAT TAA GCC GCA CA-3′). When the 800 bp target was not reachable, either a V1f-530r fragment (*c*.530 bp long) or even a V1f-UNIr fragment (*c*.350 bp long) was used. Although the V1f-UNIr fragment is short, it contained enough phylogenetic information to ascribe sequences to, at least, the species level (see "Results", Additional file 2: Table S2 and Additional file 3: Data S1). PCR products were purified and sequenced directly with the BigDye technology by Genewiz, Inc., UK, using the forward primers used in the PCR. Using Geneious 10.2.2 [53], raw sequences were edited, trimmed and checked for being microsporidian sequences *via* BlastN [54] search against the sequences available on GenBank.

Phylogeny reconstruction for microsporidians

Four types of microsporidian sequences constituted our dataset: (i) sequences newly produced from our collection of infected *G. roeselii* individuals; (ii) published *SSU* sequences representing diversity and divergence of microsporidians already found to infect European freshwater or brackish water amphipods (we did not include sequences outside Europe, e.g. the recently published parasites from Lake Baikal [34, 35]); (iii) published *SSU* sequences for microsporidians infecting other taxa, prioritising freshwater or brackish water invertebrates, when closely related amphipod sequences relative to newly produced sequences were not found; (iv) published *SSU* sequences representative of the five microsporidian clades (Clades I–V), as determined in the integrative phylogenies presented in literature [55, 56]. All sequences were aligned using MAFFT7.388 software [57], with the E-IONS-I algorithm using legacy gap penalty option, incorporated in Geneious 10.2.2 [53]. Our dataset consisted of sequences of variable lengths depending on both the success in producing new sequences (from 180 to 826 bp) and on various length of the published ones (from 300 to 1448 bp for microsporidians, and 1786 bp for the fungus *Basidiobolus ranarum* used as an outgroup). All details, including sequence length, are given in Additional file 4: Table S3. Alignments are given in Additional file 5: Data S2. As some sequences were relatively short, reducing the full dataset to a standard size would, on the one hand, allow defining haplotypes but, on the other hand, would potential induce losing phylogenetic signal. Therefore, we attributed each sequence to haplogroups, defined in such a way that sequences belonged to distinct haplogroups if they differed by one or more variable sites, generating diagnostic features (Additional file 2: Table S2), whatever sequence length. Two sequences were clustered in one haplogroup, despite variable length, based on 100% pairwise identity, therefore sharing the same diagnostic sites. A limited set of newly produced sequences could be assigned to at least two haplogroups due to a combination of reduced length and lack of diagnostic features. Only the longest sequence representing each haplogroup was used for the phylogeny reconstruction (248 to 826 bp, noted in Additional file 4: Table S3; see also alignments in Additional file 5: Data S2).

Bayesian phylogeny reconstructions were performed with MrBayes [58] incorporated in Geneious 10.2.2. The best-fitting model of nucleotide substitution was determined with JModelTest-2.1.10. [59]. This was always the General Time Reversible (GTR) model with gamma-distributed rate heterogeneity (G) and a significant proportion of invariable sites (I). Four heated chains, each 1,100,000 iterations long, sampled every 200 iterations, were run. The runs reached satisfactory effective sampling sizes (ESS > 200), and the potential scale reduction factor values equalled 1 for all parameters. The 50% majority-rule consensus tree was constructed after the removal of 10% 'burn-in' trees. Four phylogenetic trees were constructed. The first tree contained all haplogroups (i.e. sequences from this study and published sequences) using *Basidiobolus ranarum* (GenBank: AY635841) as the outgroup [55]. In this tree, we described novel parasites by conservatively using provisional names, e.g. *Microsporidia* sp. (hereafter abbreviated *Msp*) followed by the clade number (from I to V) *sensu* Vossbrinck et al. [55] and a superscript roman letter. The three other phylogenies represent detailed analyses for the already identified parasites of the microsporidian species of the genera infecting amphipods: *Nosema* [12], *Cucumispora* [9] and *Dictyocoela* [6]. *Nosema antherae* (GenBank: DQ073396), *Vavraia culicis* (GenBank: AJ252961), *Dictyocoela cavimanum* (GenBank: AJ438960) were used as outgroups for the *Nosema*, *Cucumispora* and *Dictyocoela* phylogenies, respectively. Following Grabner et al. [5], if a newly obtained sequence was > *c*.98% similar to a sequence for which a full taxonomic description was available, providing genus and species name, such name

was ascribed to the new sequence. Alignments used for building these trees are provided in Additional file 5: Data S2).

Geographical distribution of parasites and potential host specificity

In addition to Bayesian trees, we provided a map constructed in Qgis 2.18.4 to show the geographical distribution of the three genera *Nosema*, *Cucumispora*, *Dictyocoela* in *G. roeselii* (including GenBank data). We also added to these maps geographical positions of gammarids' parasite sequences found in the literature (see also Additional file 4: Table S3). We did not analyse geographical distribution of other microsporidian clades, as most of them were present only in single locations.

Host phylogeny and distribution *versus* microsporidia prevalence and phylogeny

The *cox*1 (cytochrome *c* oxidase subunit 1) results of Grabowski et al. [41, 47] were used as the backbone for *G. roeselii* phylogeography used in the present study. In addition, any individual infected by microsporidians for which the host *cox*1 sequence was not already available (e.g. additional sampling sites) was tentatively newly sequenced for *cox*1 following all the molecular procedures described by Grabowski et al. [41]. These new *cox*1 sequences were attributed, using phylogenetic reconstruction and pairwise identity, to the host MOTUs defined by Grabowski et al. [41] (Additional file 1: Table S1).

Host phylogeny and geographical distribution *versus* those of the microsporidians were assessed in two ways. First, we compared the proportion of infected populations between the three biogeographical regions: 1, 2a and 2b, using Likelihood-Ratio χ^2 or Fisher's exact test. Secondly, we tried to confront parasite phylogeny to the phylogeny of the host. The first challenge would be to find an appropriate taxonomic level for microsporidians relative to the age of diversification, as the *G. roeselii* diversification started *c.*18 Mya [41] while the phylum Microsporidia is likely to date back hundreds of Mya [60]. Microsporidian genus level (e.g. *Nosema*, *Cucumispora* and *Dictyocoela*) might be a better choice than the phylum Microsporidia to run such an analysis. However, the number of microsporidian clades at this taxonomic level was limited relative to the high number of MOTUs observed in *G. roeselii*. For these reasons, we were not able to use co-phylogenetic methods (e.g. [61]); the comparisons were therefore made by eye, by investigating how the parasite haplogroups were distributed across the host phylogeny.

Results

Overall prevalence and broad geographical distribution of microsporidian infections in *G. roeselii*

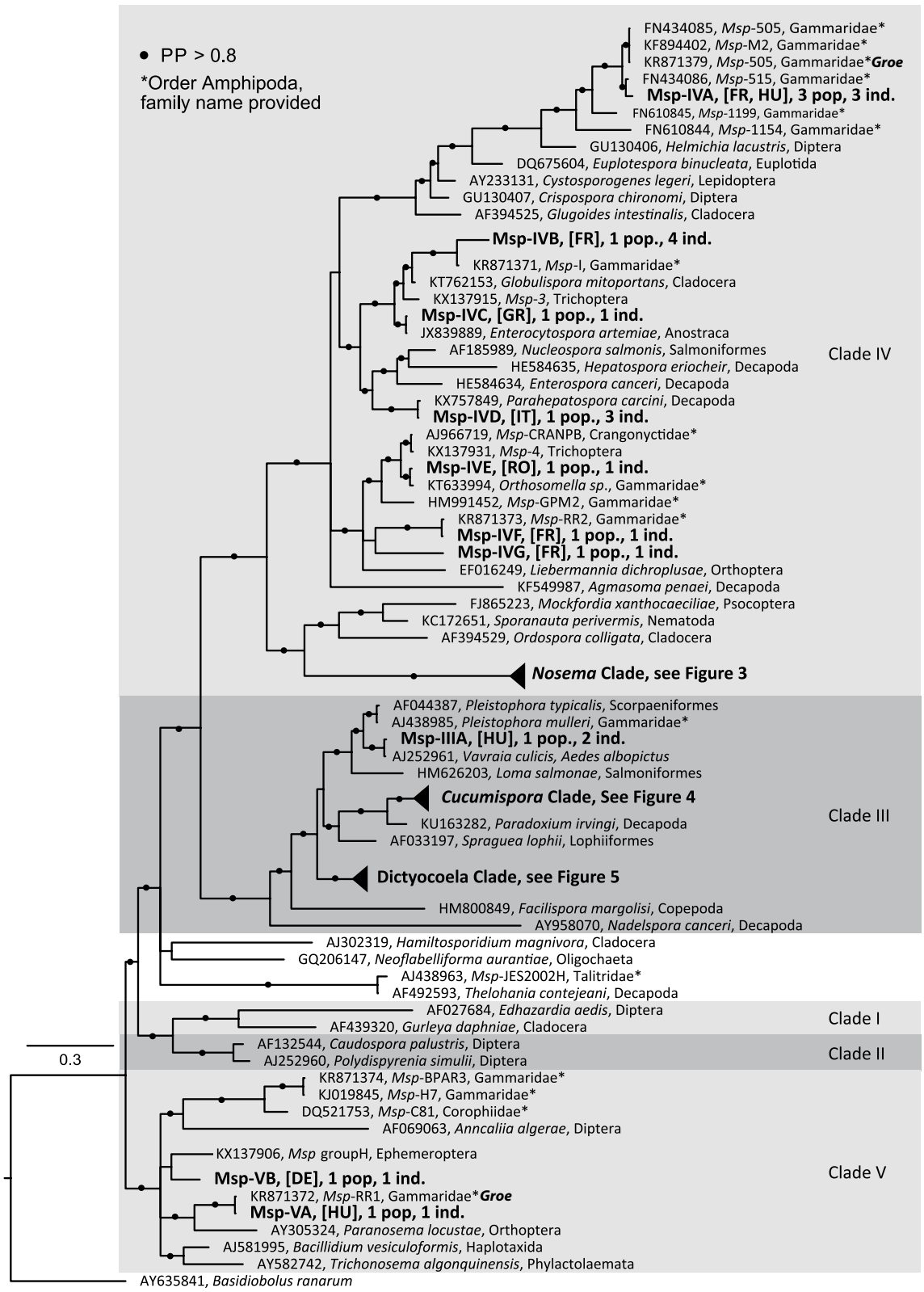
The overall prevalence of microsporidian infections in *G. roeselii* was 16.6% with 316 infected individuals out of the 1904 tested. In 51 sites (54.2%), at least one *G. roeselii* individual was found to be infected with a microsporidian parasite (Fig. 1, Additional file 1: Table S1). There was nevertheless a high variation among sites, ranging from nil to even 100% in one French population (# 97, all individuals being female, Additional file 1: Table S1). This crude pattern of microsporidian infections showed a strong contrast depending on the geographical region. Infections were detected in 4/26 sites (15.4%) in Region 1, in 12/22 sites (54.5%) in Region 2a, and in 35/46 sites (76.1 %) in Region 2b (Likelihood ratio $\chi^2 = 26.38$, $P < 0.0001$) (Fig. 1, Additional file 1: Table S1). In fact, the proportion of infected populations was lower in Region 1 compared to Region 2a or 2b (Fisher's exact test: $P = 0.006$ and $P < 0.0001$, respectively), while the Regions 2a and 2b showed a comparable proportion of infected populations (Fisher's exact test: $P = 0.095$).

Microsporidian diversity and phylogenetic position

Depending on sequencing success, the final sequence length of microsporidian *SSU* from the 316 infected hosts ranged from 180 bp (1 case) to 840 bp, and 39.4% of the sequences were ≥ 687 bp (Additional file 4: Table S3). These sequences could be ascribed to 24 microsporidian haplogroups, which themselves could be clustered into 18 species-level taxa, based on the divergence threshold of *c.*2% (Additional file 1: Table S1 and Additional file 4: Table S3). The newly generated sequences were only associated with three (III, IV and V) of the five microsporidian clades defined by Vossbrinck and Debrunner-Vossbrinck [55]. Most of these sequences (298/316, 94.3%) could be ascribed to species of three genera already known to infect gammarid hosts: *Nosema* (96/316, 30.38%), *Cucumispora* (37/316, 11.71%) and *Dictyocoela* (165/316, 52.22%) (Fig. 2, Additional file 1: Table S1 and Additional file 4: Table S3). The remaining 18 sequences (5.7%), although being rare overall, were recorded from 11 different geographical sites and accounted for 10 out of the 24 observed haplogroups (Fig. 2, Additional file 1: Table S1).

Rare infections (i.e. infections not ascribed to the genera *Nosema*, *Cucumispora* and *Dictyocoela*)

Most of the microsporidians representing rare infections in *G. roeselii* were either phylogenetically close to various microsporidians infecting gammarids (*Msp*-IVA, *Msp*-IVB, *Msp*-IVE, *Msp*-IVF, *Msp*-IVG, Msp-VA) or close to microsporidians infecting other


• PP > 0.8
*Order Amphipoda, family name provided
FN434085, *Msp*-505, Gammaridae*
KF894402, *Msp*-M2, Gammaridae*
KR871379, *Msp*-505, Gammaridae***Groe***
FN434086, *Msp*-515, Gammaridae*
Msp-IVA, [FR, HU], 3 pop, 3 ind.
FN610845, *Msp*-1199, Gammaridae*
FN610844, *Msp*-1154, Gammaridae*
GU130406, *Helmichia lacustris*, Diptera
DQ675604, *Euplotespora binucleata*, Euplotida
AY233131, *Cystosporogenes legeri*, Lepidoptera
GU130407, *Crispospora chironomi*, Diptera
AF394525, *Glugoides intestinalis*, Cladocera
Msp-IVB, [FR], 1 pop., 4 ind.
KR871371, *Msp*-I, Gammaridae*
KT762153, *Globulispora mitoportans*, Cladocera
KX137915, *Msp-3*, Trichoptera
Msp-IVC, [GR], 1 pop., 1 ind.
JX839889, *Enterocytospora artemiae*, Anostraca
AF185989, *Nucleospora salmonis*, Salmoniformes
HE584635, *Hepatospora eriocheir*, Decapoda
HE584634, *Enterospora canceri*, Decapoda
KX757849, *Parahepatospora carcini*, Decapoda
Msp-IVD, [IT], 1 pop., 3 ind.
AJ966719, *Msp*-CRANPB, Crangonyctidae*
KX137931, *Msp*-4, Trichoptera
Msp-IVE, [RO], 1 pop., 1 ind.
KT633994, *Orthosomella sp.*, Gammaridae*
HM991452, *Msp*-GPM2, Gammaridae*
KR871373, *Msp*-RR2, Gammaridae*
Msp-IVF, [FR], 1 pop., 1 ind.
Msp-IVG, [FR], 1 pop., 1 ind.
EF016249, *Liebermannia dichroplusae*, Orthoptera
KF549987, *Agmasoma penaei*, Decapoda
FJ865223, *Mockfordia xanthocaeciliae*, Psocoptera
KC172651, *Sporanauta perivermis*, Nematoda
AF394529, *Ordospora colligata*, Cladocera
***Nosema* Clade, see Figure 3**
Clade IV
AF044387, *Pleistophora typicalis*, Scorpaeniformes
AJ438985, *Pleistophora mulleri*, Gammaridae*
Msp-IIIA, [HU], 1 pop., 2 ind.
AJ252961, *Vavraia culicis, Aedes albopictus*
HM626203, *Loma salmonae*, Salmoniformes
***Cucumispora* Clade, See Figure 4**
KU163282, *Paradoxium irvingi*, Decapoda
AF033197, *Spraguea lophii*, Lophiiformes
Dictyocoela Clade, see Figure 5
HM800849, *Facilispora margolisi*, Copepoda
AY958070, *Nadelspora canceri*, Decapoda
Clade III
AJ302319, *Hamiltosporidium magnivora*, Cladocera
GQ206147, *Neoflabelliforma aurantiae*, Oligochaeta
AJ438963, *Msp*-JES2002H, Talitridae*
AF492593, *Thelohania contejeani*, Decapoda
AF027684, *Edhazardia aedis*, Diptera
AF439320, *Gurleya daphniae*, Cladocera
Clade I
AF132544, *Caudospora palustris*, Diptera
AJ252960, *Polydispyrenia simulii*, Diptera
Clade II
0.3
KR871374, *Msp*-BPAR3, Gammaridae*
KJ019845, *Msp*-H7, Gammaridae*
DQ521753, *Msp*-C81, Corophiidae*
AF069063, *Anncaliia algerae*, Diptera
KX137906, *Msp* groupH, Ephemeroptera
Msp-VB, [DE], 1 pop, 1 ind.
KR871372, *Msp*-RR1, Gammaridae****Groe***
Msp-VA, [HU], 1 pop, 1 ind.
AY305324, *Paranosema locustae*, Orthoptera
AJ581995, *Bacillidium vesiculoformis*, Haplotaxida
AY582742, *Trichonosema algonquinensis*, Phylactolaemata
Clade V
AY635841, *Basidiobolus ranarum*

(See figure on previous page.)

Fig. 2 Bayesian phylogenetic reconstruction based on partial small ribosomal subunit rDNA alignment. Labels in bold are parasites of *G. roeselii* found in the present study. These labels show the name of the parasite, the country where it was found (two letter ISO code, see Additional file 1: Table S1), the number of infected populations (=pop.), and the total number of infected individuals (=ind.). Labels with accession numbers are parasite sequences taken from GenBank. These labels show the accession number, the parasite name given in the associated publication, the order of the host (except for amphipod hosts where the family is provided; if *G. roeselii* was found infected by a haplogroup, this was indicated by '*Groe*'). Microsporidian clade numbers are as in Vossbrinck and Debrunner-Vossbrinck [49]. Branches are collapsed according to the three genera (triangle sizes not reflecting actual size): *Nosema*, *Cucumispora* and *Dictyocoela*, as further details are given in Figs. 3, 4 and 5, respectively. *Abbreviation*: PP, Bayesian posterior probability

crustaceans (*Msp*-IVC, *Msp*-IVD) (Fig. 2). Among the parasites linked to other gammarid infections, *Msp*-IVA was found in France and Hungary (populations 37, 39 and 91, Fig. 1, Additional file 1: Table S1). It was very close to *Msp*-515 infecting Irish and French populations of *G. duebeni* (GenBank: FN434086) (97.7% identity; 311 bp coverage). The *Msp*-IVB haplogroup was found in four individuals in one French population (# 89, Additional file 1: Table S1). This sequence showed between 99.4–99.7% identity with *c.*750 bp coverage to a set of sequences including *Msp*-JES2002I (GenBank: AJ438964), previously detected in *G. pulex* from Scotland [6] and *Msp*-I (GenBank: KR871371), from *G. roeselii* in Germany [5, 26, 31]. *Msp*-IVE parasite was found in one individual in population 31 from Romania and was 99.4% similar (with 327 bp coverage) to *Orthosomella* sp. (GenBank: KT633994) infecting *G. fossarum* and the subterranean amphipod *Niphargus schellenbergi* in Luxembourg [16]. *Msp*-IVF and *Msp*-IVG haplogroups were found in one host individual each, in French populations 84 and 98 (Fig. 1, Additional file 1: Table S1), respectively. The highest match for these haplogroups in GenBank was *Msp*-RR2 (GenBank: KR871373) and microsporidians found in the Ruhr region of Germany, infecting *G. pulex*, *G. fossarum* and *G. roeselii* [26], with 99.7% and 78.6% identity, respectively (coverage 334 bp). Finally, the sequence of *Msp*-VA showed 100% identity (246 bp coverage) with *M*sp-RR1 (GenBank: KR871372) previously found by Grabner et al. [5] to infect *G. pulex* and *G. roeselii* in Germany. In the microsporidian Clade III, *Msp*-IIIA was found in two individuals from the Hungarian population 46 (Additional file 1: Table S1). This haplogroup is relatively closely related to *Pleistophora* parasites (Fig. 2), notably *Pleistophora mulleri* (GenBank: AJ438985) infecting *G. duebeni celticus* (88.6% identity, coverage 388 bp) [62], but was even more closely related to *Vavraia culicis* infecting mosquitoes (99.0% identity, coverage 384 bp) (GenBank: AJ252961) (Fig. 2). Finally, within the Clade V, *Msp*-VB was found in one G. *roeselii* from Germany, population 77 (Fig. 1, Additional file 1: Table S1). This parasite had only 75.7% identity with *Msp*-RR1 infecting gammarids and 80.3% identity with *Msp*-Group H infecting Ephemeroptera (Fig. 2). Most of identified parasites have been reported in the literature to infect gammarid hosts; however, they seem to be relatively rare and geographically widespread.

Two parasites of *G. roeselii* were related to other parasites infecting crustaceans. *Msp*-IVC was found in one *G. roeselii* individual from the Greek population 12 (Fig. 1, Additional file 1: Table S1). This sequence is identical (coverage 330 bp) to the sequence of *Enterocytospora artemiae* found in *Artemia franciscana* in France, USA and Israel [63] (GenBank: JX839889). Similarly, *Msp*-IVD, infecting three individuals from northern Italy, population 60, was 99.5% similar, with a coverage of 762 bp, to *Parahepatospora carcini* infecting, the European shore crab (*Carcinus maenas*) (GenBank: KX757849) [11].

The geographical distribution of these rare infections was similar between the geographical regions (Fig. 1, Additional file 1: Table S1), with 1/26 populations infected in Region 1 (3.8%; *Msp*-IVC only), 4/22 (18.2%) in Region 2a and 6/46 populations infected in Region 2 (13.0%; all other haplogroups) (Likelihood Ratio $\chi^2 = 2.89$, $P = 0.23$).

Infections ascribed to the genus *Nosema*

A total of 96 *G. roeselii* individuals were infected by microsporidia for which partial *SSU* rDNA sequences were ascribed to the genus *Nosema* (Additional file 1: Table S1). Three haplogroups of *Nosema* were identified in our study (Fig. 3), all belonging to *Nosema granulosis* (Fig. 3). Infections with *N. granulosis* represented 30.4% of all the microsporidian infections in *G. roeselii*. *Nosema granulosis 1* was the most frequent haplogroup, infecting 76 individuals from 19 populations across 8 countries (Fig. 3). The haplogroups *N. granulosis 1* and *2* showed, respectively, 100% and 99.8% identity with *N. granulosis* already found in *G. roeselii* from France (GenBank: AY584251) (Additional file 4: Table S3) [19]. The haplogroup *N. granulosis 3* was 100% identical with *N. granulosis* infecting the subterranean amphipod *Niphargus schellenbergi* but also *G. fossarum* individuals from Luxembourg and *G. pulex* from Poland (Additional file 4: Table S3) (GenBank: KP633991 and KM657357, respectively) [16]. This haplogroup had *c.* 98% identity with *N.*

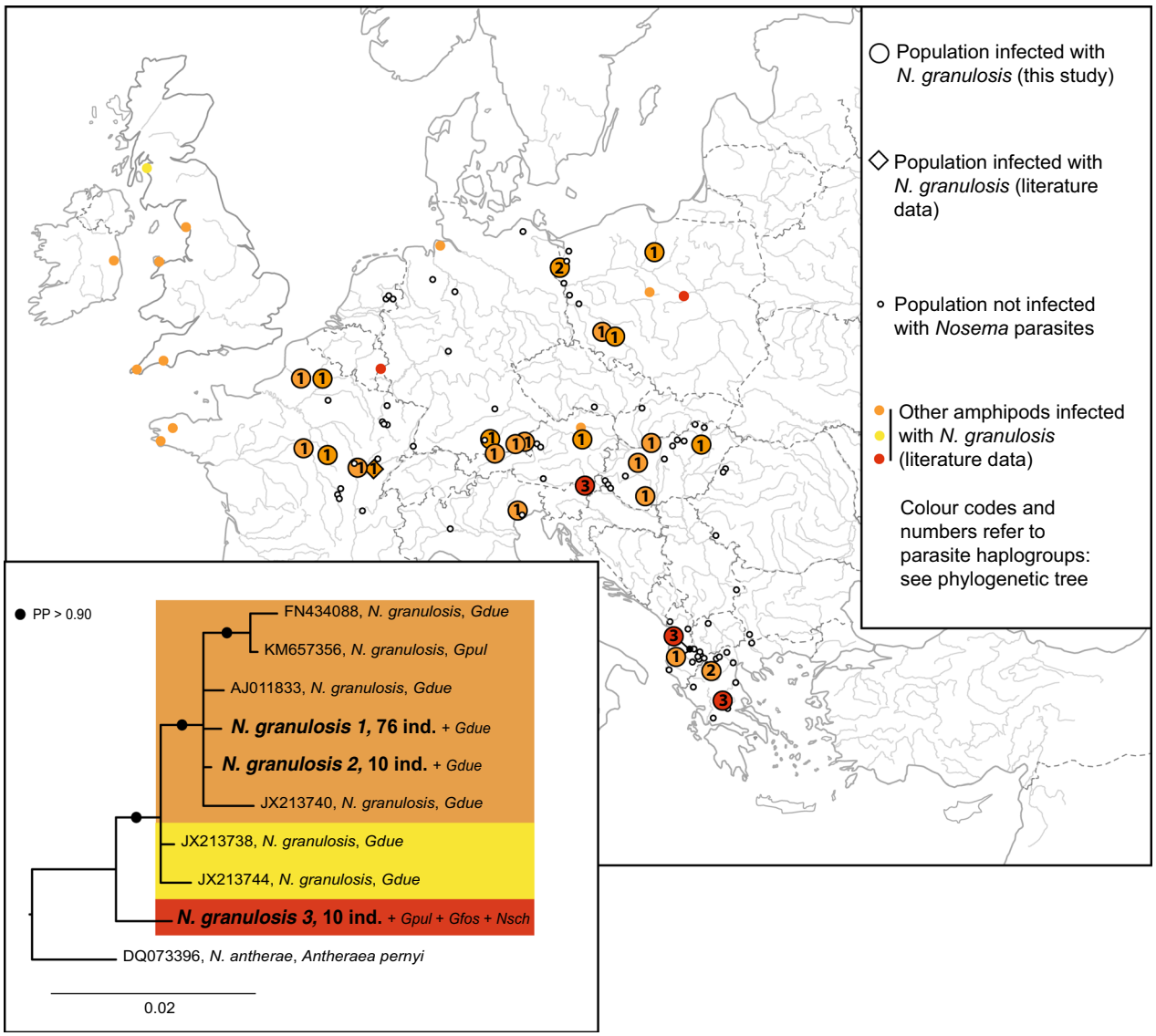


Fig. 3 Geographical distribution and phylogenetic tree for infections by microsporidia species of the genus *Nosema* in *Gammarus roeselii*. The map shows the infections found in *G. roeselii* (large colored dots for those detected in the present study; large colored diamonds for those detected in previous studies) and in other amphipods (small colored dots, see Additional file 4: Table S3 for further details). Dot colors match clade colors on the phylogenetic tree. The Bayesian phylogenetic reconstruction is based on a small ribosomal subunit rDNA alignment of *Nosema granulosis*. *Nosema antherae* was used as the outgroup and divergent lineages were ascribed a color code. Sequences from the present study are in bold and labels include haplogroup names, the total number of *G. roeselii* infected individuals (=ind.), plus other hosts found infected by the same haplogroup. Sequences from GenBank are all other *Nosema granulosis* haplogroups (Additional file 4: Table S3). Labels include the accession number, the species name given in the associated publication and the host species abbreviated names. For abbreviations of host species names see Additional file 4: Table S3. *Abbreviation*: PP, Bayesian posterior probability

granulosis haplogroups *1* and *2* and many sequences, e.g. with *N. granulosis* found in *G. duebeni*, the type-material used to describe this microsporidian species [8] (Additional file 4: Table S3). This parasite was found mostly in the *G. roeselii* diversification hotspot (Region 1), i.e. in Albania and Greece, but also in Slovenia (Region 2a), altogether in 10 infected individuals (Additional file 1: Table S1, Fig. 3).

These parasites were found in both the host diversification areas (Region 1 and Region 2a) and the recent expansion area (Region 2b). It is worth noting that the repartition of infections with *N. granulosis 1* was uneven among the host geographical regions. Only 1/26 populations (3.85%) were infected in Region 1, 5/22 (22.73%) in Region 2a and 13/46 (28.26%) in Region 2b (Likelihood ratio $\chi^2 = 7.79$, $P = 0.020$).

Infections ascribed to the genus *Cucumispora*

Three haplogroups close to the already described *Cucumispora* parasites, i.e. *C. ornata* [10], *C. dikerogammari* [9] and *C. roeselii* [20], were found in our study (Fig. 4). Indeed, one of the parasites, present in one host individual was 99.9% identical to the sequences of *C. dikerogammari* identified initially in another gammarid host, *Dikerogammarus villosus* [9, 17]. We named this haplogroup *C. dikerogammari 1* (Fig. 4, Additional file 4: Table S3). The second sequence was found in 11 individuals from 3 populations, showing 99.8% identity with *C. roeselii*, already found to infect *G. roeselii* in Poland [20]. We named this haplogroup *C. roeselii 1* (Fig. 4, Additional file 4: Table S3). Finally, one host individual was infected by a parasite showing 97.3% identity with *C. ornata* identified initially in *Dikerogammarus haemobaphes* [10]. We named this haplogroup *C. ornata 1* (Fig. 4). Although all newly generated sequences were informative enough to ascribe them to a species, 23 were too short (180–249 bp) to ascribe them to a specific haplogroup. This was the case for 11 sequences of *C. ornata*, that could belong both to *C. ornata 1* or to the KR871368 haplogroup (see Additional file 6: Data S3), and for 12 sequences of *C. dikerogammari*, that could belong both to *C. dikerogammari* 1 or to the GQ258752 haplogroup (Fig. 4, Additional file 4: Table S3). Altogether, the individuals infected with *Cucumispora* represented 11.7% of microsporidian infections found in *G. roeselii*. The three *Cucumispora* species were represented in almost equal proportions in our *G. roeselii* collection, with between 11–13 individuals infected by each species; however, *C. dikerogammari* was found only in one site in France (population 89).

Even if the overall low number of populations with individuals infected with *Cucumispora* prevented any statistical analyses, it is worth noting that species of the genus *Cucumispora* were found mostly in the recent expansion areas of *G. roeselii* (Regions 2a and 2b) (Fig. 4).

Infections ascribed to the genus *Dictyocoela*

Seven haplogroups of *G. roeselii* parasites were phylogenetically closely related to the following three species of *Dictyocoela* parasites: *D. roeselum* [19], *D. muelleri* [6] and *D. berillonum* [6] (Fig. 5, Additional file 7: Figure S1). We found also another *Dictyocoela* haplogroup that could not be assigned to any of the already described species (Fig. 5, Additional file 7: Figure S1). Overall, *Dictyocoela* spp. were the most common microsporidian parasites infecting *G. roeselii*, with 165 individuals infected in 27 populations, corresponding to 52.2 % of all microsporidian infections found.

Dictyocoela roeselum parasites were the most common and the most diverse: five haplogroups were found in 133 host individuals from 23 populations. Haplogroups *D. roeselum* 1 to 5 were 99.7% similar to the closest *D. roeselum* sequence (GenBank: MG773219, [24]). *Dictyocoela roeselum* 1, 3, 5 were found in western and northern Europe (Regions 2a and 2b), while haplogroups 2 and 4 were found in Region 2b only (Fig. 5, Additional file 1: Table S1). *Dictyocoela roeselum* infections in eight *G. roeselii* individuals were associated with short sequences (see "Methods"): we were unable to assert if they belong to *D. roeselum* 4 or 5 (7 individuals), or *D. roeselum* 1 or 5 (one individual) (Fig. 5, Additional file 1: Table S1 and Additional file 4: Table S3).

A single haplogroup (called *D. muelleri* 1) infecting *G. roeselii* showed 100% identity with the sequence used originally to identify *D. muelleri* (Terry et al. [6]; parasite found in *G. duebeni* from northern Europe). This unique haplogroup was relatively widespread in the recent expansion area of *G. roeselii* (Region 2b), infecting 28 individuals from 14 different populations in Germany, France and Poland. No *G. roeselii* infected with this parasite were found in the two diversification areas (Region 1 or 2a).

We found one population in Germany with four individuals infected with an undescribed *Dictyocoela* species. We called it *Dictyocoela* sp. *N4* in the absence of morphological data enabling full species description. This haplogroup was 97.7% similar to *Dictyocoela* sp. *N1* infecting *Echinogammarus ischnus* from Poland [24] (Fig. 5, Additional file 1: Table S1, Additional file 7: Figure S1).

Finally, in one individual from Poland, we found one haplogroup of the *D. berillonum* clade (Fig. 5). It was 99.8% similar to the type-sequence used in the first discrimination of *D. berillonum* (GenBank: AJ438957, [6]). However, this sequence was too short to be distinguished from many other closely related *D. berillonum* haplogroups (Additional file 4: Table S3, Additional file 7: Figure S1) and was individualised for the sake of clarity.

Parasite infections across *G. roeselii* phylogeny

The massive majority of microsporidian haplogroups and number of infected individuals were found in association with the *Gammarus roeselii* MOTU C (Fig. 6). Accordingly, because only a few genotypes of *G. roeselii* from this MOTU colonised north-western Europe ([41, 47]; see "Methods"), most infections were found in the host's recent expansion area, i.e. Region 2b (see above, Figs. 3, 4, 5, summarised in Fig. 6).

However, it is worth noting that *Nosema granulosis* parasites were more scattered across the *G. roeselii*

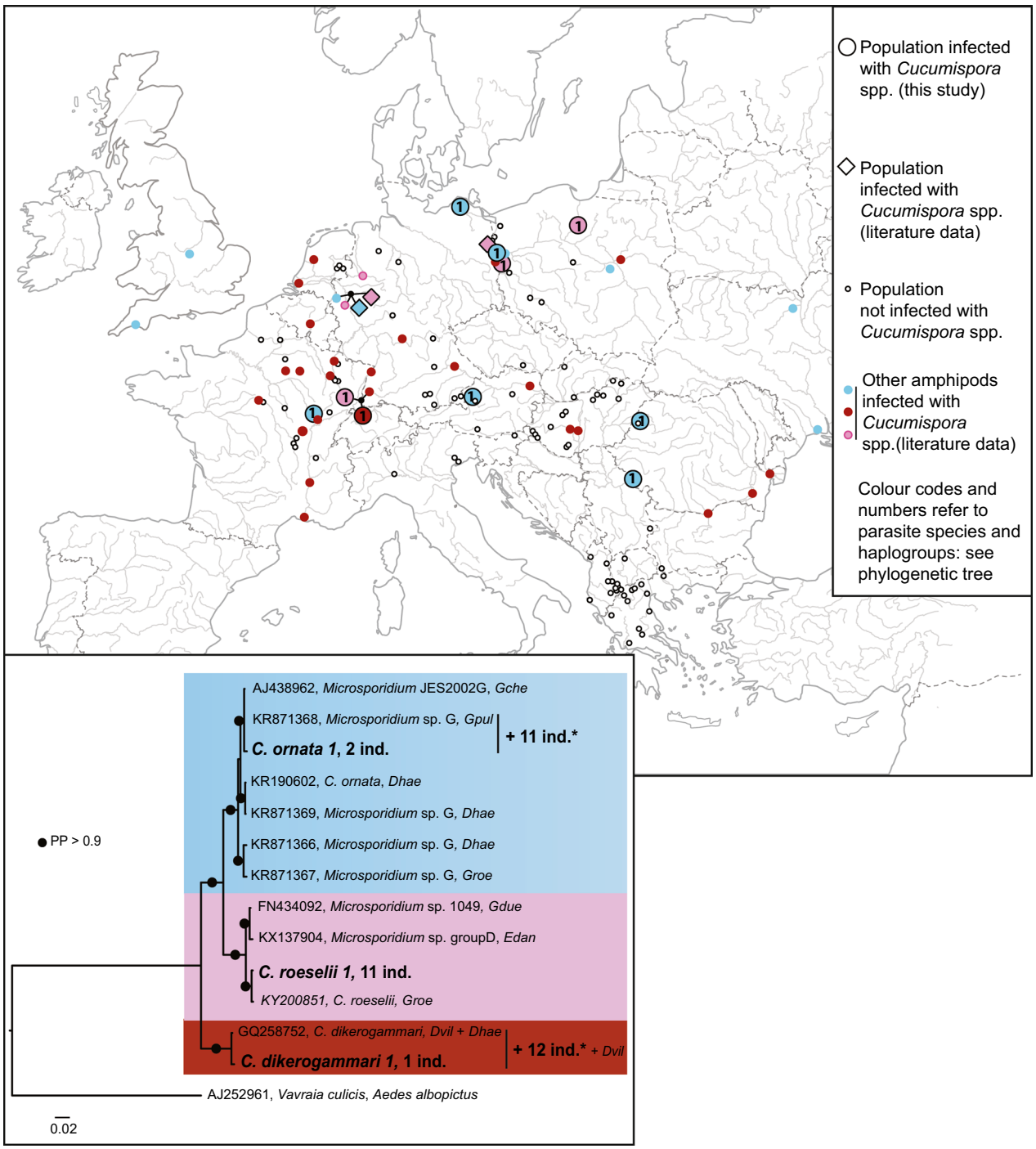
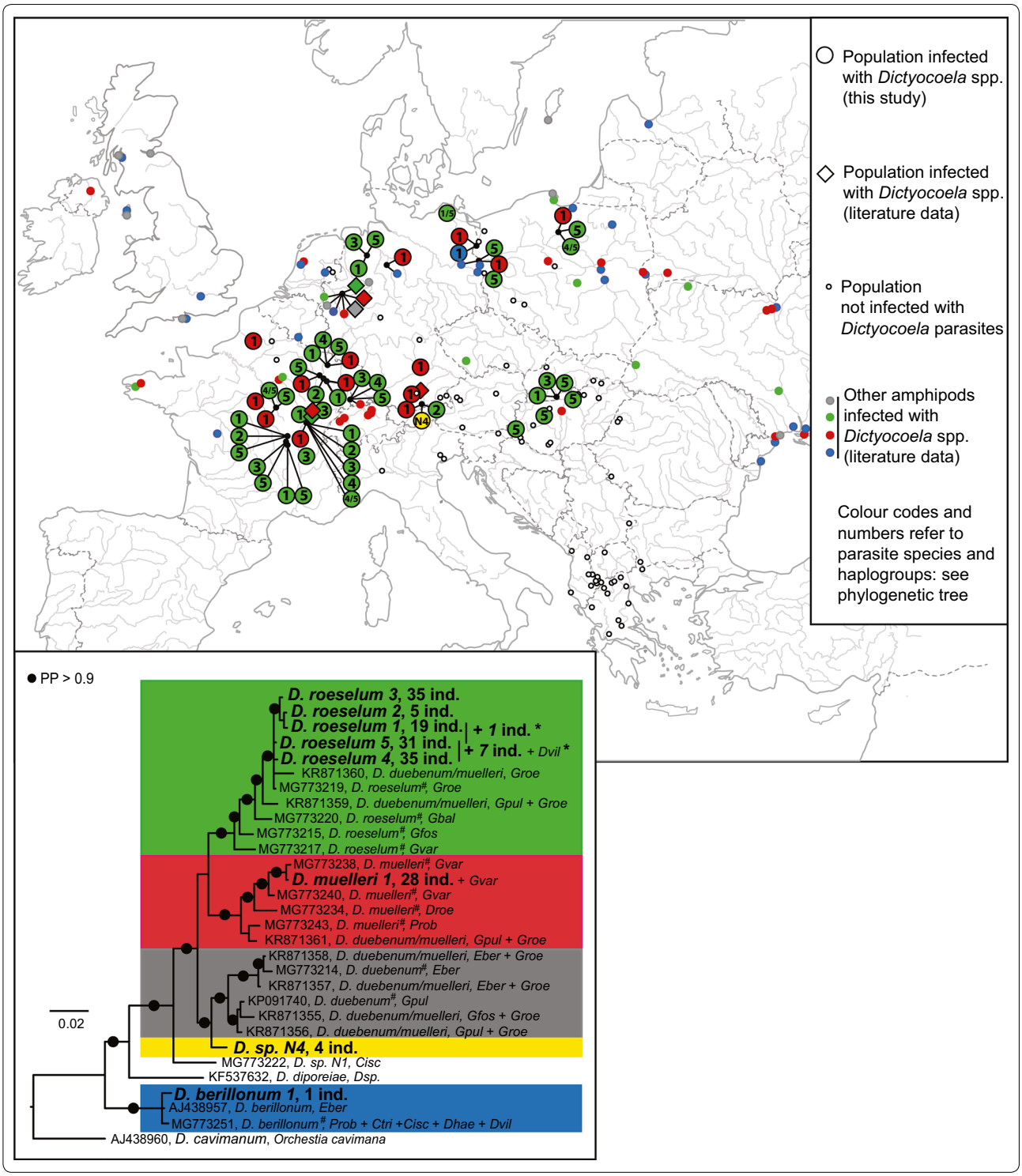


Fig. 4 Geographical distribution and phylogenetic tree for infections by microsporidia species of the genus *Cucumispora* in *Gammarus roeselii*. The map shows the infections found in *G. roeselii* (large colored dots for those detected in the present study; large colored diamonds for those detected in previous studies), and in other amphipods (small colored dots, see Additional file 4: Table S3 for further details). Dot colors match clade colors on the phylogenetic tree. The Bayesian phylogenetic reconstruction is based on a small ribosomal subunit rDNA alignment of *Cucumispora* spp. *Vavraia culicis* was used as the outgroup and divergent lineages were ascribed a color code. Sequences from the present study are in bold and labels include haplogroup names, the total number of *G. roeselii* infected individuals (=ind.), plus other hosts found infected by the same haplogroup. Sequences from GenBank represent all other *Cucumispora* haplogroups, following Bojko et al. [20] (Additional file 4: Table S3). Labels include, in this order, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated name(s). For abbreviations of host species names see Additional file 4: Table S3. *Parasites for which the sequence did not allow to distinguish their assignment between the haplogroups indicated by the vertical bar. *Abbreviation*: PP: Bayesian posterior probability



phylogeny than the other parasites. For example, *N. granulosis 2* and *3* were more frequently found in Region 1 than in Regions 2a and 2b. In Region 1, *N. granulosis 2* was associated with host MOTUs A and E, while *N. granulosis 3* was associated with MOTUs G and K. These host MOTUs, endemic to the Balkans, were deeply divergent from the widespread MOTU C. On the other hand, *N. granulosis 1* was found mostly in Regions 2a and 2b (infecting only the host MOTU C), while it was rare in Region 1, where it infected host MOTUs E and G (Fig. 6).

(See figure on previous page.)
Fig. 5 Geographical distribution and phylogenetic tree for infections by microsporidia species of the genus *Dictyocoela* in *Gammarus roeselii*. The map shows the infections found in *G. roeselii* (large colored dots for those detected in the present study; large colored diamonds for those detected in previous studies) and in other amphipods (small colored dots, see Additional file 4: Table S3 for further details). Dot colors match clade colors on the phylogenetic tree. The Bayesian phylogenetic reconstruction is based on a small ribosomal subunit rDNA alignment of *Dictyocoela spp*. *Dictyocoela cavimanum* was used as the outgroup and divergent lineages were ascribed a color code. Sequences from the present study are in bold and labels include haplogroup names, the total number of *G. roeselii* infected individuals (=ind.), plus other hosts found infected by the same haplogroup. Sequences from GenBank are only a representative panel of *Dictyocoela* diversity, divergence and host range (see Additional file 4: Table S3 and Additional file 3: Figure S1 for complete data). Labels include, in this order, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated name(s). For abbreviations of host species names see Additional file 4: Table S3. *Parasites for which the sequence did not allow to distinguish their assignment between the haplogroups indicated by the vertical bar. #Parasites used by Bacela-Spychalska et al. [24] to reassess species level phylogeny of *Dictyocoela* genus. *Abbreviation*: PP, Bayesian posterior probability

Most of the *Cucumispora* parasites were found in Region 2b, except for *C. ornata*, also found in two populations of Region 2a, all being therefore associated only with host MOTU C.

Dictyocoela roeselum 1, *3* and *5* were found mostly in Region 2b, but few host individuals belonging to four populations (#39, 42, 44, 49) were found infected by these haplogroups in Region 2a. Contrastingly, *D. roeselum 2* and *4* were found in Region 2b only. All were associated with host MOTU C. The other *Dictyocoela* spp., i.e. the frequent *D. muelleri* and the rare *Dictyocoela* sp. *N4* and *D. berillonum* were found only in Region 2b, also associated with the MOTU C.

The rare microsporidians were associated with host MOTU C in Regions 2a and 2b, except for *Msp*-IVC, associated with MOTU A in Region 1 (Fig. 6).

Discussion

Diversity of microsporidian infections in *Gammarus roeselii* across its entire range

Our study revealed that *Gammarus roeselii* is infected by several microsporidian taxa, represented by at least 24 haplogroups. Part of the newly generated sequences are relatively short, leading to possible underestimation of very closely related haplogroups that could be distinguished only with longer sequences. However, even if *SSU* rDNA is sometimes limited at resolving phylogenies in the Microsporidia at higher taxonomic levels, it has been recently shown that phylogenetic reconstruction based on this marker reflect accurately multimarker phylogenies at the generic level [24]. Thereby, even short sequences (V1f-UNIr) convey enough phylogenetic information for species-level divergence assessment. The 24 haplogroups could be clustered into 18 species-level taxa and, in part, ascribed to already described species (see below). Most of them were associated with Clades III and IV defined by Vossbrinck and Debrunner-Vossbrinck [55]: only two were associated with Clade V and none with Clade I and II.

Infections were found in around half of the investigated populations, ranging from low (prevalence of *c.*5%) to frequent (prevalence of 100%). However, low sampling at some sites may challenge prevalence estimates. More interestingly, the 18 species-level taxa could be classified into two broad categories related to infection patterns: species-level taxa with rare occurrence of infection (present in ≤ 3 populations) and species-level taxa with frequent occurrence (parasites present in more than 10 populations). The three clades with frequent occurrence belong to three microsporidian genera well identified to comprise amphipod-infecting parasites: *Nosema*, *Cucumispora* and *Dictyocoela* [6, 8–10, 24].

The rare infections in *G. roeselii* consisted of 10 parasite haplogroups. Although they represent a high diversity, their detection is challenging, because they represent a neglected parasitic fauna if sampling effort is limited [5, 18, 26, 31]. Nevertheless, most of them were phylogenetically closely related to parasites from other amphipods (sometimes sequences were identical), even sampled from different regions of Europe. This is the case for *Msp*-IVA, *Msp*-IVB, *Msp*-IVE, *Msp*-IVF and *Msp*-VA (Fig. 2). They were all infecting a single host (MOTU C), in the region recently colonised by *G. roeselii*. These parasites may represent generalists, infecting a variety of gammarids. These generalist parasites could have been acquired by *G. roeselii* by horizontal transfers while living in sympatry with local gammarid species (not investigated in the present study). Grabner et al. [5] proposed such a hypothesis for some of these parasites. However, we cannot dismiss a hypothesis that these parasites come from recent, transient, host shifts that will not sustain transmission within the new *G. roeselii* host. The rare infections *Msp*-IVG, *Msp*-IIIA and *Msp*-VB, were less closely related to other gammarid parasites (Fig. 2). They might represent rare cases of horizontal transfers from other, completely unstudied, aquatic taxa. The closest relatives of *Msp*-VB, for example, are parasites infecting aquatic insects (Fig. 2). The case of *Msp*-IIIA haplogroup

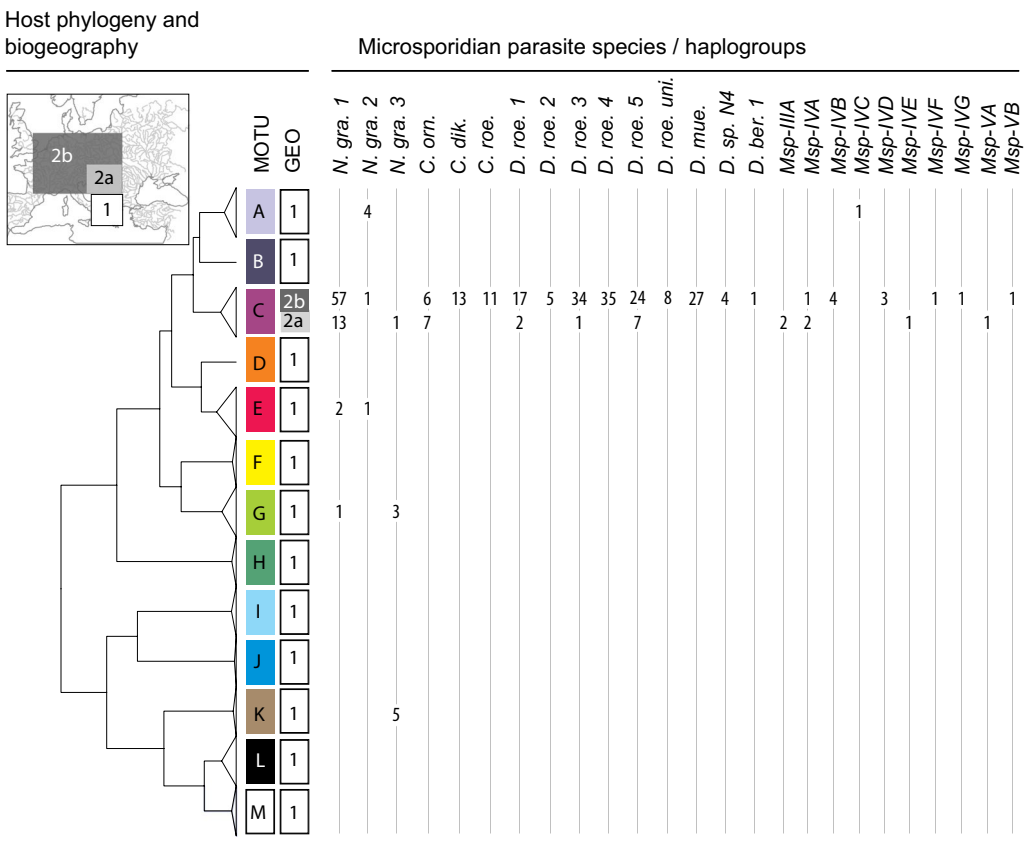


Fig. 6 Overview of microsporidian infections in *Gammarus roeselii* according to host phylogeny and biogeography. The number of infected individuals is provided for each microsporidian haplogroup or species. The host phylogeny, the names for Molecular Operational Taxonomic Units (MOTUs) represented by letters and the geographical distribution of these MOTUs (GEO, and small map) refers to Grabowski et al. [41, 47]

seems similar. It is very closely related to *Vavraia*, a genus recognised as *Pleistophora*-like microsporidia [64] associated with insects. This pattern is advocating for a transient association with *G. roeselii*, perhaps through the trophic chain, gammarids being both scavengers and predators of other members of the macroinvertebrate community [65].

The remaining microsporidian parasites detected as rare in *G. roeselii* were phylogenetically close to other crustacean microsporidians. *Msp*-IVC found at a site in continental Greece had the same sequence as *Enterocytospora artemiae* from *Artemia franciscana* present in the USA, France and Israel [63], and *Msp*-IVD found in *G. roeselii* form mainland Italy was very close to *Parahepatospora carcini* infecting *Carcinus maenas*, the European shore crab [11, 63]. How is it possible that these parasites from salt water can be found in freshwater animals? A direct horizontal transfer can be excluded, the two sites being too far from the seashore, preventing direct contact between *G. roeselii* and crabs or *Artemia*. However, these sites were located at around 100 km from Aegean or Ligurian seas, respectively (Fig. 1). We can first hypothesise a long-distance transport of microsporidian spores, e.g. by migratory shorebirds or birds associated with both marine and freshwater environments such as seagulls. These birds may have consumed crustaceans, released parasite spores with their faeces, which could have been consumed by *G. roeselii*. We cannot discriminate if these parasites are real infections or are food-borne microsporidia just passing through the gut. A second hypothesis could be that these microsporidia are generalists, infrequently found in various types of host over large geographical areas, both in fresh and saline waters.

Infections ascribed to the genus *Nosema*

Since the early work of Terry et al. [8], it is known that *Nosema granulosis* infects at least seven amphipod species. *Gammarus roeselii* is one of its hosts, since Haine et al. [19] found one haplogroup of the parasite in three French populations. This haplogroup was 100% similar to *N. granulosis 1* found in the present study. Our study extended this observation geographically, and we additionally detected two other haplogroups (*N. granulosis 2* and *3*) in Albania, Austria, Germany, Greece, Hungary,

Italy and Poland. *Nosema granulosis 1* is the most frequent haplogroup in our data. It is widely distributed in *G. roeselii*, but predominantly in Region 2b of the host range (Fig. 4). The main mode of transmission of *N. granulosis 1* in *G. roeselii* and *N. granulosis* in *G. duebeni* is the vertical one (from mother to offspring, *via* eggs). It induces sex ratio distortion in the host populations by reversing males into functional females [8, 19, 48, 66]. We may, therefore, suggest that this peculiar mode of transmission associated to overproduction of female hosts may have helped *N. granulosis 1* to colonise a large area, and even helped the host *G. roeselii* to rapidly spread in north-western Europe. This hypothesis was previously proposed to explain the invasive success of *Crangonyx pseudogracilis* colonising Europe [67]. Indeed, an excess of females generated by the infection may have increased the population dynamics of *G. roeselii*. Neither *N. granulosis 2* nor *N. granulosis 3* showed such a high frequency or large geographical distribution. Their vertical transmission and feminizing effect are therefore questionable, and more data are needed to reveal the effects they may have on the host phenotype.

All of the *N. granulosis* haplogroups infecting *G. roeselii* are scattered throughout the host phylogeny (Fig. 6), and this was the only parasite with such a pattern in our study. Notably, parasites of these haplogroups infect hosts present in the ancient diversification area (Region 1), as well as hosts present in the secondary diversification area (Region 2a). It is therefore tempting to propose that the infection by *N. granulosis* is ancient in *G. roeselii*, and that host–parasite co-diversification occurred after an initial ancestral infection. This hypothesis will be interesting to test in the future, by addressing the following issues. First, even if *SSU* rDNA is a useful marker for microsporidian phylogeny, it might not appear to be variable enough in this case for a detailed exploration of *N. granulosis* diversification history. Thus, additional markers such as RPB1 [68] could be useful to the dataset to step forward in addressing specificity. Secondly, the precise host diversification process in the secondary diversification area (Region 2a) remains to be explored in details [47]. It will notably be useful to understand if variation within the host MOTU C fits variation within *N. granulosis*. Thirdly, all *N. granulosis* haplogroups infecting *G. roeselii* are shared between this host and other amphipod hosts (*G. duebeni* for *N. granulosis 1* and *2*; *Niphargus schellenbergi*, *G. fossarum* and *G. pulex* for *N. granulosis 3*; see Fig. 5, Additional file 4: Table S3). It has been proposed that interspecific horizontal transmission may occasionally occur, which could be a survival strategy of *N. granulosis* in ephemeral habitats [69]. Horizontal transmission was also evoked to explain the presence of the same parasite haplogroup in subterranean and surface amphipods [16]. Our results are in agreement with such hypotheses. However, another hypothesis would be worth exploring, namely that infection with *N. granulosis* is very ancient in amphipods, and that these parasites co-diverged with the host species. Such a hypothesis assumes that *SSU* rDNA is not an appropriate marker for revealing such a pattern because the same haplogroups are shared by very divergent host species. Therefore, again, it requires employment of more variable markers. The multiplication of studies over an entire range of other gammarid species or studies testing vertical transmission [8, 19, 48, 66] would provide more opportunities to test these hypotheses.

Infections ascribed to the genus *Cucumispora*

We found three haplogroups within the genus *Cucumispora* [9]. Two of the sequences in our samples were almost identical to *C. dikerogammari* and *C. roeselii*. The third one was more distantly related to *C. ornata* (97.7% identity), but, following Bojko et al. [20], the similarity level suggests that it could be considered as belonging to *C. ornata*. *Cucumispora* are horizontally-transmitted parasites, virulent for their amphipod hosts. *Cucumispora dikerogammari* and *C. ornata* followed the hosts during their invasion of western Europe [10, 17]. They also have the potential to shift hosts and threaten local gammarid species [21, 70]. We hypothesise that the scattered infection pattern in *G. roeselii* with individuals infected by one clade of microsporidia may be due to interspecific horizontal transfers in sites where the original infected hosts live in sympatry with other gammarid species. This is strengthened by the fact that *Cucumispora* parasites are absent in Region 1 of *D. roeselii*, where the Ponto-Caspian hosts of *C. dikerogammari* and *C. ornata* are seldom present. The only infection with *C. dikerogammari* was identified where *G. roeselii* co-occurs with *D. villosus* (population 89) so it can be treated as the first evidence of a host shift observed in this parasite as was suggested by Bojko et al. and Bacela-Spychalska et al. [20, 21].

Infections ascribed to the genus *Dictyocoela*

With 27 populations and 165 individuals infected, *Dictyocoela* were the most abundant parasites found in our survey, confirming their status of dominant microsporidian infections in gammarids [24]. However they were found to infect only the host MOTU C in Regions 2a and 2b (Figs. 5, 6). We found eight haplogroups of *Dictyocoela*, corresponding to three previously identified species, i.e. *D. roeselum*, *D. muelleri*, *D. berillonum*, plus an unidentified one: *Dictyocoela* sp. *N4* (Fig. 5). Recently,

Bacela-Spychalska et al. [24] re-assessed *Dictyocoela* diversity based on the integrative approach combining phylogenetic, large geographical survey and ultrastructural data, and this paper provides an ideal backbone of our discussion.

Dictyocoela roeselum is the most abundant *Dictyocoela* in *G. roeselii* regarding both the total number of infected hosts and the number of populations with high prevalence. This result is analogous to previous studies, with prevalence reaching 60% in some local populations [19]. In our study, five haplogroups have been identified thus revealing this microsporidian to be the most diversified one in *G. roeselii*. This species has been described as infecting sporadically some other gammarid species such as *G. fossarum*, *G. varsoviensis*, *G. balcanicus* and *D. villosus* ([17, 24]; Additional file 4: Table S3, Additional file 7: Figure S1, Fig. 5). However, each host species is infected with a different, divergent, *D. roeselum* haplogroup [24]. Therefore, it seems that the particular *D. roeselum* haplogroups show some host specificity. *Dictycoela roeselum* was shown to be vertically transmitted in *G. roeselii* [19], strengthening this possibility, although the precise haplogroup could not be determined at that time. The exception could be the microsporidian found in single *D. villosus* individuals in two populations [17], which could be explained by the acquisition of the parasite by this predatory species, upon feeding on infected *G. roeselii*, as both the host species co-occurred in these sites. *Dictycoela roeselum* was absent in Region 1 of *G. roeselii*, but three haplogroups were present in Region 2a. This suggests that the infection by this parasite was as ancient as the secondary diversification of the host within the Pannonian basin (but younger than the primary diversification in the Balkans) and that these co-diversified parasites were carried during colonisation of the Region 2b. The presence of two supplementary haplogroups in this region suggests that parasite diversification is still ongoing in the region of recent invasion. This would assume a high nucleotide substitution rate in the parasite species (the colonisation of Region 2b was recent, probably post-glacial, see [47]), for which we have no information. Alternatively, underestimation of *D. roeselum* diversity in Region 2a, due to smaller sample size, cannot be dismissed.

Contrasting with *D. roeselum*, only a single haplogroup of *D. muelleri* was observed in our samples. It was nevertheless widely distributed, being present in 14 different populations from western and northern Europe (Region 2b). *Dictycoela muelleri* has been observed to infect numerous other gammarids (*D. haemobaphes*, *D. villosus*, *G. aequicauda*, *G. duebeni*, *G. varsoviensis* and *Pontogammarus robustoides*), sometimes at high prevalence [6, 17, 18, 24, 71]. In *G. roeselii*, this parasite was shown to use vertical transmission, and its role in sex ratio distortion was proposed [19]. Therefore, two parasites (*D. roeselum* and *D. muelleri*) with a similar life-cycle in a single host (vertical transmission) display different diversity patterns in *G. roeselii*. It is possible that *D. muelleri* only recently infected *G. roeselii* from local host species after it colonised the Region 2b, and, thanks to vertical transmission and sex ratio distortion in its host, rapidly spread throughout Europe. Alternatively, *D. roeselum* would represent a rather ancient infection for *G. roeselii*, which has co-diversified during the secondary diversification of its host.

We found one haplogroup of *D. berillonum*, highly similar to the haplogroup described from *Echinogammarus berilloni* [6]. This microsporidian species was previously found mainly in Ponto-Caspian hosts or other closely-related gammarids [5, 24, 25, 72]. Numerous *D. berillonum* infections result from successful co-invasion of the parasites alongside their host invasion [5, 72]. Moreover, the same *D. berillonum* haplogroups can infect several host species (Fig. 5, [24]). *Dictycoela berillonum*, therefore, does not show host-specificity, suggesting a high rate of horizontal transmission. The infected *G. roeselii* individual we found in Poland may, therefore, have acquired this parasite that way.

Finally, we encountered one haplogroup of uncertain phylogenetic proximity (*Dictyocoela* sp. *N4*). It was 97.7% similar to a set of sequences called *Dictyocoela* sp. *N1*, *2*, *3* [24] but also 97.4% similar to *D. duebenum* [24]. *Dictyocoela* sp. N4 is therefore a temporary name for the species, awaiting ultrastructural data for full species description. The three other *Dictycoela* sp. *N* previously described were only infecting the Ponto-Caspian *Echinogammarus ischnus* in Poland [73]. The single population where we found *Dictyocoela* sp. *N4* is situated in Germany, within the area where *E. ischnus* occurs in sympatry with *G. roeselii*. We may, therefore, infer that the presence of this parasite in the latter species originates from a recent (perhaps transient) host shift.

Comparison of microsporidian diversity in *G. roeselii* with other gammarid hosts

Our results can be compared with the only other published study concerning the whole Microsporidia community over the European geographical range of a single host species, i.e. *Gammarus duebeni* from the north-western Europe [15]. We found a lower proportion of infected populations in *G. roeselii* (51/94 in our study *vs* 32/35 in *G. duebeni*; Fisher's exact test, $P < 0.0001$) but a higher parasite haplogroup diversity (24 *vs* 11 haplogroups). This was found despite a higher numbers of populations sampled (94 *vs* 35) but a comparable sampling effort per population (20.2 ± 1.90 individuals/population

in our study *vs* 25.5 ± 4.25 in [18]; $F_{(1.132)} = 1.72$; $P = 0.19$). *Gammarus roeselii* and *G. duebeni* shared the same prevailing microsporidian genera: *Dictyocoela*, *Nosema* and *Cucumispora* (the latter group was not described yet at the time the results for *G. duebeni* were published, but the haplogroup *Msp*-1049 (GenBank: FN434092) falls within this clade. Some major differences can nevertheless be noted between the infection patterns in *G. roeselii* and *G. duebeni*. First, some infections that we qualified as "rare" in *G. roeselii* are more common in *G. duebeni*. This is the case of *Pleistophora-Vavraia*-like parasites, being present in seven populations of *G. duebeni* but only in one of *G. roeselii*. Similarly, but in larger proportions, *Msp*-505-515 were found in only three populations of *G. roeselii* (*Msp*-IVA in our study), but in 10 populations of *G. duebeni* [18]. Since parasites of this group have also been identified in other species of *Gammarus* [5, 31], it is tempting to suggest, following [5], that they are ubiquitous in *Gammarus* and may show a horizontal transmission pattern as well as low host specificity. Conversely, while relatively abundant in *G. roeselii*, *D. muelleri* was much rarer in *G. duebeni* [18]. Finally, the sharp disproportion in prevalence of *Dictyocoela* species between the two hosts is particularly interesting for understanding the evolution of microsporidians within the family Gammaridae. While *D. duebenum* was predominant in *G. duebeni*, we did not find it in *G. roeselii* in our study (however, in other studies, it was found in Germany [5], see Fig. 5); the reverse is true for *D. roeselum*. It, therefore, seems that, even if some haplogroups of these parasites were sporadically found in other gammarid species [24, 25]; these two parasites might show a certain amount of host specialisation. All these differences may be due to different co-evolutionary histories of the two host–pathogen associations (e.g. *G. roeselii* presents a higher cryptic species diversity compared to *G. duebeni* in Europe). The lack of overlap in the geographical distribution of the two host species may also prevent microsporidian host shifts, thereby leading to these contrasting patterns. Only an increasing number of studies similar to the present one and the one conducted by Krebes et al. [18], involving hosts with overlapping distributions, would allow to discriminate between these competing hypotheses.

Another study, although conducted on only a smaller part of the host geographical range, showed that the invasive Ponto-Caspian species *Dikerogammarus villosus* also harbours a substantial microsporidian parasite diversity [17]. In this species, the most abundant parasite was *Cucumispora dikerogammari* [9], the parasite being quite rare in *G. roeselii* (Fig. 4). Infections with *Nosema* and *Dictyocoela* parasites represented less than 1% in *D. villosus*, sharply contrasting with the prevalence observed in the present study (Figs. 3 and 5). It is worth noting that the geographical area studied for *D. villosus* partially overlaps with the geographical range of *G. roeselii* in central-western Europe and in the northern Balkans. This overlap is nevertheless recent, after the invasion of *D. villosus* in these areas in the last 30 years [74]. Therefore, it is probable that the differences in infection patterns reflect the different host–parasite evolutionary histories before the overlap of the geographical ranges of different hosts. The slight similarities could be due to high probabilities for interspecific transfers of parasites after this overlap [17, 22].

As noted by Pilosof et al. [75] and Wells et al. [76], the parasite assemblage of a given spreading host species often highly depends on the host–parasite network met by this species in a newly colonised area. Since *G. roeselii* is a species expanding its range, the comparison with the study of Grabner et al. [5] may help understand if *G. roeselii* shares parasitic fauna with local hosts. In Grabner et al. [5], *G. roeselii* was found to be infected with 11 microsporidian haplogroups representing four parasite species-level taxa: *D. duebenum/muelleri* complex, *Msp*-G, *Msp*-RR1, *Msp*-505. They shared these haplogroups with the local gammarid assemblages at the scale of the study (tributaries to the Rhine River, an area representing *c.*1000 km^2). It is nevertheless worth noting that some of the dominant species detected in our study were absent in *G. roeselii* in [5] (e.g. *Nosema*), indicating that it would be useful to investigate other local assemblages to compare precisely local and global parasite fauna of *G. roeselii*.

Conclusions

Microsporidian infections are common, diverse and widespread in *Gammarus roeselii* over its entire European geographical range. Two microsporidian species share infections between regions of host differentiation (Region 1 and/or 2a) and the recently colonised area (Region 2b): *Nosema granulosis* and *Dictyocoela roeselum*. For these two species, an evolutionary scenario of co-diversification with the host is a reasonable hypothesis. These patterns sharply contrasted with those of *Dictyocoela muelleri* and of the three species of *Cucumispora* parasites. In the latter, a single haplogroup per parasite species was found associated to many populations of Region 2b and therefore only to the host MOTU C. It seems parsimonious to explain these patterns as secondary acquisitions by host shifts from local gammarid species, after recent colonisation of this area by *G. roeselii*, rather than invoking an ancient infection and a secondary loss of parasites in the diversification area. Similar patterns of interspecific parasite transfers would also explain most of the rare infections, because they were close to parasites from other gammarid

species. Indeed, it is known that host shifts are more probable between phylogenetically related hosts than between unrelated hosts [75, 77]. Our data cannot distinguish between recent spillover events (i.e. transient infections) and sustained transmission within the new *G. roeselii* host, but we know from [23] that *D. muelleri* is vertically transmitted in this host and, therefore, is well established in *G. roeselii* populations.

Additional files

Additional file 1: Table S1. Microsporidian infections in the 94 populations investigated over the geographical range of *Gammarus roeselii*.

Additional file 2: Table S2. Variable sites in sequences of the three major microsporidian genera used to construct phylogenetic trees (Figs. 3, 4, 5).

Additional file 3: Data S1. Alignments based on *SSU* rDNA sequences, used for Additional file 2: Table S2.

Additional file 4: Table S3. Individual data for microsporidian infections from this study and found in GenBank (NCBI), mainly for freshwater and brackish waters amphipod species occurring in Europe.

Additional file 5: Data S2. Alignments based on *SSU* rDNA sequences, used for trees in Figs. 2, 3, 4, 5 and Additional file 1: Figure S1.

Additional file 6: Data S3. Sequences under 200 bp for which no GenBank number can be attributed.

Additional file 7: Figure S1. Phylogenetic tree for infections by microsporidians of the genus *Dictyocoela* in European freshwater amphipods, including all haplogroups. Bayesian phylogenetic reconstruction based on small ribosomal subunit rDNA alignment of *Dictyocoela spp*. *Dictyocoela cavimanum* was used as outgroup. Divergent lineages were ascribed to the same color code as Fig. 5, and follow recent reassessment of the genus taxonomy by Bacela-Spychalska et al. [24]. Sequences from the present study are in bold and labels include haplogroup names, and the hosts found infected by the haplogroup. Sequences from GenBank are all other *Dictyocoela* haplogroups (Additional file 4: Table S3). Labels include, in this order, the accession number, the microsporidia species name given in the associated publication and the species abbreviated name(s). For abbreviations of host species names: see Additional file 4: Table S3. Numbers on the branches indicates Bayesian posterior probability.

Abbreviations

SSU: small subunit ribosomal RNA gene; *Msp*: *Microsporidia* sp.; SNP: single nucleotide polymorphism; MOTU: molecular operational taxonomic unit; Mya: million years ago.

Acknowledgements

We would like to thank Tomasz Rewicz, Tomasz Mamos and Anthony Ollivier for their help in sampling, DNA extraction and PCR amplification.

Authors' contributions

AQ, RW, TR and KBS designed the experiment. KBS, MG, TR, MT and RW collected and fixed the organisms. AQ, MT, NL and RW performed molecular analyses. AQ and RW conducted the phylogenetic analyses. AQ, TR, RW, KBS and MG wrote the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

All the data are available from the laboratory of Biogeosciences, University Bourgogne Franche-Comté, Dijon, France and material is available at the Department of Invertebrate Zoology and Hydrobiology, University of Lodz, Poland, upon request. GenBank accession numbers of partial *SSU* rDNA sequences generated in this study are MK719236-MK719541.

Funding

The authors would like to acknowledge French national research agency (ANR) funding (Grant ANR-15-CE32-0006). The material was collected within the projects funded by the Polish Ministry of Science and Education: Nos. N303 579439 and N304 350139.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Université Bourgogne Franche-Comté, Laboratoire Biogéosciences, UMR CNRS 6282, 6 Boulevard Gabriel, 21000 Dijon, France. [2] Department of Invertebrate Zoology and Hydrobiology, University of Lodz, 12/16 Banacha Street, 90-237, Lodz, Poland.

Received: 14 January 2019 Accepted: 19 June 2019
Published online: 28 June 2019

References

1. Keeling PJ, Fast NM. Microsporidia: biology and evolution of highly reduced intracellular parasites. Annu Rev Microbiol. 2002;56:93–116.
2. Wittner M. Historic perspective on the microsporidia: expanding horizons. In: Wittner M, Weiss LM, editors. The microsporidia and microsporidiosis. Washington, DC: American Society for Microbiology Press; 1999. p. 1–6.
3. Dunn AM, Smith JE. Microsporidian life cycles and diversity: the relationship between virulence and transmission. Microbes Infect. 2001;3:381–8.
4. Ebert D. The epidemiology and evolution of symbionts with mixed-mode transmission. Annu Rev Ecol Evol Syst. 2013;44:623–43.
5. Grabner DS, Weigand AM, Leese F, Winking C, Hering D, Tollrian R, et al. Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. Parasit Vectors. 2015;8:419.
6. Terry RS, Smith JE, Sharpe RG, Rigaud T, Littlewood DTJ, Ironside JE, et al. Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum Microspora. Proc R Soc B Biol Sci. 2004;271:1783–9.
7. Stentiford GD, Feist SW, Stone DM, Bateman KS, Dunn AM. Microsporidia: diverse, dynamic, and emergent pathogens in aquatic systems. Trends Parasitol. 2013;29:567–78.
8. Terry RS, Smith JE, Bouchon D, Rigaud T, Duncanson P, Sharpe RG, et al. Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF) microsporidium. J Eukaryot Microbiol. 1999;46:492–9.
9. Ovcharenko MO, Bacela K, Wilkinson T, Ironside JE, Rigaud T, Wattier RA. *Cucumispora dikerogammari* n. gen. (Fungi: Microsporidia) infecting the invasive amphipod *Dikerogammarus villosus*: a potential emerging disease in European rivers. Parasitology. 2010;137:191–204.
10. Bojko J, Dunn AM, Stebbing PD, Ross SH, Kerr RC, Stentiford GD. *Cucumispora ornata* n. sp. (Fungi: Microsporidia) infecting invasive "demon shrimp" (*Dikerogammarus haemobaphes*) in the United Kingdom. J Invertebr Pathol. 2015;128:22–30.
11. Bojko J, Clark F, Bass D, Dunn AM, Stewart-Clark S, Stebbing PD, et al. *Parahepatospora carcini* n. gen., n. sp., a parasite of invasive *Carcinus maenas* with intermediate features of sporogony between the Enterocytozoon clade and other microsporidia. J Invertebr Pathol. 2017;143:124–34.
12. Nägeli N. Uber die neue krankheit der seidenraupe und verwandte organismen. Österreichische Bot Zeitschrift. 1857;15:760–1.
13. Terry RS, Smith JE, Dun AM. Impact of a novel, feminising microsporidium on its crustacean host. J Euk Microbiol. 1998;45:497–501.
14. Kelly A, Dunn AM, Hatcher MJ. Population dynamics of a vertically transmitted, parasitic sex ratio distorter and its amphipod host. Oikos. 2001;94:392–402.

15. Ironside JE, Dunn AM, Rollinson D, Smith JE. Association with host mitochondrial haplotypes suggests that feminizing microsporidia lack horizontal transmission. J Evol Biol. 2003;16:1077–83.
16. Weigand AM, Kremers J, Grabner DS. Shared microsporidian profiles between an obligate (*Niphargus*) and facultative subterranean amphipod population (*Gammarus*) at sympatry provide indications for underground transmission pathways. Limnologica. 2016;58:7–10.
17. Wattier RA, Haine ER, Beguet J, Martin G, Bollache L, Muskó IB, et al. No genetic bottleneck or associated microparasite loss in invasive populations of a freshwater amphipod. Oikos. 2007;116:1941–53.
18. Krebes L, Blank M, Frankowski J, Bastrop R. 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. 2010;10:1027–38.
19. Haine ER, Brondani E, Hume KD, Perrot-Minnot MJ, Gaillard M, Rigaud T. Coexistence of three microsporidia parasites in populations of the freshwater amphipod *Gammarus roeseli*: evidence for vertical transmission and positive effect on reproduction. Int J Parasitol. 2004;34:1137–46.
20. Bojko J, Bącela-Spychalska K, Stebbing PD, Dunn AM, Grabowski M, Rachalewski M, et al. Parasites, pathogens and commensals in the "low-impact" non-native amphipod host *Gammarus roeselii*. Parasit Vectors. 2017;10:193.
21. Bacela-Spychalska K, Wattier RA, Genton C, Rigaud T. Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. Biol Invasions. 2012;14:1831–42.
22. Bacela-Spychalska K, Rigaud T, Wattier RA. A co-invasive microsporidian parasite that reduces the predatory behaviour of its host *Dikerogammarus villosus* (Crustacea, Amphipoda). Parasitology. 2014;141:254–8.
23. Iltis C, Spataro T, Wattier R, Médoc V. Parasitism may alter functional response comparisons: a case study on the killer shrimp *Dikerogammarus villosus* and two non-invasive gammarids. Biol Invasions. 2018;20:619–32.
24. Bacela-Spychalska K, Wróblewski P, Mamos T, Grabowski M, Rigaud T, Wattier R, et al. Europe-wide reassessment of *Dictyocoela* (Microsporidia) infecting native and invasive amphipods (Crustacea): molecular versus ultrastructural traits. Sci Rep. 2018;8:8945.
25. Wilkinson TJ, Rock J, Whiteley NM, Ovcharenko MO, Ironside JE. Genetic diversity of the feminising microsporidian parasite *Dictyocoela*: new insights into host-specificity, sex and phylogeography. Int J Parasitol. 2011;41:959–66.
26. Grabner DS, Schertzinger G, Sures B. Effect of multiple microsporidian infections and temperature stress on the heat shock protein 70 (hsp70) response of the amphipod *Gammarus pulex*. Parasit Vectors. 2014;7:170.
27. Ironside JE, Alexander J. Microsporidian parasites feminise hosts without paramyxean co-infection: support for convergent evolution of parasitic feminisation. Int J Parasitol. 2015;45:427–33.
28. Madyarova EV, Adelshin RV, Dimova MD, Axenov-Gribanov DV, Lubyaga YA, Timofeyev MA. Microsporidian parasites found in the hemolymph of four Baikalian endemic amphipods. PLoS ONE. 2015;10:1–7.
29. Slothouber Galbreath JGM, Smith JE, Becnel JJ, Butlin RK, Dunn AM. Reduction in post-invasion genetic diversity in *Crangonyx pseudogracilis* (Amphipoda: Crustacea): a genetic bottleneck or the work of hitchhiking vertically transmitted microparasites? Biol Invasions. 2009;12:191–209.
30. Ironside JE, Wilkinson TJ, Rock J. Distribution and host range of the microsporidian *Pleistophora mulleri*. J Eukaryot Microbiol. 2008;55:355–62.
31. Grabner DS. Hidden diversity: parasites of stream arthropods. Freshw Biol. 2017;62:52–64.
32. Ironside JE, Smith JE, Hatcher MJ, Sharpe RG, Rollinson D, Dunn AM. Two species of feminizing microsporidian parasite coexist in populations of *Gammarus duebeni*. J Evol Biol. 2003;16:467–73.
33. Gismondi E, Rigaud T, Beisel JN, Cossu-Leguille C. Effect of multiple parasitic infections on the tolerance to pollutant contamination. PLoS ONE. 2012;7:e41950.
34. Dimova M, Madyarova E, Gurkov A, Drozdova P, Lubyaga Y, Kondrateva E, et al. Genetic diversity of Microsporidia in the circulatory system of endemic amphipods from different locations and depths of ancient Lake Baikal. PeerJ. 2018;6:e5329.
35. Ironside JE, Wilkinson TJ. Accumulation and exchange of parasites during adaptive radiation in an ancient lakering adaptive radiation in an ancient lake. Int J Parasitol. 2018;48:297–307.
36. Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol Evol. 2005;20:238–44.
37. Meyran JC, Monnerot M, Taberlet P. Taxonomic status and phylogenetic relationships of some species of the genus *Gammarus* (Crustacea, Amphipoda) deduced from mitochondrial DNA sequences. Mol Phylogenet Evol. 1997;8:1–10.
38. Witt JDS, Hebert PDN. Cryptic species diversity and evolution in the amphipod genus *Hyalella* within central glaciated North America: a molecular phylogenetic approach. Can J Fish Aquat Sci. 2000;57:687–98.
39. Lagrue C, Wattier R, Galipaud M, Gauthey Z, Rullmann JP, Dubreuil C, et al. Confrontation of cryptic diversity and mate discrimination within *Gammarus pulex* and *Gammarus fossarum* species complexes. Freshw Biol. 2014;59:2555–70.
40. Mamos T, Wattier R, Burzyński A, Grabowski M. The legacy of a vanished sea: a high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. Mol Ecol. 2016;25:795–810.
41. Grabowski M, Mamos T, Bącela-Spychalska K, Rewicz T, Wattier RA. Neogene paleogeography provides context for understanding the origin and spatial distribution of cryptic diversity in a widespread Balkan freshwater amphipod. PeerJ. 2017;5:e3016.
42. Hogg ID, Larose C, Lafontaine YD, Doe KG. Genetic evidence for a *Hyalella* species complex within the Great Lakes—St. Lawrence River drainage basin: implications for ecotoxicology and conservation biology. Can J Zool. 1998;76:1134–52.
43. Lefébure T, Douady CJ, Gouy M, Gibert J. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Mol Phylogenet Evol. 2006;40:435–47.
44. Sutherland DL, Hogg ID, Waas JR. Phylogeography and species discrimination in the *Paracalliope fluviatilis* species complex (Crustacea: Amphipoda): can morphologically similar heterospecifics identify compatible mates? Biol J Linn Soc. 2010;99:196–205.
45. Westram AMM, Baumgartner C, Keller I, Jokela J. Are cryptic host species also cryptic to parasites? Host specificity and geographical distribution of acanthocephalan parasites infecting freshwater *Gammarus*. Infect Genet Evol. 2011;11:1083–90.
46. Galipaud M, Bollache L, Lagrue C. Variations in infection levels and parasite-induced mortality among sympatric cryptic lineages of native amphipods and a congeneric invasive species: are native hosts always losing? Int J Parasitol Parasites Wildl. 2017;6:439–47.
47. Grabowski M, Krzywozniak P, Rewicz T, Mamos T, Bacela-Spychalska K, Wattier R. *Gammarus roeselii* Gervais, 1835 (Gammaridae) in western and central Europe: post-glacial colonisation or human mediated introduction? Biodivers J. 2017;8:525–6.
48. Haine ER, Motreuil S, Rigaud T. Infection by a vertically-transmitted microsporidian parasite is associated with a female-biased sex ratio and survival advantage in the amphipod *Gammarus roeseli*. Parasitology. 2007;134:1363–7.
49. QGIS Development Team. QGIS geographic information system. Open Source Geospatial Foundation Project; 2016. https://www.qgis.org/en/site/.
50. Karaman GS, Pinkster S. Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea-Amphipoda). Part I. *Gammarus pulex*-group and related species. Bijdr Dierkd. 1977;47:1–86.
51. Karaman GS, Pinkster S. Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea-Amphipoda). Part II. *Gammarus roeseli*-group and related species. Bijdr Dierkd. 1977;47:165–96.
52. Weiss LM, Zhu X, Cali A, Tanowitz HB, Wittner M. Utility of microsporidian rRNA in diagnosis and phylogeny: a review. Folia Parasitol (Praha). 1994;41:81–90.
53. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.
54. Madden T. The BLAST sequence analysis tool. NCBI handbook; 2003. p. 1–15. https://blast.ncbi.nlm.nih.gov/Blast.cgi.
55. Vossbrinck CR, Debrunner-Vossbrinck BA. Molecular phylogeny of the Microsporidia: ecological, ultrastructural and taxonomic considerations. Folia Parasitol (Praha). 2005;52:131–42.

56. Williams BAP, Hamilton KM, Jones MD, Bass D. Group-specific environmental sequencing reveals high levels of ecological heterogeneity across the microsporidian radiation. Environ Microbiol Rep. 2018;10:328–36.
57. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66.
58. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001;17:754–5.
59. Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9:772.
60. Berbee ML, James TY, Strullu-Derrien C. Early diverging Fungi: diversity and impact at the dawn of terrestrial life. Annu Rev Microbiol. 2017;71:41–60.
61. Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R. Jane: a new tool for the cophylogeny reconstruction problem. Algorithms Mol Biol. 2010;5:16.
62. Terry RS, Macneil C, Dick JTA, Smith JE, Dunn AM. Resolution of a taxonomic conundrum: an ultrastructural and molecular description of the life cycle of *Pleistophora mulleri* (Pfeiffer 1895; Georgevitch 1929). J Eukaryot Microbiol. 2003;50:266–73.
63. Rode NO, Lievens EJP, Segard A, Flaven E, Jabbour-Zahab R, Lenormand T. Cryptic microsporidian parasites differentially affect invasive and native *Artemia* spp. Int J Parasitol. 2013;43:795–803.
64. Cheney SA, Lafranchi-Tristem NJ, Bourges D, Canning EU. Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). J Eukaryot Microbiol. 2001;48:111–7.
65. Macneil C, Dick JT, Elwood RW. The trophic ecology of freshwater *Gammarus* spp. (Crustacea: Amphipoda): problems and perspectives concerning the functional feeding group concept. Biol Rev. 1997;72:349–64.
66. Jahnke M, Smith JE, Dubuffet A, Dunn AM. Effects of feminizing microsporidia on the masculinizing function of the androgenic gland in *Gammarus duebeni*. J Invertebr Pathol. 2013;112:146–51.
67. Johanna GM, Galbreath S, Smith JE, Terry RS, Becnel JJ, Dunn AM. Invasion success of *Fibrillanosema crangonycis*, n. sp., n. g.: a novel vertically transmitted microsporidian parasite from the invasive amphipod host *Crangonyx pseudogracilis*. Int J Parasitol. 2004;34:235–44.
68. Ironside JE. Diversity and recombination of dispersed ribosomal DNA and protein coding genes in Microsporidia. PLoS ONE. 2013;8:e55878.
69. Krebes L, Zeidler L, Frankowski J, Bastrop R. (Cryptic) sex in the microsporidian *Nosema granulosis*—evidence from parasite rDNA and host mitochondrial DNA. Infect Genet Evol. 2014;21:259–68.
70. Bojko J, Stentiford G, Stebbing P, Hassall C, Deacon A, Cargill B, et al. Pathogens of *Dikerogammarus haemobaphes* regulate host activity and survival, but also threaten native amphipod populations in the UK. Dis Aquat Organ. 2018. https://doi.org/10.3354/dao03195.
71. Arundell K, Dunn A, Alexander J, Shearman R, Archer N, Ironside JE. Enemy release and genetic founder effects in invasive killer shrimp populations of Great Britain. Biol Invasions. 2015;17:1439–51.
72. Green Etxabe A, Short S, Flood T, Johns T, Ford AT. Pronounced and prevalent intersexuality does not impede the 'Demon Shrimp' invasion. PeerJ. 2015;3:e757.
73. Grabowski M, Jazdzewski K, Konopacka A. Alien crustacea in polish waters—Amphipoda. Aquat Invasions. 2007;2:25–38.
74. Rewicz T, Grabowski M, MacNeil C, Bącela-Spychalska K. The profile of a 'perfect' invader the case of killer shrimp, *Dikerogammarus villosus*. Aquat Invasions. 2014;9:267–88.
75. Pilosof S, Morand S, Krasnov BR, Nunn CL. Potential parasite transmission in multi-host networks based on parasite sharing. PLoS ONE. 2015;10:e0117909.
76. Wells K, O'Hara RB, Morand S, Lessard JP, Ribas A. The importance of parasite geography and spillover effects for global patterns of host–parasite associations in two invasive species. Divers Distrib. 2015;21:477–86.
77. Davies TJ, Pedersen AB. Phylogeny and geography predict pathogen community similarity in wild primates and humans. Proc R Soc B Biol Sci. 2008;275:1695–701.

Publisher's Note

CHAPTER II. EXPLORING HOST-PARASITE ASSOCIATION BETWEEN *GAMMARUS BALCANICUS* SPECIES COMPLEX AND THE DIVERSITY OF ITS MICROSPORIDIAN PARASITES.

II.1 BACKGROUND

Microsporidia are ubiquitous parasites, infecting a wide range of hosts inhabiting various habitats (Keeling & Fast 2002). These eukaryotes are obligate unicellular endoparasites belonging to an extremely ancient and phylogenetically diverse phylum, with more than 1300 species in 160 genera (Wittner 2014). Microsporidia are now revealed to be either fungi (Corradi & Selman 2013; Hibbett *et al.* 2007) or their sister group (Capella-Gutiérrez *et al.* 2012). They exhibit different transmission strategies and virulence impact on host: i) horizontal transmission (HT) between conspecifics or between host species is linked mostly to high virulence. ii) vertical transmission (VT) from mother to eggs and offspring, is usually associated with low or supposedly no virulence, or iii) a combination of both VT and HT (Dunn & Smith 2001; Ebert 2013). Microsporidia are infecting numerous vertebrates, including humans, but they are also responsible for many diseases of insect, fish and crustaceans (Becnel & Andreadis 2014; Becnel & Takvorian 2014; Bulnheim 1975; Stentiford *et al.* 2016; Weber *et al.* 1994). Among aquatic arthropods, the freshwater amphipod crustaceans, and especially Gammaridae, are common hosts for microsporidia (for overviews see Grabner 2017; Stentiford *et al.* 2013; Weiss & Becnel 2015). The first descriptions of microsporidia infecting amphipods were from the 1930s, and then the late 1960s – early 1970s (references in Terry *et al.* (1999). Since that time, literature continually increases, with many papers recording new genera and species (Bacela-Spychalska *et al.* 2018; Bojko *et al.* 2015, 2017b; Dimova *et al.* 2018; Grabner 2017; Grabner *et al.* 2015; Haine *et al.* 2004; Ironside *et al.* 2003, 2008; Ironside 2007; Ironside & Alexander 2015; Ironside & Wilkinson 2018; Krebes *et al.* 2010, 2014; Ovcharenko *et al.* 2010; Quiles *et al.* 2019; Slothouber Galbreath *et al.* 2009; Terry *et al.* 1999; Weigand *et al.* 2016; Wilkinson *et al.* 2011; Winters & Faisal 2014; Yang *et al.* 2011). Three main microsporidian genera are commonly infecting freshwater amphipods: *Nosema* (Naegeli 1857), *Cucumispora* (Ovcharenko *et al.* 2010) and *Dictyocoela* (Terry *et al.* 2004). Those genera are known to infect many host species across Eurasia (Dimova *et al.* 2018; Grabner 2017; Hogg *et al.* 2002), but also in North America (Slothouber Galbreath *et al.* 2009; Winters *et al.* 2014). Some studies demonstrated several microsporidia species to be vertically transmitted (they infect oocytes, and are therefore present in most embryos): *Nosema granulosis*, *Dictyocoela roeselum*, *D. duebenum* and *D. muelleri* (Dubuffet *et al.* 2013; Haine *et al.* 2004; Terry *et al.* 1999). These species induce low virulence and induce sex-ratio distortion in their host populations. Indeed, the putative male host receiving the vertically transmitted parasites are reversed into functional females (Dunn & Smith 2001; Ironside & Alexander 2015; Kelly *et al.* 2002). This characteristic induces an unusual and diagnostic bias in prevalence: because these microsporidian parasites are both vertically-transmitted to most oocytes and feminize most of the hosts they infect, they induce

high prevalence in females, but are rare (if not absent) in males (Haine *et al.* 2004; Terry *et al.* 1999, 2004), while also inducing sex ratio bias in favour of females. However, it is to be noticed that some strains of *Dictyocoela duebenum* are not feminizing the host species (Ironside & Alexander 2015), indicating possible variation in selection of this trait according to the host species. *Nosema granulosis* and *Dictyocoela* spp. infect a wide range of amphipod species, particularly *Gammarus* spp. (see references cited above), but it is still not clear if each host species harbour specific parasite strains, or if they are infected by generalist lineages, or if both scenario co-exist.

In fact, surveys conducted over full host geographic ranges, *e.g*. for *Gammarus duebeni* (Krebes *et al.* 2010) and *G. roeselii*, (Quiles *et al.* 2019), concluded that some hosts share generalists microsporidian strains (probably acquired horizontally after host spill-over, even for parasite species renowned for being vertically-transmitted as *Nosema*), but some strains appeared to be host-specific. Other gammarid microsporidia seem mainly horizontally transmitted after the ingestion of infected tissues or spores (Bacela-Spychalska *et al.* 2012; Ovcharenko *et al.* 2010), and induce pathology in their hosts by infecting and replacing muscle tissues (Bojko *et al.* 2015, 2017a; Ovcharenko *et al.* 2010). This is the case of parasites belonging to the *Cucumispora* genus. Some species of *Dictyocoela* also show signs of the same type of pathology, *e.g. D. diporae* infecting an American amphipod species (Winters *et al.* 2014).

In addition to these three main genera infecting amphipods, several of microsporidian lineages were found based on molecular surveys. Most of them were observed occasionally, thanks to population screenings in amphipods. Those rare microsporidia are not yet fully described since niether anatomical nor ultrastructural descriptions are available (Grabner 2017; Grabner *et al.* 2015; Ironside *et al.* 2008; Krebes *et al.* 2010; Quiles *et al.* 2019; Terry *et al.* 2004; Wilkinson *et al.* 2011).

This study aimed to investigate the pattern of microsporidian infection in the *Gammarus balcanicus* species complex. This host was chosen as a biological model for its biogeography and diversification pattern has recently been solved (Mamos *et al.* 2014, 2016), and because, contrary to some other *Gammarus* species, its microsporidian associated fauna is virtually unexplored. *Gammarus balcanicus* populations are inhabiting mountainous areas from the eastern Carpathians through the Balkan Peninsula, to the eastern Alps (Copilaş-Ciocianu & Petrusek 2017; Mamos *et al.* 2014), although its localities are known also from the Black Sea lowlands and from Crimea. *Gammarus balcanicus* is a morphospecies characterized by high cryptic diversity, including at least 50 divergent lineages or MOTUs (Molecular Operational Taxonomic Units) of Miocene origin (Mamos *et al.* 2016). Due to its unstable geological history over the last 20 million years,

the Balkan Peninsula and the Carpathian arch have been characterized by vastly dynamic landscape remodelling, resulting in high geographical complexity (Popov *et al.* 2004). Several episodes of inland water colonization from local brackish water were observed. Present day *Gammarus balcanicus* MOTUs are locally endemic due to their habitat fragmentation and complex phylogeographical history (Copilaş-Ciocianu & Petrusek 2017; Mamos *et al.* 2014, 2016). In a caricatural way, we could say that each site harbour a single MOTU of *G. balcanicus*. *Gammarus balcanicus* species complex starts its diversification from *c.* 20 Ma, in the early Miocene in the central Balkans, partially in the shallow epicontinental sea present at that time (Mamos *et al.* 2016). This early diversification generated two major clades: the AR clade, nowadays endemic to a small area in the Rhodope Mountains in the central Balkan Peninsula, and a clade which later (*c.* 15 Ma) split into the north-eastern clade (hereafter N) and the south-western clade (Hereafter S) (Fig. 1). Subsequent diversification occurred within the north-eastern and south-western clades of *G. balcanicus*. Therefore, the present distribution of the fifty *G. balcanicus* MOTUs, presented in (Fig. 1), reflects the Miocene palaeogeography of the area (Mamos *et al.* 2016). Furthermore, diversification and geographic expansion continued following the Alpine orogeny during Miocene/Pliocene and, finally, during the Pleistocene glaciations (Fig. 1) (Mamos *et al.* 2016).

The Carpathians and the Balkans are recognized as most valuable present-day hot-spots of biodiversity and endemism, and a model system for studies upon biogeography and the evolution of numerous organisms (Blondel & Blondel 2010; Médail & Diadema 2009; Poulakakis *et al.* 2015). Notably, it is an ancient centre of diversity for some freshwater gammarid amphipods (Copilaş-Ciocianu & Petrusek 2017; Grabowski *et al.* 2017b; Mamos *et al.* 2016). Despite these valuable features, only a single paper have extensively studied host-parasite interaction, in the Balkans, on *G. roeselii* species complex (Quiles *et al.* 2019). *Gammarus roeselii* is not restricted to the Balkans, its center of old diversitication (Grabowski *et al.* 2017a), but also recently extends its range to northern and western Europe (Grabowski *et al.* 2017b). Microsporidian diversity in *G. roeselii* was high with not less than twenty species‐level taxa (Quiles *et al.* 2019). Ten microsporidian species were rare, infecting a few individual hosts in a few populations, and were mainly related to parasites from other amphipods or crustaceans, being also in most cases rare in these hosts. The main microsporidians found were widespread genera with high prevalence: *Nosema*, *Cucumispora* and *Dictyocoela*. Two contrasting host association patterns stood in *G. roeselii.* First, vertically transmitted microsporidian species such as *Nosema granulosis* and *Dictyocoela roeselum*, which share the pattern of infecting *G. roeselii* over most of its range, and are specific to this host suggesting the co‐diversification scenario. Second, horizontally transmitted microsporidia such as *Dictyocoela muelleri*, the three species of *Cucumispora*, and the

rare parasites, present only in latest colonised region of the host suggesting a recent acquisitions from local host species (Quiles *et al.* 2019). These two scenario could be also present in *G. balcanicus* as far as microsporidia are a very ancient group; some horizontally transmitted microsporidia could probably have already infect *Gammarus roeselii* and *G. balcanicus* prior to their speciation and diversification.

Actually, only scarce studies identified microsporidia parasites in G. balcanicus (Bacela-Spychalska *et al.* 2018; Bojko & Ovcharenko 2018). The microsporidia found were *Octospora* sp. (Bojko & Ovcharenko 2018), *Thelohania muelleri* (=*Dictyocoela muelleri* (Terry *et al.* 2004) and *Dictyocoela roeselum* (Accession: MG773218; MG773220; MG773221) (Bacela-Spychalska *et al.* 2018), observed in the Carpathian populations. The phylogenetic reconstruction of the genus *Dictycoela* assessed by Bacela-Spychalska *et al*. (2018) presents several lineages within the *D. roeselum* part of the tree, where *G. balcanicus* and *G. roeselii* are closely related (Bacela-Spychalska *et al.* 2018). *Dictyocoela roeselum* is known to be vertically transmitted microsporidium and was suspected of co-diversification within *G. roeselii* (Quiles *et al.* 2019). Following, this early comparison, *D. roeselum* species seems to show some host specificity.

The objectives of this study are to explore microsporidian infections and to analyse the parasite diversity in the host *G. balcanicus*, over its geographic range and through its extremely divergent and numerous host MOTUs, by addressing several questions.

First, we clarified the infection pattern and specificity of the detected microsporidia, by exploring if microsporidia lineages infecting *G. balcanicus*, or also found in other Gammaridae. Special attention was given to comparing parasites of *G. balcanicus* with those previously encountered in *G. roeselii,* especially as the present ranges of the two species overlap in the Balkans (Quiles *et al.* 2019).

Second, we looked for phylogeographic information in the ecological infection patterns, leading to possible host-parasite evolutionary scenarios. Parasite clades could be present only in restricted geographic areas, following the phylogeographic diversification pattern of the host (*i.e.* numerous endemic entities). Such a finding would mean possible diversification and speciation matching that of the host (suggesting potential host-parasite co-diversifications), a pattern that could be strengthened by vertical transmission of parasites. Parasite clades could be found infecting the whole geographic range of the host. Such a pattern would mean a very ancient host-parasite association, prior to the early diversification of the host.

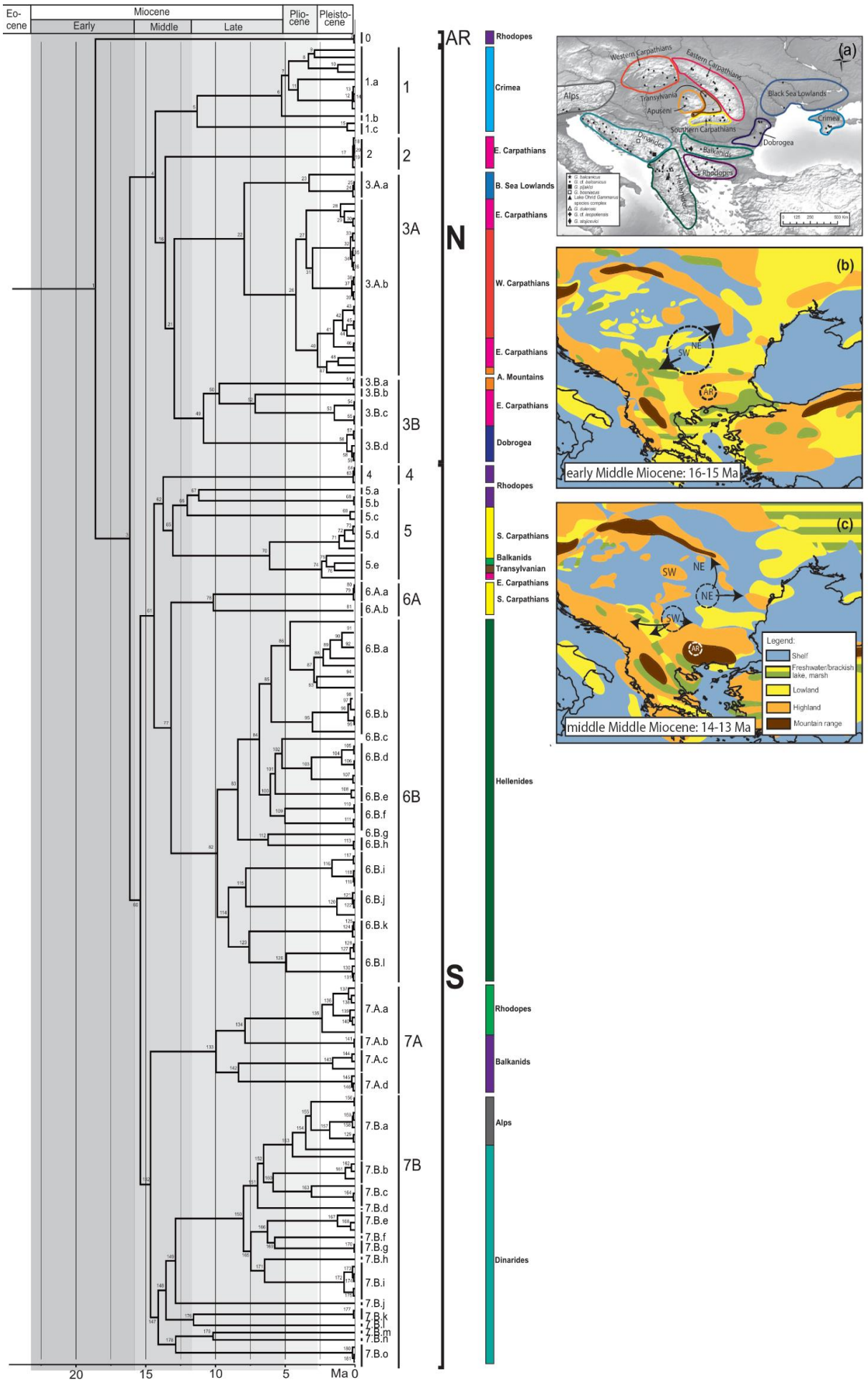

Eo-cene
Miocene
Early
Middle
Late
Plio-cene
Pleisto-cene
20
15
10
5
Ma 0
AR
N
S
1
2
3A
3B
4
5
6A
6B
7A
7B
Rhodopes
Crimea
E. Carpathians
B. Sea Lowlands
E. Carpathians
W. Carpathians
E. Carpathians
A. Mountains
E. Carpathians
Dobrogea
Rhodopes
S. Carpathians
Balkanids
Transylvanian
E. Carpathians
S. Carpathians
Hellenides
Rhodopes
Balkanids
Alps
Dinarides
(a)
Western Carpathians
Eastern Carpathians
Alps
Transylvania
Apuseni
Black Sea Lowlands
Crimea
Southern Carpathians
Dobrogea
Dinarides
Balkanids
Rhodopes
Hellenides
(b)
NE
SW
AR
early Middle Miocene: 16-15 Ma
(c)
SW
NE
NE
SW
AR
middle Middle Miocene: 14-13 Ma
Legend:
Shelf
Freshwater/brackish lake, marsh
Lowland
Highland
Mountain range

Fig. 1. Maximum clade credibility chronogram generated using Bayesian inference and based on the full multimarker data from Fig. S4 of Mamos *et al.* (2016). Grouping of lineages: 3 major clades: AR = Ancient Rhodopean, N = North-eastern, S = South-western. N and S were further subdivided into four (1, 2, 3A and 3B) and six (4, 5, 6A, 6B, 7A, 7B) clades respectively. Extra small letter (a, b, c...) represent part of the post middle Miocene diversification). Bar colours refer to the delimitation of physiographical regions on (a). (a) Sampling localities with delimitation of main physiographical regions. Circles and arrows indicate, respectively, putative geographical positions of common ancestors and possible routes of colonization projected on palaeomaps (redrawn after Popov *et al.* 2004) showing key geological events: (b) early Middle Miocene – 16–15 Ma. (c) middle Middle Miocene 14–13 Ma.

Third, since nothing is known about dynamic of microsporidian infections in *G. balcanicus*, we explored a hypothetical bias of prevalence between male and female hosts. Following Terry *et al.* (2004), finding a high prevalence in females and almost no infection in males could be a first step toward the evidence of vertical transmission associated with a feminizing effect.

II.2 METHODS

II.2.1 Sampling

Gammarus balcanicus were collected during several sampling campaigns between 2004 and 2016. Individuals were caught using hand nets and kick-sampling method, at 88 sites in 13 countries, covering the full distribution of *G. balcanicus* in Europe *i.e.* the Balkan peninsula (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Kosovo, Macedonia, Montenegro) extending to the north-eastern part of Italy, the Carpathian Arch (Hungary, Romania, Serbia, Slovakia, Ukraine) but also to the Black Sea lowlands Crimea and Turkey (Table S1). Sites were plotted on a map (Fig. 2) using Qgis 2.18.4 (QGIS Development Team 2009). All individuals were immediately fixed in 96% ethanol at the sampling site and stored at room temperature after returning to the laboratory. Amphipods were identified to the species level using morphological characters described in available keys (*e.g.* Karaman & Pinkster 1977a, 1977b).

Samples used in the present study correspond to the *G. balcanicus* species complex samples as sed in Mamos *et al.* (2016). All the specimens were stored at the Department of Invertebrate Zoology and Hydrobiology, University of Lodz, Poland.

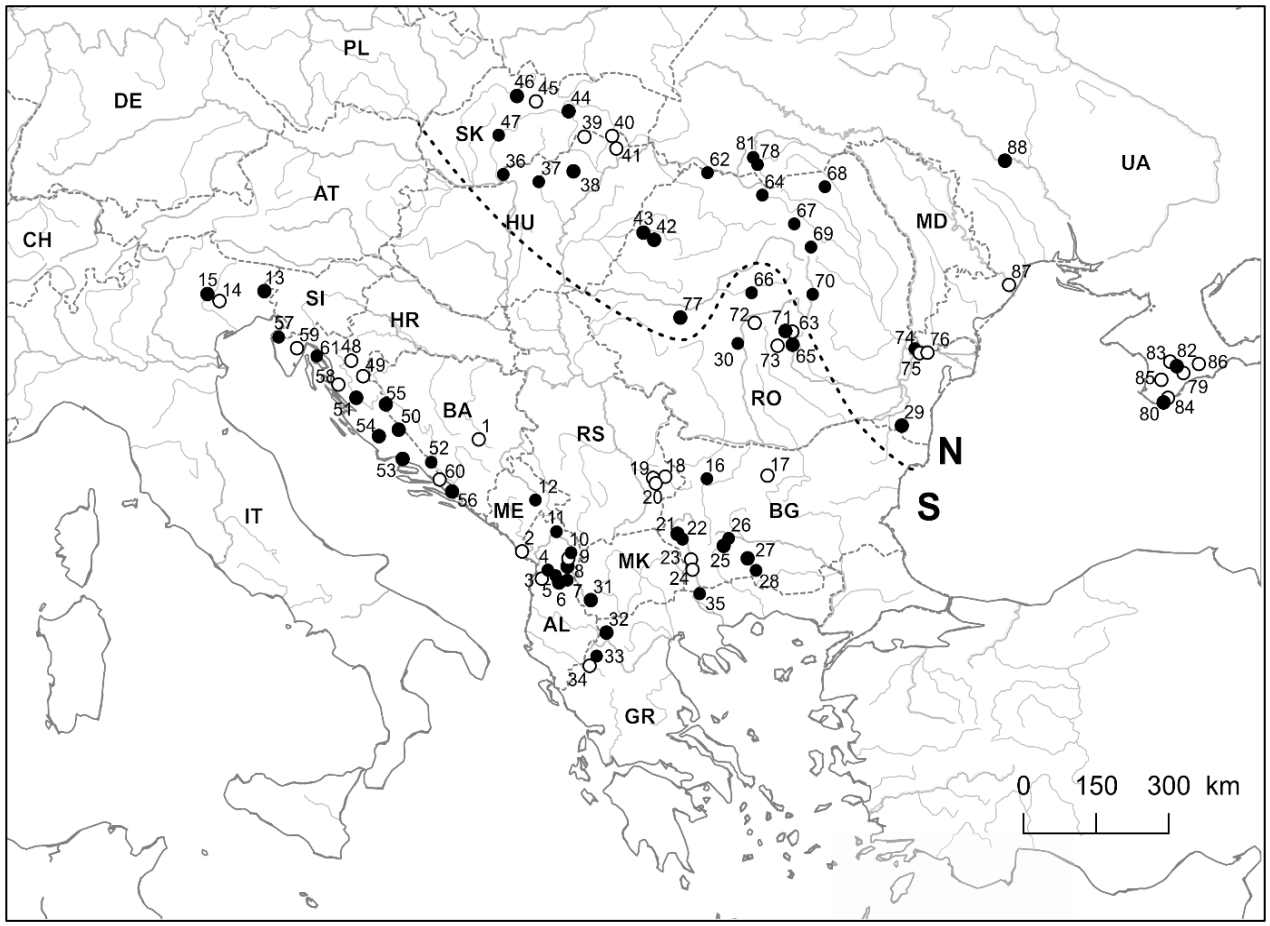


Fig. 2. *Gammarus balcanicus* sampling sites. White dots refer to sites with no microsporidian infection. Black dots refer to sites where at least one infection was found in a *G. balcanicus* individual and for which a sequence was obtained. Sites are designed by numbers (1-88), Table S1 for details (*e.g.* sampling sizes, GPS coordinates). The two letter codes refer to ISO code of countries (Table S1). The dashed line delimits the geographic distribution of the so called N and S clade (see text for details).

II.2.2 Host dissection and total DNA extraction

Each gammarid was dissected under binoculars taking *c.* 2mm^3 of host tissue (including muscles and gonads), from the thoracic segments 6 and 7. For microsporidia are intracellular parasites, their DNA was co-extracted with host DNA. Among the 2255 individuals used in the present study, DNA of 1202 male individuals was already extracted within the study by Mamos *et al.* (2016). The material was supplemented with 1053 newly extracted females (up to 24 per population, when possible) in order to assess prevalence according to sex (Table S1). Altogether, Region N and S accounted, respectively, 907 and 1348 individuals (Fig. 2, Table S1 & S2). The DNA extraction was performed using either (i) standard phenol-chloroform protocol or (ii) Biobasic EZ-10 96 Well Plate Genomic DNA Isolation Kit for Animal Sample and eluted in 100

µl of TE (pH 8). The DNA samples were kept at 4 °C until amplification and subsequently at -20 °C for long-term storage.

II.2.3 Molecular screening for microsporidians

All 2255 individuals were screened for the presence of microsporidia following the strategy by Quiles *et al.* (2019), using a short (*c.* 350 bp long) diagnostic fragment of the small ribosomal subunit (*SSU* rDNA) marker. The microsporidia-specific primer V1f (forward) (5'-CAC CAG GTT GAT TCT GCC TGA C-3') (Weiss *et al.* 1994) paired with UNIr (reverse) (5'-TCA GGC TCC CTC TCC GGA AT-3') (Quiles *et al.* 2019) was used. The use of this short fragment maximised the ability to detect the presence of microsporidians even in case of low infection intensity. As negative and positive controls in PCR reactions, we used, respectively, water and microsporidian DNA (*Dictyocoela roeselum*). The PCR conditions and visualization of PCR products were as described in Quiles *et al.* (2019).

For individuals with positively diagnosed microsporidian infections (see PCR screening above) our objective, as in Quiles *et al.* (2019), was to sequence the *c*. 800 bp long fragment of the *SSU* rRNA gene matching the 5' part either as one or two overlapping fragment. When the 800 bp long was not attainable we used, either a V1f-530r fragment (*c*. 530 bp long) or even a V1f-UNIr fragment (*c*. 350bp long), that contained enough phylogenetic information to attribute sequences to the species level without any ambiguity (results, Table S2 and Data S1). PCR products were purified and sequenced directly with the BigDye technology by Genewiz, Inc., UK, using the forward primers from PCR. Using Geneious 10.2.2 (Kearse *et al.* 2012), raw sequences were edited, trimmed and confirm for being microsporidian sequences using BlastN (Madden 2003) search on available sequences from GenBank.

II.2.4 Phylogeny reconstruction for microsporidians

Our dataset is composed of four types of microsporidian sequences: (i) newly produced sequences from the infected *G. balcanicus* individuals; (ii) literature *SSU* sequences representing diversity and divergence of microsporidians found to infect European freshwater or brackish water amphipods (sequences outside Europe were not included, *e.g.* the published parasites from Lake Baikal (Dimova *et al.* 2018; Ironside & Wilkinson 2018), except when being the only close relative available); (iii) literature *SSU* sequences of microsporidians infecting other taxa, prioritising freshwater or brackish water invertebrates, when newly produced sequences where not related to any of those found in amphipods; (iv) published *SSU* sequences representative of the five microsporidian clades

(Clades I-V), as determined in the latest integrative phylogenies of microsporidia (Vossbrinck *et al.* 2014; Williams *et al.* 2018). Sequences were aligned using MAFFT7.388 software (Katoh 2002; Katoh & Standley 2013), with the E-IONS-I algorithm using the legacy gap penalty option, incorporated in Geneious 10.2.2 (Kearse *et al.* 2012). Our dataset contains sequences of different lengths depending on both the newly produced sequences (from 75 to 817 bp) and on various length of the published ones (from 140 to 1448 bp for microsporidians, and 1786 bp for the fungus *Basidiobolus ranarum* used as an outgroup in Fig. 3). All details, including sequence length, are given in Table S2. As some sequences were relatively short (*c*. 5% below 200pb), reducing the full dataset to a standard size would, on the one hand, allow defining haplotypes but, on the other hand, would potentially induce losing phylogenetic signal. Therefore, following the strategy described by Quiles *et al.* (2019) we attributed each sequences to a haplogroup. Haplogroup is defined in such a way that sequences belonging to distinct haplogroups harboured at least one or more variable sites, generating diagnostic features (Table S3, Data S2), no matter the sequence length. In other words, sequences were gather in one haplogroup, despite variable length, based on 100% pairwise identity. Few newly produced sequences could not be assigned to only one haplogroup due to reduced length of the sequences and lack of diagnostic features. The longest sequence of each haplogroup was used for the phylogeny reconstruction (304 to 816 bp, noted in Table S2).

Phylogenetic reconstructions were build using only Bayesian methodology with MrBayes (Huelsenbeck & Ronquist 2001) integrated in Geneious 10.2.2. The best-fitting model of nucleotide substitution was determined with JModelTest-2.1.10. (Darriba *et al.* 2012). This was always the General Time Reversible (GTR) model with gamma-distributed rate heterogeneity (G) and a significant proportion of invariable sites (I). Four heated chains, each 1,100,000 iterations long, sampled every 200 iterations, were run. The runs reached satisfactory effective sampling sizes (ESS > 200), and the potential scale reduction factor values equalled 1 for all parameters. The 50% majority-rule consensus tree was constructed after the removal of 10% 'burn-in' trees. Six Bayesian phylogenetic trees were constructed. The first tree contained all haplogroups (*i.e.* sequences from this study and published sequences) using *Basidiobolus ranarum* (GenBank: AY635841) as the outgroup (Vossbrinck *et al.* 2014) (Fig. 3). In this tree, we described novel parasites by conservatively using provisional names, *e.g.* Microsporidia sp. (hereafter abbreviated Msp) followed by the clade number (from I to V) sensu Vossbrinck *et al.* (2014) and a superscript Roman letter. The five other phylogenies represent detailed analyses for the already identified

parasites of the microsporidian genera infecting amphipods: *Nosema* (Naegeli 1857) (Fig. 5) and *Dictyocoela* (Terry *et al.* 2004) (Fig.5 - 8). *Nosema antheraea* (GenBank: DQ073396) (Fig. 5), *Dictyocoela cavimanum* (GenBank: AJ438960) (Fig. 6) and *Dictyocoela sp.* N1 (GenBank: MG773222) (Fig. 8) were used as outgroups for the *Nosema* and *Dictyocoela* spp. phylogenies, respectively. Sequence for which a full taxonomic description was available from literature data were used to provide genus and species name to the new sequences. Such name was ascribed, if the genetic distance between the newly obtained sequence and the type sequence of already described of the species was below 2% pairwise identity, following Grabner *et al.* (2015).

II.2.5 Geographical distribution of parasites

Along with Bayesian reconstructed trees, we provided maps (Qgis 2.18.4) to display the geographical distribution of each haplogroup but also sequences of gammarids' parasites found in the literature for which sampling locations were available (Fig. 5, 6 – 8, Table S2). The only exceptions, were the very rare microsporidian clades, due to only few individuals and populations infected.

II.2.6 Parasite infection according to host phylogeny and distribution *versus* microsporidia prevalence and phylogeny

The host phylogeny reconstructed by Mamos *et al.* (2016) using multimarker haplotypes (cytochrome oxidase subunit I, 16S ribosomal RNA and 28S ribosomal RNA), was used as the backbone for *G. balcanicus* phylogeography in the present study. We used from Mamos *et al.* (2016) all the DNA previously extracted for which all individuals were males associated to a defined MOTU in the host phylogeny rebuild in this article. The newly extracted females were not sequenced for host clade attribution. We used information from males at a given sampling site from *Mamos et al*. (2016) to assign females to a single MOTU. The sites where two host MOTU were present is limited to three (1, 54 and 64), with only one (64) harbouring two infected females being the only case showing ambiguity. The host tree presented in Fig. 9-10 is a simplified presentation of the phylogeny reconstructed by Mamos *et al.* (2016) (as presented in Fig. 1), meaning that the diversity potentially present within the 50 MOTUs was collapsed to triangles, which sizes do not reflect actual diversity and intra-MOTU divergence. Investigation on host-parasite co-phylogenies is challenging due to the different evolutionary histories of host and parasite. First, the age of diversification of *G. balcanicus* started *c.* 19 Ma (Mamos *et al.* 2016) while that of the phylum Microsporidia is likely to be hundreds of Ma old (Berbee *et al.* 2017). Microsporidian genus level (*e.g. Nosema* and *Dictyocoela*) might be a better choice than the phylum to run such an investigation. However, the number of microsporidian clades at this

taxonomic level was limited relative to the high number of MOTUs observed in *G. balcanicus*. Furthermore, some of these host MOTU have not been tested due to the lack of individuals from this specific clade (6.A.a, 6.B.c, 6.B.e, 6.B.f, 6.B.g, 7.A.d). For these reasons, we were not able to use methods for analysing co-phylogeny (*e.g.* Conow *et al.* 2010); the relation were made by eye, trying to resolve how the parasite haplogroups were linked to host MOTUs.

II.2.7 Parasite infections according to host sex

The proportion of infected host individuals according to sex was assessed in three ways. First, we compared the proportion of infected males vs. females at the microsporidia species level using Fisher's exact test. Second, we compared this proportion between populations for the same microsporidia species using Likelihood-Ratio test. Third, we compared this proportion within populations, for different parasite haplogroups.

II.3 RESULTS

II.3.1 The overall prevalence and broad geographic distribution of microsporidian infections in *G. balcanicus*

The overall prevalence of Microsporidian infection in *G. balcanicus* was 10.91% with 245 infected individuals out of 2255 individuals tested. Fifty-six sites out of 88 sites (63.63%) contained at least one *G. balcanicus* individual infected with a microsporidian parasite (Fig. 2, Table S1). Among sites with infections, a wide range of prevalence was nevertheless observed, ranging from 2.08% to up to 83.33%, being 19.88% on average (Table S1). Sites with infection appeared to be spread all over the study area (Fig. 2).

II.3.2 Microsporidia overall diversity and phylogenetic position

As sequencing success was variable, sequences length for the 245 new partial microsporidia SSU ranged from 75 bp to 817 bp, with only 25.61% of the sequences bellow 300 bp in size (Table S2). These sequences clustered in 55 microsporidian haplogroups, which themselves clustered in 17 species-level-taxa, as based on a divergence threshold of *c*. 2% (Table S1-2). Among the 5 clades (named I-V) defined in the microsporidian phylogeny of Vossbrinck and Debrunner-Vossbrinck (2005) theses sequences fitted with 3 clades (III, IV and V) (Fig. 3). Most of the new sequences (215/245 – 87.4%), representing 35 haplogroups could be ascribed to two genera: *Nosema* (76/245 – 30.9%) representing 4 haplogroups and *Dictyocoela* (139/245 – 56.5%) representing 31 haplogroups (Table S1-2).

The remaining 30 sequences (12.60%), representing 20 haplogroups (36.4%), could be clustered in 12 species-level taxa and represent the “rare” infections.

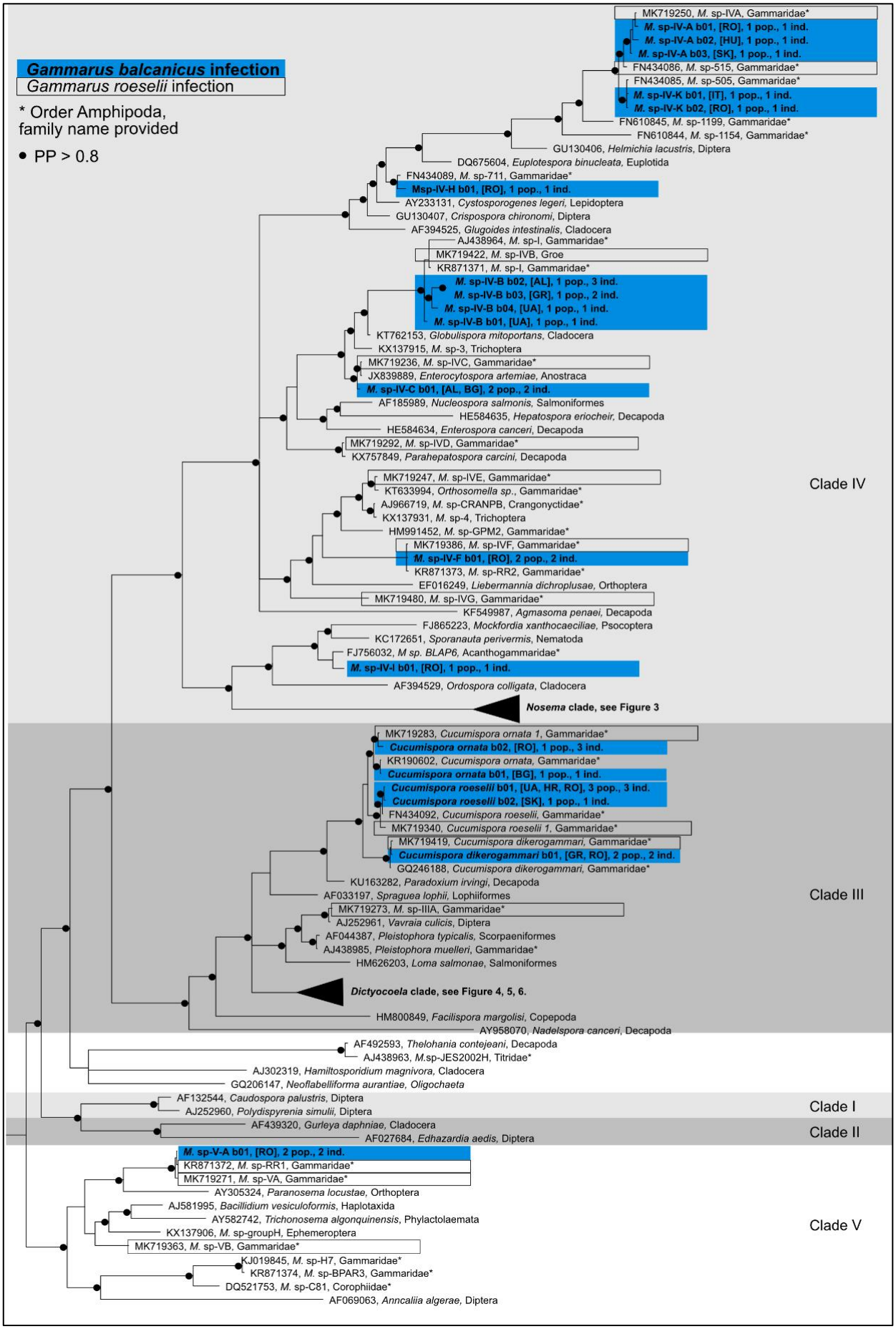

Gammarus balcanicus infection
Gammarus roeselii infection
* Order Amphipoda, family name provided
• PP > 0.8
MK719250, M. sp-IVA, Gammaridae*
M. sp-IV-A b01, [RO], 1 pop., 1 ind.
M. sp-IV-A b02, [HU], 1 pop., 1 ind.
M. sp-IV-A b03, [SK], 1 pop., 1 ind.
FN434086, M. sp-515, Gammaridae*
FN434085, M. sp-505, Gammaridae*
M. sp-IV-K b01, [IT], 1 pop., 1 ind.
M. sp-IV-K b02, [RO], 1 pop., 1 ind.
FN610845, M. sp-1199, Gammaridae*
FN610844, M. sp-1154, Gammaridae*
GU130406, Helmichia lacustris, Diptera
DQ675604, Euplotespora binucleata, Euplotida
FN434089, M. sp-711, Gammaridae*
Msp-IV-H b01, [RO], 1 pop., 1 ind.
AY233131, Cystosporogenes legeri, Lepidoptera
GU130407, Crispospora chironomi, Diptera
AF394525, Glugoides intestinalis, Cladocera
AJ438964, M. sp-I, Gammaridae*
MK719422, M. sp-IVB, Groe
KR871371, M. sp-I, Gammaridae*
M. sp-IV-B b02, [AL], 1 pop., 3 ind.
M. sp-IV-B b03, [GR], 1 pop., 2 ind.
M. sp-IV-B b04, [UA], 1 pop., 1 ind.
M. sp-IV-B b01, [UA], 1 pop., 1 ind.
KT762153, Globulispora mitoportans, Cladocera
KX137915, M. sp-3, Trichoptera
MK719236, M. sp-IVC, Gammaridae*
JX839889, Enterocytospora artemiae, Anostraca
M. sp-IV-C b01, [AL, BG], 2 pop., 2 ind.
AF185989, Nucleospora salmonis, Salmoniformes
HE584635, Hepatospora eriocheir, Decapoda
HE584634, Enterospora canceri, Decapoda
MK719292, M. sp-IVD, Gammaridae*
KX757849, Parahepatospora carcini, Decapoda
MK719247, M. sp-IVE, Gammaridae*
KT633994, Orthosomella sp., Gammaridae*
AJ966719, M. sp-CRANPB, Crangonyctidae*
KX137931, M. sp-4, Trichoptera
HM991452, M. sp-GPM2, Gammaridae*
MK719386, M. sp-IVF, Gammaridae*
M. sp-IV-F b01, [RO], 2 pop., 2 ind.
KR871373, M. sp-RR2, Gammaridae*
EF016249, Liebermannia dichroplusae, Orthoptera
MK719480, M. sp-IVG, Gammaridae*
KF549987, Agmasoma penaei, Decapoda
FJ865223, Mockfordia xanthocaeciliae, Psocoptera
KC172651, Sporanauta perivermis, Nematoda
FJ756032, M sp. BLAP6, Acanthogammaridae*
M. sp-IV-I b01, [RO], 1 pop., 1 ind.
AF394529, Ordospora colligata, Cladocera
Nosema clade, see Figure 3
Clade IV
MK719283, Cucumispora ornata 1, Gammaridae*
Cucumispora ornata b02, [RO], 1 pop., 3 ind.
KR190602, Cucumispora ornata, Gammaridae*
Cucumispora ornata b01, [BG], 1 pop., 1 ind.
Cucumispora roeselii b01, [UA, HR, RO], 3 pop., 3 ind.
Cucumispora roeselii b02, [SK], 1 pop., 1 ind.
FN434092, Cucumispora roeselii, Gammaridae*
MK719340, Cucumispora roeselii 1, Gammaridae*
MK719419, Cucumispora dikerogammari, Gammaridae*
Cucumispora dikerogammari b01, [GR, RO], 2 pop., 2 ind.
GQ246188, Cucumispora dikerogammari, Gammaridae*
KU163282, Paradoxium irvingi, Decapoda
AF033197, Spraguea lophii, Lophiiformes
MK719273, M. sp-IIIA, Gammaridae*
AJ252961, Vavraia culicis, Diptera
AF044387, Pleistophora typicalis, Scorpaeniformes
AJ438985, Pleistophora muelleri, Gammaridae*
HM626203, Loma salmonae, Salmoniformes
Dictyocoela clade, see Figure 4, 5, 6.
HM800849, Facilispora margolisi, Copepoda
AY958070, Nadelspora canceri, Decapoda
Clade III
AF492593, Thelohania contejeani, Decapoda
AJ438963, M.sp-JES2002H, Titridae*
AJ302319, Hamiltosporidium magnivora, Cladocera
GQ206147, Neoflabelliforma aurantiae, Oligochaeta
AF132544, Caudospora palustris, Diptera
AJ252960, Polydispyrenia simulii, Diptera
Clade I
AF439320, Gurleya daphniae, Cladocera
AF027684, Edhazardia aedis, Diptera
Clade II
M. sp-V-A b01, [RO], 2 pop., 2 ind.
KR871372, M. sp-RR1, Gammaridae*
MK719271, M. sp-VA, Gammaridae*
AY305324, Paranosema locustae, Orthoptera
AJ581995, Bacillidium vesiculoformis, Haplotaxida
AY582742, Trichonosema algonquinensis, Phylactolaemata
KX137906, M. sp-groupH, Ephemeroptera
MK719363, M. sp-VB, Gammaridae*
KJ019845, M. sp-H7, Gammaridae*
KR871374, M. sp-BPAR3, Gammaridae*
DQ521753, M. sp-C81, Corophiidae*
AF069063, Anncaliia algerae, Diptera
Clade V

Fig. 3. Bayesian phylogenetic reconstruction based on microsporidian partial small ribosomal subunit rDNA alignment. Labels highlighted in bold text and blue background are parasites infecting *G. balcanicus*. Labels for these parasites correspond to haplogroups and include in the following order: taxa name, two letters ISO codes for country (Table S1), number of infected populations (= pop.) and total number of infected individuals (= ind.). Taxa name is either binomial name for fully described taxa, or M. sp (*i.e.* Microsporidium sp) followed by clade number as in Vossbrinck & Debrunner-Vossbrinck (2005) for undescribed species. Other parasite sequences were taken from Genbank. Labels include: Genbank accession numbers, taxa name as given in publication, the order name of the host (except for amphipod hosts where the family is provided). Labels highlighted in black frames are parasites of *G. roeselii*. For *Nosema* and *Dictyocoela* branches are collapsed (triangle sizes not reflecting evolutionary distance) and detailed phylogeny are given in Fig. 5 and 5-8. PP : Bayesian Posterior Probabilities. The outgroup used was *Basidiobolus ranarum* AY635841 (not represented in the tree).

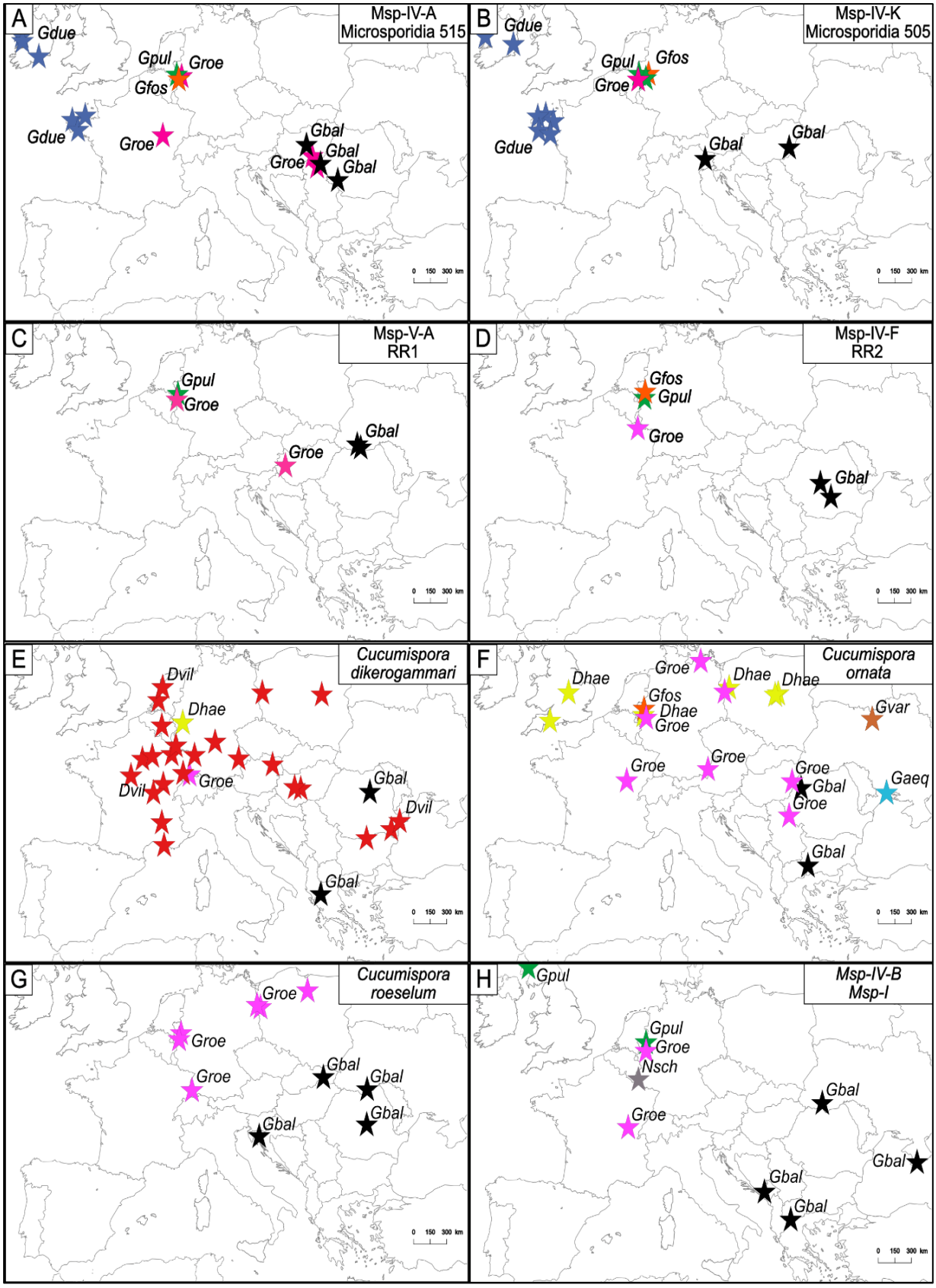

A
Msp-IV-A
Microsporidia 515
Gdue
Gpul
Groe
Gfos
Gdue
Groe
Gbal
Gbal
Groe
Gbal
B
Msp-IV-K
Microsporidia 505
Gdue
Gpul
Gfos
Groe
Gdue
Gbal
Gbal
C
Msp-V-A
RR1
Gpul
Groe
Gbal
Groe
D
Msp-IV-F
RR2
Gfos
Gpul
Groe
Gbal
E
Cucumispora
dikerogammari
Dvil
Dhae
Dvil
Groe
Gbal
Dvil
Gbal
F
Cucumispora
ornata
Dhae
Groe
Dhae
Dhae
Gfos
Dhae
Groe
Gvar
Groe
Groe
Groe
Gbal
Gaeq
Groe
Gbal
G
Cucumispora
roeselum
Groe
Groe
Groe
Gbal
Gbal
Gbal
Gbal
H
Msp-IV-B
Msp-I
Gpul
Gpul
Groe
Nsch
Groe
Gbal
Gbal
Gbal
Gbal
0 150 300 km

Fig. 4. Geographic distribution of some rare microsporidia taxa of *G. balcanicus*, showing their occurrence in other gammarid species over Europe. Each map (A-G) refers to the parasite taxa presented in the top-right inset. Each star is representing a sampling site where microsporidia infected amphipod. Each Amphipods is associated with a colour. Taxa name are as follows: Dhae = *Dikerogammarus haemobaphes*, Dvil = *D. villosus*, Gaeq = *Gammarus aequicauda*, Gbal = *G. balcanicus*, Gdue = *G. duebenum*, Gfos = *G. fossarum*, Gpul = *G. pulex*, Groe = *G. roeselii*, Gvar = *G. varsoviensis*.

II.3.3 Rare infections (*i.e.* infections not ascribed to the genera *Nosema* or *Dictyocoela*)

All the rare infections were parasites known also to infect other gammarid species. Among them, eight were not fully described taxa.

The three haplogroups of M. *sp*-IV-A were found each in a single individual and population (Fig. 3, Table S1-2), located in Hungary, Romania and Slovakia (pop. 38, 42 and 46, respectively; Fig. 2, Fig. 4A, Table S1-2). All these 3 haplogroups were phylogenetically close (98.4-99.3% identity; 122-139 bp coverage) to M. *sp*-515 (GenBank: FN434086) infecting Irish and French populations of *G. duebeni* (Krebes *et al.* 2010) and to M. *sp*-IV-A (GenBank: MK719250) from French and Hungarian populations of *G. roeselii* (Quiles *et al.* 2019) (II: Fig. 3, Fig. 4A, Table S1-2). Other phylogenetically close sequences were also detected in *G. pulex, G. fossarum* and *G. roeselii* in Germany (Fig 3A, Table S2, (Grabner 2017; Grabner *et al.* 2015). The two haplogroups of M. *sp*-IV-K (Fig. 3, Table S1) were found each in a single individual and population (Fig. 3, Table S1-2), from Italy and Romania (pop. 13 and pop. 42, respectively; Fig. 2, Fig. 4B, Table S1-S2). These two haplogroups were phylogenetically close (99.4% identity; 322pb coverage) to M. *sp*-505 (GenBank: FN434085) infecting Irish and French populations of *G. duebeni* (Krebes *et al.* 2010). Additionally, phylogenetically close sequences were also detected in *G. pulex, G. fossarum* and *G. roeselii* in Germany (Fig. 4B, Table S2, (Grabner 2017; Grabner *et al.* 2015)). M. *sp*-IV-H (Fig. 3, Table S1) was found in one individual from Romania (pop. 43, Fig. 2 and Fig. 4, Table S1-S2) and was phylogenetically close (98.9% identity; 179 coverage) to *M. sp*-711 (GenBank: FN434089) previously found in *G. duebeni* (Krebes *et al.* 2010). The 4 haplogroups of M. *sp*-IV-B (Fig. 3, Table S1) were found infecting a total of 7 individuals from 4 populations in Albania, Greece and Ukraine (pop. 10, 32, 80 and 81, respectively, Fig. 2, Fig. 4H, Table S1). Identities were on average of 97.1% (with *c.* 323 bp coverage) between these haplogroups (M. *sp*-IV-B-b02, M. *sp*-IV-B-b03, M. *sp*-B-IV-B-b04), and 98.2% (with *c.* 718 bp coverage) between M. *sp*-IV-B-b01 and M. *sp*-I (GenBank: KR871371) and M. *sp*-IV-B (GenBank: MK719422), both microsporidian being detected in *G. roeselii* in few populations in Germany (Grabner 2017; Grabner *et al.* 2015) and in one French population (Quiles *et al.* 2019) (Fig. 4H). Additionally, phylogenetically close sequences were also detected in *G. pulex* in Germany (Grabner *et al.* 2015)

and Scotland (Terry *et al.* 2004), but also in *Niphargus schellenbergi* in Luxemburg (Weigand *et al.* 2016) (Fig. 4H, Table S2). M. *sp*-IVC was found in 2 individuals from Albania and Greece (Pop. 4 and 25, respectively, Fig. 2, Fig. 3). This sequence is closely related (98.2% identity; 190pb coverage) to the sequence of a microsporidia infecting *G. roeselii* from Greece (GenBank: MK719236) (Quiles *et al.* 2019), which was itself 100% similar to the sequences of *Enterocytospora artemiae* found infecting *Artemia franciscana* from France, USA and Israel (GenBank: JX839889) (Rode *et al.* 2013). The only M. *sp*-IV-F haplogroup (Fig. 3, Table S1) was found in 2 individuals from 2 populations in Romania (pop. 65 and 66, Fig. 2, Fig. 4D, Table S1). The sequence was identical to M. *sp*-RR2 (GenBank: KR871373, coverage 313pb) found in the Ruhr region of Germany, infecting *G. pulex*, *G. fossarum* and also *G. roeselii* (Grabner *et al.* 2015) and M. sp-IVF (GenBank: MK719386, 313 pb coverage) also found in two French populations infecting *G. roeselii* (Quiles *et al.* 2019). One haplogroup M. *sp*-IV-I (Fig. 3) infecting 1 individual from Romania (pop. 68, Fig. 2) was closest related to an unpublished microsporidia infecting an amphipod species in Baikal lake (GenBank: FJ756032) (92.4% identity; 532pb coverage) (Qiu *et al.* unpublished). Finally, MspV-A, the only microsporidian taxa in clade V infecting *G. balcanicus*, was found as a single haplogroup infecting 2 individuals in two populations in Romania (pop. 78 and 81). It was close to M. *sp*-RR1 infecting *G. pulex* (GenBank: KR871372) (99.8% identity; 814pb coverage) (Grabner 2015). A phylogenetically closely related parasite was also found in *G. roeselii* (GenBank: MK719271) (99.6% identity; 247pb coverage) (Quiles *et al.* 2019).

Three of the rare infections could be ascribed to fully described taxa, all 3 being species of the genus *Cucumispora* which belonging to Clade III: *C. dikerogammari* (Ovcharenko 2010), *C. roeselii* (Bojko 2017) and *C. ornata* (Bojko 2015).

One haplogroup was associated with *C. dikerogammari* (Fig. 3, Table S1) and was found infecting 2 individuals from 2 populations from Greece and Romania (pop. 33 and 69, respectively, (Fig. 2). This haplogroup showed identical sequences to *C. dikerogammari* (Genbank: GQ258752) initially identified in *Dikerogammarus villosus* host (Ovcharenko 2010, Wattier 2007). It is to be noticed that *C. dikerogammari* b01 is also 100% similar to *C. dikerogammari* 1 (Fig. 3, Table S2) found in *G. roeselii* in France (Quiles *et al.* 2019). *Cucumispora dikerogammari* is frequent and widespread in *D. villosus* in Europe (Ovcharenko *et al.* 2010; Wattier *et al.* 2007) but is also known to infect *D. haemobaphes*, notably in Germany (Bojko *et al.* 2017c) (Fig. 4E). Two haplogroups were ascribed to *C. roeselii* (Fig. 3, Table S1), infecting a total of 4 individuals, each individual coming from a single population from Slovakia, Hungary, Romania and Austria (pop. 44, 61, 66 and 81, respectively, Fig. 2). The 2 haplogroups were phylogenetically very close (99.5-99.7%

identity; 394-393pb coverage) with *C. roeselii* (GenBank: FN434092) used for the full description of the species (Bojko *et al.* 2017a). *Cucumispora roeselii* was also previously found to infect *G. roeselii* in Poland (Bojko *et al.* 2017a; Quiles *et al.* 2019) and Germany (Grabner 2017; Quiles *et al.* 2019) (Fig. 4G Table S2). Finally, two haplogroups were ascribed to *C. ornata*, one being present in 3 individuals from 1 population in Hungary (pop. 43, Fig. 2 and Fig. 4F), the other one with 1 individual from 1 population in Bulgaria (pop. 21, Fig. 2, Fig. 4F). The 2 haplogroups were phylogenetically very close (97.3-100% identity 253-241pb coverage) with *C. ornata* (GenBank: KR190602) used in the full description of the species (Bojko *et al.* 2015). *Cucumispora ornata* was also found in a variety of gammarid taxa all other Europe (*e.g. G. roeselii*, *G. fossarum, G. varsoviensis, G. aequicauda, D. haemobaphes* (Fig. 4F, Table S2, (Bojko *et al.* 2015, 2017a, 2017c; Quiles *et al.* 2019)).

Trying to summarize data about all these rare microsporidia is not easy. We can nevertheless roughly organise them in 2 groups: 1) the rare microsporidia for which large geographic coverage and high number of other gammarid hosts were known from the literature (Msp-IV-A, *C. dikerogammari* and *C. ornata*) and for which *G. balcanicus* only extends the host taxonomic spectrum but barely the geographic range and 2) all the other microsporidia, having scarce literature records often with few or even very few geographic records, for which *G. balcanicus* not only extend the host taxonomic spectrum but also significantly extend the geographic distribution of these microsporidia taxa.

II.3.4 Infections ascribed to the genus *Nosema*

Seventy-six infections of *G. balcanicus* individuals (30.7% of infections) fall within 4 haplogroups ascribed to *Nosema granulosis* (Fig. 5), a fully described species (Terry *et al.* 1999). *Nosema granulosis* is the only *Nosema* known to infect Amphipoda and it have been detected in many gammarid taxa (*e.g.* Dunn *et al.* 2006; Grabner 2017; Haine *et al.* 2004; Kelly *et al.* 2002; Quiles *et al.* 2019; Terry *et al.* 2004; Weigand *et al.* 2016). These 4 haplogroups were restricted to *G. balcanicus*, while 9 other haplogroups were associated with other gammarids (Fig. 5).

The first haplogroup associated with *G. balcanicus, i.e. N. granulosis* b01 (highlighted in red in Fig. 5) was found to infect 2 individuals, one in Montenegro and one in Romania (Fig. 5, Table S1-2). The phylogenetically closest haplogroup to *N. granulosis* b01 was a haplogroup

known to infect *G. roeselii*, but also the subterranean amphipod *Niphargus schellenbergi*, *G. fossarum* and *G. pulex*.

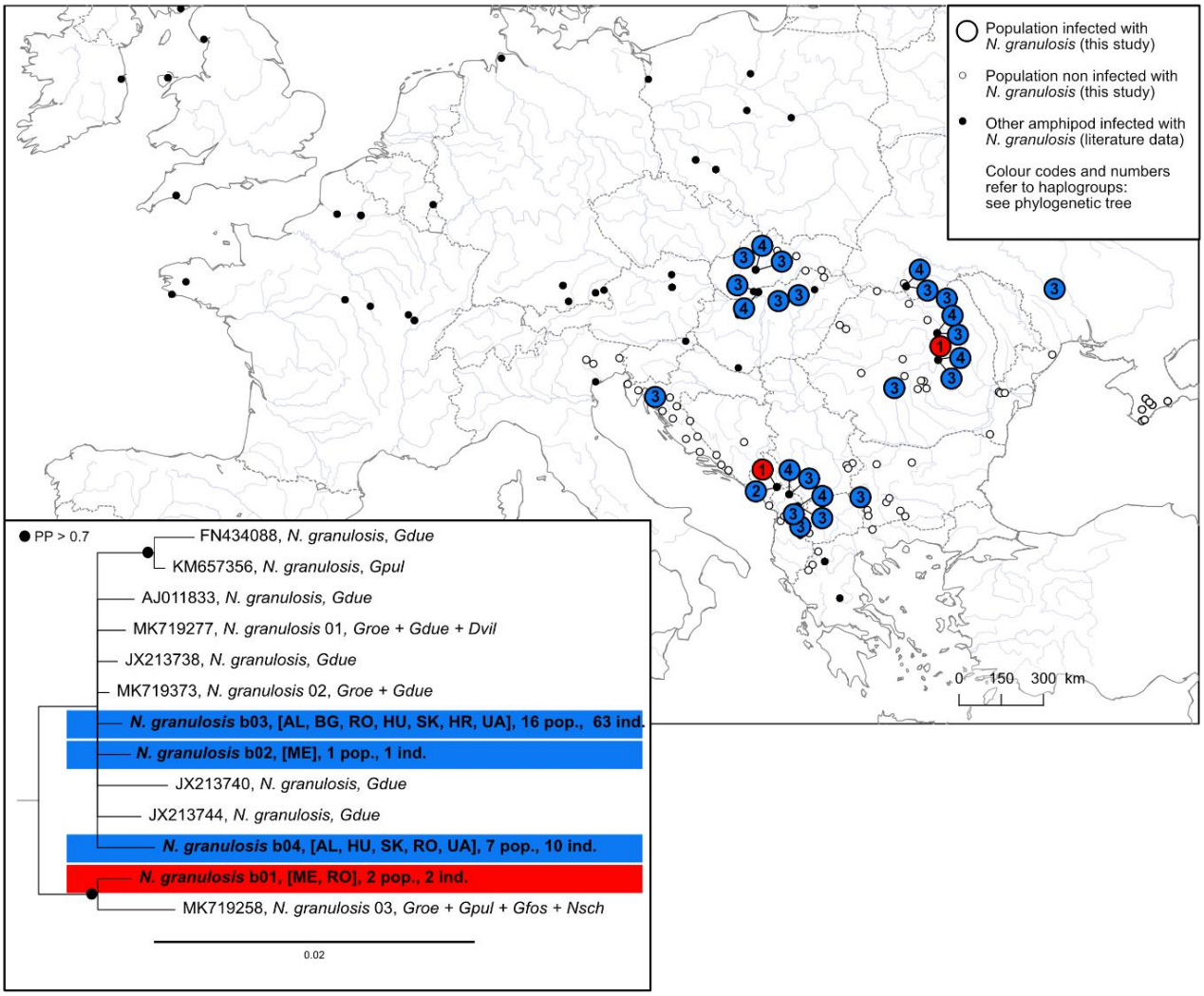


Fig. 5. Geographic distribution and phylogenetic tree for infections by *Nosema granulosis* infecting *Gammarus balcanicus*. Large coloured dots represent infections found in *G. balcanicus*, haplogroup b01 and b02-04 being red and blue, respectively. Small black dots are infections reported in the literature in other amphipods (Table S2 for further details). The Bayesian phylogenetic reconstruction is based on small ribosomal subunit rDNA. *Nosema antherae* (DQ073396) was used as outgroup (not shown on the tree). Sequences from the present study are in bold, highlighted in colored boxes matching color code on map. Labels include haplogroup names, two letters ISO codes for countries (Table S1), number of infected populations (= pop.), total number of infected individuals (= ind.). Sequences from Genbank are representative of other *Nosema granulosis* haplogroups (Table S2). Labels include the accession number, the microsporidia species name as given in the associated publication (haplogroup name (01-03) being additionally provided when the host is G. roeselii) and the host species abbreviated names (Dvil = *Dikerogammarus villosus*, Gdue = *Gammarus duebeni*, Gfos = *G. fossarum*, Gpul = *G. pulex*, Groe = *G. roeselii*, Gvar = *G. varsoviensis*, Nsch = *Niphargus schellenbergi*. PP: Bayesian Posterior Probability.

The 3 remaining *G. balcanicus* associated haplogroups, *N. granulosis* b02-b04 (highlighted in blue in Fig. 5), were closely related phylogenetically. Haplogroup *N. granulosis* b03 was the

most frequent haplogroup, with 63 individuals, and was widespread, being associated with 16 populations across 7 countries (Fig. 5, Table S1-2). Haplogroup *N. granulosis* b04 was also quite frequent and widespread, found in10 individuals, 7 populations across 5 countries (Fig. 5, Table S1-2). Only the last haplogroup, *N. granulosis* b04 was restricted to 1 individual from Macedonia. These 3 haplogroups *N. granulosis* b02-04 were phylogenetically close to a set of microsporidia known to be associated with a wide range of gammarid species all over Europe (Fig. 5, Table S2). It is to be noticed that while the haplogroups *N. granulosis* b01 and *N. granulosis* b02-04 were more distantly related, the two sets of haplogroups overlap in geographic coverage. Even if *N. granulosis* b01 is rare, it is both observed in the southern and northern part of *G. balcanicus* distribution range (Fig. 5)

II.3.5 Infections ascribed to the genus *Dictyocoela*

Thirty-one haplogroups of *G. balcanicus* parasite could be ascribed to 4 fully described species of *Dictyocoela* parasites: *Dictyocoela roeselum* (Bacela-Spychalska *et al.* 2018; Haine *et al.* 2004), *D. muelleri* (Terry *et al.* 2004), *D. duebenum* (Terry *et al.* 2004) and *D. berillonum* (Terry *et al.* 2004) (Fig. 6, Table S1-2). It is to be noticed that none of these 31 haplogroups was shared with any other gammarid species, while these 4 species are known to infect many other gammarid hosts (Bacela-Spychalska *et al.* 2018). Overall, *Dictyocoela* spp. were the most common microsporidian parasites infecting *G. balcanicus* with 139 individuals infected in 45 populations, corresponding to 57.1% of all microsporidian infections found (Table S1). However, the 4 species were not equally represented as for both haplogroup diversity and prevalence (see below).

Dictyocoela duebenum was found in 36 individuals in 11 populations (Fig. 7, Table S1-2) all over the range of *G. balcanicus,* although 7 populations were in the southern part of the distribution (Fig. 7). The 3 haplogroups *D. duebenum* b01 to b03 were in average 99.7% identical (at least 384pb coverage), being themselves 99.6% identical (at least 391pb coverage) to the sequences initially provided with the description the species (GenBank: MG773214) (Bacela-Spychalska *et al.* 2018). One individual could not be attributed to specific haplogroups (either b01 or b03) due to too short sequence (See Methods) (Fig. 7, Table S2). *Dictyocoela duebenum* is a species known to harbour a vast haplogroup diversity with 25 haplogroups (additionally to the 3 new associated with *G. balcanicus*) (Fig. 6). These haplogroups are associated with a broad gammarid host spectrum all over northern and western Europe, although especially frequently associated with *G. duebeni* in north-eastern Europe (Ironside & Alexander 2015; Krebes *et al.* 2010).

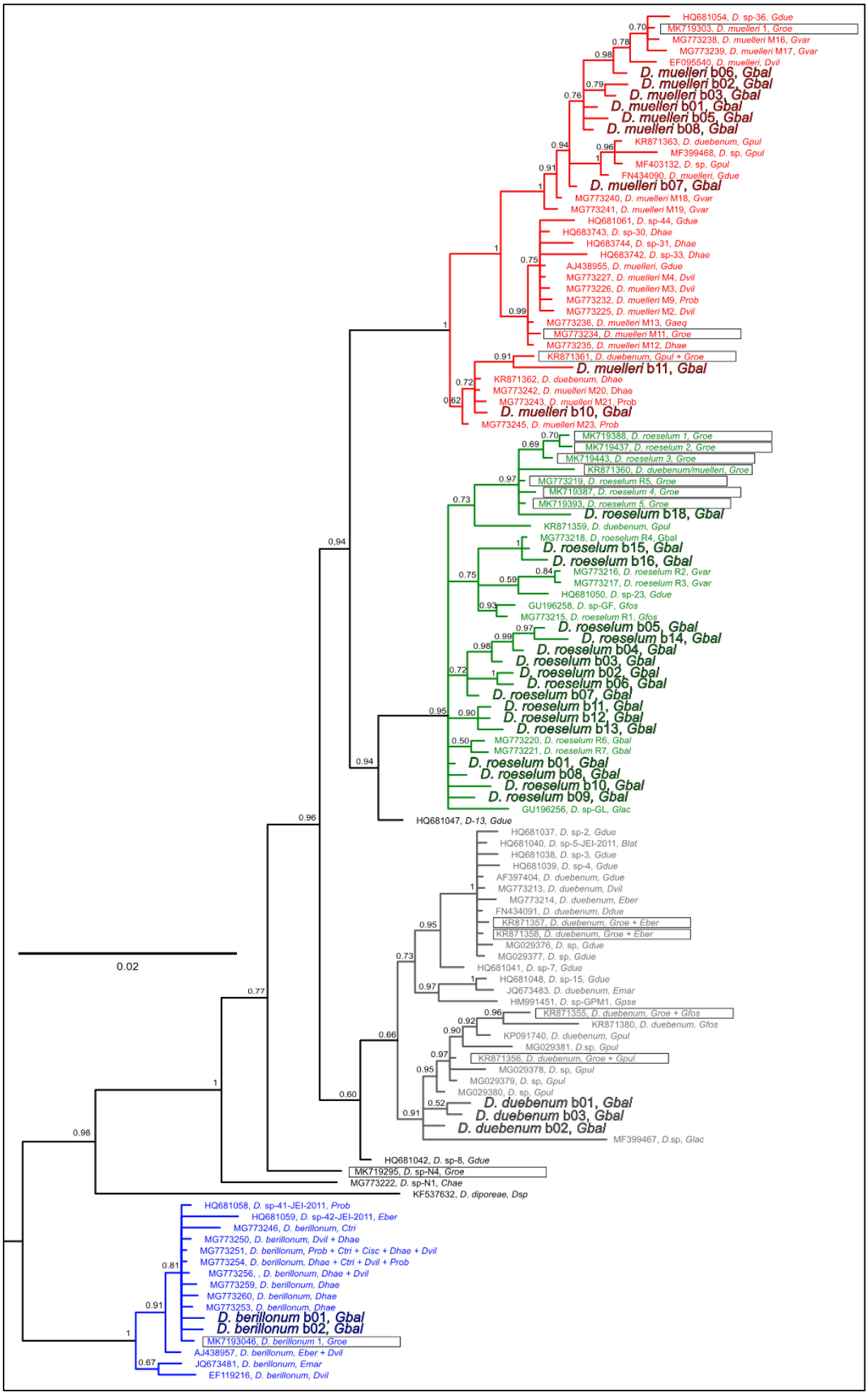

HQ681054, D. sp-36, Gdue
MK719303, D. muelleri 1, Groe
MG773238, D. muelleri M16, Gvar
MG773239, D. muelleri M17, Gvar
EF095540, D. muelleri, Dvil
D. muelleri b06, Gbal
D. muelleri b02, Gbal
D. muelleri b03, Gbal
D. muelleri b01, Gbal
D. muelleri b05, Gbal
D. muelleri b08, Gbal
KR871363, D. duebenum, Gpul
MF399468, D. sp, Gpul
MF403132, D. sp, Gpul
FN434090, D. muelleri, Gdue
D. muelleri b07, Gbal
MG773240, D. muelleri M18, Gvar
MG773241, D. muelleri M19, Gvar
HQ681061, D. sp-44, Gdue
HQ683743, D. sp-30, Dhae
HQ683744, D. sp-31, Dhae
HQ683742, D. sp-33, Dhae
AJ438955, D. muelleri, Gdue
MG773227, D. muelleri M4, Dvil
MG773226, D. muelleri M3, Dvil
MG773232, D. muelleri M9, Prob
MG773225, D. muelleri M2, Dvil
MG773236, D. muelleri M13, Gaeq
MG773234, D. muelleri M11, Groe
MG773235, D. muelleri M12, Dhae
KR871361, D. duebenum, Gpul + Groe
D. muelleri b11, Gbal
KR871362, D. duebenum, Dhae
MG773242, D. muelleri M20, Dhae
MG773243, D. muelleri M21, Prob
D. muelleri b10, Gbal
MG773245, D. muelleri M23, Prob
MK719388, D. roeselum 1, Groe
MK719437, D. roeselum 2, Groe
MK719443, D. roeselum 3, Groe
KR871360, D. duebenum/muelleri, Groe
MG773219, D. roeselum R5, Groe
MK719387, D. roeselum 4, Groe
MK719393, D. roeselum 5, Groe
D. roeselum b18, Gbal
KR871359, D. duebenum, Gpul
MG773218, D. roeselum R4, Gbal
D. roeselum b15, Gbal
D. roeselum b16, Gbal
MG773216, D. roeselum R2, Gvar
MG773217, D. roeselum R3, Gvar
HQ681050, D. sp-23, Gdue
GU196258, D. sp-GF, Gfos
MG773215, D. roeselum R1, Gfos
D. roeselum b05, Gbal
D. roeselum b14, Gbal
D. roeselum b04, Gbal
D. roeselum b03, Gbal
D. roeselum b02, Gbal
D. roeselum b06, Gbal
D. roeselum b07, Gbal
D. roeselum b11, Gbal
D. roeselum b12, Gbal
D. roeselum b13, Gbal
MG773220, D. roeselum R6, Gbal
MG773221, D. roeselum R7, Gbal
D. roeselum b01, Gbal
D. roeselum b08, Gbal
D. roeselum b10, Gbal
D. roeselum b09, Gbal
GU196256, D. sp-GL, Glac
HQ681047, D-13, Gdue
HQ681037, D. sp-2, Gdue
HQ681040, D. sp-5-JEI-2011, Blat
HQ681038, D. sp-3, Gdue
HQ681039, D. sp-4, Gdue
AF397404, D. duebenum, Gdue
MG773213, D. duebenum, Dvil
MG773214, D. duebenum, Eber
FN434091, D. duebenum, Ddue
KR871357, D. duebenum, Groe + Eber
KR871358, D. duebenum, Groe + Eber
MG029376, D. sp, Gdue
MG029377, D. sp, Gdue
HQ681041, D. sp-7, Gdue
HQ681048, D. sp-15, Gdue
JQ673483, D. duebenum, Emar
HM991451, D. sp-GPM1, Gpse
KR871355, D. duebenum, Groe + Gfos
KR871380, D. duebenum, Gfos
KP091740, D. duebenum, Gpul
MG029381, D.sp, Gpul
KR871356, D. duebenum, Groe + Gpul
MG029378, D. sp, Gpul
MG029379, D. sp, Gpul
MG029380, D. sp, Gpul
D. duebenum b01, Gbal
D. duebenum b03, Gbal
D. duebenum b02, Gbal
MF399467, D.sp, Glac
HQ681042, D. sp-8, Gdue
MK719295, D. sp-N4, Groe
MG773222, D. sp-N1, Chae
KF537632, D. diporeae, Dsp
HQ681058, D. sp-41-JEI-2011, Prob
HQ681059, D. sp-42-JEI-2011, Eber
MG773246, D. berillonum, Ctri
MG773250, D. berillonum, Dvil + Dhae
MG773251, D. berillonum, Prob + Ctri + Cisc + Dhae + Dvil
MG773254, D. berillonum, Dhae + Ctri + Dvil + Prob
MG773256, , D. berillonum, Dhae + Dvil
MG773259, D. berillonum, Dhae
MG773260, D. berillonum, Dhae
MG773253, D. berillonum, Dhae
D. berillonum b01, Gbal
D. berillonum b02, Gbal
MK7193046, D. berillonum 1, Groe
AJ438957, D. berillonum, Eber + Dvil
JQ673481, D. berillonum, Emar
EF119216, D. berillonum, Dvil
0.02

Fig. 6. Bayesian phylogenetic reconstruction based on small ribosomal subunit rDNA for the microsporidian genus *Dictyocoela* infecting amphipods *Dictyocoela cavimanum* (AJ438960) was used as outgroup (not shown on the tree). Four taxa were ascribed a color following recent reassessment of the genus taxonomy by Bacela-Spychalska *et al.* (2017) *i.e. D. muelleri* (red), *D. roeselum* (green), *D. duebenum* (grey) and *D. berillonum* (blue). Haplogroups from the present study are in bold. Sequences from Genbank are representative of other Dictyocoela haplogroups (Table S2). Labels include, in this order, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated name(s). Blat = *Brandtia latissima*, Cisc = *Chaetogammarus ischnus*, Ctri = C. *trichiatus*, Dhae = *Dikerogammarus haemobaphes*, Dvil = *D. villosus*, Dsp = *Diporeia sp.*, Eberi = *Echinogammarus berilloni*, Gaeq = *Gammarus aequicauda*, Gdue = *G. duebeni*, Gfos = *G. fossarum*, Glac = *G. lacustris*, Gpul = *G. pulex*, Groe = *G. roeselii*, Gvar = *G. varsoviensis*, Prob = *Pontogammarus robustoides*. Or For abbreviations of host species names: Table S2. Haplogroups infecting G. roeselii are highlighted by a black frame. Numbers on the branches indicate Bayesian Posterior Probabilities.

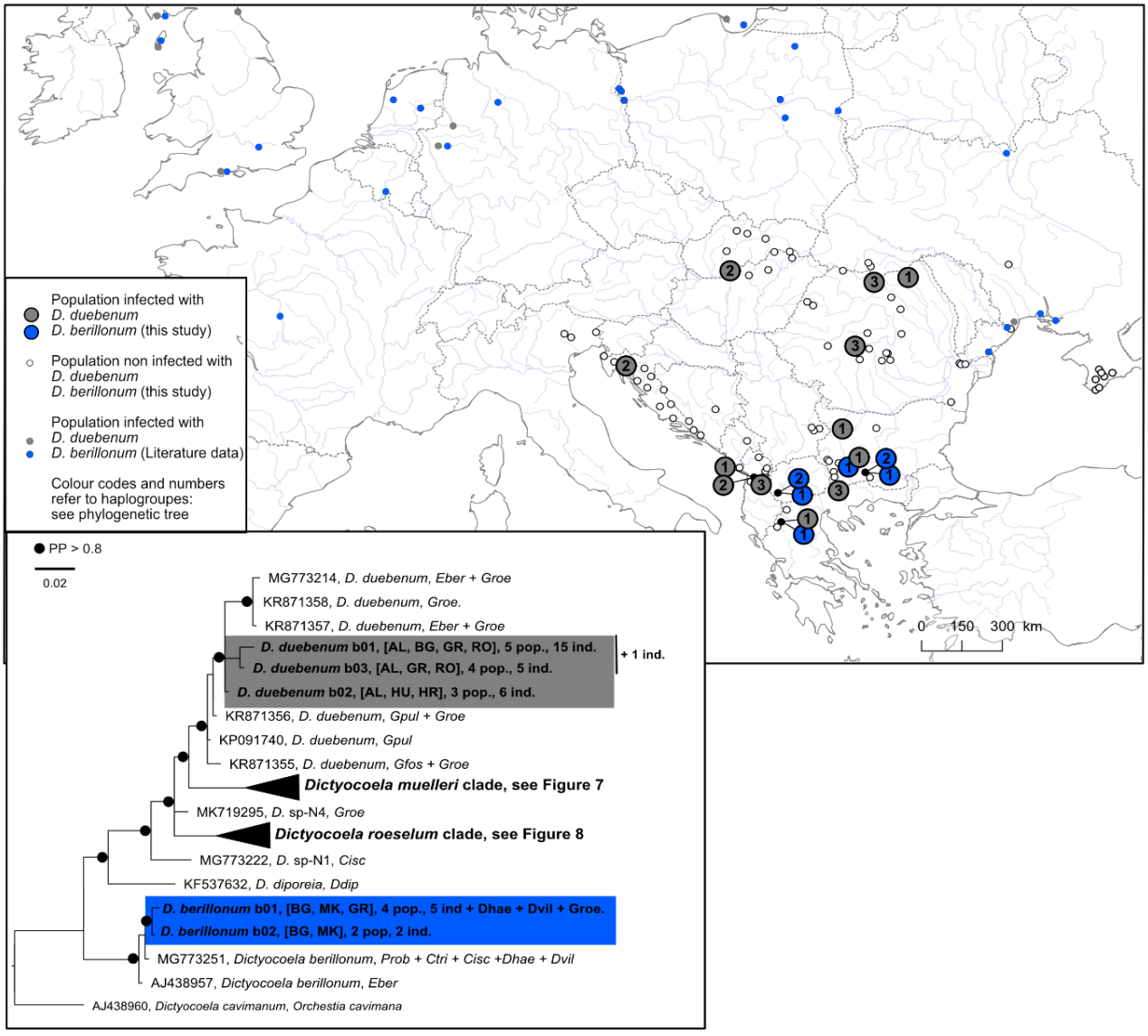


Fig. 7. Geographic distribution and phylogenetic tree for infections by microsporidia of the genus *Dictyocoela duebenum* and *Dictyocoela berillonum* in *Gammarus balcanicus*. Large coloured dots represent infections found in *G. balcanicus*, from *Dictyocoela duebenum* and *Dictyocoela berillonum*, haplogroups b01-03 and haplogroup

b01-02 in grey and blue respectively. Small coloured dots represent infection found in previous studies in other amphipods (Table S2 for further details). The Bayesian phylogenetic reconstruction is based on small ribosomal subunit rDNA. Triangles represents other *Dictyocoela* genera. *Dictyocoela cavimanum*, (AJ438960) was used as outgroup for this tree. Sequences from the present study are in bold, highlighted in colored boxes matching color code on map. Labels include haplogroup names, two letters ISO codes for countries (Table S1), number of infected populations (= pop.), total number of infected individuals (= ind.). Sequences from Genbank are all other *Dictyocoela duebenum* and *Dictyocoela berillonum* haplogroups (Table S2). Labels include, respectively, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated names (Table S2). PP : Bayesian Posterior Probability.

Dictyocoela berillonum was found in 7 individuals from 4 populations (Fig. 7, Table S1-2) only in the southern part of the *G. balcanicus* distribution (Fig. 7, Table S1). The 2 haplogroups of *D. berillonum* b01 and b02 were 99.7% identical [(348 bp coverage, being themselves 99.6-100% identical (at least 391 bp coverage)] to the type sequence used in the first determination of *D. berillonum* (Genbank: AJ438957, Terry *et al.* 2004). *Dictyocoela berillonum* is also a species known to harbour substantial haplogroup diversity with 13 haplogroups (excluding the 3 new associated with *G. balcanicus*), with a poor phylogenetic resolution (Fig. 6). This parasite is mostly associated with Ponto-Caspian hosts (*e.g. D. villosus*, *D. haemobaphes*, *Pontogammarus robustoides*), all invasive in northern, western and eastern Europe (Fig. 6).

Dictyocoela muelleri was found in 48 individuals in 17 populations in Romania, Ukraine and Croatia, so both in the N and S parts of *G. balcanicus* range (Fig. 8, Table S1-S2). A total of 9 haplogroups were detected in our study, adding to the 29 already known haplogroups in other gammarid species (Fig. 8, Table S2). Along with sequences from the literature and according to their phylogenetic proximity, these 9 haplogroups were tentatively clustered in 3 sets (Fig. 8). One set (highlighted in red in Fig. 8) included haplogroups *D. muelleri* b01, 02, 03, 08, 05, and 07. This set is the most represented with 39/48 (81.25%) individuals infected among *D. muelleri* infections (Fig. 8). These haplogroups present a geographic pattern with haplogroup *D. muelleri* b07 and b08 in the S part of the distribution of *G. balcanicus* (Fig. 8, Table S1-2), while haplogroups *D. muelleri* b01, 02, 03, 05 were present only in the N part of the host range (Fig. 8, Table S1-2). Haplogroup *D. muelleri* b06 (highlighted in green in Fig. 8) was found only in 1 Romanian population (Fig. 8, Table S1-2) with two individuals infected (Fig. 8). The third group (highlighted in yellow, in Fig. 8) included haplogroup *D. muelleri* b10 and b11 found in 7 individuals in four Romanian populations (Fig. 8, Table S1-2), *i.e.* the N part of the distribution *G. balcanicus* (Fig. 8). *Dictyocoela muelleri* b11 formed a set with sequences found in *G. pulex*. *Dictyocoela muelleri*

b10 is phylogenetically close to microsporidia infection found in *D. haemobaphes* and *P. robustoides* (Fig. 8).

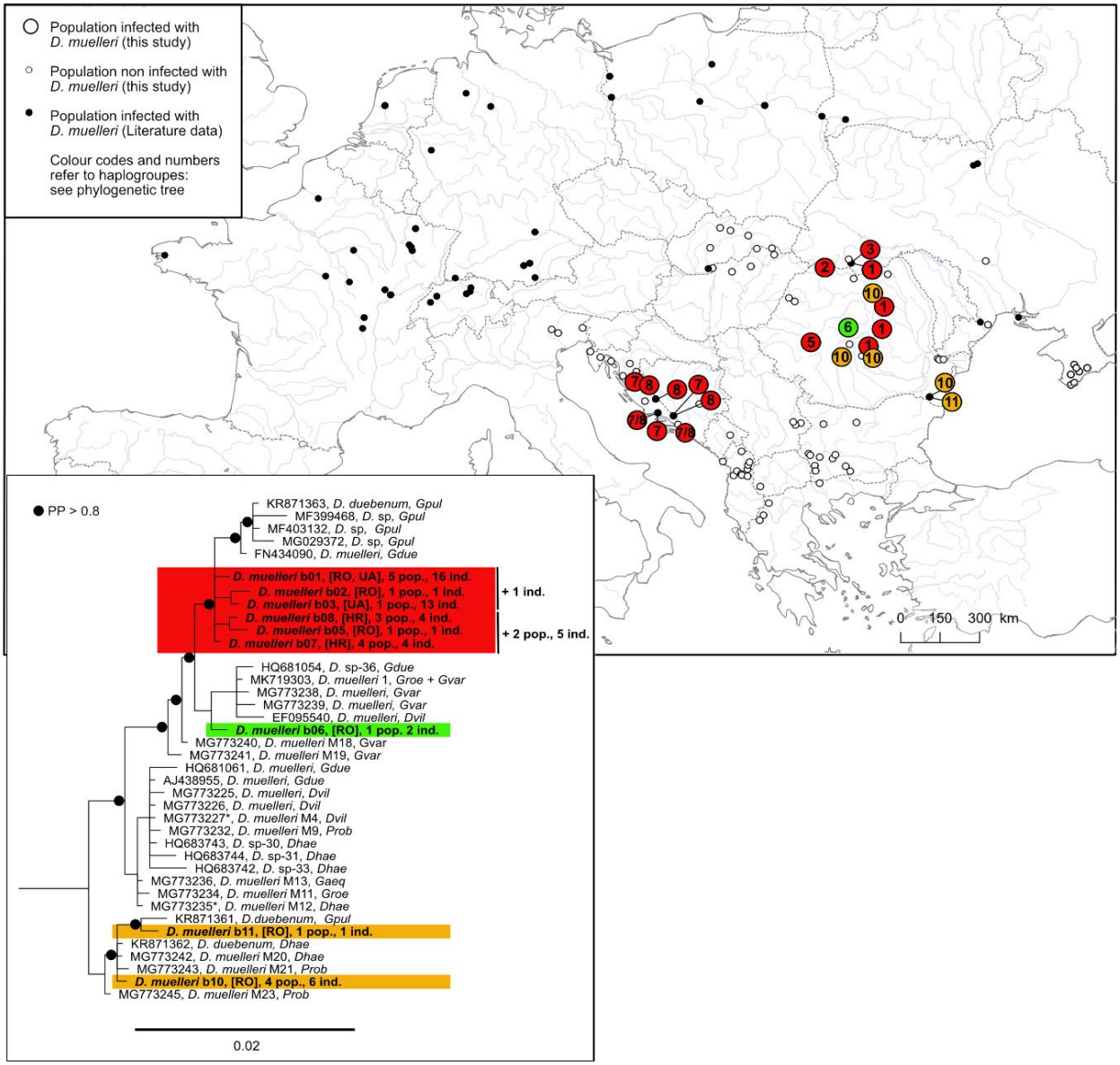


Fig. 8. Geographic distribution and phylogenetic tree for infections by microsporidia of the species *Dictyocoela muelleri* in *Gammarus balcanicus*. Large coloured dots represent infections found in *G. balcanicus*, from *Dictyocoela muelleri*, haplogroups b01-11. Small coloured dots represent infection found in previous studies in other amphipods (Table S2 for further details). The Bayesian phylogenetic reconstruction is based on small ribosomal subunit rDNA. *Dictyocoela* sp N1, (MG773222) was used as outgroup for this tree. Sequences from the present study are in bold, highlighted in colored boxes matching color code on map. Labels include haplogroup names, two letters ISO codes for countries (Table S1), number of infected populations (= pop.), total number of infected individuals (= ind.). Sequences from Genbank are all other *Dictyocoela muelleri* haplogroups (Table S2). Labels include, respectively, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated names (Table S2). PP : Bayesian Posterior Probability.

Dictyocoela roeselum associated with *G. balcanicus* was almost as common and widespread as *D. muelleri*, infecting 46 individuals in 23 populations all over the range of *G. balcanicus* (Fig. 6, Fig. 9, Table S1-2). However, *D. roeselum* was more diverse, accounting for 17 haplogroups in our data set. In addition, 3 other haplogroups were already reported for *G. balcanicus* (MG773218, *D. roeselum* R4, MG773221, *D. roeselum* R7, and MG773220 *D. roeselum* R6, Fig. 9, (Bacela-Spychalska *et al.* 2018) and 17 additional haplogroups were already known to infect other gammarids. Altogether, *D. roeselum* appears to be extremely diverse, with a total of 40 haplogroups (Fig. 6). Haplogroup *D. roeselum* b18 (highlighted in green in Fig. 9) was associated with only one individual found in Hungary; this haplogroup was 100% identical to individuals found infecting *G. roeselii* (Genbank: AY584252 (Bacela-Spychalska *et al.* 2018; Haine *et al.* 2004). Haplogroups *D. roeselum* b16 and b15 (highlighted in yellow in Fig. 9), and *D. roeselum* b11, b12 and b13 (highlighted in red in Fig. 9) are two groups sharing the features of only infecting individuals from the north-western part of the *G. balcanicus* range. Four individuals could be attributed to either haplogroups b16, b15 or MG773218, due to short length sequences (See Methods). Individuals belonging to haplogroups *D. roeselum* b03, b04, b05, b14 (highlighted in pink in Fig. 9) represent a set only found in one population (pop. 57, Fig. 2) in north Croatia. Haplogroups *D. roeselum* b02 and b06 (highlighted in dark blue in Fig. 9) were also found in only one Italian population (pop. 15, Fig. 2). The haplogroup *D. roeselum* b07 was found infecting four individuals in 3 populations in Croatia and Greece, along the eastern coast of the Adriatic sea (Fig. 9). Finally, the remaining haplogroups *D. roeselum* b01, b08, b09, b10 (in white in Fig. 9), which phylogenetic position was not resolved, were present both in northern and southern parts of *G. balcanicus* range. It is nevertheless worth noting that haplogroups b09 and b10 were only found in the Crimea peninsula.

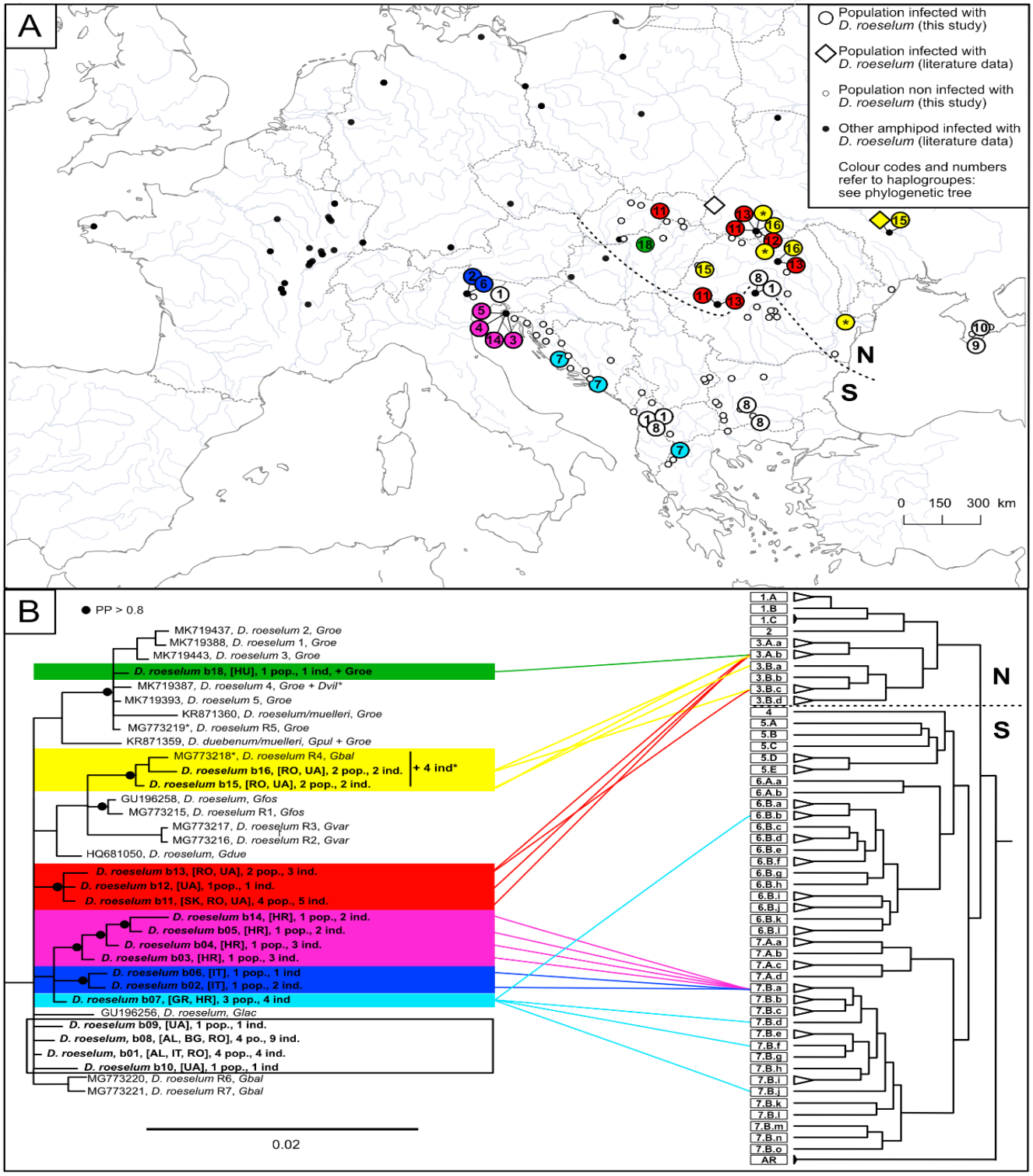


Fig. 9. Geographic distribution and phylogenetic tree for infections by microsporidia of the species *Dictyocoela roeselum* in *Gammarus balcanicus*. A. Large coloured dots represent infections found in *G. balcanicus*, for *Dictyocoela roeselum* haplogroups b01-18 detected in the present study. Large diamond show infection detected in *G. balcanicus* in previous studies. Small coloured dots represent infection found in previous studies in other amphipods (Table S2 for further details). B. Host-parasite co-phylogeny. The parasite phylogeny is shown on the left. This Bayesian phylogenetic reconstruction is based on small ribosomal subunit rDNA. *Dictyocoela* sp N1, (MG773222) was used as outgroup for this tree. Sequences from the present study are in bold, highlighted in

colored boxes matching color code on map. Labels include haplogroup names, two letters ISO codes for countries (Table S1), number of infected populations (= pop.), total number of infected individuals (= ind.). Sequences from Genbank are all other *Dictyocoela roeselum* haplogroups (Table S2). Labels include, respectively, the accession number, the microsporidia species name given in the associated publication and the host species abbreviated names (Table S2). PP: Bayesian Posterior Probability. The host phylogenetic tree is shown on the right. The tree was redrawn from Mamos *et al*., 2016. Small boxes correspond to Motu numbers. Dotted line corresponds to ancient host divergence separating gammarids from north-eastern (N) and south-western (S) regions.

II.3.6 Parasite infections across *G. balcanicus* phylogeny

Linking microsporidia and *G. balcanicus* phylogenetic diversity is challenging for the "rare" microsporidia haplogroups, as the rarity itself challenge any quantitative analysis. First, it is to be pointed out that most (14 out of 30) infected individuals associated with 11 haplogroups of rare microsporidia were found in association with the host MOTU 3.A.b (Fig. 10). The remaining 16 other infected individuals being associated with a total of 9 other host MOTUs (Fig. 10). Second, it is also to be pointed out that some haplogroups *e.g. Cucumispora roeselum* b01 could be both associated with host MOTU 3.A.b, but also distantly related MOTUs 5.E and 7.B.g.

Nosema granulosis parasite was mostly (*c.* 40%) found infecting host Motu 3.A.b. All *N. granulosis* haplogroups (except, of course, the singleton haplogroup b02) were infecting in equal proportion the two clades N and S that emerged from the early diversification of the host (34 infections in the N clade vs 42 infections in the S clade) (Fig. 11). The pattern of distribution of *D. muelleri* infections followed a somewhat similar pattern (*i.e.* presence in both N and S host clades). However, only two parasite haplogroups were found sharing these two host clades (b01 and b10). Other *N. granulosis* haplogroups were either present only in N host clade (*N. granulosis* b02, b03, b05, b11) or in the S host clade (*N. granulosis* b07 and b08) (Fig. 10).

Contrastingly, *D. berillonum* were found infecting only the S clade of the host phylogeny (Fig. 11). Similarly, *D. duebenum* were predominantly infecting host individuals belonging to the S clade, mostly in host MOTU 6.B.l, with nevertheless two notable exceptions (Fig. 11).

The distribution of the variation found in *D. roeselum* was different from all other parasites. While globally *D. roeselum* was found infecting both the N and S host clades (16 infections in the N clade; 30 infections in the S clade), most of the groups of *D. roeselum* variants belonged to a peculiar host clade, or were even restricted to a given host MOTU (Fig. 9). Three of these groups were restricted to the N host clade: *D. roeselum* b18, b15-b16 and b11-b12-b13 (the latter group is restricted to the host MOTU3.A.b, but scattered in diverse populations). On the other side, the group b03-b04-b05-b14 was restricted to a single population and the single host MOTU 7.B.a

belonging to the host clade S. The very same pattern, but in different populations, was found for the parasite group b02-b06 (Fig. 9). Similarly, the haplogroup *D. roeselum* b07 was found in a single host MOTU (7.B.j) of the S host clade, but scattered in 3 populations (Fig. 9). The parasite haplogroups b09 and b10 were also restricted to the Crimean Peninsula and therefore to host MOTUs 3.A.b and 1.A, both belonging to the N host clade. Similarly, haplogroups b01 and b08 were only found infecting hosts from different populations belonging to the S clade (MOTUs 6.B.l, 6.B.d, 7.B.a and 5.E for b01, and MOTUs 6.B.k, 5.B and 5.E for b08) but also AR for b08 a host clade being also present in the south (Fig. 9, Table S1). However, our phylogenetic reconstruction did not allow to resolved phylogenetic affinities within these parasite haplogroups as they are all part of a basal multifurcation, and their relatedness relative to other parasites is questionable. Therefore, the whole *D. roeselum* diversity can be separated into two sets of parasites that seem to parallel the early diversification of the host (N and S).

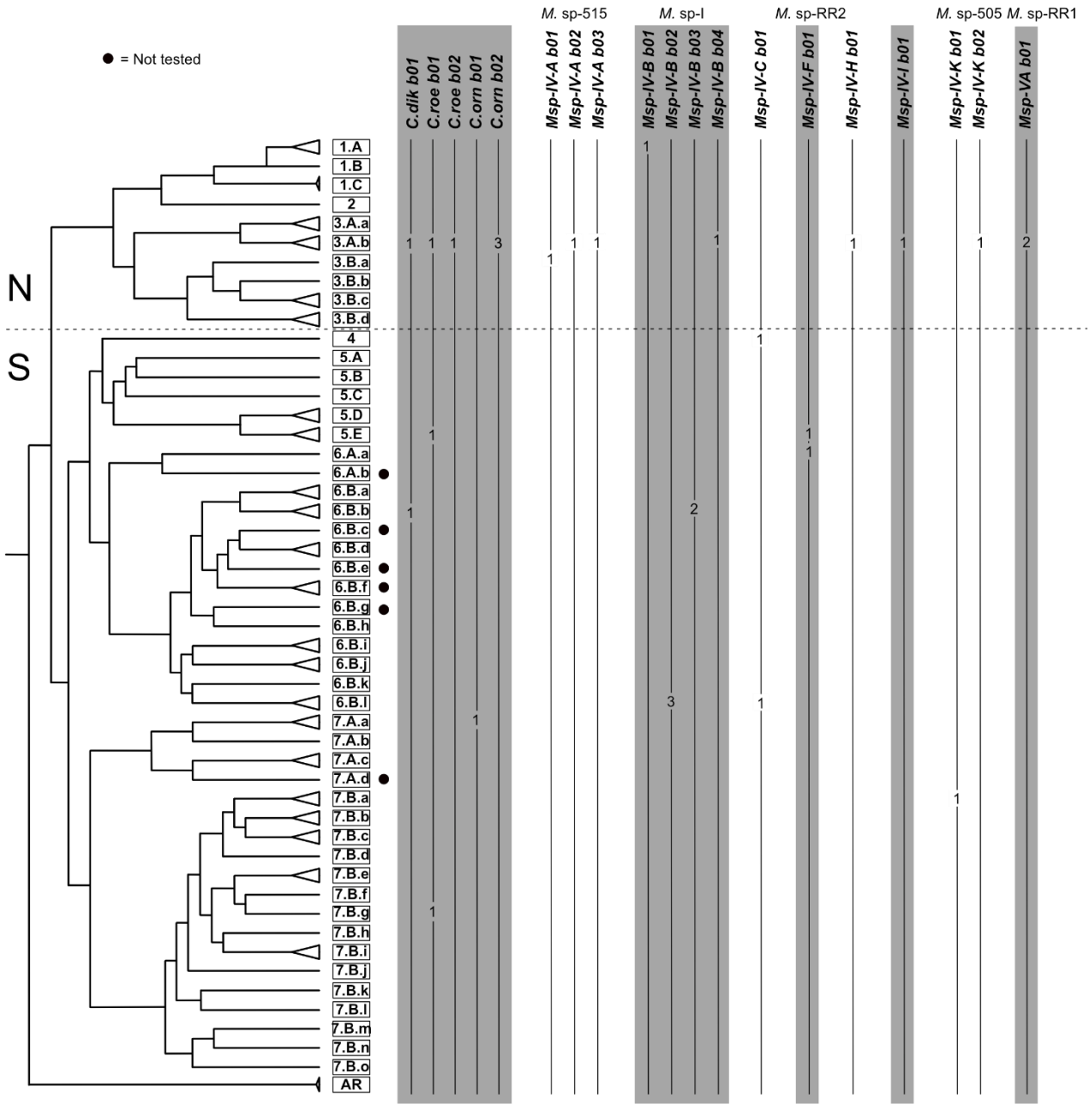


Fig. 10. Overview of microsporidian infections in Gammarus balcanicus according to the host phylogeny, for rare microsporidian infections. The number of infected individuals are provided for each microsporidian haplogroup or species. The host phylogeny and names for Molecular Operational Taxonomic Units (MOTUs) were adapted Mamos *et al.* (2016). Dotted line corresponds to ancient host divergence separating gammarids from north-eastern (N) and south-western (S) regions.

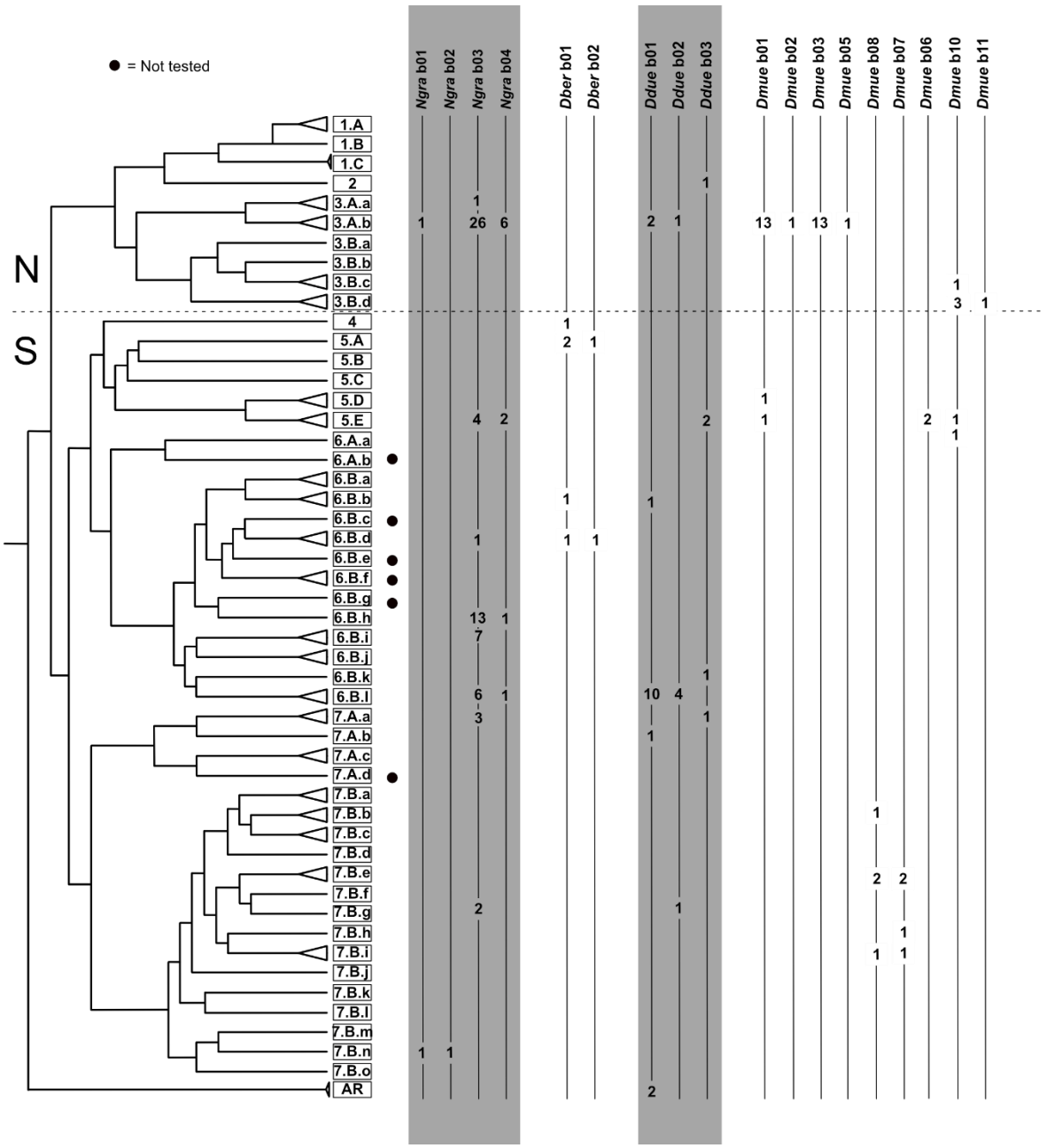


Fig. 11. Overview of microsporidian infections in *Gammarus balcanicus* according to the host phylogeny, for *Nosema granulosis*, *Dictyocoela berillonum*, *D. duebenum* and *D. muelleri*. The number of infected individuals are provided for each microsporidian haplogroup or species. The host phylogeny and the names for Molecular Operational Taxonomic Units (MOTUs) were adapted Mamos *et al.* (2016). Dotted line corresponds to ancient host divergence separating gammarids from north-eastern (N) and south-western (S) regions.

II.3.7 Parasite infections according to host sex

Since infections by vertically-transmitted and feminizing microsporidia are characterized in gammarids by female-biased infections (Haine *et al.* 2004; Terry *et al.* 2004), we investigated this prevalence disequilibrium for *Nosema* and *Dictyocoela* parasites of *G. balcanicus*.

The majority of microsporidia species were found predominantly in males when considering together all host populations, the only exception being *D. duebenum*, where males were as infected as females (Table 1; Table S2). Nevertheless, all microsporidian parasites were found at overall relatively low prevalence in both male and female hosts (maximum *c.* 7.5 %) (Table 1).

Table 1: Microsporidia prevalence in *G. balcanicus* according to host sex

	Males	Females	*P* Fisher Exact test (two-tailed)
Total tested	733	718	
N. granulosis	54 7.37%	20 2.79%	**<0.0001**
D. berillonum	4 0.005%	0	N.D.
D. duebenum	13 1.77%	14 1.95%	0.8478
D. muelleri	37 5.05%	17 2.37%	**0.0080**
D. roeselum	33 4.5%	13 1.81%	**0.0040**

N.D. : not done (inappropriate sample size)

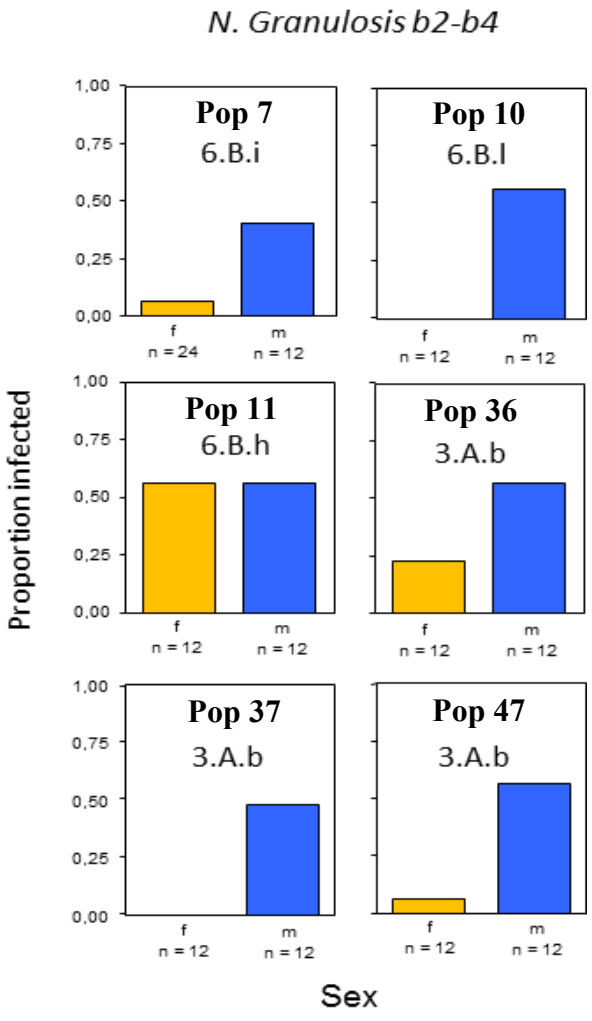
However, when going into details, this apparent homogeneous pattern is hiding a more heterogeneous point of view. For the *Nosema granulosis* close haplogroups b2-b4 (the haplogroup b1 is too rare to be analysed), the prevalence was found to be variable between populations (Fig. 7, Table S1). In two populations, *N. granulosis* infected as many males as females (population 11: Fisher exact test, P = 1 ; population 36: Fisher exact test, P = 0.21). In other populations, females were less infected than males (Fig. 12).

Table 2: Logistic regression analysis. *N. granulosis* prevalence in *G. balcanicus* according to population and host sex (adjusted with the Firth method). In the analysis only populations with > 5 infected individuals were used. (Fig. 7)

Source of variation	d.f.	Likelihood Chi 2	P
Population	5	14.40	0.0133
Sex	1	30.82	<0.0001
Population*Sex	5	10.25	0.0684

These six populations (7, 10, 11, 36, 37 and 47) infected by *N. granulosis b2-b4* harboured *G. balcanicus* belonging to different MOTUs, but there was no discernible pattern of sex-biased infection according to host MOTU (Fig. 11), albeit more populations should be sampled to be affirmative.

Fig. 12. *Nosema granulosis* prevalence (haplogroups b2 to b4) in *G. balcanicus,* according to populations and host sex. Populations presented here are those who harbour at least 5 individuals infected — allowing statistical analysis. The population code is provided, as well as the name of the host MOTU (Table S1).



For *Dictyocoela muelleri*, we found an intriguing infection pattern within the population 78 in Ukraine. By considering the parasite species level, as many males as females were found infected. However, after discriminating the parasite haplogroups, *D. muelleri* haplogroup b3 was found to infect females only, while haplogroup b1 was found mostly in males.

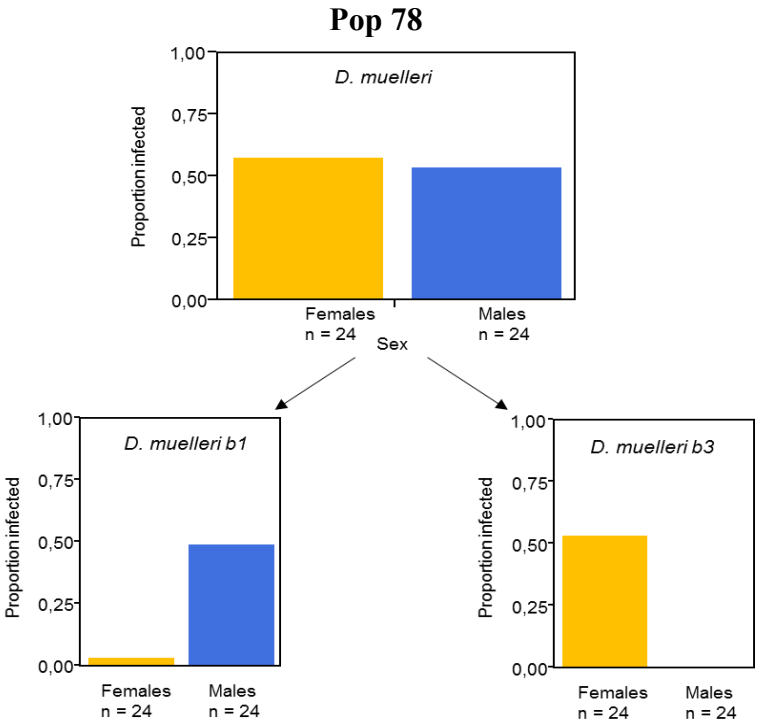


Fig. 13. *Dictyocoela muelleri* prevalence in males and females of *G. balcanicus* in the population 78. Parasites are considered at the species level (upper graph), or after separation of the two parasite haplogroups found in this population (lower graphs).

This highlight the importance of considering parasites at the haplogroup level. The host MOTU in this population is 3.B.a, but can even be restricted to a narrow genetic background being associated to a single BOLD BIN (ACQ2375). For its part, the even repartition of *Dictyocoela duebenum* in *G. balcanicus* of the population 4 in Albania was found whatever the parasite taxonomic level considered (species or haplogroup) (all Fisher exact test $P > 0.10$) (Table 1).

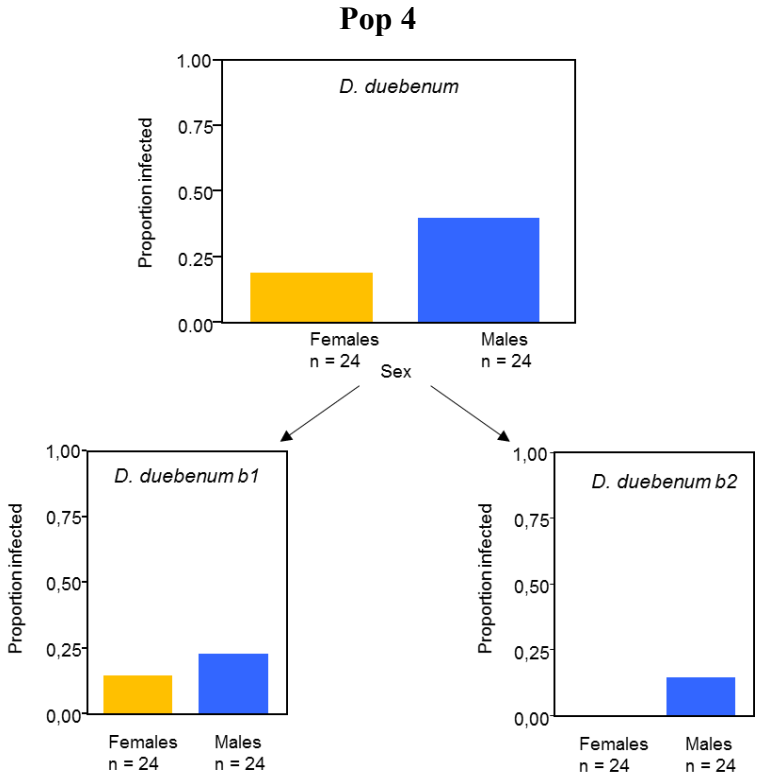


Fig. 14. *Dictyocoela duebenum* prevalence in males and females of *G. balcanicus* in the population 4, when considering parasites at the species level (upper graph), or after separation of the two parasite haplogroups found in this population (lower graphs).

II.4 DISCUSSION

II.4.1 Diversity of microsporidian infections in *Gammarus balcanicus*.

Before the present study, *G. balcanicus* was only seldomly reported as host for microsporidia, with four individuals found infected by parasites belonging to the *Dictyocoela* genus (Bacela-Spychalska *et al.* 2018). Therefore, thanks to our comprehensive sampling over its entire geographic range, our study is the first to highlight, that *Gammarus balcanicus* species complex is infected by 55 microsporidian haplogroups belonging to 17 species-level-taxa. Infections were found in half of the investigated populations, ranging from low prevalence (*c.* 2%) to high prevalence (up to 83.33%). Most of them were associated with Clades III and IV as defined by Vossbrinck and Debrunner-Vossbrinck (2005): only one was assigned to Clade V and none to Clade I and II.

We classified these microsporidia infections into two broad categories related to infection patterns: species-level taxa with rare occurrence (present in ≤ 4 populations) and species-level taxa with frequent occurrence (parasites present in more than 10 populations). Two genera were predominantly found infecting *G. balcanicus*: *Nosema* (Naegeli 1857) and *Dictyocoela* (Bacela-Spychalska *et al.* 2018; Terry *et al.* 1999, 2004).

Rare infections in *G. balcanicus* consisted of 17 parasite haplogroups. Most of these rare microsporidia were phylogenetically closely related to other amphipod parasites. The only exception was microsporidia *M.* sp-IV-I sequence, which is distant to any sequence associated with publication on microsporidia. The closest sequence found in GenBank nevertheless consisted of unpublished microsporidia from Baikal lake, found in *Acanthogammarus lappaceus* host, with 92.4% of identity under 342pb. Therefore, the origin of *M.* sp-IV-I found in a Romanian *G. balcanicus* remains obscure. It could be a parasite transiently acquired from infected sympatric aquatic invertebrate understudied by now, as pointed out by Grabner (2017). Indeed, even if microsporidians are frequently found in the freshwater environment, Stentiford *et al.* (2013) predicted that many thousands of microsporidian taxa remain undescribed in aquatic hosts because they are able to infect a vast taxonomic range of hosts, and because of the relative lack of pathogen census of aquatic organisms.

However, *M.* sp-IV-I is an exception in our study. All other rare microsporidia infecting *G. balcanicus* were very close, if not 100% identical, to already known gammarid microsporidia. They represent a “neglected parasitic fauna” in most of the previous studies (Grabner 2017; Grabner *et al.* 2014, 2015; Krebes *et al.* 2010). Their detection is challenging, and investigating their diversity is possible only associated with extensive sampling either in large scale studies

(Quiles *et al.* 2019) or very local studies (Grabner 2017; Grabner *et al.* 2014, 2015). Most of these rare parasites of *G. balcanicus* share the common characteristics of expanding the geographic range of the microsporidian clade which they belong to (Fig. 4). This is the case for the sister clades *M.* sp-IV-A and *M.* sp-IV-K (Fig. 4A,B), for *M.* sp-V-A (Fig. 4C), *M.* sp-IV-F (Fig. 4D), *Cucumispora roeselum* (Fig. 4G) and *M.* sp-IV-B (Fig. 4H). Most of these parasites were first detected in north-western or north-central Europe and were first thought to be restricted to a single host. For example, *Microsporidium* 515 and 505, close to *M.* sp-IV-A and *M.* sp-IV-K, were initially detected in several populations of *G. duebeni* (Krebes *et al.* 2010). Closely-related microsporidia were then found in *G. fossarum* and *G. pulex* (Grabner 2017) as well as in *G. roeselii* (Quiles *et al.* 2019), extending the host range progressively. Our study therefore expands these observations to central Europe. Similarly, Microsporidia RR1 and RR2 (close to *M*. sp-V-A and *M*. sp-IV-F, respectively) were initially found in a small area of the German Ruhr region (Grabner *et al.* 2015). Quiles *et al.* (2019) then found closely related parasites in *G. roeselii*, but the microsporidian sequences found in *G. balcanicus* considerably extend eastward the geographic range of these parasite clades. A similar pattern was also found for *M*. sp-I parasites. It was first detected in north-western Europe in *G. roeselii* and *G. pulex* (Grabner 2017; Grabner *et al.* 2015; Quiles *et al.* 2019) but also in *Niphargus schellenbergi* (Weigand *et al.* 2016). Here we show that closely related *M.* sp-IV-B is present in *G. balcanicus* in Greece, Albania and Ukraine. Finally, *Cucumispora roeselum*, initially detected in Poland and Western Europe in *G. roeselii* (Bojko *et al.* 2017a; Grabner *et al.* 2015; Quiles *et al.* 2019) is now found in *G. balcanicus* across the whole Balkan region (Fig. 4G). The accumulation of such data now suggests that all these parasites are ancient infections in the European gammarids. Our findings highlight that different haplogroups of these parasites often infect different host species (Fig. 3) suggesting that they may have diversified following the host diversification. However, their phylogenetic closeness more likely also suggest that they are parasites exploiting these crustaceans as generalists, and infect local host species. The parasite genetic differentiation might therefore be due to a geographical differentiation. The problem with these parasites is that their biology is totally unknown preventing further interpretations. All but one have only been detected using molecular tools (like in the present study), with only *C. roeselum* being complemented by an anatomical description (Bojko *et al.* 2017a). It is neither known if they are pathogens, nor their primary transmission way (vertical *vs*. horizontal).

The biology of two other rare microsporidia infecting *G. balcanicus* is better known. Indeed, *Cucumisprora dikerogammari* and *C. ornata* are known to be parasites pathogenic to two hosts, (*Dikerogammarus villosus* and *D. haemobaphes*, respectively) and use mainly horizontal

transmission to infect their hosts (Bojko *et al.* 2017a; Ovcharenko *et al.* 2010). They also have the potential to shift hosts (though in rare cases) and threaten local gammarid species (Bacela-Spychalska *et al.* 2012; Bojko *et al.* 2017a; Quiles *et al.* 2019). The scattered infection of *Cucumispora dikerogammari* in *G. balcanicus* suggests that these infections were probably acquired by interspecific horizontal transfers at sympatric sites. This hypothesis is enhanced by the fact that *C. dikerogammari* is known to mainly infect the invasive species *Dikerogammarus villosus* (Bacela-Spychalska *et al.* 2012; Bojko *et al.* 2017a; Grabner *et al.* 2015; Ovcharenko *et al.* 2010). Therefore, the expanding range of this host species might explain why other local gammarid species can be infected after their arrival. However, the presence of *C. dikerogammari* in a *G. balcanicus* in a Romanian river and in a Greek site is very puzzling, since *D. villosus* is known to be present only in the Romanian part of the Danube and not in this region. With our discovering in *G. balcanicus*, *C. ornata* is now known to infect six gammarid species all over Europe (Fig. 4F) (Bojko *et al.* 2015, 2017a, 2017c). Therefore the scenario proposed for *C. dikerogammari* could also apply if we consider the invasive *D. haemobaphes* being the initial host in the case of *C. ornata* (Bojko *et al.* 2015, 2017a, 2017c).

The two remaining rare microsporidia infecting *G. balcanicus* are more ambiguous. The *M.* sp-IV-C haplogroup's closest relative parasite sequence is from *Enterocytospora artemiae* found in *Artemia franciscana* present in the USA, France, and Israel (Rode *et al.* 2013), but also in *G. roeselii* in a Greece (Quiles *et al.* 2019). *M.* sp-IV-C was found in a Romanian site which is more than 500 km away from the seashore. It is still quite intriguing that these parasites from salt-water can be found in freshwater animals hundreds of kilometre from the sea. We can hypothesis that this infection can come from unstudied aquatic insects or crustaceans that have possibility to live in waters of various salinity (*e.g.* chironomids, cladocerans or copepods). In Quiles *et al.* (2019) study we hypothesis a long-distance transport of microsporidian spores by, *e.g.*, migratory shorebirds, to explain infection of the freshwater *G. roeselii* by a salted-water parasite. This new gammarid host and site finding might be explained by the same hypothesis, but here the distance from the seashore is much larger. It was previously described that some microsporidian species may persist outside their host and may still be horizontally transmitted after spore desiccation (Vizoso *et al.* 2005). We hypothesis that it should be the case for *E. artemiae,* also taking into account that the life cycle of the *Artemia* sp. host is well known to produce resistant thick-shelled eggs (cysts) in case of drying environment. This host eggs can remain in a dormant state, dried for a number of years. This resistance to an extremely dry condition should have been develop through evolution by parasite *E. artemiae*, following coevolution with their hosts, explaining the possibility

for long-lasting travels and possibilities for infecting hosts outside their current habitats. This would nevertheless imply that these parasites are not specific to *Artemia* hosts.

The last parasite, *M.* sp-IV-H is very close to *M.* sp-711 previously found in *Gammarus duebeni* (Krebes *et al.* 2010). Very little is known about this symbiont since our record is only the second one for this microsporidia clade.

II.4.2 Infections ascribed to *Nosema* spp.

Nosema granulosis is known to infect at least seven amphipod species (Haine *et al.* 2004; Ironside 2013; Krebes *et al.* 2010; Terry *et al.* 1999; Wang & Chen 2007; Weigand *et al.* 2016).This study provides the first record of *Nosema granulosis* infecting *G. balcanicus* host.

Four different *Nosema granulosis* haplogroups (b01 to b04) have been detected in *G. balcanicus* host, and seems to be specific to this host, since none of the sequences from this study was identical to *N. granulosis* sequences found infecting other gammarids, although they were phylogenetically close. The four *Nosema* haplogroups were present in both the so-called N and S host clades, which are ca 15-16 Ma old (Fig. 11). This suggests, as already observed for *G. roeselii* (Quiles *et al.* 2019), that the infection with *N. granulosis* seems to be ancient in *G. balcanicus*. As *N. granulosis* is also present in many other gammarid species, such infection could possibly be as old as the early diversification of the genus *Gammarus* itself or even the family Gammaridae. This hypothesis will be interesting to test in the future in a broader phylogenetic host taxa sampling. However, for microsporidia, the SSU rDNA seems not to convey enough informative phylogenetic content to provide a resolved phylogeny at such low taxonomic level (Fig. 5). Thus, additional markers such as RPB1 (Ironside *et al.*, 2013) could be useful to step forward in addressing parasite specificity and host-parasite co-diversification history.

While many *Nosema* species are virulent, horizontally-transmitted parasites (Sprague *et al.* 1992), *Nosema granulosis* is known to be vertically transmitted *via* the eggs of infected females (Dunn & Smith 2001). Previous studies showed that the infection causes little pathogenicity (Kelly *et al.* 2001; Terry *et al.* 1997, 1998). *Nosema granulosis* is also known to induce sex ratio distortion in the populations of *Gammarus duebeni* and *G. roeselii*, by reversing males into functional females (Haine *et al.* 2004, 2007; Jahnke *et al.* 2013; Terry *et al.* 1999). Vertical transmission is associated *via* female hosts because the oocytes are germinal cells containing large cytoplasm which will be also the egg cytoplasm (Hurst & Majerus 1993). Because of feminization, each egg receiving parasites from the mother will develop into the female, and as a result most females are infected (Haine *et al.* 2004; Terry *et al.* 1999). Only a few males may be infected (Haine *et al.* 2004), presumably because of incomplete penetrance of the parasite's feminizing trait or because

few individuals may resist feminization. Therefore, finding female biased ratio of infection in the population is an indirect sign for a vertically-transmitted, feminizing, parasite (Terry *et al.* 1999). We investigated this disequilibrium and found surprisingly a statistically global excess of *N. granulosis* infection in males, with nevertheless variability among population or between host MOTUs (Fig. 12). At least, we never found excesses of infections in females. It is therefore doubtful that *N. granulosis* haplogroups observed in *G. balcanicus* induce sex ratio distortion in this host species. Its vertical transmission is also very questionable, contrary to all previous description of *N. granulosis* (Dunn *et al.* 2006; Kelly *et al.* 2002; Rodgers-Gray *et al.* 2004; Weedall *et al.* 2006).

The excess of infection in males can be explained by various hypotheses: First, it has been found that gammarid males invest less in immunity than females (Rigaud & Moret 2003). Males can therefore be less resistant than female to microsporidia infections, provided that these infections are not vertically-transmitted and have an infectious mode of transmission. Second, gammarid males are always larger than female and consequently need more food uptake. This may enhance the probability of spore ingestion of horizontally transmitted microsporidia during feeding. Furthermore, gammarid males are described to be more cannibalistic then females (Dick 1995), which may enhance the probability of ingesting infected conspecifics (MacNeil *et al.* 2003). Moreover, individuals of *G. duebeni* parasitized by the microsporidia *Pleistophora mulleri* were more likely to be cannibalised by conspecifics (MacNeil *et al.* 2003). Therefore, the combination of higher food uptake and higher cannibalism rate in males, may explain their higher infection rate, especially if infected animals are more prone to be cannibalized. It remains unknown, however, if *G. balcanicus* infected by *N. granulosis* haplogroup b02-b04 are more sensitive to cannibalism compared to uninfected animals.

II.4.3 Infections ascribed to *Dictyocoela* spp.

With 46 populations and 139 individuals infected, infections associated with *Dictyocoela* were the most abundant found in *G. balcanicus*, confirming one more time their status of dominant microsporidian infections in gammarids (Bacela-Spychalska *et al.* 2018; Grabner *et al*. 2015; Quiles *et al.* 2019). We found 31 haplogroups of *Dictyocoela*, corresponding to four previously identified and fully described species (Fig. 6): *D. duebenum* (Fig. 7)*, D. berillonum* (Fig. 7)*, D. muelleri* (Fig. 8)*, D. roeselum* (Fig. 9).

We found 3 haplogroups of *D. duebenum*, a parasite species initially found in *G. duebeni* (Hogg *et al.* 2002) and two haplogroups of *D. berillonum,* a parasite species initially described in *Echinogammarus berilloni* (Terry *et al.* 2004). Both *D. duebenum* and *D. berillonum* are

microsporidia species reported in the literature to infect gammarid species living in north-western Europe (Bacela-Spychalska *et al.* 2018; Grabner *et al.* 2015; Green Etxabe *et al.* 2015; Wilkinson *et al.* 2011). Interestingly, *D. duebenum* was not found in *G. roeselii*, despite being sampled extensively in south-eastern Europe and especially in the Balkans (Quiles *et al.* 2019). On the opposite, the three *D. duebenum* haplogroups found in *G. balcanicus* were distributed homogeneously across the range of the host. Our results highlight the presence of *G. duebenum* in southern – central Europe for the first time. *Dictyocoela duebenum* was described to be vertically-transmitted and distort sex ratio in *Gammarus duebeni duebeni, G. tigrinus* and *E. berilloni* host species (Terry *et al.* 2004). We found no bias in prevalence between sexes, so it seems improbable that *D. duebenum* do bias sex ratio in *G. balcanicus* host (Fig. 14, Table 1).

The two *D. berillonum* haplogroups detected in *G. balcanicus* were found only in the southern part of the host range (Fig. 7), being associated specifically with host MOTUs 4-6 (Fig. 11). Although more geographically restricted than *D. duebenum* for *G. balcanicus*, the present study significantly increases the southward range of *D. berillonum* compared to the literature. It is to be noted, that even if *D. berillonum* have been detected in range-wide analysis of *G. roeselii* by Quiles *et al.* (2019), it was present only in one individual in Poland. *Dictycoela berillonum* was not found showing female-biased distributions in any of its hosts (Terry *et al.* 2004). Unfortunately, we could not test this hypothesis with *G. balcanicus* since we found too few infected individuals for statistical analysis.

Dictyocoela muelleri species was first described infecting *Gammarus duebeni* (Terry *et al.* 2004). In several papers, *D. muelleri* was referred to as a species complex with *D. duebenum* (*e.g. D. muelleri*/*D. duebenum* or *D. sp*) (Grabner *et al.* 2015; Ironside & Wilkinson 2018; Wilkinson *et al.* 2011). Thanks to the recent reassessment of *Dictyocoela* genus made by Bacela-Spychalska *et al.* (2018), we were able to address each SSU sequences to the correct species, without any ambiguity (Fig. 6). Synthesizing the abundant literature reported since its first description *Dictyocoela muelleri* can be described as: 1) harbouring numerous haplogroups, 2) each but one infecting a single host species (last one infecting only two), 3) as being associated with a wide range of amphipod species, in genera such as *Gammarus*, *Pontogammarus* and *Dikerogammarus* (Bacela-Spychalska *et al.* 2018; Krebes *et al.* 2010; Quiles *et al.* 2019; Wattier *et al.* 2007; Wilkinson *et al.* 2011) and 4) all being located in north-eastern Europe.

Our study, adds *Gammarus balcanicus* as another infected taxon, extending southward the range of *D. muelleri* (excluding the southern Balkans), dramatically increases haplogroup diversity while still following the rule of being host species complex restricted (Fig. 8). In addition, *D. muelleri* in *G. balcanicus*, with a total of 54 individuals infected in 17 populations is quantitatively

the most represented *Dictyocoela* species (Fig. 8). Finally, *D. muelleri* haplogroups associated with *G. balcanicus* were scattered across *D. muelleri* phylogeny.

Out of these features, could some host-parasite (*G. balcanicus*-*D. muelleri*) specificity and old association be therefore suggested? (Fig. 8). The interpretation is challenging and one must carefully take into account the relative frequencies of each haplogroup, the associated number of infected populations and the connection with the host diversification history.

For example, haplogroup *D. muelleri* b10 is distantly related to other haplogroups associated with *G. balcanicus,* suggesting an overall old association. Although not very frequent (4 individuals) it is present in four populations, suggesting specificity. On the other hand, the haplogroup b10 is associated with 4 host subclades, 3.B.c, 3.B.d, 5.e and 6.A.a, that diverged from *c*. 13 to *c*. 15 Ma, challenging the idea of a co-diversification. In addition, haplogroup b10 is molecularly very close to *D. muelleri* found infecting two Ponto-Caspian amphipods, *Dikerogammarus haemobaphes* and *Pontogammarus robustoides*, both species being known to occur in the same region as *G. balcanicus*. Therefore, a recent acquisition by horizontal transfer from the local fauna seems a realistic scenario.

Haplogroup *D. muelleri* b11, being present in a single individual in the same region as haplogroup b10, and also molecularly close to a parasite found in Ponto-Caspian hosts, is also a good candidate for a horizontal transfer from the local fauna.

Haplogroup *D. muelleri* b06, given its phylogenetic position, rarity (1 pop and 2 individuals) deficiencies formulating a scenario. It can be another case of horizontal transfer from the local fauna.

For all *D. muelleri* b10, b11 and b06, testing the hypothesis of horizontal transient transfers from sympatric hosts through, would obviously need the screening of the local fauna.

The remaining six *D. muelleri* haplogroups (b01, b02, b03, b05, b07, b08) could be considered at first sight to form a set based on molecular proximity, being also, as a set, both most abundant (81.25%) and widespread geographically (Fig. 8). Based on these features, one could reasonably hypothesize this *D. muelleri* set of haplogroups as being specific infections of *G. balcanicus* host. However, first, there is an apparent dichotomy in geographic distribution of the haplogroups, as haplogroups b07, b08 are restricted to a few populations in the south-west range of *G. balcanicus* (being associated to four 7.B MOTUs.) while the others are scattered within the north-east range, in western Carpathian (Fig. 8). In addition, while haplogroups 02, 03, 05 are restricted to the host MOTU 3.A.b, the haplogroup b01 was also found in host MOTU 3.A.b, 5.D

and 5.E, which diverged from 3.A.b more than 15 Ma (clades N and S on Fig. 11). Such results are puzzling. All the haplogroups of this set are molecularly close, part of a polytomy, while one would expect some clear dichotomies to appear, paralleling host diversification history. The low-resolution power of our genetic marker might be the reason why the phylogenetic relationships among these haplogroups is not solved, so the precise divergence history of this set of haplogroups remains obscure. Exploring further the phylogenetic relationship within this set of haplogroups might benefit from applying in the future potentially more informative marker such as RPB1 (Brandon Matheny *et al.* 2002; Nocybe *et al.* 2002; Xu & Zhou 2010). The haplogroup b01 associated with two divergent hosts MOTUs would be a good candidate for horizontal transfer.

In literature, *D. muelleri* was described to have a female-biased distribution in *Gammarus roeselii* but not in *Gammarus duebeni celticus* (Terry *et al.* 2004). Our result for *D. muelleri* infecting *G. balcanicus* showed a sex bias in prevalence, but toward males. This male-biased prevalence can be explained by the three hypotheses previously described for *N. granulosis*. However, we found an intriguing pattern of prevalence according to host sex in the host population 78. Considering the parasite species level, as many males as females were found infected. However, after discriminating the parasite haplogroups, *D. muelleri* haplogroup b03 was found to infect females only, while haplogroup b01 was found mostly in males. This highlight the importance of considering parasites at the haplogroup level to understand host-parasite co-evolution, but this apparent specialization of parasite strains according to host sex is confusing. While it is possible to imagine that one *D. muelleri* strain (here b03) specialized in vertical transmission and feminization, like in *Gammarus roeselii* (Haine *et al.* 2004, 2007), it is difficult to imagine a strain specialized in horizontal transmission almost only by the male pathway. One hypothesis can nevertheless be proposed for this unusual parasite prevalence pattern. It is known that the presence of some parasite or symbiont species confers protection to their hosts from other pathogenic infections (Haine 2008). This is particularly true for vertically-transmitted symbionts such as *Wolbachia* which confer protection for their insect hosts against a range of pathogens including bacteria, viruses, nematodes and the malaria parasite (Ye *et al.* 2013) or symbiotic bacteria of the decapod crustacean, *Palaemon macrodactylus*, which can chemically defend shrimp embryos from a pathogenic fungus (Gil-Turnes *et al.* 1989). It could therefore be possible that *D. muelleri* haplogroup b03 infecting only females can protect this host gender from infected male with *D. muelleri* b01. This would suppose that *D. muelleri* haplogroup b03 (♀) is a vertically-transmitted feminizing parasite, while *D. muelleri* haplogroup b01 (♂) is a horizontally-transmitted pathogenic parasite. Vertically-transmitted microsporidia were shown to confer protection against an acanthocephalan macroparasite in *Gammarus roeselii* (Haine *et al.* 2005).

This protection was not, nevertheless, against superinfection, but against the effects of this macro-parasite. Clearly, this "protection" hypothesis in *Gammarus balcanicus* requires further experimental tests.

Dictyocoela roeselum is the second most abundant *Dictyocoela* found in our survey infecting 50 *G. balcanicus* individuals. This microsporidian was firstly described infecting *Gammarus roeselii* (Haine *et al.* 2004) and shown to be vertically transmitted (Haine *et al.* 2004). Since then, this microsporidia species have been found infecting many *Gammarus* species, but at low rates (*G. pulex, G. fossarum, G. varsoviensis, G. duebenum, G. lacustris*) (Bacela-Spychalska *et al.* 2018; Quiles *et al.* 2019; Wattier *et al.* 2007) (Fig. 9, Table S2). *Dictyocoela roeselum* is the most diversified microsporidium species found in this study with 17 different haplogroups. The total number of infected hosts and the number of populations with high prevalence is analogous to previous studies on *G. roeselii* (Haine *et al.* 2004; Quiles *et al.* 2019).

Within *D. roeselum* tree (Fig. 9), we reinforce the previous findings that each host species is infected with a different *D. roeselum* haplogroup (Bacela-Spychalska *et al.* 2018; Quiles *et al.* 2019). Hence, our findings strengthened the previous hypothesis of a probable host specificity within *D. roeselum* species complex (Haine *et al.* 2004; Quiles *et al.* 2019).

The biogeography of *D. roeselum* showed a strict dichotomy between haplogroups (or groups of closely-related haplogroups forming discrete and well-supported clades) present in the south-western part of the range of the host and those present in the north-eastern part of the range (Fig. 9). The microsporidian haplogroups *D. roeselum* b11-b12-b13 (highlighted in red in Fig. 9), *D. roeselum* b15-b16 (in yellow), and *D. roeselum* b18 (green) is associated to the host clade N (Fig. 9) which is restricted to the north-eastern part of the range of *G. balcanicus.* The haplogroups *D. roeselum* b02-b06 (dark blue), *D. roeselum* b07 (light blue) and *D. roeselum* b03-b04-b05-b14 (purple) are present in the south-western host clades only (Fig. 9). This suggests an ancient infection by this parasite, before the early diversification of the host, and a subsequent diversification within each host clade N and S, following the host diversification. The poor phylogenetic resolution between the different parasite clades did not allow to go further in our interpretation, but it is worth noting that some of the different *D. roeselum* haplogroups were found within a single host MOTU (*e.g.* MOTU 7.B.a, Fig. 9), but also within single populations. Indeed, the 4 closely-related haplogroups *D. roeselum* b03, b04, b05 and b14 were found in a single population, as well as *D. roeselum* b02 and b06, *D. roeselum* b15 and b16 and the *D. roeselum* b11, b12 and b13. All these observations plead for several recent independent local diversifications of *D. roeselum*. Such diversifications would be compatible with some degrees of local host-parasite co-evolutions. Finally, some haplogroups *D. roeselum* b01, b08, b09, and b10 (in white

on Fig. 9) were missing molecular information to be topologically supported in the tree. *D. roeselum* b09 and b10 were restricted to the Crimean Peninsula (and the associated host MOTUs, see Table S1), while *D. roeselum* b01 and b08 spread in the southern hots MOTUs / geographic area (Fig. 9, Table S1). Another, more variable, molecular marker for the parasite, , would be supportive in the future to solve the possible host-parasite co-phylogeny within each N and S host clades.

Dictyocoela roeselum was observed in eggs of *G. roeselii* (therefore are vertically-transmitted) and infect significantly more females than males, which was interpreted as feminization (Haine *et al.* 2004, 2005). In *G. balcanicus*, such a pattern was not found, and a vertical transmission associated with feminization is therefore unlikely in this host species. A pathogenic, horizontally-transmitted parasite would be much more consistent with the observation that several haplogroups recently evolved within populations to override the resistance, which would be an indication of a red-queen process between host genotypes resisting to pathogenic parasites and parasite genotypes overriding this resistance (Ebert 2008; Lively *et al.* 1990). The excess of parasites in males may be explained by the same hypothesis as already proposed for *N. granulosis*.

The new microsporidia sequences found in *G. balcanicus* extend substantially the range of many microsporidia species once found in *G. roeselii*. *Gammarus roeselii* and *G. balcanicus* have overlapping geographical ranges in the Balkans. Nevertheless, only five (18, 35, 37, 39 and 43) out of 88 populations were presenting cases of sympatry between *G. roeselii* and *G. balcanicus.* The two species do not share the same habitat. *Gammarus balcanicus* is usually found in higher altitudes, living more upstream river than *G. roeselii,* but with some contacts still possible. Nevertheless, no microsporidia haplogroups o were shared between the two host species. We can hypothesise that microsporidia infection, especially for the preferentially vertically transmitted microsporidia, can be as old as the early diversification of the genus *Gammarus* or even to the level of the Gammaridae family. Some other species have infected *Gammarus* host more recently *e.g.* “rare” microsporidia (*M.* sp 505, 515, RR1, RR2) previously found in *G. roeselii* and other Amphipoda without any clue for host-parasite evolutionary scenario. Finally, some sporadic microsporidia such as *Enterocytospora artemiae* found in both hosts seem to be transient associations in both host species.

II.5 CONCLUSIONS

Infection with microsporidia in *Gammarus balcanicus* is common, widespread and highly diverse. Most of the microsporidia taxa were previously found in other closely related and well-

studied gammarids such as *G. roeselii* or *G. duebeni* (Krebes *et al.* 2010; Quiles *et al.* 2019). The dominant species found infecting predominantly *G. roeselii* and *G. duebeni* were the same in *G. balcanicus*, with noticeable exceptions (*e.g.* M. sp. 505 or 515, widespread in *G. duebeni*, but rare in *G. balcanicus*). However, microsporidia haplogroups are not mixing between *G. balcanicus* and other host species, and often form host-specific clades in the microsporidia phylogenetic tree. Moreover, for three abundant microsporidia species, *Nosema granulosis*, *Dictyocoela muelleri* and *Dictyocoela roeselum*, the pattern of parasite genetic diversification may be explained to a certain extent by the pattern of ancient host diversification. This hypothesis is strengthened by the observations that *G. balcanicus* diversification led to emergence of highly endemic MOTUs (that can be seen as "species") (Mamos *et al.* 2016).

Such patterns contrasted with the microsporidian haplogroups that were found identical across all range of *G. balcanicus*. These microsporidia showed no link to the evolutionary history of the host. Almost all of the rare microsporidia (*e.g. Cucumispora* spp.) or *Dictyocoela duebenum* and *D. berillonum* presented this pattern. Most of them were found in a few individuals and some of them were found associated with many other host species. For these microsporidia, the most reasonable hypothesis for their presence in *G. balcanicus* is acquisitions through horizontal transmission, by host-shift from local gammarid species. Indeed, host-shifts are more probable between phylogenetically-related hosts than between unrelated ones (Davies & Pedersen 2008; Pilosof *et al.* 2015). An alternative hypothesis is a spill-over event (*i.e.* transient infections) from an infected prey passing through the gut of the studied gammarids in which the infection is not sustainable. Our findings are congruent with the previous findings in *G. roeselii* (Quiles *et al.* 2019) but also with local assemblages, where microsporidia haplogroups are shared between different hosts (Grabner *et al.* 2015). However, the same kind of pattern was found for the abundant *Nosema granulosis*: each of the four haplogroups was scattered throughout the host geographical range. Here rare spill-over events cannot explain this pattern. Therefore, perhaps the rate of molecular evolution is not the same between the different microsporidian species, or selective pressures due to different life-histories led to the contrasted observed patterns. For example, as already proposed, pathogenic effects of *Dictyocoela roeselum* might explain the high local within-population diversification rate of this parasite, because of a red-queen process, while a more benign parasite effect would explain why *N. granulosis* did not diversify locally. A better explanation for the phylogeographic pattern of *N. granulosis* variation would be an ancient diversification, before the host diversification, followed by persistence of parasites in most of the host populations during this diversification (*i.e.* weak phenomenon of parasite loss), and in parallel a phase of stasis at the level of parasite molecular evolution.

SUPPLEMENTARY MATERIAL

Chap. II, Table S1. Microsporidian infections in the 88 populations investigated over the geographical range of *Gammarus balcanicus*.

Chap. II, Table S2. Individual data for microsporidian infections from this study and found in GenBank (NCBI), mainly for freshwater and brackish waters amphipod species occurring in Europe.

Chap. II, Table S3. Variable sites in sequences of the microsporidian species used to construct phylogenetic trees (Figs. 2; 4-9).

Chap. II, Data S1. Alignments based on SSU rDNA sequences, used for trees in Figs. 2; 4-9.

Chap. II, Data S2. Alignments based on SSU rDNA sequences, used for Additional file 2: Table S2.

All suplementary material are available at:
https://cloud.u-bourgogne.fr:443/index.php/s/XP9ZW8wnr87kP7z

or scanning the following QR code:

REFERENCES

Bacela-Spychalska, K., Wattier, R. A., Genton, C., & Rigaud, T. (2012). Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. *Biological Invasions,* 14(9), 1831–1842.

Bacela-Spychalska, K., Wróblewski, P., Mamos, T., … Ovcharenko, M. (2018). Europe-wide reassessment of *Dictyocoela* (Microsporidia) infecting native and invasive amphipods (Crustacea): Molecular versus ultrastructural traits. *Scientific Reports*, 8(1), 1–16.

Becnel, J. J., & Andreadis, T. G. (2014). Microsporidia in Insects. Microsporidia: *Pathogens of Opportunity*: First Edition, 521–570.

Becnel, J. J., & Takvorian, P. M. (2014). Checklist of Available Generic Names for Microsporidia with Type Species and Type Hosts. Microsporidia: *Pathogens of Opportunity*: First Edition, 671–686.

Berbee, M. L., James, T. Y., & Strullu-Derrien, C. (2017). Early Diverging Fungi: Diversity and Impact at the Dawn of Terrestrial Life. *Annual Review of Microbiology*, 71(1), 41–60.

Blondel, J. (Jacques), & Blondel, J. (Jacques). (2010). The Mediterranean region : biological diversity in space and time, *Oxford University Press*.

Bojko, J., Bącela-Spychalska, K., Stebbing, P. D., … Stentiford, G. D. (2017a). Parasites, pathogens and commensals in the "low-impact" non-native amphipod host *Gammarus roeselii. Parasites & Vectors*, 10(1), 193.

Bojko, J., Clark, F., Bass, D., … Stentiford, G. D. (2017b). *Parahepatospora carcini* n. gen., n. sp., a parasite of invasive *Carcinus maenas* with intermediate features of sporogony between the Enterocytozoon clade and other microsporidia. *Journal of Invertebrate Pathology*, 143, 124–134.

Bojko, J., Dunn, A. M., Stebbing, P. D., Ross, S. H., Kerr, R. C., & Stentiford, G. D. (2015). *Cucumispora ornata* n. sp. (Fungi: Microsporidia) infecting invasive "demon shrimp" (*Dikerogammarus haemobaphes*) in the United Kingdom. *Journal of Invertebrate Pathology*, 128, 22–30.

Bojko, J., & Ovcharenko, M. (2018). Pathogens and other symbionts of the Amphipoda: a review of their taxonomic diversity and pathological significance. *Diseases of Aquatic Organisms*. doi:10.3354/dao03321

Bojko, J., Stentiford, G., Stebbing, P., … Dunn, A. (2017c). Pathogens of *Dikerogammarus haemobaphes* regulate host activity and survival, but also threaten native amphipod populations in the UK. *Diseases of Aquatic Organisms*, View(January), 1–16.

Brandon Matheny, P., Liu, Y. J., Ammirati, J. F., & Hall, B. D. (2002). Using RPB1 sequences to improve phylogenetic inference among mushrooms (Inocybe, Agaricales). *American Journal of Botany*, 89(4), 688–698.

Bulnheim, P. (1975). Microsporidian Infections of Amphipods with Special Reference to Host-Parasite Relationships : *A Review. Marine Fisheries Review*, 37, 39–45.

Capella-Gutiérrez, S., Marcet-Houben, M., & Gabaldón, T. (2012). Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. *BMC Biology,* 10(1), 47.

Conow, C., Fielder, D., Ovadia, Y., & Libeskind-Hadas, R. (2010). Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms for Molecular Biology* : AMB, 5(1), 16.

Copilaş-Ciocianu, D., & Petrusek, A. (2017). Phylogeography of a freshwater crustacean species complex reflects a long-gone archipelago. *Journal of Biogeography*, 44(2), 421–432.

Corradi, N., & Selman, M. (2013). Latest progress in microsporidian genome research. *Journal of Eukaryotic Microbiology,* 60(3), 309–312.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). JModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772.

Davies, T. J., & Pedersen, A. B. (2008). Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings of the Royal Society B: Biological Sciences*, 275(1643), 1695–1701.

Dick, J. T. A. (1995). The cannibalistic behaviour of two *Gammarus* species (Crustacea: Amphipoda). *Journal of Zoology*, 236(4), 697–706.

Dimova, M., Madyarova, E., Gurkov, A., … 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.

Dubuffet, A., Smith, J. E., Solter, L., Alejandra Perotti, M., Braig, H. R., & Dunn, A. M. (2013). Specific detection and localization of microsporidian parasites in invertebrate hosts by using in situ hybridization. *Applied and Environmental Microbiology*, 79(1), 385–388.

Dunn, A. M., Hogg, J. C., & Hatcher, M. J. (2006). Transmission and burden and the impact of temperature on two species of vertically transmitted microsporidia. *International Journal for Parasitology*, 36(4), 409–414.

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

Ebert, D. (2008). Host-parasite coevolution: Insights from the Daphnia-parasite model system. *Current Opinion in Microbiology*, 11(3), 290–301.

Ebert, D. (2013). The epidemiology and evolution of symbionts with mixed-mode transmission. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 623–643.

Gil-Turnes, M. S., Hay, M. E., & Fenical, W. (1989). Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. *Science*, 246(4926), 116–118.

Grabner, D. S. (2017). Hidden diversity: parasites of stream arthropods. *Freshwater Biology,* 62(1), 52–64.

Grabner, D. S., Schertzinger, G., & Sures, B. (2014). Effect of multiple microsporidian infections and temperature stress on the heat shock protein 70 (hsp70) response of the amphipod *Gammarus pulex*. *Parasites & Vectors*, 7(1), 170.

Grabner, D. S., Weigand, A. M., Leese, F., … Sures, B. (2015). Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. *Parasites and Vectors*, 8(1), 1–4.

Grabowski, M., Krzywozniak, P., Rewicz, T., Mamos, T., Bacela-spychalska, K., & Wattier, R. (2017a). *Gammarus roeselii* Gervais , 1835 (Gammaridae) in Western and Central Europe : post-glacial colonisation or human mediated introduction ? *Biodiversity Journal,* 8(2), 525–526.

Grabowski, M., Mamos, T., Bącela-Spychalska, K., Rewicz, T., & Wattier, R. A. (2017b). Neogene paleogeography provides context for understanding the origin and spatial distribution of cryptic diversity in a widespread Balkan freshwater amphipod. *PeerJ*, 5, e3016.

Green Etxabe, A., Short, S., Flood, T., Johns, T., & Ford, A. T. (2015). Pronounced and prevalent intersexuality does not impede the 'Demon Shrimp' invasion. *PeerJ*, 3, e757.

Haine, E. R. (2008). Symbiont-mediated protection. *Proceedings of the Royal Society B: Biological Sciences*, 275(1633), 353–361.

Haine, E. R., Boucansaud, K., & Rigaud, T. (2005). Conflict between parasites with different transmission strategies infecting an amphipod host. *Proceedings of the Royal Society B: Biological Sciences*, 272(1580), 2505–2510.

Haine, E. R., Brondani, E., Hume, K. D., Perrot-Minnot, M. J., Gaillard, M., & Rigaud, T. (2004). Coexistence of three microsporidia parasites in populations of the freshwater amphipod *Gammarus roeseli*: Evidence for vertical transmission and positive effect on reproduction. *International Journal for Parasitology*, 34(10), 1137–1146.

Haine, E. R., Motreuil, S., & Rigaud, T. (2007). Infection by a vertically-transmitted microsporidian parasite is associated with a female-biased sex ratio and survival advantage in the amphipod *Gammarus roeseli. Parasitology*, 134(Pt 10), 1363–1367.

Hibbett, D. S., Binder, M., Bischoff, J. F., … Zhang, N. (2007). A higher-level phylogenetic classification of the Fungi. *Mycological Research*, 111(5), 509–547.

Hogg, J. C., Ironside, J. E., Sharpe, R. G., Hatcher, M. J., Smith, J. E., & Dunn, a M. (2002). Infection of *Gammarus duebeni* populations by two vertically transmitted

microsporidia; parasite detection and discrimination by PCR-RFLP. *Parasitology*, 125(Pt 1), 59–63.

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

Hurst, G. D. D., & Majerus, M. E. N. (1993). Why do maternally inherited microorganisms kill males? *Heredity*, 71(1), 81–95.

Ironside, J. E. (2007). Multiple losses of sex within a single genus of Microsporidia. *BMC Evolutionary Biology,* 7(1), 48.

Ironside, J. E. (2013). Diversity and recombination of dispersed ribosomal DNA and protein coding genes in Microsporidia. *PLoS ONE*, 8(2). doi:10.1371/journal.pone.0055878

Ironside, J. E., & Alexander, J. (2015). Microsporidian parasites feminise hosts without paramyxean co-infection: Support for convergent evolution of parasitic feminisation. *International Journal for Parasitology*, 45(6), 427–433.

Ironside, J. E., Smith, J. E., Hatcher, M. J., Sharpe, R. G., Rollinson, D., & Dunn, A. M. (2003). Two species of feminizing microsporidian parasite coexist in populations of *Gammarus duebeni*. *Journal of Evolutionary Biology,* 16(3), 467–473.

Ironside, J. E., & Wilkinson, T. J. (2018). Accumulation and exchange of parasites during adaptive radiation in an ancient lake. *International Journal for Parasitology*, 48(3–4), 297–307.

Ironside, J. E., Wilkinson, T. J., & Rock, J. (2008). Distribution and host range of the microsporidian *Pleistophora mulleri*. *Journal of Eukaryotic Microbiology*, 55(4), 355–362.

Jahnke, M., Smith, J. E., Dubuffet, A., & Dunn, A. M. (2013). Effects of feminizing microsporidia on the masculinizing function of the androgenic gland in *Gammarus duebeni*. *Journal of Invertebrate Pathology*, 112(2), 146–151.

Karaman, G. S., & Pinkster, S. (1977a). Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea-Amphipoda). Part II. Gammarus roeseli-group and related species. Bijdragen Tot de Dierkunde, 47(1), 165–196.

Karaman, G. S., & Pinkster, S. (1977b). Freshwater Gammarus species from Europe, North Africa and adjacent regions of Asia (Crustacea - Amphipoda) Part I. *Gammarus pulex*-group and related species. *Bijdragen Tot De Dierkunde*, 47(1), 1–86.

Katoh, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780.

Kearse, M., Moir, R., Wilson, A., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28(12), 1647–1649.

Keeling, P. J., & Fast, N. M. (2002). Microsporidia: Biology and Evolution of Highly Reduced Intracellular Parasites. *Annual Review of Microbiology*, 56(1), 93–116.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2001). Population dynamics of a vertically transmitted, parasitic sex ratio distorter and its amphipod host. *Oikos*, 94(3), 392–402.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2002). Incomplete feminisation by the microsporidian sex ratio distorter, *Nosema granulosis*, and reduced transmission and feminisation efficiency at low temperatures. *International Journal for Parasitology*, 32(7), 825–831.

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. *Infection, Genetics and Evolution*, 10(7), 1027–1038.

Krebes, L., Zeidler, L., Frankowski, J., & Bastrop, R. (2014). (Cryptic) sex in the microsporidian *Nosema granulosis* - Evidence from parasite rDNA and host mitochondrial DNA. *Infection, Genetics and Evolution*, 21, 259–268.

Lively, C. M., Craddock, C., & Vrijenhoek, R. C. (1990). Red Queen hypothesis supported by parasitism in sexual and clonal fish. *Nature*, pp. 864–866.

MacNeil, C., Dick, J. T. A., Hatcher, M. J., Fielding, N. J., Hume, K. D., & Dunn, A. M. (2003). Parasite transmission and cannibalism in an amphipod (Crustacea). *International Journal for Parasitology*, 33(8), 795–798.

Madden, T. (2003). The BLAST Sequence Analysis Tool. In The NCBI Handbook, *National Center for Biotechnology Information (US),* pp. 1–15.

Mamos, T., Wattier, R., Burzyński, A., & Grabowski, M. (2016). The legacy of a vanished sea: A high level of diversification within a European freshwater amphipod species complex driven by 15 Ma of Paratethys regression. *Molecular Ecology*, 25(3), 795–810.

Mamos, T., Wattier, R., Majda, A., Sket, B., & Grabowski, M. (2014). Morphological vs. Molecular delineation of taxa across montane regions in Europe: The case study of *Gammarus balcanicus* Schäferna, (Crustacea: Amphipoda). *Journal of Zoological Systematics and Evolutionary Research,* 52(3), 237–248.

Médail, F., & Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, 36(7), 1333–1345.

Naegeli, K. (1857). Über die neue Krankheit der Seidenraupe und verwandte Organismen. *Bot Ztg*, 15, 760–761.

Nocybe, I., Garicales, A., Atheny, P. B. R. M., Iu, Y. A. J. L., Mmirati, J. O. F. A., & All, B. E. D. H. (2002). U Sing Rpb1 Sequences To Improve Phylogenetic, 89(4), 688–698.

Ovcharenko, M., 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(2), 191–204.

Pilosof, S., Morand, S., Krasnov, B. R., & Nunn, C. L. (2015). Potential parasite transmission in multi-host networks based on parasite sharing. *PLoS ONE,* 10(3), e0117909.

Popov, S. V., Rögl, F., Rozanov, A. Y., Steininger, F. F., Shcherba, I. G., & Kovac, M. (2004). Lithological-paleogeographic maps of paratethys. CFS Courier Forschungsinstitut Senckenberg, Schweizerbart'sche Verlagsbuchhandlung.

Poulakakis, N., Kapli, P., Lymberakis, P., … Mylonas, M. (2015). A review of phylogeographic analyses of animal taxa from the aegean and surrounding regions. *Journal of Zoological Systematics and Evolutionary Research,* 53(1), 18–32.

QGIS Development Team. (2009). QGIS {Geographic} {Information} {System}. *Open Source Geospatial Foundation.* Retrieved from http://qgis.osgeo.org

Quiles, A., Bacela-Spychalska, K., Teixeira, M., … 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. *Parasites and Vectors*, 12(1), 1–20.

Rigaud, T., & Moret, Y. (2003). Differential phenoloxidase activity between native and invasive gammarids infected by local acanthocephalans: Differential immunosuppression? *Parasitology*, 127(6), 571–577.

Rode, N. O., Lievens, E. J. P., Segard, A., Flaven, E., Jabbour-Zahab, R., & Lenormand, T. (2013). Cryptic microsporidian parasites differentially affect invasive and native *Artemia* spp. *International Journal for Parasitology*, 43(10), 795–803.

Rodgers-Gray, T. P., Smith, J. E., Ashcroft, A. E., Isaac, R. E., & Dunn, A. M. (2004). Mechanisms of parasite-induced sex reversal in *Gammarus duebeni* . *International Journal for Parasitology*, 34(6), 747–753.

Slothouber Galbreath, J. G. M., Smith, J. E., Becnel, J. J., Butlin, R. K., & Dunn, A. M. (2009). Reduction in post-invasion genetic diversity in *Crangonyx pseudogracilis* (Amphipoda: Crustacea): A genetic bottleneck or the work of hitchhiking vertically transmitted microparasites? *Biological Invasions*, 12(1), 191–209.

Sprague, V., Becnel, J. J., & Hazard, E. I. (1992). Taxonomy of phylum microspora. *Critical Reviews in Microbiology*, 18(5–6), 285–395.

Stentiford, G. D., Becnel, J. J., Weiss, L. M., … Solter, L. (2016). Microsporidia – Emergent Pathogens in the Global Food Chain. *Trends in Parasitology*, 32(4), 336–348.

Stentiford, G. D., Feist, S. W., Stone, D. M., Bateman, K. S., & Dunn, A. M. (2013). Microsporidia: Diverse, dynamic, and emergent pathogens in aquatic systems. *Trends in Parasitology*, 29(11), 567–578.

Terry, R. S., Dunn, A. M., & Smith, J. E. (1997). Cellular distribution of a feminizing microsporidian parasite: A strategy for transovarial transmission. *Parasitology*, 115(2), 157–163.

Terry, R. S., Smith, J. E., Bouchon, D., … Dunn, A. M. (1999). Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF) microsporidium. *Journal of Eukaryotic* Microbiology, 46(5), 492–499.

Terry, R. S., Smith, J. E., & Dun, A. M. (1998). Impact of a novel, feminising microsporidium on its crustacean host. *J. Euk. Microbiol*., (5), 497–501.

Terry, R. S., Smith, J. E., Sharpe, R. G., … Dunn, A. M. (2004). Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum Microspora. *Proceedings of the Royal Society B: Biological Sciences*, 271(1550), 1783–1789.

Vizoso, D. B., Lass, S., & Ebert, D. (2005). Different mechanisms of transmission of the microsporidium *Octosporea bayeri*: A cocktail of solutions for the problem of parasite permanence. *Parasitology*, 130(5), 501–509.

Vossbrinck, C. R., & Debrunner-Vossbrinck, B. A. (2005). Molecular phylogeny of the Microsporidia: Ecological, ultrastructural and taxonomic considerations. *Folia Parasitologica,* 52(1–2), 131–142.

Vossbrinck, C. R., Debrunner-Vossbrinck, B. A., & Weiss, L. M. (2014). Phylogeny of the Microsporidia. *In Microsporidia: Pathogens of Opportunity: First Edition, Chichester, UK*: John Wiley & Sons, Inc., pp. 203–220.

Wang, W., & Chen, J. (2007). Ultrastructural study on a novel microsporidian, *Endoreticulatus eriocheir* sp. nov. (Microsporidia, Encephalitozoonidae), parasite of Chinese mitten crab, *Eriocheir sinensis* (Crustacea, Decapoda). *Journal of Invertebrate Pathology,* 94(2), 77–83.

Wattier, R. A., Haine, E. R., Beguet, J., … Rigaud, T. (2007). No genetic bottleneck or associated microparasite loss in invasive populations of a freshwater amphipod. *Oikos,* 116(11), 1941–1953.

Weber, R., Bryan, R. T., Schwartz, D. A., & Owen, R. L. (1994). Human microsporidial infections. *Clinical Microbiology Reviews*, 7(4), 426–461.

Weedall, R. T., Robinson, M., Smith, J. E., & Dunn, A. M. (2006). Targeting of host cell lineages by vertically transmitted, feminising microsporidia. *International Journal for Parasitology,* 36(7), 749–756.

Weigand, A. M., Kremers, J., & Grabner, D. S. (2016). Shared microsporidian profiles between an obligate (*Niphargus*) and facultative subterranean amphipod population (*Gammarus*) at sympatry provide indications for underground transmission pathways. *Limnologica*, 58, 7–10.

Weiss, L. M., & Becnel, J. J. (2015). Microsporidia: pathogens ofopportunity. *Ames: John Wiley & Sons*, (JANUARY 2013). doi:DOI 10.1002/9781118395264.fmatter

Weiss, L. M., Zhu, X., Cali, A., Tanowitz, H. B., & Wittner, M. (1994). Utility of microsporidian rRNA in diagnosis and phylogeny: *A review*. *Folia Parasitologica*, 41(2), 81–90.

Wilkinson, T. J., Rock, J., Whiteley, N. M., Ovcharenko, M. O., & Ironside, J. E. (2011). Genetic diversity of the feminising microsporidian parasite *Dictyocoela*: New insights into host-specificity, sex and phylogeography. *International Journal for Parasitology*, 41(9), 959–966.

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. *Environmental Microbiology Reports*, 10(3), 328–336.

Winters, A. D., & Faisal, M. (2014). Molecular and ultrastructural characterization of *Haplosporidium diporeiae* n. sp., a parasite of *Diporeia* sp. (Amphipoda, Gammaridea) in the Laurentian Great Lakes (USA). *Parasite*, 7(1), 343.

Winters, A. D., Fitzgerald, S., Brenden, T. O., Nalepa, T., & Faisal, M. (2014). Spatio-temporal dynamics of parasites infecting *Diporeia* spp. (Amphipoda, Gammaridae) in southern Lake Michigan (USA). *Journal of Invertebrate Pathology*, 121, 37–45.

Wittner, M. (2014). Historic Perspective on the Microsporidia: Expanding Horizons. In M. Wittner & L. M. Weiss, eds., The Microsporidia and Microsporidiosis, Washington: *ASM Press*, pp. 1–6.

Xu, J., & Zhou, Z. (2010). Improving phylogenetic inference of microsporidian *Nosema antheraeae* among *Nosema* species with RPB1 , -and -tubulin sequences. *African Journal of Biotechnology,* 9(46), 7900–7904.

Yang, G., Short, S., Kille, P., & Ford, A. T. (2011). Microsporidia infections in the amphipod, *Echinogammarus marinus* (Leach): Suggestions of varying causal mechanisms to intersexuality. *Marine Biology*, 158(2), 461–470.

Ye, Y. H., Woolfit, M., Rancès, E., O'Neill, S. L., & McGraw, E. A. (2013). Wolbachia-Associated Bacterial Protection in the Mosquito *Aedes aegypti*. *PLoS Neglected Tropical Diseases*, 7(8). doi:10.1371/journal.pntd.0002362

CHAPTER III. *NOSEMA GRANULOSIS* (MICROSPORIDIA) INFECTING *GAMMARUS BALCANICUS* AND *G. ROESELII* (AMPHIPODA): A PRELIMINARY RE-ASSESSMENT OF HOST-PARASITE RELATIONSHIP BASED ON RPB1 GENE

III.1 BACKGROUND

Since the mid-XIXth century, Microsporidia, a large group of obligate intracellular microbial eukaryotic parasites, has aroused interest in both primary and applied sciences (Franzen 2008; Mathis 2000; Vávra & Lukeš 2013). In that context, the microsporidian genus *Nosema* stands in a seminal place for many reasons. The first-ever microsporidia formally taxonomically described was *Nosema bombycis* (Naegeli 1857). This parasite was slightly earlier presented by Guérin-Méneville (1849-1850) as the agent causing the "pepper disease" ("pébrine" in French) ravaging silkworms (*Bombyx mori*) industry in southern Europe. Later, Pasteur (1870) studied its aetiology. The phylum Microsporidia was finally established by Balbiani (1882) based on the description of *Nosema bombycis* by Naegeli (1857). Since these times, infections caused by microsporidia belonging to the genus *Nosema* were shown to have a significant impact on other industries using insects. For example, *Nosema apis* and *N. ceranae* are both microsporidia species infecting honey bee (*Apis mellifera*), causing substantial intestine injuries, leading to the collapse of bee colonies (Fries 2010; Higes *et al.* 2006). Other *Nosema* species were studied as a biological control agent for pest insects. For example, *Nosema locustae* have been studied to control migratory locusts grasshoppers in Africa (Lockwood *et al.* 1999) or more recently, *Nosema pyrausta*, to regulate the population of the European corn borer, *Ostrinia nubilalis*, in synergy with conventional insecticide (Lewis *et al.* 2009). Many of the interactions between *Nosema* parasites and their hosts are complex and are showing some degrees of host-specificity, indicating a high degree of host-parasite coevolution (Ironside 2007; Quiles *et al.* 2019). The genome of the few economy-threatening *Nosema* spp has been sequenced for a better understanding of their biology (*e.g. Nosema bombycis* (Xu *et al.* 2006), *Nosema ceranae* (Cornman *et al.* 2009)*, Nosema apis* (Chen *et al.* 2013)).

Among the genus *Nosema*, *N. granulosis* appears to be a peculiar parasite. This species is not infecting insects as almost all known *Nosema*, but has been first detected and described in the amphipod crustacean *Gammarus duebeni* (Terry *et al.* 1999). *Nosema granulosis* has been found in six other amphipod species (Haine *et al.* 2004; Ironside 2013; Krebes *et al.* 2010; Ku *et al.* 2007; Quiles *et al.* 2019; Chapter II of this thesis). The genetic diversity for SSU rDNA described for *N. granulosis* suggests some degree of host-specificity (see chap I and II of this thesis). The most striking characteristics of *N. granulosis*, compared to other *Nosema*, is that some strains are using almost only vertical transmission in their life-cycle: the infected mothers transmit the parasites to their offspring through the eggs (Dunn & Smith 2001). Vertical transmission occurs only *via* female hosts because the eggs are containing enough cytoplasm in which the microsporidia can live and be transmitted, which is not the case for male gametes (Hurst & Majerus

1993). Previous studies have shown that these vertically-transmitted infections cause little pathogenicity to their hosts (Haine *et al.* 2004; Kelly *et al.* 2001; Terry *et al.* 1997, 1998). Some strains of *N. granulosis* are also known to cause female-biased sex ratio in populations of *G. duebeni* and *G. roeselii* (Haine *et al.* 2004; Ironside *et al.* 2003; Ironside & Alexander 2015; Terry *et al.* 2004). Indeed, they turn males into functional females (Haine *et al.* 2004, 2007; Jahnke *et al.* 2013; Terry *et al.* 1999). *Nosema granulosis* manipulates host sex by preventing androgenic gland differentiation and androgenic gland hormone production, which coordinates male sexual differentiation (Rodgers-Gray *et al.* 2004). As a consequence, only a few males were found to be infected by *N. granulosis* in the French populations of *G. roeselii* studied by Haine *et al.* (2004). Nevertheless, some males were infected, and this finding may be due to incomplete parasite's feminizing ability or because some individuals may resist feminization. Adding to the fact that *N. granulosis* induces no pathology to their gammarid hosts, its presence may even have some beneficial effects. In *G. roeselii*, infected females breed earlier in the reproductive season than the uninfected ones (Haine *et al.* 2004), and it is known that early reproduction increases the number of broods produced in a given reproductive season (Pockl, 1993). The infection also provides a survival advantage to *G. roeselii* females, compared to the uninfected ones (Haine *et al.* 2007). Because of this, female excess and positive effect on reproduction, the populations infected with these *N. granulosis* strains should have higher dynamics and may help the invasive host *G. roeselii* in colonizing new territories. Supporting this hypothesis, Quiles *et al.* (2019, see also Chapter I of this thesis) found that the *N. granulosis* strain known to induce feminization in *G. roeselii* is associated with the only host genotype that invaded Western Europe after the last glaciation. However, in *G. balcanicus* host, we observed more infected males than females, contrary to all previous descriptions of *N. granulosis* (see Chapter II of this thesis). This new finding is thus questioning all the previous statements on the generality of vertical transmission and feminization in *N. granulosis* (Dunn *et al.* 2006; Kelly *et al.* 2002; Rodgers-Gray *et al.* 2004; Weedall *et al.* 2006).

Most studies on microsporidia are using the small rDNA subunit (SSU) to screen host for infections, identify microsporidian parasites at the species level and reconstruct phylogenies. This proved to be very useful for phylogenetic reconstructions at higher taxonomic levels and distinguishing major clades of microsporidia (*e.g.* genera) (Vossbrinck *et al.* 2014). In addition, SSU also proved to be useful at resolving phylogenetic relationship at lower taxonomic levels, *e.g.* between species, as clearly exemplified by Bacela-Spychalska *et al.* (2018) for the genus *Dictyocoela*.

Nevertheless, SSU clearly showed limitations for intra-specific phylogeography, at least for *N. granulosis* (*e.g.* Quiles *et al.* 2019). On one hand, combined data from the seven gammarid host species infected show that the number of haplogroups observed in *N. granulosis* is relatively large (13 haplogroups). On the other hand, these haplogroups differ only by few substitutions, resulting in low resolution of phylogenetic relationships (see Fig. 1). However, it is to be noted that the informative content of the SSU at this taxonomic level was not nil. For example, possible host–parasite specificity was pointed out for *G. balcanicus* as this species harbour three specific haplogroups. However, globally, the low molecular divergence of this marker at this scale is challenging any firm conclusion about the evolutionary history of *N. granulosis* and both *G. roeselii* (Quiles *et al.* 2019 and Chapter I of this thesis) or *G. balcanicus* (Chapter II of this thesis their gammarid hos). For example, the three *N. granulosis* haplogroups detected for *G. roeselii* are shared with other host species (*e.g. N. granulosis* 01 found in *G. roeselii* is shared with *G. duebeni* and *Dikerogammarus villosus*) (Fig. 1). Therefore, a variable markers is needed to unravel intra-species phylogeography of *N. granulosis*, allowing a more fruitful analysis of host-parasite associations.

The use of the large rDNA subunit (LSU) and/or the internal transcribed spacer (ITS) located between SSU and LSU has improved resolution in microsporidia phylogeny (Bacela-Spychalska *et al.* 2018; Hogg *et al.* 2002; Ryan & Kohler 2010; Stentiford *et al.* 2016; Terry *et al.* 2004; Winters & Faisal 2014). Unfortunately, it was already challenging to obtain 800 bp sequences from SSU for all the positively diagnosed microsporidia individuals (see the previous chapters).

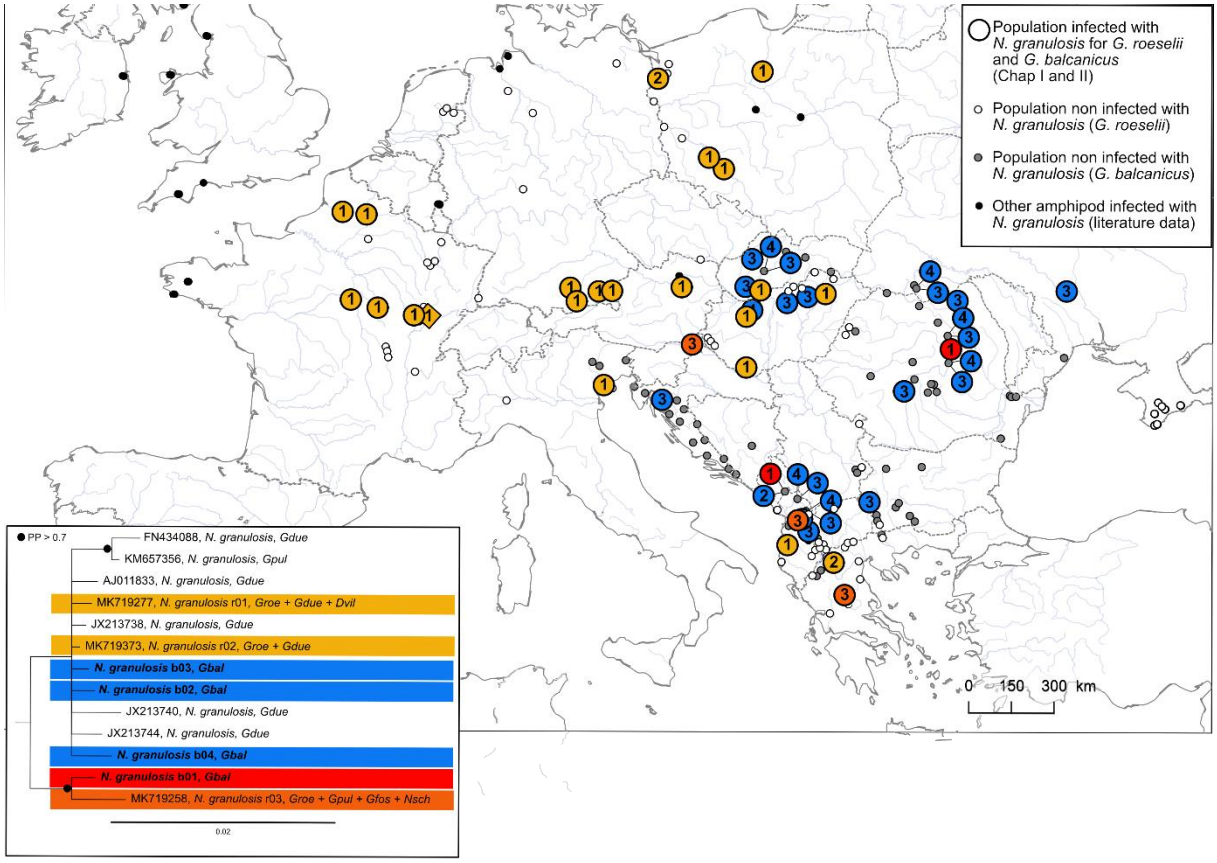


Fig. 1. Geographic distribution and phylogenetic tree for infections by *Nosema granulosis* in gammarids over Europe. Large coloured dots represent infections found in: (i) *Gammarus roeselii*; haplogroup r01-r02 and r03 being highlighted in light orange and in dark orange, respectively, (ii) *G. balcanicus*; haplogroup b01 and b02-04 being highlighted in red and blue, respectively. Small white and grey dots are *G. roeselii* and *G. balcanicus* population not infected with *N. granulosis*, respectively. Small black dots are infections reported in the literature in other amphipods (see Table S1 for further details). The Bayesian phylogenetic reconstruction is based on small ribosomal subunit (SSU) rDNA. *Nosema antherae* (DQ073396) was used as outgroup (not shown on the tree). Sequences from Genbank are representative of other *N. granulosis* haplogroups (see Table S1). Labels include the accession number, the microsporidia species name as given in the associated publication (haplogroup name (01-04) being additionally provided when the host is *G. roeselii* (r01 to r03) or *G. balcanicus* (b01 to b04)) and the host species abbreviated names (Dvil = *Dikerogammarus villosus*, Gdue = *G. duebeni*, Gfos = *G. fossarum*, Gpul = *G. pulex*, Nsch = *Niphargus schellenbergi*). PP: Bayesian Posterior Probability.

Coding genes were used to propose host-parasite evolution scenario between *Apis mellifera* and *Nosema apis* and *N. ceranae* parasites (Gómez-Moracho *et al.* 2014; Maside *et al.* 2015a). RPB1 gene, notably, has been used in different studies for improving our knowledge on the biology of microsporidia (*e.g.* Brandon Matheny *et al.* 2002; Hirt *et al.* 1999; Hopper *et al.* 2016; Ironside 2007; Maside *et al.* 2015b; Pretto *et al.* 2018). RPB1 is coding for the large subunit of the RNA polymerase II, an enzyme for all mRNA synthesis in eukaryotes (Cramer *et al.* 2001). The RPB1 gene displays a high level of synonymous variations and can be a suitable gene candidate to

unravel microsporidia intra-species level phylogenetic relationships (Andolfatto 2001; Leffler *et al.* 2012). This gene has been used for inferring relationships among microsporidian taxa, more particularly among the lepidopteran *Nosema* (Gisder & Genersch 2013; Hirt *et al.* 1999; Ku *et al.* 2007; Kyei-Poku & Sokolova 2017; Liu *et al.* 2015; Vavra *et al.* 2006) and microsporidians paraistising other insects (Tokarev *et al.* 2019). It has also been used in a few studies of *Nosema* parasitising Amphipoda (Cheney *et al.* 2001; Ironside 2007). This gene proved useful helping to understand events of sex loss within *Nosema/Vairimorpha* complex (Ironside 2007) or clarifying the phylogenetic position of *N. antheraeae* (silkmoth-infecting species) among other species of *Nosema* (Xu & Zhou 2010). In addition, the RPB1 gene was also used to distinguish *N. apis* from *N. ceranae*, whereas the use of SSU rRNA gave ambiguous results (Gisder & Genersch 2013).

The aim of this study was to deepen our understanding of the evolutionary history of *Nosema granulosis* infecting *Gammarus* species; by using the RPB1 gene. We aimed to: (i) verify the previous phylogenetic affinities obtained with SSU rDNA, (ii) explore if there is genetic variation within the haplogroups detected with the SSU rDNA, (iii) determine more precisely the phylogenetic relationships between the different parasites haplogroups and (iv) propose host-parasite evolutionary history scenarios to explain the diversity and co-bio-geographical pattern observed between *N. granulosis* in the two host species.

We used *N. granulosis* individuals detected previously in the extensively-studied host species *Gammarus roeselii* and *G. balcanicus* (Quiles *et al.* 2019, Chapters I and II of this thesis), and compared them with those known from *G. duebeni* (Ironside *et al.* 2008), the only other gammarid for which RPB1 data are available.

III.2 METHODS

III.2.1 Host sampling

Most of the data obtained in the present study focused on *Nosema granulosis* infecting two amphipod species, *Gammarus roeselii* and *G. balcanicus*. The same samples as in Chapter I (Quiles *et al.* 2019) and Chapter II were used. In short, a total of 1904 individuals for *G. roeselii* and 2255 individuals for *G. balcanicus* were sampled, from 94 sites in 19 countries and from 88 sites in 13 countries, respectively. In both cases, sampling covered the whole range distributions of both hosts. As noted earlier, the evolutionary history of both *G. roeselii* and *G. balcanicus* have recently been investigated (Grabowski *et al.* 2017b, 2017a; Mamos *et al.* 2016). Both hosts are characterised by extensive cryptic diversity with *c*. 13 and 50 highly divergent phylogenetic lineages (Molecular Operational Taxonomy Units, MOTUs), respectively. These species have an overlapping range in the Balkans, but *G. roeselii* populations are also widely distributed across

Central and Western European freshwater ecosystems. This last morphospecies diversified mostly over Miocene in the Balkan Peninsula (Grabowski *et al.* 2017b), and only one of its MOTUs (MOTU C) expanded postglacially its geographical range in northern and western Europe (Grabowski *et al.* 2017a). Contrastingly, MOTUs of *Gammarus balcanicus* are locally endemic due to their habitat fragmentation and complex phylogeographic history (Copilaş-Ciocianu & Petrusek 2017; Mamos *et al.* 2014, 2016). *Gammarus balcanicus* appeared to start its diversification in the early Miocene in the central Balkans (Mamos *et al.* 2016). This early diversification generated two major clades split into the north-eastern and the south-western Balkan (see chapter II).

Other gammarids used in the present study were selected from the literature as being infected by *N. granulosis,* as based on SSU rRNA (Table S1, see also Table S2 Chapter II), including other *Gammarus* species*, i.e. G. duebeni*, *G. pulex*, *G. fossarum*, but also *Dikerogammarus villosus* and *Niphargus schellenbergi*. Sequences of the RPB1 gene are only available from *G. duebeni* (Table S1).

III.2.2 *Nosema granulosis*: infection status, haplogroup diversity and phylogeny based on ribosomal DNA.

All raw data, as well as all analyses for *N. granulosis*, are the same as in chapters I and II. In short, for *G. roeselii* and *G. balcanicus*, part of the small ribosomal subunit (*SSU* rDNA) was used for both molecular screening (specific PCR primers for Microsporidia) and assignment to *Nosema granulosis* (based on BLAST results of sequenced PCR products against sequences available in Genbank). Furthermore, these sequences, in addition to other *N. granulosis* sequences from literature, were used for both: (i) definition of haplogroup diversity (based on diagnostic SNPs) and (ii) phylogenetic reconstruction. Results are summarized in Fig. 1 (see also Fig. 3 in Chapter I and Fig. 4 in Chapter II).

III.2.3 Partial amplification and sequencing of *N. granulosis* RPB1

The RNA polymerase II largest subunit (RPB1) gene is known to be a nuclear single-copy 4818bp long gene , with no intron (1,606 codons) in *Vairimorpha necatrix*, a close relative to *Nosema* (Hirt *et al.* 1999, GenBank reference: AF060234).

First of all, we tested all already available primers found in the literature (Cheney *et al.* 2001; Hirt *et al.* 1999; Ironside 2007; Stiller & Hall 1997) trying to amplify RPB1 gene from our *N. granulosis* samples. Since the results from these amplifications were low, we decided to design new primers for PCR amplification.

Our strategy for primer design was as follow. Our starting point was the unpublished RPB1 5' partial sequence associated with one individual of *G. roeselii* from France (River Ouche population), known to be infected by *N. granulosis* SSU rDNA haplogroup r01. This 2445bp long sequence (referred as NGRA010103_000010.1) was obtained from the sequencing of the *N. granulosis* genome by Alexandre CORMIER and Richard CORDEAU from Laboratory "Ecologie et Biologie des Interactions" (UMR CNRS 7267, Poitiers). Second, we retrieved 14 sequences from Genbank: 4 sequences from *N. granulosis* infecting *G. duebeni* (GenBank: DQ996233, JX213746, JX213747, JX213748), 8 sequences of *N. bombycis* (GenBank: EOB12667, DQ996231, JX213750, JX213751, JX213752, JX213753, JX213754, JX213755) and 2 sequences are from *N. anthereae* (GenBank: PRJNA183977). The two later *Nosema* species are known, based on SSU RNA data, to be the closest relative to *N. granulosis* (Kyei-Poku & Sokolova 2017; Liu *et al.* 2012). Therefore, our final data set includes 15 sequences, allowing identification of both variable regions of interest for phylogenetic reconstructions and of the conserved regions suitable for primer design. Four sets of degenerated primers were designed using Geneious 10.2.2 (Kearse *et al.* 2012), targeting five overlapping fragments (named F1-5) of *c.* 500-600 bp each, spanning the 2445 bp of the NGRA010103_000010.1 sequence (Fig. 2 and Table 1).

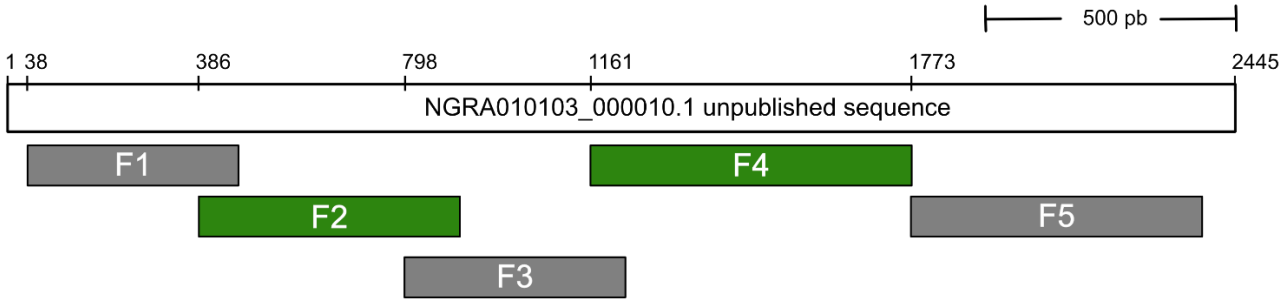


Fig. 2. Schematic view of fragments (F1-5) targeting five PCR amplification of parts of the RNA polymerase II largest subunit gene (RPB1). These partially overlapping fragments are referred from the 5' part of the gene, *e.g.* F1 fragment starting at position 38pb.

Table 1. Primers targeting PCR amplification of five fragments (F1-5) of the RNA polymerase II largest subunit (RPB1) gene. See Fig. 2 for a schematic representation of amplicon mapping.

Fragment	Primer name	Primer Sequence (5′- 3′)	Position NGRA010103	Amplicon size *(c.* bp)
F1	F1f	TCC GTT GAT ACR AAG AGC	38	420
	F1r	CTG AGT TRT CTT CTC CTT TC	462	
F2	F2f	GKT GTG GRA ATA AAC AGC	386	520
	F2r	TCT ACT CTC TTM CCC ATA AG	909	
F3	F3f	TGG ACA ACC WCA AGC YCT	798	441
	F3r	TGT GGC CCA TCA TAG ACA T	1239	
F4	F4f	GAA AGA CAC ATG CAG RAT G	1161	640
	F4r	TTC CWG ACA TGA TYT CTC C	1803	
F5	F5f	GAG GAG ARA TCA TGT CWG	1773	580
	F5r	CTG TTC TCT ACA AAC CC	2356	

PCRs were performed in a volume of 30 µl, containing 2.5 mM MgCl2, 0.5 units of 5 PRIME HotMaster Taq DNA polymerase (Qiagen, Hilden, Germany), 0.2 µM dNTPs (MP Biomedicals Europe, Illkirch, France), 0.2 µM each of forward and reverse primers (Eurofins Genomics, Ebersberg, Germany) and 2 ng DNA template. Amplification conditions were as follows: an initial denaturing phase at 94°C for 2 min, 35 cycles at 94°C for 20s, fragment specific annealing temperature (Ta) 50°C for both F2 and F4 for 20 s and 65°C for 30 s and a final extension at 65°C for 5 min. All individuals positively diagnosed to be infected with *N. granulosis* as based on *SSU* rDNA results (see above) were tentatively amplified.

The PCR products were purified and sequenced directly with the BigDye technology by Genewiz, Inc., DE, using the forward primers from PCR. Using Geneious 10.2.2 (Kearse *et al.* 2012), raw sequences were checked for being microsporidian RPB1 sequences, via BlastN search (Madden 2003) against the sequences available in GenBank. Each sequence was edited by eye, and clear double peaks were noted following the International Union of Pure and Applied Chemistry "IUPAC" degenerate nucleotide code. Such double peaks corresponded to the superposition of single peaks, seen in linkage disequilibrium, in other haplogroups. Therefore, one could interpret the observed pattern either as double infections with *Nosema* individuals harbouring different haplogroups or as single infections by a heterozygote individual. As earlier pointed out, it is to be noted that RPB1 gene is known to be nuclear and single-copy gene in *Vairimorpha necatrix*, a close relative to *Nosema* (Hirt *et al.* 1999) which is compatible with heterozygosity detection using direct sequencing.

Fragments F2 and F4 (the only ones that were successfully amplified, see results) were trimmed to a final maximum size of 456 and 557bp, respectively. Sequences were aligned using MAFFT7.388 software (Katoh 2002; Katoh & Standley 2013) with the E-IONS-I algorithm using the legacy gap penalty option, incorporated in Geneious 10.2.2 (Kearse *et al.* 2012), using also the

translated versions in protein as a guideline. The sequence AF060234, from *Vairimorpha necatrix*, the only known microsporidian complete gene sequence from a close relative available to our knowledge, proved very useful in that context. None of the haplogroups was associated with the presence of stop codon, a rough proxy of the sequence not being associated with pseudo-genes.

Some haplogroups showed three short deletions in the F4 fragment (Fig. 4). Alignment around these deletions proved to be uncertain (even, *e.g.* taking care of maintaining alignment consistent with a frame allowing correct translation into proteins), leading to the decision of not taking them, as well as the deletion areas themselves, into account in the final alignment for phylogenetic reconstruction. However, the qualitative information of the presence of such deletions in some haplogroups was kept in mind as itself it may convey valuable phylogenetic information. The regions deleted for the analysis are shown in Fig. 4. Finally, as for some individuals sequences were produced for both F2 and F4, while others only for F2 or F4, our dataset is 'composite'.

III.2.4 Phylogeny reconstruction for microsporidians

Our dataset is composed of : (i) newly produced microsporidian RPB1 sequences of *N. granulosis* infecting *G. roeselii* and *G. balcanicus* individuals and (ii) RPB1 sequences of *N. granulosis* from the literature found infecting *G. duebeni*. The outgroup used was *Nosema antheraeae* (GenBank: HQ215550) (Xu & Zhou 2010). All details, including fragment used and sequence length, are given in Table S1.

As sequence lengths were heterogeneous, precise definition of haplotypes was not possible, but we were able to attribute each sequence to a specific haplogroup. As defined in Quiles *et al.* (2019), haplogroups were build when sequences differed at least by one or more variable sites, generating diagnostic features whatever sequence length. Two sequences were clustered in one haplogroup, despite the variable length, based on 100% pairwise identity, sharing the same diagnostic sites. Only one sequence (population 4 = GR25-18, *N. granulosis* r03) could not be assigned to a given RPB1 haplogroup, due to reduced sequence size and lack of diagnostic features between RPB1 haplogroups B1 and B2 (see results, Fig. 5). Only the longest sequences representing each haplogroup were used for the Bayesian phylogeny reconstruction (384 to 917 bp, Table S1). Missing data were coded as "N". Haplogroup alignment used for building the tree is provided in Data S1. Compared to SSU rRNA, the amount of nucleotidic variation at RPB1 was extremely important, although being almost only synonymous when verified at the protein level (see Fig 4. B).

As a coarse proxy to test if the variation we observed was not only associated with nucleotide substitutions corresponding to the third codon (involving possibly homoplasy and saturation), phenetic p distances were estimated with MEGA 7 (Kumar *et al.* 2016) using either all three nucleotides within codons or only the two first nucleotides (Table. S3). Although the distances were shorter using only the two first nucleotides, they were not nil, being in addition congruent with the distances based on all three-nucleotides. Therefore, all three nucleotide positions within a codon were used for phylogenetic analyses. Neither ka/ks ratio values calculations (aiming at detected types of selection), nor saturation test (aiming at detecting of saturation associated to homoplasy) were performed on the present data. However it is to be pointed out that Ironside *et al.* (2007) showed low values of ka/ks ratio for all pairwise comparison he performed (a feature interpreted as a sign of purifying selection acting the same way on all parasites tested), as well as the absence of saturation. We are fully aware that these features (in Ironside *et al.* 2007) do not presume the absence of directional selection on one or the other new haplogroup or the presence of saturation in our dataset (see discussion).

Phylogenetic reconstructions were build using Bayesian methods in MrBayes (Huelsenbeck & Ronquist 2001) integrated into Geneious 10.2.2. The best-fitting model of nucleotide substitution was determined with JModelTest-2.1.10. (Darriba *et al.* 2012). We used the General Time Reversible (GTR) model with gamma-distributed rate heterogeneity (G) and a significant proportion of invariable sites (I). Four heated chains, each 1,100,000 iterations long, sampled every 200 iterations, were run. The runs reached satisfactory effective sampling sizes (ESS > 200), and the potential scale reduction factor values equalled 1 for all parameters. The 50% majority-rule consensus tree was constructed after the removal of 10% 'burn-in' trees.

III.3 RESULTS

Based on SSU rRNA sequences, 172 host individuals were known to be infected by *N. granulosis*, 96 being *G. roeselii* and 76 *G. balcanicus* (Table S1, or see also Quiles *et al.* 2019 and Chapter II).

Among the five primer sets designed in this study, three sets (for fragments F1, F3 and F5) were unable to amplify the test panel representing the *N. granulosis* genetic diversity present in both *G. roeselii* or *G. balcanicus* infected individuals.

For the two remaining sets (*i.e.* F2 and F4) amplification success was more promising on the tested panel. However, as applied on the 172 infected host individuals, the success rate was variable, including no success (97 ind., 56.4%), success for one or the other fragments (48 ind., 27.9%) and success for both fragments (27 ind., 15.7%). Therefore, we sequenced a total of 75 *N.*

granulosis individuals for RPB1 partial gene marker, 51 and 24 parasites being associated with *G. roeselii* and *G. balcanicus* host, respectively. Geographic distribution of the parasites for which the RPB1 sequences were obtained is shown in Fig. 3. Host ranges were covered for both *G. roeselii* and *G. balcanicus*.

RPB1 sequences were obtained for all the parasite SSU rDNA haplogroups from *G. roeselii* (*N. granulosis* r01, r02 and r03), but only two (*N. granulosis* b03 and b04) out of four for *G. balcanicus* (no sequence for *N. granulosis* b01 and b02) (Fig. 5).

Moreover, very few *N. granulosis* infecting amphipods from the literature were sequenced for RPB1. Only two sequences found in GenBank, DQ996233 and JX213746, matched the RPB1 area that we were able to sequence. They were both from *N. granulosis* infecting *G. duebeni* in Wales (Ironside *et al.* 2007, 2013) and were identical. Therefore, only DQ996233 was used for our analysis (Fig. 3, 4, 5).

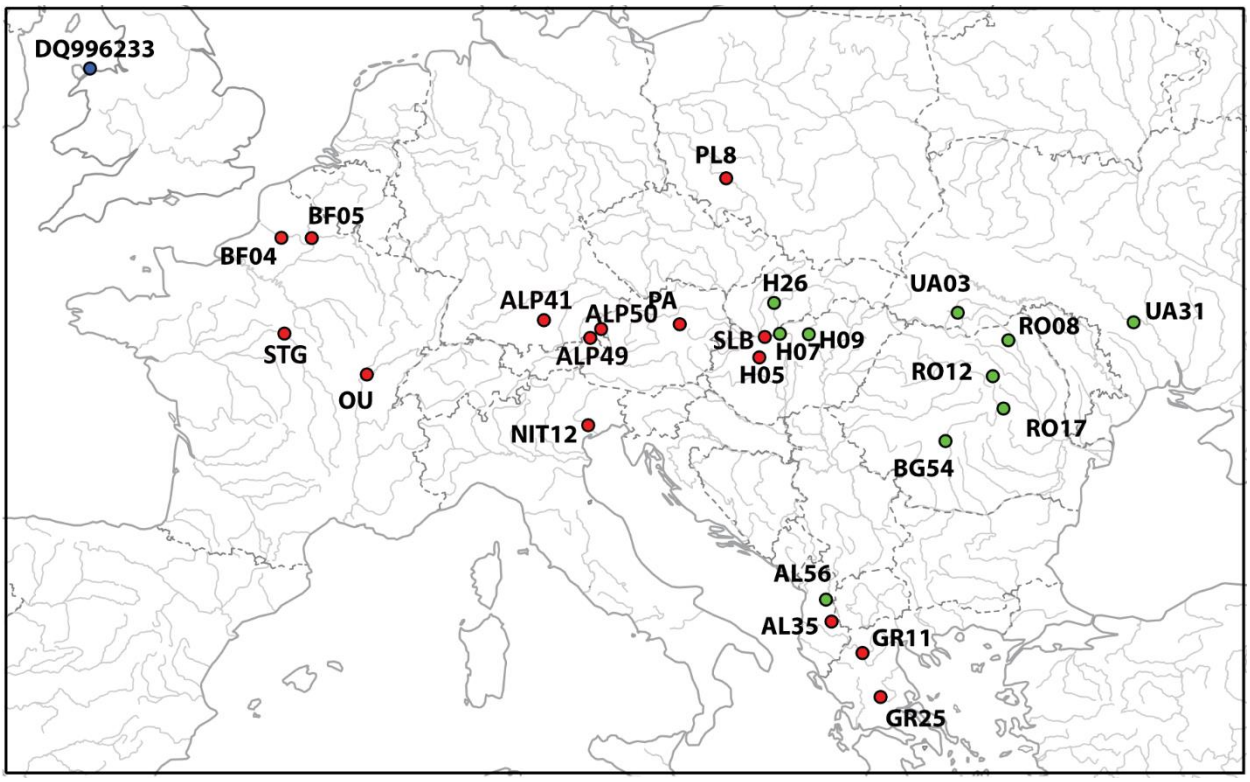


Fig. 3. Geographic distribution of populations infected by *Nosema granulosis* parasites for which RPB1 sequences were obtained, either newly produced or collected from the literature, for host species of the genus *Gammarus*. Red and green dots correspond to infections associated with *Gammarus roeselii* and *G. balcanicus*, respectively. The blue dot is the infection found in *Gammarus duebeni*, published by Ironside (2007, 2013). The names are the population abbreviations (see Table. S1) or GenBank accession number for the published sequence.

The alignment of the RPB1 marker sequences from *N. granulosis* (Table S3) revealed six haplogroups (Fig. 4A), being based almost exclusively on synonymous substitutions for fragment F2 but including both synonymous and nonsynonymous substitutions for fragment F4 (Fig. 4B).

RPB1 partial sequences allowed an improvement of the phylogeny compared to the rDNA SSU marker (Fig. 5). The RPB1 tree based on nine haplogroups was composed of three well defined and supported clades, each including three close haplogroups (Fig. 5). The first clade, A (highlighted in blue in Fig. 5.B), included 3 haplogroups (A1, A2, A3) associated to SSU haplogroups *N. granulosis* b03 and b04 infecting *G. balcanicus* (Fig. 5.A). The second clade, B (highlighted in green in Fig. 5.B), was also composed of three haplogroups (B1, B2, B3), all associated to SSU haplogroup *N. granulosis* r03 infecting *G. roeselii* (Fig. 5.A). The third clade, C (highlighted in orange in Fig. 5.B), included three RPB1 haplogroups (C1, C2, C3), each corresponding to one SSU haplogroup (Fig. 5.A). The C1 RPB1 haplogroup corresponded to one of the SSU haplogroups infecting *G. duebeni* (Ironside 2007). The C2 and C3 RPB1 haplogroups corresponded to *N. granulosis* r02 and r01 SSU haplogroups, respectively, infecting *G. roeselii* (Fig. 5.B).

For *Nosema granulosis* haplogroups based on the SSU rDNA, the mean pairwise identity was 98.8% (outgroup excluded), illustrating low overall divergence (Fig. 5.A). Based on the RPB1 gene, the mean pairwise identity was 84.2 % (Fig. 5.B). For example, comparing individuals GR11-06 (belonging to the *N. granulosis* r02 SSU haplogroup) and AL56-03 (belonging to the *N. granulosis* b03 SSU haplogroup) yielded 4 nucleotides differences for 714 sites with the SSU rDNA, but 60 nucleotides differences for 612 sites using the RPB1 gene (RPB1 haplogroups A1 and C2, respectively).

All the haplogroups of clade C showed three short deletions in the F4 fragment (deletions of 27, 15 and 15 bp, respectively) (Fig. 4). We considered these deletions as a derived character as both individuals associated to clade A and B in *N. granulosis* and *Nosema antheranaeae*, *N. bombycis* sister clades (not shown in our results) are showing nucleotide sequence for these areas.

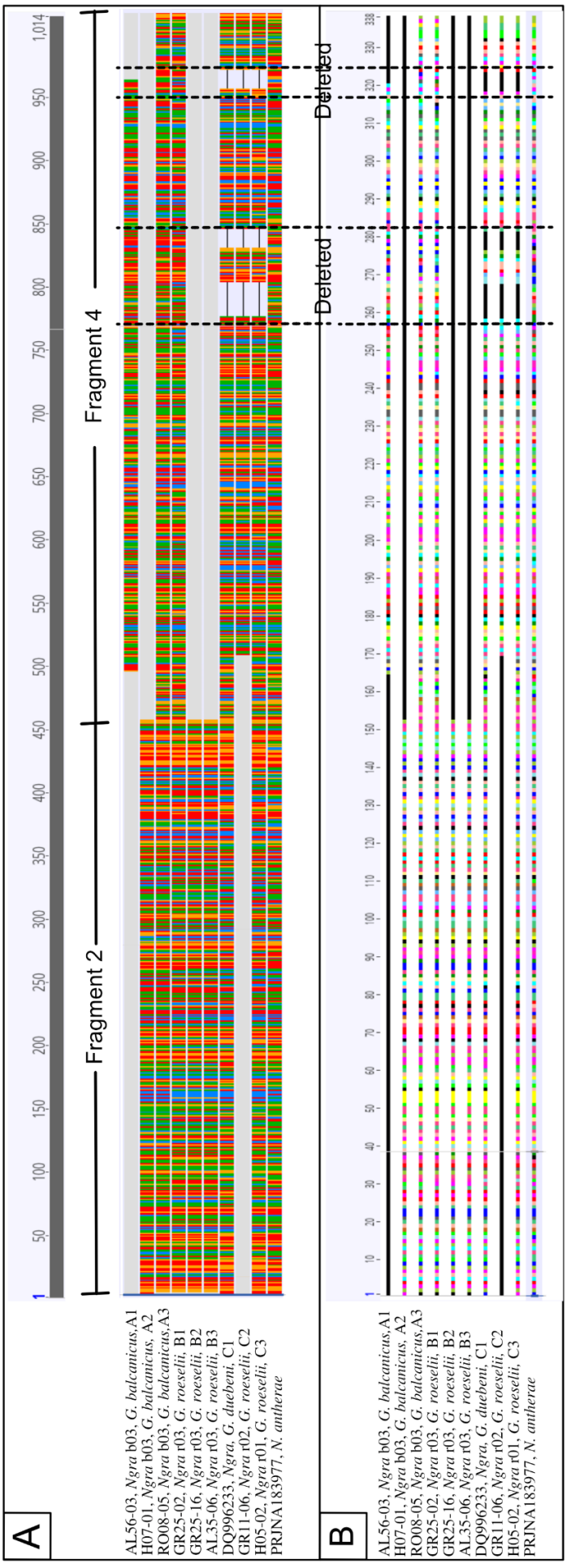


Fig. 4. A. Overview of the nucleotide alignment for the RNA polymerase II largest subunit (RPB1) gene haplogroups, for two fragments (F2 and F4). Grey color stands for either absence of PCR product or shorter sequences for a given fragment. Areas delimited by vertical dotted lines were not taken into account in the final alignment. B. Overview of the amino-acid alignment. Black color stands for absence of amino acides.

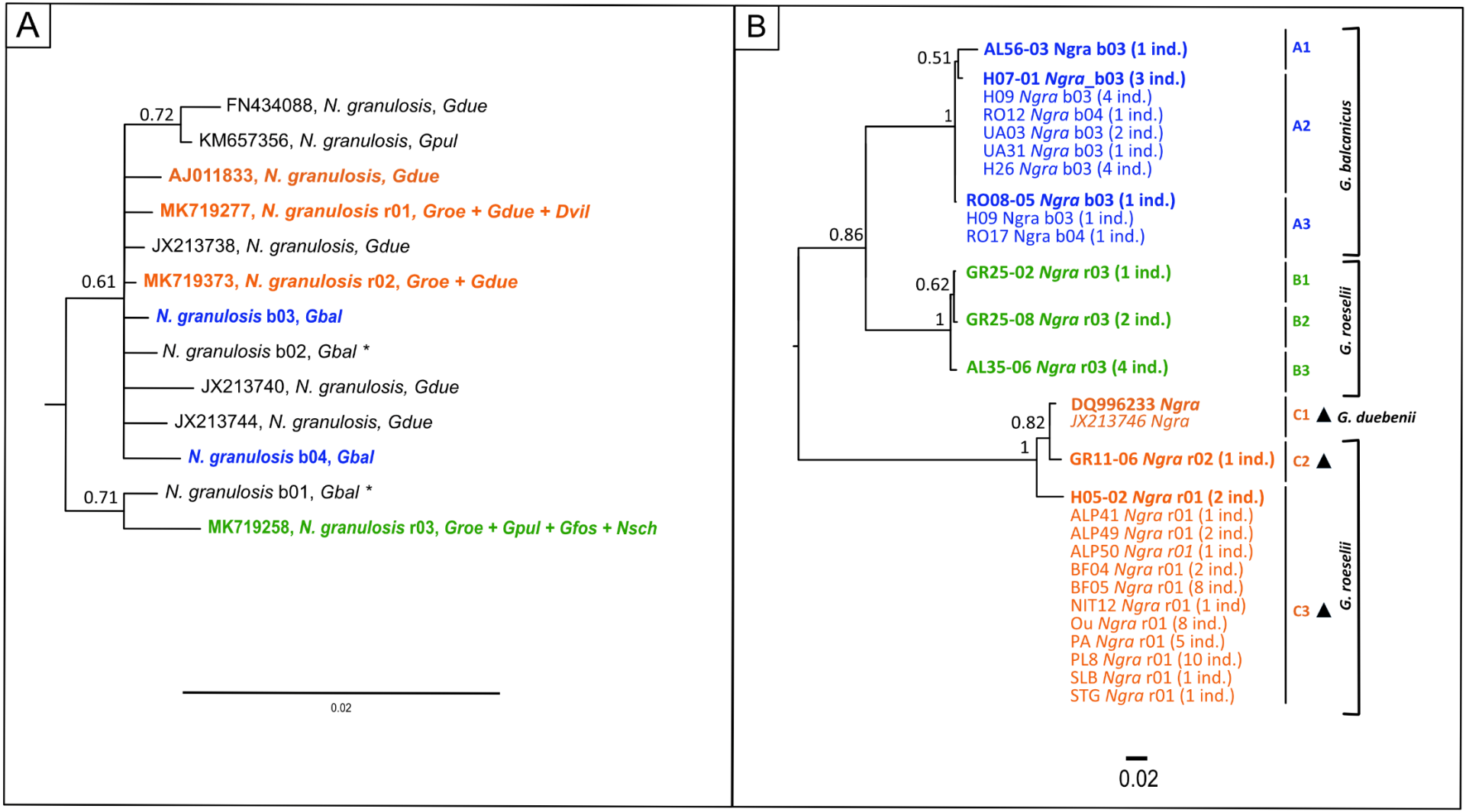

A
0.72
FN434088, N. granulosis, Gdue
KM657356, N. granulosis, Gpul
AJ011833, N. granulosis, Gdue
MK719277, N. granulosis r01, Groe + Gdue + Dvil
JX213738, N. granulosis, Gdue
0.61
MK719373, N. granulosis r02, Groe + Gdue
N. granulosis b03, Gbal
N. granulosis b02, Gbal *
JX213740, N. granulosis, Gdue
JX213744, N. granulosis, Gdue
N. granulosis b04, Gbal
0.71
N. granulosis b01, Gbal *
MK719258, N. granulosis r03, Groe + Gpul + Gfos + Nsch
0.02
B
0.51
AL56-03 Ngra b03 (1 ind.)
H07-01 Ngra_b03 (3 ind.)
H09 Ngra b03 (4 ind.)
1
RO12 Ngra b04 (1 ind.)
UA03 Ngra b03 (2 ind.)
UA31 Ngra b03 (1 ind.)
H26 Ngra b03 (4 ind.)
RO08-05 Ngra b03 (1 ind.)
H09 Ngra b03 (1 ind.)
0.86
RO17 Ngra b04 (1 ind.)
0.62
GR25-02 Ngra r03 (1 ind.)
1
GR25-08 Ngra r03 (2 ind.)
AL35-06 Ngra r03 (4 ind.)
0.82
DQ996233 Ngra
JX213746 Ngra
1
GR11-06 Ngra r02 (1 ind.)
H05-02 Ngra r01 (2 ind.)
ALP41 Ngra r01 (1 ind.)
ALP49 Ngra r01 (2 ind.)
ALP50 Ngra r01 (1 ind.)
BF04 Ngra r01 (2 ind.)
BF05 Ngra r01 (8 ind.)
NIT12 Ngra r01 (1 ind)
Ou Ngra r01 (8 ind.)
PA Ngra r01 (5 ind.)
PL8 Ngra r01 (10 ind.)
SLB Ngra r01 (1 ind.)
STG Ngra r01 (1 ind.)
A1
A2
A3
G. balcanicus
B1
B2
B3
G. roeselii
C1
G. duebenii
C2
C3
G. roeselii
0.02

Fig. 5. Bayesian phylogenetic reconstructions of *Nosema granulosis* parasites. (A) Based on partial sequences of the SSU rDNA (see Chapter II). *Nosema antherae* (DQ073396) was used as outgroup (not shown on the tree). Labels include, in this order, the accession number when available, the haplogroup name, the abbreviated name of the hosts (Dvil = *Dikerogammarus villosus*, Gbal = *Gammarus balcanicus*, Gdue = *G. duebeni*, Gfos = *G. fossarum*, Gpul = *G. pulex*, Groe = *G. roeselii*, Nsch = Niphargus *schellenbergi*). Stars indicated SSU haplogroups for which we did not retrieve RPB1 sequences.(B) Based on the partial sequences of the RPB1 gene. *Nosema antherae* (PRJNA183977) was used as outgroup (not shown on the tree). Labels include, in this order, the population name (Table S1), the abbreviated SSU haplogroup name (*N. granulosis* r01-r03, *N. granulosis* b3, b4) and the number of host individuals infected in the population (ind.). Labels in bold indicate sequences used to build the tree. Labels not in bold indicate other members of the haplogroup. A, B, C indicated RPB1 clades names, and the additional number indicated haplogroups names (*e.g.* A1 or B2). Triangles indicated haplogroups for which three small deletions are present. (A & B) : Coloured names are haplogroups that can be found in both trees. Sequences from Genbank are representative of other *Nosema granulosis* haplogroups (see Table S1). Numbers on the trees are Bayesian Posterior Probability. Scale bar indicated 2% substitutions per nucleotide site. The tree is not including six heterozygote sequences individuals that can be attributed each to two haplogroups.

The genetic distance between these three clades was high, much higher than the distances obtained with the rDNA SSU (Fig. 5). In two clades, the use of the RPB1 gene allowed to find more genetic diversity than previously found with the SSU. In particular, three RPB1 haplogroups can be found within each of the SSU haplogroups *N. granulosis* r03, b03 and b04. This variation was found among populations, but also even within the same population (populations 39 (HO9) and population 4 (GR25)). Intriguingly, while the SSU was able to distinguish two haplogroups *N. granulosis* b03 and *N. granulosis* b04, the RPB1 partial sequences did not distinguish these two different SSU haplogroups. Sequences assigned to *N. granulosis* b03 or b04 were attributed to the RPB1 haplogroups A2 and A3. These findings contrast with those for the SSU haplogroup *N. granulosis* r01, for which no variation was found using RPB1, despite a large number of individuals sequenced (Fig. 5.B), and a wide geographic area investigated (Fig. 3). Indeed, this C3 RPB1 haplogroup was found in 12 populations and 45 individuals across Europe, compared, for example, to the B clade, where 3 haplogroups were found in 7 individuals belonging to two populations. The absence of variation found in the RPB1 sequences of the SSU haplogroup *N. granulosis* r01 is therefore remarkable.

III.4 DISCUSSION

III.4.1 Design of new RPB1 primers

We tried to amplify the RPB1 gene of *N. granulosis* infecting *G. roeselii* and *G. balcanicus* for a better understanding of the *Nosema granulosis* diversity. We first used primers already designed from literature (Ironside 2007, 2013), but they all failed to amplify our samples (or

generated amplifications that were unspecific or that we were unable to be sequenced). Since these previous primers were designed to compare very distantly-related *Nosema spp.* (*N. granulosis* from *Gammarus duebeni*, but also *N. apis*, *N. trichoplusiae*, *Vairimorpha cheracis, V. disparis, V. necatrix,* etc.), our unsuccessful attempts may be due to the very high level of degeneration in these primers that prevented their binding with our DNA sequences. This feeling is strengthened by the fact that Ironside (2013) redraw specific primers for sequencing each of its parasite species. We, therefore, re-designed primers, using a restrained set of *Nosema* sequences, as close as possible to *N. granulosis* and including one of our parasite RPB1 thanks to genome-sequencing data (A. Cormier & R. Cordaux Person. Communication). This was a necessary strategy, instead of designing primers specific to our *N. granulosis,* since we did not know a priori the level of variation hidden behind our other SSU haplogroups. We suspected that a certain amount of variation would exist, at least between the SSU haplogroups *N. granulosis* b01 and *N. granulosis* r03, on one side, and the other haplogroups on the other side, since there was divergence at the level of SSU rDNA sequences (Fig. 1). In addition, many papers in the literature have shown substantial variation in RPB1 in *Nosema* (*e.g.* Pretto *et al.* 2018). Even with such a reduced sequence set, the primers we designed have degenerated nucleotides, and only two sets were successful (F2 and F4). Even with these successful primer sets, we were able to amplify and sequence only 76 individuals out of 172 individuals identified using the SSU rDNA marker. This relatively low level of success could be explained by the probable rarity of parasite DNA among host DNA. Indeed, microsporidia DNA cannot be separated from the host DNA, and *N. granulosis* parasites are present in relatively low numbers (Haine *et al.* 2004; Terry *et al.* 1999). It seems possible that this scarcity of parasite DNA target explains the lack of annealing of our primers. In addition, while SSU rDNA is known to be multicopy gene, RPB1 is likely to be a nuclear single-copy gene (Hirt *et al.* 1999).

Thereby, for RPB1, we were unable to analyse all the diversity found using SSU gene (Fig. 5). Out of 7 haplogroups found in *G. roeselii* and *G. balcanicus* using SSU, we were unable to amplify 2 of them with RPB1: *N. granulosis* b01 and b02, both from *G. balcanicus*.

III.4.2 Variation within N. granulosis using RPB1 gene and phylogenetic analysis

We were able to define three clades in *N. granulosis* infecting gammarids, each comprising three haplogroups.

Haplogroups A1 to A3 corresponded to SSU haplogroups *N. granulosis* b03 and b04. It is worth noting that, while *N. granulosis* b03 and b04 were discrete haplogroups using the SSU rDNA marker, this was not the case using the RPB1 marker. Indeed, both b03 and b04 parasites were found to be either A2 or A3 haplogroups. This, therefore, could highlight the first discrepancy

between SSU and RPB1 markers. At least for RPB1, the fact that some individuals are presenting both haplogroups could temptingly be interpreted as heterozygosity.

The B1-B3 RPB1 haplogroups were all corresponding to the *N. granulosis* r03 SSU haplogroup. Therefore, RBP1 allowed discriminating more variation than SSU rDNA in these groups of parasites. This is in contrast with the C3 RPB1 haplogroup corresponding to the *N. granulosis* r01 SSU haplogroup, where no variation was observed. This C3 haplogroup was closely related to C2, also infecting the host *Gammarus roeselii*, and to C1, infecting *G. duebeni*.

Compared to the SSU-generated tree, the RPB1 tree allowed to elongate the branches of our phylogenetic tree, mainly because the genetic distance was higher among RPB1 haplogroups. However, the topology support of this RPB1 tree is not extremely strong, in particular relationships between B haplogroups and the other ones.

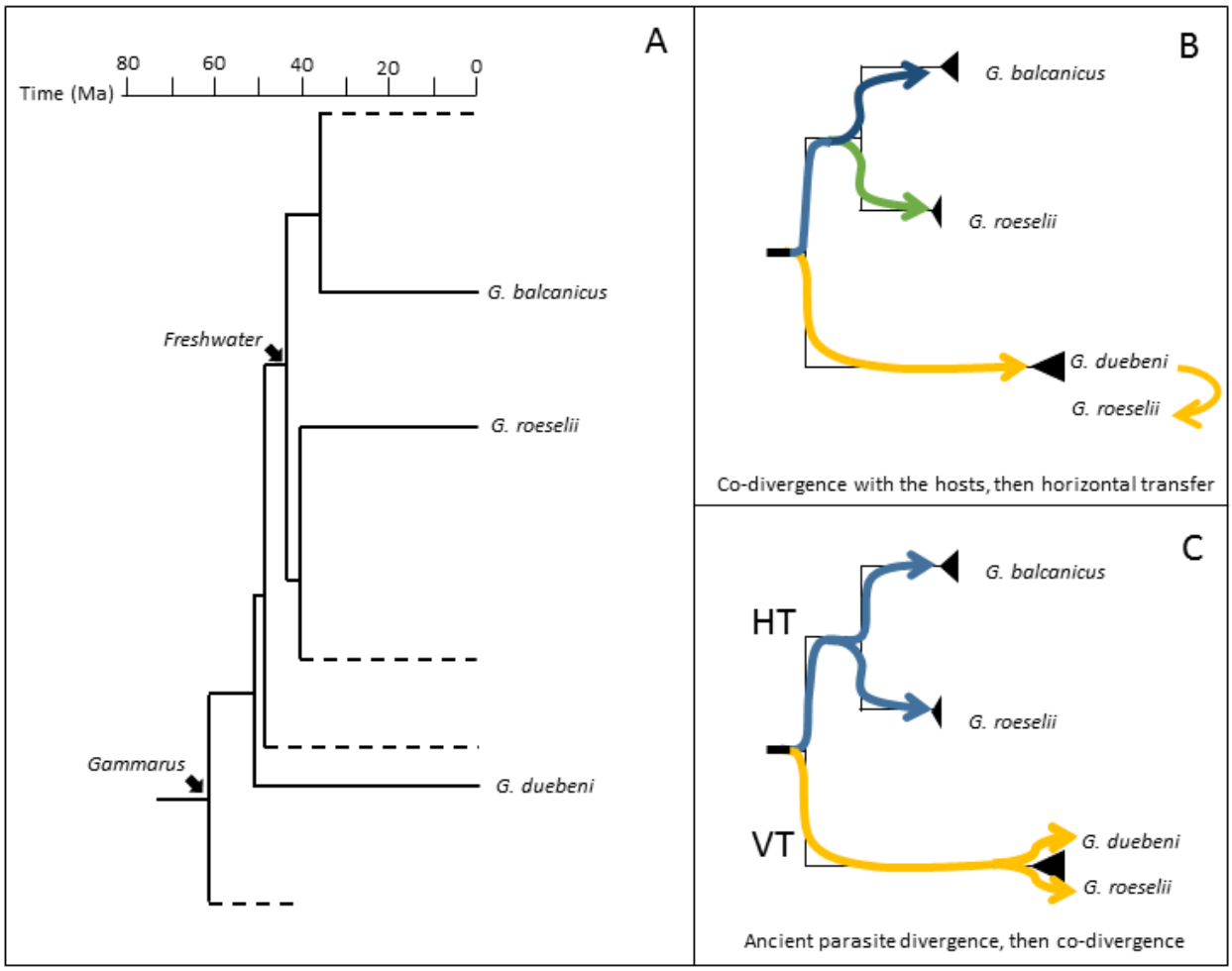
The tree topologies of SSU rRNA and RPB1 gene are not consistent. Using SSU rRNA gene haplogroups r01, r02 from *G. roeselii* and b03, b04 from *G. balcanicus* were in the same "clade" (though with poor support, the "clade" designation is perhaps abusive), showing 99.3% similarity. Using RPB1 gene these four parasites haplogroups are now clearly separated in two, well supported, clades with 94.2% similarity. This result is congruent with the study by Pretto *et al.* (2018) in which the topology of both tree SSU rRNA and RPB1 are not congruent for microsporidian parasites infecting crayfish. In addition, this observation is not surprising owing the poor support of our SSU phylogenetic reconstruction and the low sequences divergence. Due to the higher variation in the RPB1 gene, we can propose that the RPB1 tree corresponds more accurately to the between-strains parasite relationships within the parasite species. However, as the present study can be considered as being preliminary, RPB1 data are opening more doors than offering definitive conclusions (see next paragraph, but also conclusions and perspectives).

III.4.3 *N. granulosis* : two independent evolutionary histories?

Assuming that the tree topology described above is acceptable, two main clades are found within *N. granulosis* (Fig. 6). One is including the A and B haplogroups, and one clade is including the C haplogroups. Parasites from the clade A-B are *N. granulosis* infecting *G. roeselii* and *G. balcanicus*, while parasites from the second clade are infecting *G. roeselii* and *G. duebeni* (Ironside 2007, 2013; Quiles *et al.* 2019). This dichotomy does not fit at all the host phylogeny (Fig. 6), and we can, therefore, reject a simple co-phylogenetic pattern explaining a parasite differentiation according to the host differentiation. Indeed, freshwater gammarid groups including *G. roeselii* and *G. balcanicus* diverged from the group of saline gammarids, including *G. duebeni*, around 50 Ma (Hou *et al.* 2011) (summarized in Fig. 6A). Therefore, the closely-related parasites infecting

by *G. roeselii* and *G. duebeni* either issued from an infection older than 50 Ma (Fig. 6C), or is the result of a more recent horizontal transfer (Fig. 6B).

Fig. 6. Simplified phylogeny of the gammarids (A) compared to simplified phylogeny of *N. granulosis* (B, C) for building scenarios of *N. granulosis* evolutionary histories. A. Host simplified evolutionary history, redraw from (Hou *et al.* 2011). B. *N. granulosis* evolutionary scenario where divergence in the parasites is following that of the hosts, and where *N. granulosis* infection in *G. roeselii* is the result of a recent horizontal transfer from *G. duebeni*. C. *N. granulosis* evolutionary scenario where parasite diverged prior the diversification of the hosts. Two clades of microsporidia with contrasted transmission strategies (Horizontal transmission, HT, or Vertical transmission, VT) infected an ancestor of the genus *Gammarus*, Each of them then co-diverged following host divergence. This last scenario suppose that HT parasites were lost in *G. duebeni* and VT parasites were lost in *G. balcanicus*.



The former hypothesis seems the more probable to us for the following reasons (see also the discussion in (Quiles *et al.* 2019)). First, a recent horizontal transfer between *G. duebeni* and *G. roeselii* seems improbable because of the weak overlapping of *G. duebeni* and *G. roeselii* ranges. The spread of *G. roeselii* towards north-western Europe is recent, so the present weak overlap with the *G. duebeni* distribution was even weaker before the present days. In addition, the C2

haplogroup is found in the Greek populations of *G. roeselii*, associated to an ancient lineage of this host species (Quiles *et al.* 2019), at the very opposite of the *G. duebeni* geographical range. Second, two haplogroups are infecting *G. roeselii*, the C3 haplogroup being less closely related to the C1 *G. duebeni* haplogroup. Two horizontal transfers therefore seems improbable. This would mean that infections in gammarids by *Nosema granulosis* result from two independent evolutionary histories, leading to the present pattern (Fig. 6C). The clade A-B would, instead, have evolved in *G. roeselii* and *G. balcanicus* following a common ancestor infection (perhaps ancestral to the freshwater gammarid group only) and the clade C would have evolved from another ancestral infection (perhaps ancestral to the whole *Gammarus* group), but would have been lost in *G. balcanicus* (Fig. 6C). In addition to these phylogeographic argument, a biological argument, developed in the following paragraph, plead in favour of these two independent evolutionary histories *in Nosema granulosis* (Fig. 6).

Within the A-B clade, parasites infecting *G. balcanicus* and those infecting *G. roeselii* are showing some variation on their RPB1 gene, with 3 haplogroups each. This variation is found between *G. roeselii* populations, namely between the Greek population number 4 (GR25) and the Albanian population number 22 (AL35) populations for B haplogroups, and between the nine *G. balcanicus* populations from Albania, Hungary, Ukraine and Romania, for the A haplogroups (Fig. 5). Variation is also found within populations: Population number 4 harbours RPB1 B1 and B2 haplogroups in *G. roeselii*, while population 39 (H09) harbours RPB1 A2 and A3 haplogroups in *G. balcanicus*. The same haplogroups have even been observed co-occurring in some individuals, a sign of either double infection or and heterozygous status. The second hypothesis is more tempting as RPB1 is known to be a nuclear single-copy gene in *V. necatrix* (Hirt *et al.* 1999). This situation contrasts with the pattern of variation found for the parasites of the C clade infecting *G. roeselii*. We will not discuss the haplogroup C2 for which we have only one individual and therefore no possibility for describing variation. For the haplogroup C3, parasites were found infecting 45 individuals in 12 populations (from Hungary, Germany, Slovakia, Austria, Poland, France and Italy), without any variation. As discussed earlier (Chapter I of this thesis), the C3 haplogroup (SSU *N. granulosis* r01 haplogroup) is infecting the host MOTU C that did spread and invaded all Europe recently in its evolutionary history (Grabowski *et al.* 2017; Quiles *et al.* 2019). *Nosema granulosis* infections belonging to this haplogroup were previously described to rely on efficient transovarial transmission and feminization to exploit host populations (Haine *et al.* 2004, 2007 ; Terry *et al.* 2004). These *N. granulosis* C3, therefore, experience bottlenecks at each generation of host reproduction, because the number of parasites is restricted to few dozen within each egg (Haine *et al.* 2004). Such recurrent bottlenecks may explain the observed absence of

parasite diversity (Nei *et al.* 1975), even at a broad geographic scale. Therefore, the absence of variation on RPB1 in these parasites (despite the high variation found at the scale of our whole data set) fits well the hypothesis that this haplogroup is vertical transmitted, feminizing and do not show any virulence against its host, allowing both to invade and spread in new environments (Drake 2003). This hypothesis is strengthened by the ‘’closeness’’ of this haplogroup with the C1 haplogroup infecting *Gammarus duebeni* parasites are also known to induce feminisation after vertical transmission (Dunn & Smith 2001). However, this ‘’closeness’’ is challenging to understand, knowing the hosts having diverged ca 50 Ma ago. One would expect more divergence between C1 and C2-C3. We can however argue, that under a typical scenario of very strong bottleneck for parasites in both host species, any variation would be suppressed by drift.

By now, no study has been made on the biology, pathology and transmission pathways of A and B haplogroups of *N. granulosis*. Nevertheless, because of the observed variation on RPB1, and by contrast with the data observed on the C3 haplogroup, we can hypothesis that they do not use vertical transmission as day-to-day pathway for their life cycle. Several parasite genotypes infecting the same populations or very close populations of the same host can highlight the commitment of genetic diversity to arm-race between parasite and their host (Decaestecker *et al.* 2007; Lively *et al.* 1990; Routtu & Ebert 2015). This would mean that the RPB1 haplogroups A and B of *N. granulosis* are horizontally-transmitted and virulent to their hosts, hypotheses that remain to be tested experimentally. Some host individuals presented *N. granulosis* infections with RPB1 signatures of two haplogroups of the clade A. We were unable to discriminate if these observations are due to heterozygous individuals or to bi-infections involving two haplogroups in the same individuals. Verifying the second hypothesis would strengthen the horizontal nature of the transmission of these parasites.

Globally, therefore, our results allowed us to propose that two evolutionary trajectories may occur early within *Nosema granulosis* infecting gammarids, with some parasite variants inducing pathogeny to their hosts and being horizontally transmitted and other variants that evolved vertical transmission and feminization (Fig. 6C).

III.5 CONCLUSIONS AND PERSPECTIVES

Our study revealed that the use of sequences of the RPB1 gene, even partial, help to better understand *N. granulosis* evolution and host-parasite interactions. As already proposed in chapter II, we reiterate our hypothesis that *N. granulosis* infected members of the genus *Gammarus* at the very beginning of its differentiation process. Furthermore, we hypothesize that, following this ancient infection, two evolutionary trajectories were selected. On one side, some *N. granulosis*

haplogroups seems to used vertical transmission and showed a weak genetic diversity; on the other side, a sister clade seems to used horizontal transmission and retained more genetic diversity (Fig. 6).

These hypotheses-scenarios necessitate many further explorations, either by the improvement of the molecular analysis methods and/or by running further experiments.

For the molecular analyses, an increased taxon sampling for RPB1 sequences would help to challenge our current tree topology and associated scenarios. First, even if we were unfortunately unsuccessful up to now to get RPB1 sequences for *G. balcanicus* individuals harbouring haplogroups b01 and b02 SSU rDNA, more DNA should be available soon for further PCR attempts. Second, it will be essential to extend the number of *N. granulosis* RPB1 already present in our SSU rDNA tree sequences to those detected in other amphipod hosts. Notably, it will be essential to add sequences from the microsporidia detected in *Gammarus pulex* (Ironside & Alexander 2015)*, G. fossarum* and *Niphargus schellenbergi* (Weigand *et al.* 2016), that belong to the very same SSU rDNA haplogroup than our *N. granulosis* r03. If our hypothesis is correct, they should show different RPB1 sequences than those found in *G. roeselii*. Furthermore, RPB1 sequence corresponding to the parasite found in *G. duebeni* and *G. pulex*, but belonging to the separated small SSU rDNA "clade" (sequences FN434088 and KM657356) should be different also. Third, the RPB1 sequences from other closely related Microsporidia spp. (*e.g.* other *Nosema/Vairimorpha)* but infecting insect, could be included in the analyses (*e.g.* Pretto *et al.* 2018).

Finally, the biology of the newly discovered variants must be investigated. This is important, to verify, for example, if the RPB1 C2 haplogroup is a vertically transmitted parasite, as we hypothesized, and if all the A and B haplogroups are virulent and horizontally transmitted.

SUPPLEMENTARY MATERIAL

Chap. III, Table S1. *N. granulosis* individuals infection from Quiles *et al.* (2019), & Chap II, and found in Genbank (NCBI), mainly for fresh and brackish waters amphipoda species present in Europe.

Chap. II, Table S2. Variable sites in *Nosema granulosis* RPB1 sequences used to construct phylogenentic trees. Names in bold are microsporidia used for phylogeny. D – deletion, N – data not available.

Chap. II, Table S3. P-distance based on paiwise deletion for *Nosema granulosis* RPB1 patial sequences, calculated for the three positions of the codons (I), positions 1 and 2 only (II).

All suplementary material are available at:
https://cloud.u-bourgogne.fr:443/index.php/s/XP9ZW8wnr87kP7z

or scanning the following QR code:

REFERENCES

Andolfatto, P. (2001). Adaptive hitchhiking effects on genome variability. *Current Opinion in Genetics and Development*, 11(6), 635–641.

Bacela-Spychalska, K., Wróblewski, P., Mamos, T., … Ovcharenko, M. (2018). Europe-wide reassessment of *Dictyocoela* (Microsporidia) infecting native and invasive amphipods (Crustacea): Molecular versus ultrastructural traits. *Scientific Reports,* 8(1), 1–16.

Balbiani, G. (1882). Sur les microsporidies ou psorospermies des articules. C. R. Acad. Sci., 95(95), 1168–1171.

Brandon Matheny, P., Liu, Y. J., Ammirati, J. F., & Hall, B. D. (2002). Using RPB1 sequences to improve phylogenetic inference among mushrooms (Inocybe, Agaricales). American *Journal of Botany*, 89(4), 688–698.

Chen, Y. ping, Pettis, J. S., Zhao, Y., … Evans, J. D. (2013). Genome sequencing and comparative genomics of honey bee microsporidia, *Nosema apis* reveal novel insights into host-parasite interactions. *BMC Genomics*, 14(1), 451.

Cheney, S. A., Lafranchi-Tristem, N. J., Bourges, D., & Canning, E. U. (2001). Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). *The Journal of Eukaryotic Microbiology*, 48(1), 111–7.

Copilaş-Ciocianu, D., & Petrusek, A. (2017). Phylogeography of a freshwater crustacean species complex reflects a long-gone archipelago. *Journal of Biogeography*, 44(2), 421–432.

Cornman, R. S., Chen, Y. P., Schatz, M. C., … Evans, J. D. (2009). Genomic analyses of the microsporidian *Nosema ceranae*, an emergent pathogen of honey bees. *PLoS Pathogens*, 5(6). doi:10.1371/journal.ppat.1000466

Cramer, P., Bushnell, D. A., & Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 ångstrom resolution. *Science*, 292(5523), 1863–1876.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). JModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772.

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., … De Meester, L. (2007). Host-parasite "Red Queen" dynamics archived in pond sediment. *Nature*, 450(7171), 870–873.

Drake, J. M. (2003). The paradox of the parasites: Implications for biological invasion. *Proceedings of the Royal Society B: Biological Sciences*, 270(SUPPL. 2), 133–135.

Dunn, A. M., Hogg, J. C., & Hatcher, M. J. (2006). Transmission and burden and the impact of temperature on two species of vertically transmitted microsporidia. *International Journal for Parasitology*, 36(4), 409–414.

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

Franzen, C. (2008). Microsporidia: A Review of 150 Years of Research. *The Open Parasitology Journal,* 2, 1–34.

Fries, I. (2010). Nosema ceranae in European honey bees (*Apis mellifera*). *Journal of Invertebrate Pathology*, 103(SUPPL. 1), S73–S79.

Gisder, S., & Genersch, E. (2013). Molecular differentiation of *Nosema apis* and *Nosema ceranae* based on species-specific sequence differences in a protein coding gene. *Journal of Invertebrate Pathology*, 113(1), 1–6.

Gómez-Moracho, T., Maside, X., Martín-Hernández, R., Higes, M., & Bartolomé, C. (2014). High levels of genetic diversity in *Nosema ceranae* within *Apis mellifera* colonies. *Parasitology,* 141(4), 475–481.

Grabowski, M., Krzywozniak, P., Rewicz, T., Mamos, T., Bacela-spychalska, K., & Wattier, R. (2017a). *Gammarus roeselii* Gervais , 1835 (Gammaridae) in Western and Central Europe : post-glacial colonisation or human mediated introduction ? *Biodiversity Journal*, 8(2), 525–526.

Grabowski, M., Mamos, T., Bącela-Spychalska, K., Rewicz, T., & Wattier, R. A. (2017b). Neogene paleogeography provides context for understanding the origin and spatial distribution of cryptic diversity in a widespread Balkan freshwater amphipod. *PeerJ,* 5, e3016.

Guémn-Méneville M.-T.-F. (1850). Études sur les maladies des vers à soie. Observations sur la composition intime du sang chez les insectes et surtout chez les vers à soie en santé et en maladie, et sur la transformation des éléments vivants des globules de ce sang en rudiments du végétal. *Bulletin de La Société Nationale et Centrale d'agriculture de France*, 2, 251–259.

Haine, E. R., Brondani, E., Hume, K. D., Perrot-Minnot, M. J., Gaillard, M., & Rigaud, T. (2004). Coexistence of three microsporidia parasites in populations of the freshwater amphipod *Gammarus roeseli*: Evidence for vertical transmission and positive effect on reproduction. *International Journal for Parasitology,* 34(10), 1137–1146.

Haine, E. R., Motreuil, S., & Rigaud, T. (2007). Infection by a vertically-transmitted microsporidian parasite is associated with a female-biased sex ratio and survival advantage in the amphipod *Gammarus roeseli*. *Parasitology*, 134(Pt 10), 1363–1367.

Higes, M., Martín, R., & Meana, A. (2006). *Nosema ceranae*, a new microsporidian parasite in honeybees in Europe. *Journal of Invertebrate Pathology*, 92(2), 93–95.

Hirt, R. P., Logsdon, J. M., Healy, B., Dorey, M. W., Doolittle, W. F., & Embley, T. M. (1999). Microsporidia are related to Fungi: Evidence from the largest subunit of RNA polymerase II and other proteins. *Proceedings of the National Academy of Sciences*, 96(2), 580–585.

Hogg, J. C., Ironside, J. E., Sharpe, R. G., Hatcher, M. J., Smith, J. E., & Dunn, a M. (2002). Infection of *Gammarus duebeni* populations by two vertically transmitted microsporidia; parasite detection and discrimination by PCR-RFLP. *Parasitology*, 125(Pt 1), 59–63.

Hopper, J. V., Huang, W. F., Solter, L. F., & Mills, N. J. (2016). Pathogenicity, morphology, and characterization of a *Nosema fumiferanae* isolate (Microsporidia: *Nosematidae*) from the light brown apple moth, *Epiphyas postvittana* (Lepidoptera: *Tortricidae*) in California. *Journal of Invertebrate Pathology*, 134, 38–47.

Hou, Z., Sket, B., Fiser, C., & Li, S. (2011). Eocene habitat shift from saline to freshwater promoted Tethyan amphipod diversification. *Proceedings of the National Academy of Sciences*, 108(35), 14533–14538.

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

Hurst, G. D. D., & Majerus, M. E. N. (1993). Why do maternally inherited microorganisms kill males? *Heredity,* 71(1), 81–95.

Ironside, J. E. (2007). Multiple losses of sex within a single genus of Microsporidia. *BMC Evolutionary Biology*, 7(1), 48.

Ironside, J. E. (2013). Diversity and recombination of dispersed ribosomal DNA and protein coding genes in Microsporidia. *PLoS ONE*, 8(2). doi:10.1371/journal.pone.0055878

Ironside, J. E., & Alexander, J. (2015). Microsporidian parasites feminise hosts without paramyxean co-infection: Support for convergent evolution of parasitic feminisation. *International Journal for Parasitology*, 45(6), 427–433.

Ironside, J. E., Dunn, A. M., Rollinson, D., & Smith, J. E. (2003). Association with host mitochondrial haplotypes suggests that feminizing microsporidia lack horizontal transmission. *Journal of Evolutionary Biology*, 16(6), 1077–1083.

Ironside, J. E., Wilkinson, T. J., & Rock, J. (2008). Distribution and host range of the microsporidian *Pleistophora mulleri*. *Journal of Eukaryotic Microbiology*, 55(4), 355–362.

Jahnke, M., Smith, J. E., Dubuffet, A., & Dunn, A. M. (2013). Effects of feminizing microsporidia on the masculinizing function of the androgenic gland in *Gammarus duebeni* . *Journal of Invertebrate Pathology*, 112(2), 146–151.

Katoh, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780.

Kearse, M., Moir, R., Wilson, A., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28(12), 1647–1649.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2001). Population dynamics of a vertically transmitted, parasitic sex ratio distorter and its amphipod host. *Oikos*, 94(3), 392–402.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2002). Incomplete feminisation by the microsporidian sex ratio distorter, *Nosema granulosis*, and reduced transmission and feminisation efficiency at low temperatures. *International Journal for Parasitology*, 32(7), 825–831.

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. *Infection, Genetics and Evolution,* 10(7), 1027–1038.

Ku, C. T., Wang, C. Y., Tsai, Y. C., Tzeng, C. C., & Wang, C. H. (2007). Phylogenetic analysis of two putative *Nosema* isolates from Cruciferous Lepidopteran pests in Taiwan. *Journal of Invertebrate Pathology*, 95(1), 71–76.

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

Kyei-Poku, G., & Sokolova, Y. Y. (2017). The microsporidium *Nosema disstriae* (Thomson 1959): Fine structure and phylogenetic position within the *N. bombycis* clade. *Journal of Invertebrate Pathology*, 143, 90–103.

Leffler, E. M., Bullaughey, K., Matute, D. R., … Przeworski, M. (2012). Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLoS Biology*, 10(9). doi:10.1371/journal.pbio.1001388

Lewis, L. C., Bruck, D. J., Prasifka, J. R., & Raun, E. S. (2009). *Nosema pyrausta*: Its biology, history, and potential role in a landscape of transgenic insecticidal crops. *Biological Control*, 48(3), 223–231.

Liu, H., Pan, G., Li, T., Huang, W., Luo, B., & Zhou, Z. (2012). Ultrastructure, chromosomal karyotype, and molecular phylogeny of a new isolate of microsporidian *Vairimorpha* sp. BM (Microsporidia, *Nosematidae*) from *Bombyx mori* in China. *Parasitology Research*, 110(1), 205–210.

Liu, T., Xu, J., Luo, B., … Zhou, Z. (2015). *Nosema* sp. PM-1, a new isolate of microsporidian from infected *Papilio machaon* Linnaeus, based on ultrastructure and molecular identification. *Acta Parasitologica*, 60(2), 330–336.

Lively, C. M., Craddock, C., & Vrijenhoek, R. C. (1990). Red Queen hypothesis supported by parasitism in sexual and clonal fish. *Nature,* pp. 864–866.

Loocwood, J. A., Bomar, C. R., & Ewen, A. L. B. (1999). The history of biological control with *Nosema locustae*: Lessons for locust management. *Insect Science and Its Application*, 19(4), 333–350.

Madden, T. (2003). The BLAST Sequence Analysis Tool. In *The NCBI Handbook, National Center for Biotechnology Information* (US), pp. 1–15.

Mamos, T., Wattier, R., Burzyński, A., & Grabowski, M. (2016). The legacy of a vanished sea: A high level of diversification within a European freshwater amphipod species complex driven by 15 Ma of Paratethys regression. *Molecular Ecology*, 25(3), 795–810.

Mamos, T., Wattier, R., Majda, A., Sket, B., & Grabowski, M. (2014). Morphological vs. Molecular delineation of taxa across montane regions in Europe: The case study of *Gammarus balcanicus* Schäferna, (Crustacea: Amphipoda). *Journal of Zoological Systematics and Evolutionary Research*, 52(3), 237–248.

Maside, X., Gómez-Moracho, T., Jara, L., … Bartolomé, C. (2015a). Population genetics of *Nosema apis* and *Nosema ceranae*: One host (apis mellifera) and two different histories. *PLoS ONE*, 10(12), 1–21.

Maside, X., Gómez-Moracho, T., Jara, L., … Bartolomé, C. (2015b). Population genetics of *Nosema apis* and *Nosema ceranae*: One host (*apis mellifera*) and two different histories. *PLoS ONE*, 10(12), 1–21.

Mathis, A. (2000). Microsporidia: emerging advances in understanding the basic biology of these unique organisms. *International Journal for Parasitology*, 30(7), 795–804.

Naegeli, K. (1857). Über die neue Krankheit der Seidenraupe und verwandte Organismen. *Bot Ztg*, 15, 760–761.

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, 29(1), 1.

Pasteur, L. (1870). Etudes sur la maladie des vers a soie. Moyen pratique assuré de la combattre et d'en prévenir le retour. *Gauthier-Villars*, 2.

Pretto, T., Montesi, F., Ghia, D., … Fea, G. (2018). Ultrastructural and molecular characterization of *Vairimorpha austropotamobii* sp. nov. (Microsporidia: *Burenellidae*) and *Thelohania contejeani* (Microsporidia: *Thelohaniidae*), two parasites of the white-clawed crayfish, *Austropotamobius pallipes* complex (D. *Journal of Invertebrate Pathology*, 151(November 2017), 59–75.

Quiles, A., Bacela-Spychalska, K., Teixeira, M., … 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. *Parasites and Vectors,* 12(1), 1–20.

Rodgers-Gray, T. P., Smith, J. E., Ashcroft, A. E., Isaac, R. E., & Dunn, A. M. (2004). Mechanisms of parasite-induced sex reversal in *Gammarus duebeni*. *International Journal for Parasitology*, 34(6), 747–753.

Routtu, J., & Ebert, D. (2015). Genetic architecture of resistance in Daphnia hosts against two species of host-specific parasites. *Heredity*, 114(2), 241–248.

Ryan, J. A., & Kohler, S. L. (2010). Virulence is context-dependent in a vertically transmitted aquatic host-microparasite system. *International Journal for Parasitology*, 40(14), 1665–1673.

Stentiford, G. D., Becnel, J. J., Weiss, L. M., … Solter, L. (2016). Microsporidia – Emergent Pathogens in the Global Food Chain. *Trends in Parasitology*, 32(4), 336–348.

Stiller, J. W., & Hall, B. D. (1997). The origin of red algae: Implications for plastid evolution. *Evolution Communicated by Robert Haselkorn,* 94(April), 4520–4525.

Terry, R. S., Dunn, A. M., & Smith, J. E. (1997). Cellular distribution of a feminizing microsporidian parasite: A strategy for transovarial transmission. *Parasitology*, 115(2), 157–163.

Terry, R. S., Smith, J. E., Bouchon, D., … Dunn, A. M. (1999). Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF) microsporidium. *Journal of Eukaryotic Microbiology*, 46(5), 492–499.

Terry, R. S., Smith, J. E., & Dun, A. M. (1998). Impact of a novel, feminising microsporidium on its crustacean host. *J. Euk. Microbiol*., (5), 497–501.

Terry, R. S., Smith, J. E., Sharpe, R. G., … Dunn, A. M. (2004). Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum *Microspora*. *Proceedings of the Royal Society B: Biological Sciences*, 271(1550), 1783–1789.

Tokarev, Y. S., Timofeev, S. A., Malysh, J. M., … Dolgikh, V. V. (2019). Hexokinase as a versatile molecular genetic marker for Microsporidia. *Parasitology*. doi:10.1017/S0031182018001737

Vavra, J., Hylis, M., Vossbrinck, C. R., … 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*). *Journal of Eukaryotic Microbiology*, 53(4), 292–304.

Vávra, J., & Lukeš, J. (2013). Microsporidia and 'The Art of Living Together.' *In Advances in Parasitology*, Vol. 82, pp. 253–319.

Vossbrinck, C. R., Debrunner-Vossbrinck, B. A., & Weiss, L. M. (2014). Phylogeny of the Microsporidia. In Microsporidia: *Pathogens of Opportunity: First Edition, Chichester, UK: John Wiley & Sons, Inc*., pp. 203–220.

Weedall, R. T., Robinson, M., Smith, J. E., & Dunn, A. M. (2006). Targeting of host cell lineages by vertically transmitted, feminising microsporidia. *International Journal for Parasitology*, 36(7), 749–756.

Weigand, A. M., Kremers, J., Grabner, D. S., … S., G. (2016). Shared microsporidian profiles between an obligate (*Niphargus*) and facultative subterranean amphipod population (*Gammarus*) at sympatry provide indications for underground transmission pathways. *Limnologica*, 58, 7–10.

Winters, A. D., & Faisal, M. (2014). Molecular and ultrastructural characterization of *Haplosporidium diporeiae* n. sp., a parasite of *Diporeia* sp. (Amphipoda, Gammaridea) in the Laurentian Great Lakes (USA). *Parasite*, 7(1), 343.

Xu, J., Pan, G., Fang, L., … Xiang, Z. (2006). The varying microsporidian genome: Existence of long-terminal repeat retrotransposon in domesticated silkworm parasite *Nosema bombycis*. *International Journal for Parasitology*, 36(9), 1049–1056.

Xu, J., & Zhou, Z. (2010). Improving phylogenetic inference of microsporidian *Nosema antheraeae* among *Nosema* species with RPB1 , -and -tubulin sequences. *African Journal of Biotechnology*, 9(46), 7900–7904.

CHAPTER IV. GENERAL CONCLUSION AND PERSPECTIVES

Chapters I and II presented patterns of Microsporidian infections, as based on SSU rDNA, in two *Gammarus* species, *G. roeselii* and *G. balcanicus,* over their full geographic ranges and being based on large sampling efforts (each *c*. 100 sites and 2000 individuals). Only one other paper by Krebes *et al.* (2010) on *G. duebeni* could matched these features with 35 sites and 898 samples.

The SSU rDNA marker revealed both efficient at detecting variants (haplogroups) of microsporidia and assigning a specimen to a given parasite species-level taxa. Microsporidian diversity was high in both host species, with 24 and 54 haplogroups in *G. roeselli* and *G. balcanicus*, respectively. These haplogroups clustered into 18 species-level taxa, most of them being shared between the two host species. Nine of these taxa were known from the literature to be already associated with *G. roeselii* (Bojko *et al.* 2017; Grabner *et al.* 2015; Haine *et al.* 2004). In *G. balcanicus*, only few species were known (Bacela-Spychalska *et al.* 2018; Bojko & Ovcharenko 2018), illustrating the usefulness of large sampling effort associated to my PhD work in assessing the extend of microcrosporidian species diversity in a given host.

Globally, this part of my PhD work points out:

i) the benefits of large sampling effort (*c*. 2000 individuals for each host) at allowing to detect taxa with rare occurrence.

ii) the benefits of appraising host-parasite association at full host range scale, this range being wide (continental scale).

iii) the large taxonomic diversity of Microsporidia infecting *G. roeselli* and *G. balcanicus*, and more widely gammarids.

iv) that these parasite taxa are often shared by many host taxa, not necessarily closely related, illustrating the complexity of the evolutionary history of host-parasite association.

v) that however, many microsporidia haplogroups within a given species are host-specific, suggesting specificity.

vi) the limits of SSU rDNA with low phylogenetic content reached at low taxonomic level in some taxa, notably limiting clear appraisal of host-parasite association within *G. roeseli* and *G. balcanicus* species-complexes. However, the confrontation between hosts and parasites

phylogeography allowed us to propose some degrees of co-diversification within each host species-complex, which remains to be confirmed.

However, the most striking output of these studies is that almost all of these microsporidia taxa were previously detected in other gammarids, mainly within the genus *Gammarus*, but also in other genera of Gammaridae. The species already clearly recognised as a being associated with Amphipods were: *Nosema granulosis*, *Dictyocoela roeselum, D. muelleri, D. roeselum, D. duebenum*, *D. berillonum, Cucumispora roeselum, C. ornata, Microsporidium* sp 515 and *Microsporidium* sp 505 (See Table 1 in the introduction). Many times, my results increased host taxonomic spectrums and extended geographic ranges (often widely), but we did not find any new frequent parasite species linked to one of these two European host species, despite extensive investigations. Indeed, the almost entire geographic distribution of these gammarids and the totality of the genetic (cryptic) diversity were investigated. These parasites may be frequent in a given host species, but their prevalence greatly vary between hosts. Since, at the moment, no *Dictyocoela*, no *Cucumispora* and no *Microsporidia* of the 505-515 group were found outside amphipod hosts, we may propose the hypothesis that these parasite groups are specific to this host group. This would mean ancient infections. More hots species are now needed to be investigated intensively to confirm our feeling that in numerous cases there could have been co-diversification scenarios among these microsporidian groups and their *Gammarus* hosts. Some other taxa were known to be extremely rare, having scarce literature records often with few or even very few geographic records and being not fully described (*e.g. Microsporidium* sp RR1, *Microsporidium* sp RR2, *Microsporidium* sp I) (*e.g.* (Grabner *et al.* 2015)). My PhD work either extend host taxonomic spectrum and/or deeply extend geographic ranges for these taxa. It clearly allowed a reappraisal for such taxa, changing their status from puzzling anecdotic association (proposed to result from transient spill over) to potentially overlooked established associations for amphipods. The data for *Cucumispora roeselum* are similarly improved. This parasite have been fully described recently infecting *G. roeselii* (Bojko *et al.* 2017a), but it was based on very few geographic records. My PhD clearly extend geographically its range, but also add *G. balcanicus* to its host spectrum. Finally, two Microsporia taxa, *Cucumispora dikerogammari* and *C. ornata* were aknowledged in the litterature as being primarily associated with two Ponto-Caspian invasive species, *Dikerogammarus vilosus* and *D. haemobaphes* (Bojko *et al.* 2015; Wattier *et al.* 2007). My PhD work present records of both parasites species in both host species: spill over, recent acquisition, older unsuspected association? Even if the first scenarion is favoured, the question remains open.

Among the most frequent parasite species detected in *G. roeselii* and *G. balcanicus*, a few were clearly established in the litterature as vertically-transmitted and having a feminizing effect in their hosts : *Nosema granulosis* (in both *Gammarus duebeni* and *G. roeselii*) and *Dictyocoela duebenum* (in *G. duebeni*) (Haine *et al.* 2004, 2007; Ironside & Alexander 2015; Kelly *et al.* 2002; Terry *et al.* 1999). We were unable to confirm these hypotheses in *Gammarus balcanicus*. Indeed, no clear female-biased infection has been detected. It therefore seems, and the use of the RPB1 gene as new more variable marker confirmed this feeling (see below), that some peculiar strains of these suposably feminizing parasites are not transmitted vertically and not feminizing at all. This has been previously suggested for *Dictyocoela duebenum* infecting *Gammarus pulex* (Ironside & Alexander 2015), but my PhD work strengthen and expand this observation to *Nosema granulosis* infecting both *G. balcanicus* and *G. roeselii*. For the later host, however, a within-species variation would occur since at least two parasite strains co-exist (see below).

The chapter III, of this thesis clearly showed out the limits of SSU rDNA in providing any conclusive phylogenetic content at fine taxonomic level. Focusing on *N. granulosis* only, RPB1 markers proved at lot more useful in that matter. SSU rDNA was known to present 3 haplogroups of *Nosema granulosis* for both *G. roeselii* and *G. balcanicus*.

Although fully acknowledging that results are still preliminary, mainly associated with incomplete Gammarid and parasite taxa samplings, RPB1 data allowed both to clearly identify divergent clades and to propose a scenario involving two evolutionary trajectories followed by the parasite *N. granulosis* through the evolution of its hosts. Within *N. granulosis*, vertical transmission seem to have been selected for some lineages, showing few genetic diversity, when in others, horizontal transmission seem to have been selected, showing more genetic diversity and molecular divergence. These hypotheses necessitate many further explorations, either by the improvement of the molecular analysis methods, increasing taxon sampling and/or by running further experiments. For example, the biology of the newly discovered variants must be investigated to find if the hypothesis about the vertically transmitted clade and the horizontally transmitted clade is confirmed. These methods could be supplemented by in situ hybridization techniques (FISH) and quantitative PCR that proved to be efficient in investigating the biology of intracellular parasites (*e.g.* Dubuffet *et al.* 2013; Genty *et al.* 2014; Unckless *et al.* 2009). Similarly, the effects on female fecundity and reproductive strategies (number and timing of egg-laying, for example) should be evaluated. These reproductive traits could be compared between naturally infected or microinjected females and healthy females (*e.g.* Bacela-Spychalska *et al.* 2014; Hatcher *et al.* 2005). The effect on the sex-ratio of the host could be evaluated by correlating

the parasite vertically transmitted to the sex-ratio of the offspring, for both naturally infected and experimentally infected females.

Globally, this part of my PhD work points out:

i) the limits of SSU rDNA reached low phylogenetic content at low taxonomic level.

ii) the high potential, on the opposite, of RPB1 at revealing diversity AND divergence, highlighting the fact that within what is recognised as one species level taxa by partial SSU rDNA can encompass divergent lineages.

iii) that polymorphism for VT vs HT trait could supposedly exist within what is recognised as one species level taxa by partial SSU rDNA.

iv) combining both RPB1 divergence and differences in their genetic variation, one could be tempted to reassess taxomonic status of *Nosema granulosis* as a single species.

As a final note, further study on parasite assemblage of other gammarids host species both local and full-range scale will be needed to successfully unravel what appear more and more to be complex host-parasite network, both in space and time. Such studies should include many *Gammarus* species, but also species from other genera within the Gammaridae family. Involving several taxonomic and temporal scales might help to understand host-parasite specificity as well as local and global parasite fauna. However, as most species of amphipods are highly cryptic species complexes, these studies will fully fruitful only if cryptic diversity and phylogeographic histories these hosts are also available, to get enough information allowing co-divergence to be assess at species-complexes level. In addition, RPB1 proved useful for *Nosema granulosis*, but is only one gene. In order to avoid presenting gene story of but truly a species story, adding more genes would be required. Luckily, other marker *e.g.* tubulins (Ku *et al.* 2007; Liu *et al.* 2015), Hexokinase (Tokarev *et al.* 2019) or microsporidian polar tube proteins genes (PTPs, (Gómez-Moracho *et al.* 2014; Maside *et al.* 2015) were sucessful developed, althouh on *Nosema* species infecting insects. Beyond *Nosema*, such markers could be used in other microsporidian species infecting gammarids for a better understanding of their co-evolutionary histories. No doubt that the recent sequencing of the genomes of three micropsoridia infecting *G. roeselii* (*N. granulosis*, *Dictyocoela roeselum* and *D. mulleri*; Alexandre Cormier & Richard Cordaux, Comm. Person.) will help in this matter.

REFERENCES

Bacela-Spychalska, K., Rigaud, T., & Wattier, R. A. (2014). A co-invasive microsporidian parasite that reduces the predatory behaviour of its host *Dikerogammarus villosus* (Crustacea, Amphipoda). *Parasitology*, 141(2), 254–258.

Bacela-Spychalska, K., Wróblewski, P., Mamos, T., … Ovcharenko, M. (2018). Europe-wide reassessment of *Dictyocoela* (Microsporidia) infecting native and invasive amphipods (Crustacea): Molecular versus ultrastructural traits. *Scientific Reports*, 8(1), 1–16.

Bojko, J., Bącela-Spychalska, K., Stebbing, P. D., … Stentiford, G. D. (2017). Parasites, pathogens and commensals in the "low-impact" non-native amphipod host *Gammarus roeselii*. *Parasites & Vectors*, 10(1), 193.

Bojko, J., Dunn, A. M., Stebbing, P. D., Ross, S. H., Kerr, R. C., & Stentiford, G. D. (2015). *Cucumispora ornata* n. sp. (Fungi: Microsporidia) infecting invasive "demon shrimp" (*Dikerogammarus haemobaphes*) in the United Kingdom. *Journal of Invertebrate Pathology*, 128, 22–30.

Bojko, J., & Ovcharenko, M. (2018). Pathogens and other symbionts of the Amphipoda: a review of their taxonomic diversity and pathological significance. *Diseases of Aquatic Organisms*. doi:10.3354/dao03321

Dubuffet, A., Smith, J. E., Solter, L., Alejandra Perotti, M., Braig, H. R., & Dunn, A. M. (2013). Specific detection and localization of microsporidian parasites in invertebrate hosts by using in situ hybridization. *Applied and Environmental Microbiology*, 79(1), 385–388.

Genty, L. M., Bouchon, D., Raimond, M., & Bertaux, J. (2014). Wolbachia infect ovaries in the course of their maturation: Last minute passengers and priority travellers? *PLoS ONE*. doi:10.1371/journal.pone.0094577

Gómez-Moracho, T., Maside, X., Martín-Hernández, R., Higes, M., & Bartolomé, C. (2014). High levels of genetic diversity in *Nosema ceranae* within *Apis mellifera* colonies. *Parasitology*, 141(4), 475–481.

Grabner, D. S., Weigand, A. M., Leese, F., … Sures, B. (2015). Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. *Parasites and Vectors*, 8(1), 1–4.

Haine, E. R., Brondani, E., Hume, K. D., Perrot-Minnot, M. J., Gaillard, M., & Rigaud, T. (2004). Coexistence of three microsporidia parasites in populations of the freshwater amphipod

Gammarus roeseli: Evidence for vertical transmission and positive effect on reproduction. *International Journal for Parasitology*, 34(10), 1137–1146.

Haine, E. R., Motreuil, S., & Rigaud, T. (2007). Infection by a vertically-transmitted microsporidian parasite is associated with a female-biased sex ratio and survival advantage in the amphipod *Gammarus roeseli*. *Parasitology*, 134(Pt 10), 1363–1367.

Hatcher, M. J., Hogg, J. C., & Dunn, A. M. (2005). Local adaptation and enhanced virulence of *Nosema granulosis* artificially introduced into novel populations of its crustacean host, *Gammarus duebeni* . *International Journal for Parasitology*, 35(3), 265–274.

Ironside, J. E., & Alexander, J. (2015). Microsporidian parasites feminise hosts without paramyxean co-infection: Support for convergent evolution of parasitic feminisation. *International Journal for Parasitology*, 45(6), 427–433.

Kelly, A., Dunn, A. M., & Hatcher, M. J. (2002). Incomplete feminisation by the microsporidian sex ratio distorter, *Nosema granulosis*, and reduced transmission and feminisation efficiency at low temperatures. *International Journal for Parasitology*, 32(7), 825–831.

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. *Infection, Genetics and Evolution*, 10(7), 1027–1038.

Ku, C. T., Wang, C. Y., Tsai, Y. C., Tzeng, C. C., & Wang, C. H. (2007). Phylogenetic analysis of two putative *Nosema* isolates from Cruciferous Lepidopteran pests in Taiwan. *Journal of Invertebrate Pathology*, 95(1), 71–76.

Liu, T., Xu, J., Luo, B., … Zhou, Z. (2015). Nosema sp. PM-1, a new isolate of microsporidian from infected *Papilio machaon* Linnaeus, based on ultrastructure and molecular identification. *Acta Parasitologica*, 60(2), 330–336.

Maside, X., Gómez-Moracho, T., Jara, L., … Bartolomé, C. (2015). Population genetics of *Nosema apis* and *Nosema ceranae*: One host (apis mellifera) and two different histories. *PLoS ONE,* 10(12), 1–21.

Terry, R. S., Smith, J. E., Bouchon, D., … Dunn, A. M. (1999). Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF) microsporidium. *Journal of Eukaryotic Microbiology*, 46(5), 492–499.

Tokarev, Y. S., Timofeev, S. A., Malysh, J. M., … Dolgikh, V. V. (2019). Hexokinase as a versatile molecular genetic marker for Microsporidia. *Parasitology*. doi:10.1017/S0031182018001737

Unckless, R. L., Boelio, L. M., Herren, J. K., & Jaenike, J. (2009). Wolbachia as populations within individual insects: Causes and consequences of density variation in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, 276(1668), 2805–2811.

Wattier, R. A., Haine, E. R., Beguet, J., … Rigaud, T. (2007). No genetic bottleneck or associated microparasite loss in invasive populations of a freshwater amphipod. *Oikos*, 116(11), 1941–1953.

RÉSUMÉ

Titre: Histoire évolutive des symbioses entre les microsporidies et leurs hôtes amphipodes: contribution de l'étude de deux hôtes sur leurs aires de repartion géographiques.

Mots clés: Symbioses, phylogénie, phylogéographie, amphipodes, hôte parasite, microsporidies.

Résumé: Les microsporidies sont des endoparasites obligatoires, ils utilisent deux modes de transmission pour expoiter leurs hotes. Certaines espèces utilisent la transmission verticale et d'autres la transmission horizontale. La première favorise la co-spéciation et la spécificité hôte-parasite, tandis que la seconde favorise les échanges de parasites entre espèces hôtes. Les amphipodes d'eau douce sont des hôtes privilégiers pour de nombreuses espèces de microsporidies, mais aucun schéma général de spécificité hôte-parasite ou de co-diversification n'émerge des recherches menées depuis trente ans.

Dans ce travail, la diversité des infections microsporidiennes, sur la base des données SSU rdNA, ont été évaluées dans deux complexes d'espèces de *Gammarus*, *G. roeselii* et *G. balcanicus*, sur l'ensemble de leur aire de répartition géographique, en s'appuyant sur des efforts d'échantillonnage importants (environ 100 sites et 2 000 individus). L'objectif de ce doctorat est (i) d'explorer la diversité microsporidienne présente dans les deux hôtes et leurs relations phylogénétiques; (ii) vérifier si la phylogéographiques de l'hôte peut expliqué l'association hôte-parasite (co-diversifications ou récents aquisition a partir de la faune locale); (iii) proposer des scénarios d'histoire évolutive hôte-parasite pour expliquer la diversité et le schéma co-bio-géographique observé chez les deux espèces hôtes chez *N. granulosis*.

Le marqueur SSU d'ADNr a révélé un grand nombre de variants chez ces deux hôtes (c'est-à-dire 24 et 54 haplogroupes de, respectivement), regroupés en 18 taxons d'espèces, presque tous partagés entre les deux espèces hôtes. De nombreux haplogroupes de microsporidia appartenant à une espèce parasite donnée sont spécifiques à l'hôte, ce qui suggère un certain niveau de codiversification à cette échelle. Au sein de chacun des complexes d'espèces hôtes, la confrontation des phylogéographies des hôtes et des parasites suggère également un certain degré de co-diversification. Ceci reste néanmoins à confirmer, principalement parce que l'ADNr SSU

atteint ses limites en termes d'informations phylogénétiques à cette échelle plus fine. Il est frappant de constater que presque tous ces taxons de microsporidia mis en évidence dans ce travail avaient déjà été détectés chez d'autres gammaridés. Certains étaient déjà des taxons parasites clairement identifiés associés aux amphipodes: *Nosema granulosis, Dictyocoela roeselum, D. muelleri, D. roeselum, D. duebenum, D. berillonum, Cucumispora roeselum, C. ornata, Microsporidium sp 515* et *Microsporidium sp 505*). Mes résultats ont élargis les spectres taxonomiques d'hôte et les répartitions géographiques de ces parasites (souvent très étendues). Certains autres taxons étaient connus pour être extrêmement rares, et ne sont pas complètement décrits. Mon travail de doctorat a également étendu le spectre taxonomique d'hôte et / ou élargi considérablement les aires géographiques de ces taxons. Cela permet une réévaluation de ces taxons, faisant passer leur statut d'associations anecdotiques à des associations potentiellement bien établies, mais jusque là négligées, chez les amphipodes.

Parmi les espèces de parasites les plus fréquemment détectées chez *G. roeselii* et *G. balcanicus*, quelques-unes étaient clairement établies dans la littérature comme étant à transmission verticale et ayant un effet féminisant sur leurs hôtes. Chez *Gammarus balcanicus*, pour la seule de ces espèces pour laquelle des tests statistiques ont pu être conduits, nous n'avons pas pu confirmer cette caractéristique, ce qui montre qu'une espèce donnée de microsporidie ne peut pas être considérée comme étant « à transmission verticale » sur l'ensemble de son spectre d'hôte.

En outre, en se concentrant sur *Nosema granulosis*, l'utilisation d'un marqueur moléculiare supplémentaire, à savoir RPB1, a permis: i) d'identifier clairement les clades divergents au sein de cette espèce (alors que l'ADNr SSU n'était pas concluant) et ii) de suggérer que certaines souches particulières de ces parasites ne sont pas transmises verticalement chez certaines espèces d'hôtes, voire au sein d'une même espèce. Ainsi, chez *G. roeseli*, un polymorphisme pourrait exister, une lignée étant associée à une transmission verticale, l'autre à une transmission horizontale. Ces résultats sont encore préliminaires. Ils sont en effet associés à un échantillonnage incomplet des taxons d'hôtes gammaridés et nécessitent des expériences supplémentaires pour confirmer nos hypothèses.

UBFC
UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ

FACULTY OF BIOLOGY
AND ENVIRONMENTAL
PROTECTION
University of Lodz

ABSTRACT

Tytuł: Ewolucja symbiozy pomiędzy mikrosporydiami i ich żywicielami, skorupiakami obunogimi: badania dwóch żywicieli w ich zasięgu geograficznym.

Słowa kluczowe: symbioza, filogeneza, filogeografia, skorupiaki obunogie, żywiciel-pasożyt, Mikrosporydia

Streszczenie: Mikrosporydia to obligatoryjne pasożyty wewnętrzne, wykorzystujące swoich żywicieli do transmisji pionowej lub poziomej. Podczas gdy pierwszy sposób transmisji może promować kospecjację i najczęściej jest specyficzny względem żywiciela, ten drugi może promować zdolność od infekowania wielu gatunków żywicieli. Słodkowodne skorupiaki obunogie są żywicielami wielu gatunków mikrosporydiów, ale nie jest znany ogólny wzorzec swoistości i kodywersyfikacji żywicieli.

W mojej pracy doktorskiej oceniono stopień zakażenia mikrosporydiami, zidentyfikowane za pomocą rDNA SSU, w dwóch kompleksach gatunków z rodzaju *Gammarus*: *G. roeselii* i *G. balcanicus*, w ich całkowitym zasięgu geograficznym (ok. 100 stanowisk i 2000 osobników dla każdego gatunku) w celu (i) oceny różnorodności mikrosporydiów zidentyfikowanych u tych dwóch żywicieli, jak i ich relacji filogenetycznych; (ii) przetestowania, czy historia filogeograficzna żywiciela mogła mieć wpływ na zespół żywiciel-pasożyt (kodywersyfikacja lub niedawne nabycie pasożyta od lokalnej fauny); (iii) zaproponowanie scenariusza historii ewolucji układu żywiciel- pasożyt w celu wyjaśnienia różnorodności i kobiogeograficznego wzoru obserwowanego u dwóch gatunków żywicieli na przykładzie *N. granulosis*.

Na podstawie markera SSU rDNA wykazano dużą liczbę wariantów mikrosporydiów (tj. liczba wyznaczonych haplogrup wyniosła odpowiednio 24 i 54), zgrupowanych w 18 gatunków, z których prawie wszystkie są wspólne dla dwóch gatunków żywicieli. Jednak wiele haplogrup w obrębie danego gatunku pasożyta jest specyficznych dla gatunku żywiciela, co sugeruje kodywersyfikację żywiciel-pasożyt. W obrębie każdego kompleksu gatunków badanych obunogów, podczas gdy porównanie filogenezy żywiciela i pasożyta mogła sugerować pewien stopień kodywersyfikacji, stwierdzono, że wzorce te muszą zostać potwierdzone poprzez bardziej

szczegółowe badania, ponieważ użyty marker SSU rDNA niesie ograniczoną informację filogenetyczną.

Uderzające jest to, że prawie wszystkie ze zidentyfikowanych mikrosporydiów zostały wcześniej wykryte u innych kiełży (Gammaroidea), głównie w rodzaju *Gammarus*. Niektóre były uznane za mikrosporydia wyraźnie związane z obunogami:: *Nosema granulosis*, *Dictyocoela roeselum*, *D. muelleri*, *D. roeselum*, *D. duebenum*, *D. berillonum*, *Cucumispora roeselum*, *C. ornata*, *C. dikerogammari*, *Microsporidium* sp. 515 i *Microsporidium* sp. 505). W wielu przypadkach wyniki mojej pracy znacząco zwiększyły spektrum taksonomiczne żywiciela jak i rozszerzyły zasięg geograficzny (często szeroko). Niektóre taksony pasożytów były niezwykle rzadkie, posiadając rzadkie rekordy w literaturze, często z niewielką lub nawet bardzo niewielką liczbą danych geograficznych, dodatkowo nie zostały opisane. Moja praca doktorska albo poszerza spektrum taksonomiczne żywiciela i / lub znacząco poszerza zasięg geograficzny dla tych taksonów. Umożliwiło to dokonanie ponownej oceny takich taksonów, zmieniając ich status z anegdotycznego zespołu pasożytów na potencjalnie przeoczone stałe zespoły związane z obunogami.

Spośród najczęstszych gatunków pasożytów wykrytych u *G. roeselii* i *G. balcanicus*, kilka było opisanych jako przenoszone pionowo i mających działanie feminizujące na swoich zywicieli. W przypadku *G. balcanicus*, jedynego gatunku, dla którego można było wykonać taka analizę, nie byliśmy w stanie potwierdzić tej obserwacji, pokazując, że pojedynczy gatunek mikrosporydiów nie może być uważany za całkowicie pionowo przeniesiony w całym spektrum gospodarza.

Ponadto, badania N. granulosis z zastosowaniem dodatkowego markera, tj. RPB1, pozwoliło na: i) wyraźne zidentyfikowanie odrębnych kladów (podczas gdy SSD rDNA nie było jednoznaczne) oraz ii) stwierdzenie, że niektóre linie tych rzekomo feminizujących pasożytów nie są przenoszone pionowo i nie maja efektu feminizującego. Ponadto, w przypadku *G. roeselii* może istnieć polimorfizm genetyczny względem dróg przenoszenia, gdzie czym jedna linia filogenetyczna jest związana z transmisją pionową, a druga z transmisją poziomą. Te wyniki są nadal wstępne, głównie z powodu niepełnego pokrycia zróżnicowania obserwowanego u żywiciela i pasożyta i wymagają dalszych badań dla potwierdzenia naszej hipotezy.

UBFC
UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ

FACULTY OF BIOLOGY
AND ENVIRONMENTAL
PROTECTION
University of Lodz

ABSTRACT

Title: Evolutionary histories of symbioses between microsporidia and their amphipod hosts : contribution of studying two hosts over their geographic ranges.

Keywords: Symbioses, Phylogeny, Phylogeography, Amphipods, Host-Parasite, Microsporidia

Abstract: Microsporidia are obligate endoparasites, exploiting their hosts with either vertical or horizontal transmission. While the former may promote co-speciation and host-specificity, the latter may promote shifts between host species. Freshwater amphipods are hosts for many microsporidian species, but no general pattern of host specificity and co-diversification is known.

In my PhD work microsporidian infections, identified with SSU rDNA, were assessed in two *Gammarus* species complexes, *G. roeselii* and *G. balcanicus* , over their full geographic ranges (each *c*. 100 sites and 2000 individuals) in aim of (i) exploring the microsporidian diversity present in both hosts and their phylogenetic relationships; (ii) testing if the host phylogeographic history might have impacted host-parasite association (co-diversifications or recent host-shifts from local fauna); (iii) proposing the host-parasite evolutionary history scenarios to explain the diversity and co-bio-geographical pattern observed in the two host species between using N. granulosis as a model.

The SSU rDNA marker revealed a high number of microsporidian variants (*i.e.* haplogroups, 24 and 54, respectively), clustered into 18 species-level taxa, almost all being shared between the two host species. However, many microsporidian haplogroups within a given parasite species are host-specific, suggesting host-parasite co-variation. Within each host species-complex, while the confrontation between hosts and parasites phylogeography suggested some degrees of co-diversification, these patterns remain to be confirmed, mainly as SSU rDNA reached its limits in phylogenetic information content in that matter.

Strikingly, almost all of these microsporidia taxa were previously detected in other gammarids, mainly within the genus *Gammarus*, but also in other genera of Gammaridae. Some were already clearly recognised parasite taxa associated with amphipods: *Nosema granulosis*, *Dictyocoela roeselum, D. muelleri, D. roeselum, D. duebenum*, *D. berillonum, Cucumispora*

roeselum, C. ornata, C. dikerogammari, Microsporidium sp 515 and *Microsporidium* sp 505). Many times, my results increased host taxonomic spectrums and extended geographic ranges (often widely). Some other taxa were known to be extremely rare, having scarce literature records often with few or even very few geographic records and being not fully described. My PhD work either extend host taxonomic spectrum and/or deeply extend geographic ranges for these taxa. It allowed a reappraisal for such taxa, changing their status from puzzling anecdotic association to potentially overlooked established associations for amphipods.

Among the most common parasite species detected in *G. roeselii* and *G. balcanicus*, a few were clearly established in the literature as vertically transmitted and having a feminizing effect on their hosts. In *G. balcanicus,* the only species for which it could be logistically tested*,* we were unable to confirm this feature, showing that a single microsporidia species cannot be considered as totally vertically-transmistted through its entire host spectrum.

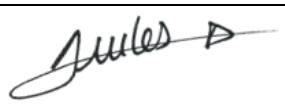
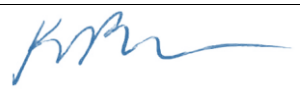
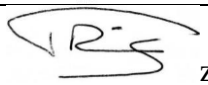
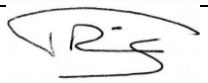
Furthermore, focusing on *N. granulosis*, the use of an additional marker, i.e. RPB1, allowed: i) to clearly identify divergent clades (while SSU rDNA was unconclusive) and ii) suggested that some peculiar strains of these supposedly feminizing parasites are not transmitted vertically and not feminizing at all. In addition, for *G.roeseli*, a within-parasite-species polymorphism could exist relative the transmission routes, one lineage being associated with vertical transmission, the other with horizontal transmission. These results are still preliminary, mainly associated with incomplete gammarid host and parasite taxa sampling, and need running further experiments to confirm our hypotheses.

10 October, 2019

AUTORSHIP STATEMENT

In agreement with all co-authors of the publication entitled: "*Microsporidian infections in the species complex* Gammarus roeselii *(Amphipoda) over its geographical range: evidence for both host–parasite co-diversification and recent host shifts."*

published by Parasites & Vectors, 12:327

I declare that my contribution in this work is: author	Contribution (and in %)	signature
Adrien Quiles	designed the experiment, performed molecular analyses, conducted the phylogenetic and statistical analyses, wrote the manuscript (35%)	
Karolina Bacela-Spychalska	designed the experiment, collected and fixed the organisms, wrote the manuscript (10%)	
Maria Teixeira	performed molecular analyses (10%)	
Nicolas Lambin	performed molecular analyses (5%)	 z up.
Michal Grabowski	collected and fixed the organisms, wrote the manuscript (10%)	
Thierry Rigaud	designed the experiment, conducted statistical analyses, wrote the manuscript (15%)	
Rémi André Wattier	designed the experiment, collected and fixed the organisms, performed molecular analyses, conducted the phylogenetic analyses, wrote the manuscript (15%)