Master's thesis

NTNU
Norwegian University of Science and Technology
NTNU University Museum
Department of Natural History

Valentina Tyukosova

Integrative taxonomy and species delimitation of water mites: the case of the *Lebertia porosa* (Acari, Parasitengona, Hydrachnidia) species complex

Master's thesis in Ecology, Behaviour, Evolution, and Biosytematics
Supervisor: Torbjørn Ekrem
Co-supervisor: Elisabeth Stur, Reinhard Gerecke
May 2021

NTNU
Norwegian University of
Science and Technology

Valentina Tyukosova

Integrative taxonomy and species delimitation of water mites: the case of the *Lebertia porosa* (Acari, Parasitengona, Hydrachnidia) species complex

Master's thesis in Ecology, Behaviour, Evolution, and Biosytematics
Supervisor: Torbjørn Ekrem
Co-supervisor: Elisabeth Stur, Reinhard Gerecke
May 2021

Norwegian University of Science and Technology
NTNU University Museum
Department of Natural History

NTNU
Norwegian University of
Science and Technology

Integrative taxonomy and species delimitation of water mites: the case of the *Lebertia porosa* (Acari, Parasitengona, Hydrachnidia) species complex

Abstract

Water mites form a highly diverse taxon of Acari unified by adaptations to aquatic life. Despite a high number of already described species, more than half as many remain unknown to science. Their ecology makes them good candidates for use as bioindicators, but the high number of undescribed species impedes this potential. *Lebertia porosa* is a water mite species described in 1900 from southern Norway. DNA barcode data sorted the specimens morphologically identified as *L. porosa* (s.l.) into ten genetic clusters (BINs), indicating the presence of a species complex. In this project, I investigated the species boundaries within the *L. porosa* complex by applying integrative taxonomy. DNA was extracted from 46 specimen collected across Norway, including the type locality. The COI mitochondrial DNA marker and 18S and 28S ribosomal nuclear DNA markers were sequenced, edited, and used to delimit species using multiple methods (GMYC, ASAP, ABGD, PTP, and BPP). The samples were also checked for the presence of the bacterium *Wolbachia* – a known manipulator of arthropod lineages. Slide-mounted specimens were studied and measured to detect morphological differences. The results indicate the presence of seven genetically and morphologically distinguishable clades. A single *Wolbachia* strain was detected among specimens from two clades, indicating no deducible pattern of interference from the bacteria. Three clades were identified from the type locality. One of these was assigned the nominal species *Lebertia porosa* based on the original species description.

Contents

List of Figures

List of Tables

Introduction

Water Mites

Hydrachnidia, formerly also known as Hydracarina or Hydrachnellae, is a monophyletic taxon of eight superfamilies collectively known as true water mites (Cook, 1974; Dabert, et al., 2010; Smit, 2020). The taxon is part of the cohort Parasitengona (Prostigmata, Trombidiformes) and is characterized by adaptations to a fully aquatic life. Water mites can be found in most types of freshwater environment including lakes, streams, springs, bogs, and interstitial waters. Around 6000 species are currently described globally and at least 4000 more have been predicted as yet to be described (Cantonati et al., 2006; Di Sabatino et al., 2008; Gerecke et al., 2018).

Water mites have a complicated lifecycle comprising of five or six stages: prelarva – which has not been observed in all species, larva, protonymph, deutonymph, tritonymph and adult. Of these, the protonymphal and tritonymphal stages are pupa-like resting phases while the adults and deutonymphs prey on the eggs and aquatic larvae of insects (Bottger, 1976; Martin & Gerecke, 2009). The water mite larvae are often parasites of adult flying aquatic insects such as Diptera (e.g., Chironomidae) or Odonata (Ilvonen et al., 2018; Martin et al., 2010). This interaction is important in aiding the dispersal of water mites, as hosts capable of flight can carry them across unconnected habitats (Bilton et al., 2001; Bohonak, 1999).

Water mites have a potential to become good bioindicators for freshwater environments (Callisto et al., 2011; Growns, 2001). While species preferring lentic environments may be too robust to respond acutely to pollution, the species inhabiting streams and springs have proven to be much more sensitive to factors such as temperature, oxygenation, and chemical composition (Di Sabatino et al., 2000; 2003). Especially the inactive pupal stages make many species apt at acting as indicators of water quality, since the conditions they demand are often much more specific than those required for the active stages (Goldschmidt, 2016). Moreover, since water mites have coevolved with other invertebrates and require their presence to complete their life cycle, they can also indicate the presence of these species (Walter & Proctor, 2013). One major obstacle stands in the way of a practical application of water mites for this purpose – the high number of undescribed species (Callisto et al., 2011; Montes-Ortiz & Elías-Gutiérrez 2020).

Lebertia porosa Thor, 1900

The water mite species *Lebertia porosa* belongs to the subgenus *Pilolebertia* within the family Lebertiidae (Lebertioidea, Hydrachnidia). This Holarctic species was first described by Sig Thor in 1900, from a small stream near the church of Vanse, Lister peninsula, southern Norway. In the most recent revision of Lebertiidae, the species is listed with 27 junior synonyms (Gerecke, 2009). Previously unpublished DNA barcode data from Norway suggests that *L. porosa*, rather than being a single species, is a complex of morphologically similar species. This would not be unprecedented as other species complexes have been previously identified within the Lebertiidae (Blattner et al., 2019; Pešić et al., 2017). The DNA barcoded specimens from Norway have been assigned 10

separate barcode identification numbers (BINs) in the Barcode of Life Data (BOLD) Systems database. Based on this data they have been pre-emptively assigned into seven tentative clades which had the potential to represent species.

Species concepts and species delimitation

While a species can be considered one of the most fundamental concepts in biology, the question of how it should ultimately be defined is still debated (e.g., Fišer et al., 2018; Freudenstein et al., 2017; Hong, 2020). The primary species concept, that of a metapopulation evolving along a single lineage, is universally accepted, but cannot be used for species delimitation. To do this, secondary, or operational, species concepts are necessary (De Queiroz, 2007). These concepts rely on different criteria, such as heritable morphological characteristics, breeding compatibility, or genetic differences. These different criteria can sometimes result in the delimitation of different number of species within the same group of organisms. This can be resolved by applying integrative taxonomy, i.e., considering all the data as additional evidence corroborating delimitation rather than as standalone alternatives (De Queiroz, 1998; 2007).

Morphological analysis has a long history of use in biology (Simpson, 1951). This method does not require much beyond a microscope, literature, and practice, but is costly in terms of time and, depending on the group studied, may require a high level of expertise. Moreover, it cannot be applied to species with no observable morphological differences. Preservation poses an obstacle, as characteristics such as scent, colour, and shapes of soft tissues might degrade very fast. Many species, including *L. porosa* were originally described based on this method alone. Modern technological advances in delimitation and imaging methods allow us to search for previously overlooked morphological characteristics when revising the species divisions.

Genetic delimitation has vastly increased in popularity with the increasing ease of access to molecular data through DNA extraction and sequencing. It requires more expensive laboratory equipment but is less time consuming and, depending on the studied organisms, potentially simpler and less prone to human error than the morphological analysis. Moreover, by analysing genetic data of a population it is possible to reveal a diversity of morphologically cryptic species. Various mathematical methods have been developed in order to delimit species on the basis of molecular data. Each model makes a different set of assumptions and their accuracy changes depending on the studied system. Parameters such as speciation rate, divergence time, population size, interspecific distance, amount of geneflow, etc., are not uniform among organism groups but can affect the results obtained by various methods. The conclusions about the best model fits are not transferrable across organism groups with different values for these parameters. If some of the parameter values are unknown or uncertain, multiple models should be used to verify the delimitation results (Carstens et al., 2013).

The errors resulting from the model assumptions and assumption violations can manifest themselves in two ways: overestimation or underestimation of the number of operational taxonomic units (OTUs). The General Mixed Yule Coalescent (GMYC) method is prone to

overestimation due to its inability to cope with high, but still biologically plausible, speciation rates (Dellicour & Flot, 2015; Esselstyn et al., 2012; Talavera et al., 2013). Automatic Barcode Gap Discovery (ABGD) and Assemble Species by Automatic Partitioning (ASAP) rely on genetic distances, which makes them much less expensive in terms of computing power than methods that include phylogenetic reconstruction (Puillandre et al., 2021). ABGD and ASAP function on the assumption that intraspecific distances are generally smaller than interspecific distances (Puillandre et al., 2012; 2021). Both methods have some tendency to underestimate the number of OTUs (Pentinsaari & Vos, 2017; Puillandre et al., 2021). Bayesian Phylogenetics and Phylogeography (BPP) and Poisson tree processes (PTP) assess the posterior probability of species being delimited. The many Markov Chain Monte Carlo (MCMC) iterations run by these models require time and computing power, but despite a small tendency for underestimation in PTP, the results tend to be very robust and close to the 'true' number of species (Camargo et al., 2012; Pentinsaari & Vos, 2017). However, these methods do not fully account for migration and the results become less accurate in the presence of gene flow between the OTUs, PTP more so than BPP (Luo et al., 2018; Zhang et al., 2011).

Wolbachia

Wolbachia is a genus of congenital bacteria infecting many arthropod species. Although no published record of *Wolbachia* infections within Hydrachnidia seems to exist, they have been recorded within other Acari groups, such as spider mites (Tetranychidae) or predatory mites (Phytoseiidae) (Breeuwer & Jacobs, 1996). *Wolbachia* infections are also common among the aquatic insect species serving as hosts and prey to water mites (Sazama et al., 2017). This could be a potential source of infection as other mite species have been implicated in horizontal transfer of *Wolbachia* (Brown & Lloyd, 2015; Cordaux et al., 2001).

Wolbachia reproduce in a manner similar to cytoplasmic elements and are usually spread vertically by infected females onto their offspring. The bacteria maximize their reproductive success by employing multiple strategies such as induced parthenogenesis, male killing, male feminization, and induced cytoplasmic incompatibility between individuals not infected by the same strain (Werren et al., 1995; 2008). These strategies may lead to skewed sex ratios or speciation due to forced assortative mating caused by postzygotic reproductive barriers. The presence of multiple strains of *Wolbachia* in a single arthropod species can also result in a pattern of multiple divergent mitochondrial lineages that are discordant with evolutionary lineages in nuclear markers (Jiang et al., 2018; Sucháčková Bartoňová et al., 2021; Whitworth et al., 2007). Since the initial observation of multiple genetic clusters in *L. porosa* is based on the mitochondrial Cytochrome C oxidase subunit I (COI) genetic marker used for DNA barcoding in animals, it is important to check for the presence and pattern of *Wolbachia* strains infecting the water mites.

Aims and Hypothesis

In this study I aim to investigate the species boundaries within the *L. porosa* species complex by analysing both morphological and molecular characters and applying various delimitation methods to mitochondrial and nuclear DNA markers. The main hypothesis of this study is that the genetic

clusters observed in the COI DNA barcode data of *L. porosa* (s.l.) represent multiple morphologically distinct species which are also recognizable as separate genetic lineages by nuclear markers and are unaffected by potential infections of *Wolbachia*. Alternatively, the DNA barcode clusters may be supported by the nuclear markers but indistinguishable morphologically. In such a case the species could still be delimited through molecular methods but would be considered cryptic. It is also possible that the clustering in the DNA barcode data will not be supported by the nuclear markers and will instead follow a pattern consistent with detected *Wolbachia* strains.

Methods

Field collection

Water mites were collected from lentic and lotic environments during the summers (July-August) of 2014-2020 across Norway (Figure 1), and near Lake Baikal, Russia. The presumable type locality of the species, a stream near the Church of Vanse at Lister Peninsula, Norway, was included among the sampling sites. Full overview of localities and collection dates can be found in Supplement 1. Specimens were collected by kick sampling and drift nets. Captured specimens were preserved in 96% ethanol to prevent the deterioration of DNA.

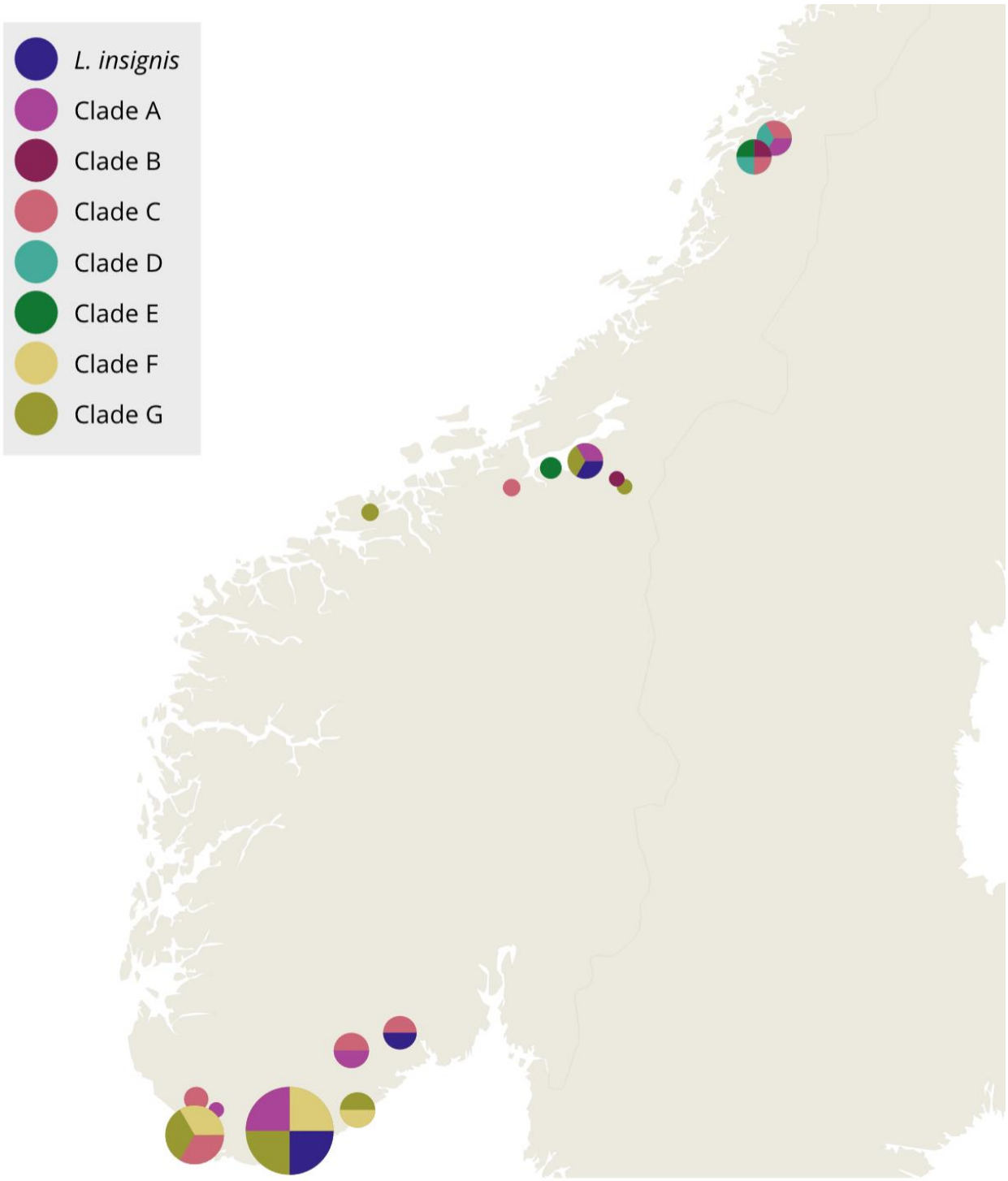

Figure 1: Collection sites for Norwegian specimens used in this study.

Molecular Methods

The COI DNA barcode fragment from previously barcoded specimens was obtained through collaboration with the Norwegian Barcode of Life (NorBOL). This data was based on the DNA-extraction of leg tissues, PCR, and bi-directional sequencing following standard protocols at the Canadian Centre for DNA Barcoding at the University of Guelph, Canada.

For samples analysed at NTNU, DNA was extracted non-destructively from the specimens using Qiagen DNeasy Blood and Tissue kit following the instructions for animal tissue samples. Specimens were removed from ethanol and dried on a piece of filter paper before lysis. Incubation in lysing buffer and proteinase-K was done overnight for approximately 14 hours. In the last step only 100 µl of buffer AE was used for elution. Last step was not repeated. Empty lysed skins were washed in water and 70% ethanol and resuspended in 90% ethanol for later morphological analysis.

PCR amplification was performed for COI mitochondrial DNA and 18S and 28S ribosomal nuclear DNA of water mites (Table 1). These markers are commonly used to delimit species or build phylogenies within Hydrachnidia (e.g., Blattner et al., 2019; Dabert, et al., 2010). For water mite DNA PCR, reactions in volume of 20 µl were prepared, consisting of 10 µl of PCR Master Mix 2X (Qiagen), 0.4 µl of 10 µM primer, and 4 µl of the extracted DNA, with the remaining volume topped up with HyClone Molecular Biology-Grade Water. The PCR cycling conditions for all reactions were: 1 cycle (5 min at 96°C), 35 cycles (30 sec at 95°C, 1 min at 50°C, 1 min at 72°C) 1 cycle (5 min at 72°C) (Dabert et al., 2016; Sazama et al., 2017). Amplification was verified by electrophoresis on a 1.5% agarose gel. Multiple primers for various *Wolbachia* markers were tested but all except for 16S DNA (Table 1) failed to amplify or did not produce readable sequences. For full list of tested *Wolbachia* primers see Supplement 2. The PCR protocol for *Wolbachia* 16S was the same as for water mite DNA except 10 µl of Multiplex PCR Master Mix 2X (Qiagen) was used instead of 10 µl of PCR Master Mix 2X (Qiagen) and the annealing temperature was lowered to 45°C.

Before sequencing, amplified samples were chemically purified with 4 µl of ExoProStar (Cytiva) to remove excess reagents and unwanted residues in the PCR products. The mixture was incubated in a thermocycler at 37°C for 15 min and 80°C for 15 min. DNA concentration and purity was checked using NanoDrop 2000 spectrophotometer (Thermo scientific). After this, 2 µl of 10 µM primers were added to the samples before shipping to Eurofins Genomics using the PlateSeq Kit Mix and PlateSeq Kit DNA. The sequences were stored in the BOLD database.

Table 1: Primers used for amplification of DNA fragments.

Primer	Direction	Marker	Sequence (5'-3')	Original publication
Leb_F	F	COI	CAA ACC AYA AAG AYA TTG GAA C	(Blattner et al., 2019)
Leb_R	R		CGA AGA ATC AAA ATA RRT GTT G	
28SHy_F	F	28S	AGT ACC GTG AGG GAA AGT TG	(Blattner et al., 2019)
28SHy_R	R		GGC AGG TGA GTT GTT ACA CA	
18Sfw	F	18S	CTT GTC TCA AAG ATT AAG CCA TGC A	(Dabert, et al., 2010)
rev960	R		GAC GGT CCA AGA ATT TCA C	
Wspecf	F	16S	AGC TTC GAG TGA AAC CAA TTC	(Werren & Windsor, 2000)
Wspecr	R		GAA GAT AAT GAC GGT ACT CAC	

Editing, Alignment and phylogenetic analysis

Forward and reverse sequences were examined, aligned, and edited according to their chromatograms using the MEGA X (Version 10.1.7) software (Kumar et al., 2018). In some samples for the 28S rDNA marker, the quality of the forward reading sequence was too low and only the reverse sequence was used for further analysis. The identity of the water mite fragments was verified using the Megablast algorithm in GenBank. *Wolbachia* 16S fragments were classified through the Silva database. The COI sequences for all specimens were aligned using ClustalW algorithm with gap opening penalty of 15 and gap extension penalty of 6.66. The 18S and 28S sequences were aligned with the MUSCLE algorithm. Gap penalties used were -400 for opening and 0 for extension for 18S and -400 for opening and -200 for extension in 28S. The reliability of the alignments was assessed using the GUIDANCE2 algorithm (Sela et al., 2015). The best substitution model fit was assessed based on the BIC parameter using the 'Find Best DNA/Protein Models (ML)' function in MEGA X and maximum likelihood trees were constructed for each marker with complete gap deletion and 500 bootstrap replicates (Hall, 2013). The substitution models used were Hasegawa-Kishino-Yano with gamma distribution and invariant sites for COI and Kimura-2 parameter for 18S and 28S.

Delimitation and Network analysis

Multiple molecular delimitation methods were applied to the dataset. A single threshold GMYC was performed in R Studio (4.0.2) using the Splits package (Ezart et al., 2013; R Core Team, 2020). Ultrametric ML trees for the analysis were prepared in MEGA X using the 'Compute timetree' function. Single marker PTP was performed at https://species.h-its.org/ptp/ using 100000 MCMC generations and burn-in of 0.1 (Zhang et al., 2013). ABGD and ASAP were performed at https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html and https://bioinfo.mnhn.fr/abi/public/asap/asapweb.html respectively, using the Kimura-2-Parameter option (Puillandre et al., 2012; 2021). The command line version of BPP (A10) using all three markers was run for 500000 MCMC generations with burn-in of 0.1 using the COI tree as a guide. The inverse gamma priors for the θ and τ parameters were $\alpha = 3$, $\beta = \{0.001, 0.002, 0.01, 0.02\}$. Due to the lack of data regarding θ and τ parameters for Hydrachnidia, various combinations of the β values were tested

for both θ and τ (Flouri et al., 2018; Yang, 2015). A TCS haplotype network of the COI sequence data was constructed in PopART (Clement et al., 2000, 2002; Leigh & Bryant, 2015).

Morphological Analysis

Specimens were initially sorted under a stereo microscope and identified to species level by Reinhard Gerecke. Some were subsequently slide-mounted in Hoyer's fluid or Glycerine Jelly. The selected specimens were dissected by leading a horizontal cut around the coxal shield to separate it from the dorsal integument. Legs were detached from the coxal plates and positioned to allow for lateral view. Both palps were detached from the gnathosomal bay and included on the slide. Male specimens were slide-mounted with a separated genital skeleton. As many specimens as possible were left undissected for better long-term preservation.

Slide-mounted specimens from each genetically determined clade were examined using Leica compound microscope (LEICA DM 6000 B) and measurements were taken using the Leica Application Suite. The number of specimens examined depended on the number of available barcoded specimens in each clade (Table 2). Remaining specimens kept by Reinhard Gerecke at the University of Tubingen were used to confirm the observed morphological patterns. The revision of Lebertiidae by Gerecke (2009) was used as a source for the anatomical terminology.

Table 2: The total number of DNA barcoded specimens per clade and the number of specimens used in the morphological analysis.

	Clade A	**Clade B**	**Clade C**	**Clade D**	**Clade E**	**Clade F**	**Clade G**
Barcoded specimens	21	3	27	7	9	11	38
Examined specimens	12	3	13	6	4	5	14

The following features were measured: the dorsal length and maximum height of segments 2-5 of legs and palps, total length and maximum height for segment six of legs (Figure 2-i), the length of each acetabulum, the distance between the inner margins of leg four insertions and the length and width of margin one and two on the coxal shield (Figure 2-ii), the length of claws on the final segments of legs from hinge to tip, from tip to the centre of the curvature, and from the centre of the curvature to the hinge (Figure 2-iii), The numbers and positions of salient setae were noted with emphasis on the number of swimming setae, the distances between the long setae on segment 3 of the palp (Figure 2-iv), and the length and width of the distomedial peg-like seta of segment 4 of the palp. Additionally, the width and length of the gnathosomal bay was measured by creating a line between the openings of the glands at each side and then drawing a perpendicular line from the central fusion seam of the coxal shield (Figure 2-ii).

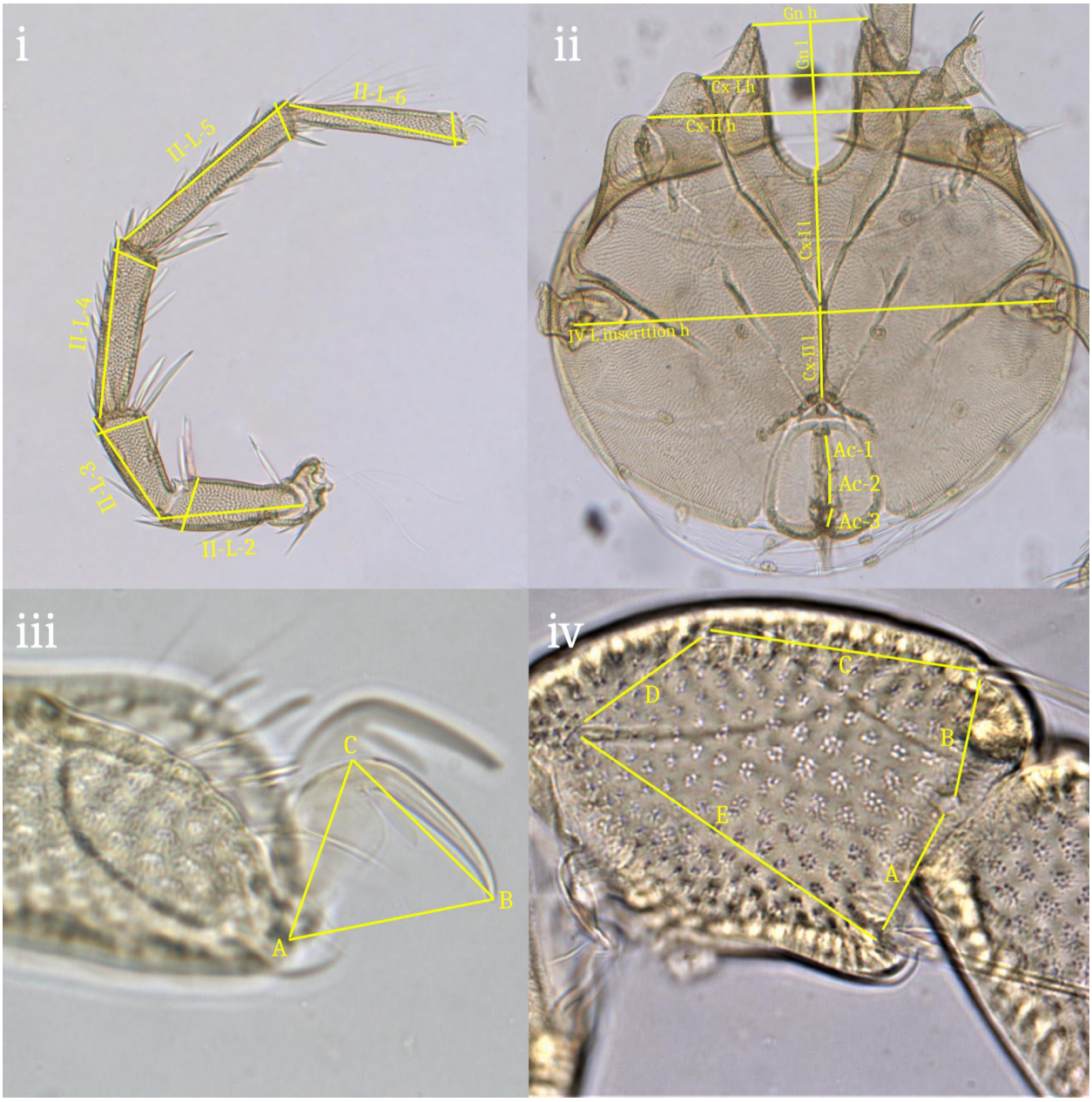

Figure 2: Positions and shorthand labelling for morphological measurements. The anatomical labelling follows Gerecke (2009). Small letters l and h stand for length and height. i) An example of segment length and height measurements for legs and palps. L indicates leg, palps are labelled P. The roman numerals refer to the order of legs starting from the front of the body. The final number stands for the order of segments starting from the most proximal. ii) Measurements across the coxal shield. Coxal margins are marked Cx, Gn stands for gnathosomal bay. Acetabula are marked Ac. iii) The dimensions of the claw. The centre of curvature (C) was considered to be the point opposite of the smaller claw. iv) The distances among the long setae on the third segment of the palp.

Results

Sequences of sufficient quality to be used for all three water mite DNA fragments were obtained from 47 specimens which included 41 specimens of various *L. porosa* clusters, 5 specimens of *L. insignis* – a putatively closely related species, and one representative of an undetermined species of the subgenus *Lebertia* (s. str.) collected near Lake Baikal as a more distant outgroup. In the final alignments the COI sequences had total lengths of 638 base pairs (bp) with 225 variable sites and 210 parsimony informative sites. The 18S sequences were 776 bp long with only 53 variable sites and 42 parsimony informative sites. The 28S sequences were the longest with 950 bp but less variable than the COI sequences, with only 220 variable sites and 177 parsimony informative sites. All sequences were lightly AT biased with the highest average ratio in COI (64.9%). The highest AT bias in the COI marker was observed on the third codon position (82%).

Trees

The results of phylogenetic analysis show that the sequences from individual specimens are always assigned to the same clade or clade fusion. The clades A, B, and C are grouped together across all three trees but are only supported as sister groups on a higher node in one phylogeny. Clades D and E are also grouped together and are supported as sister groups across all three phylogenies. *Lebertia insignis* is always among the inner brackets and never the closest to the chosen distant outgroup and this position is supported in two of the phylogenies.

In the tree constructed from the COI sequences (Figure 3), all seven clades of *L. porosa* are clearly separated by large genetic distances. The genetic distances between the sequences within each clade are very short except for clade C, which is separated into three well supported subgroups. The individual clades have high bootstrap support values on the nodes that unify all their sequences. Of the higher nodes, only the one grouping, clades D and E, has bootstrap value larger than 75 and many others have support value lower than 50.

The tree constructed from the 18S sequences (Figure 4) presents lower genetic distances between individual clades and within clades, reflecting the lowest ratio of variable sites. Clades D and E are fused with high bootstrap support on their joined node. Clades A and B are also fused and form a wedge separating clade C into two subgroups. This separation is not supported as the node bootstrap value is less than 50. One higher node has received a high bootstrap value within this tree. This node separates the inner branches into two large groups: clades A-E and clade G and *L. insignis*.

The 28S phylogeny (Figure 5) also presents all seven *L. porosa* clades. Except for clade E, all clades with multiple specimens are supported by more than 75% of the bootstrap replicates. Clade C is divided into two supported subgroups but the genetic distances among the sequences in each clade are otherwise very short. Three higher nodes received support value higher than 75 – one grouping clades D and E, one grouping clades A, B, and C, and one separating the inner branches into groups of clade A, B, C and G and clade D, E and *L. insignis*. The relative genetic distance between clades D and E is much shorter for this marker than for the COI phylogeny.

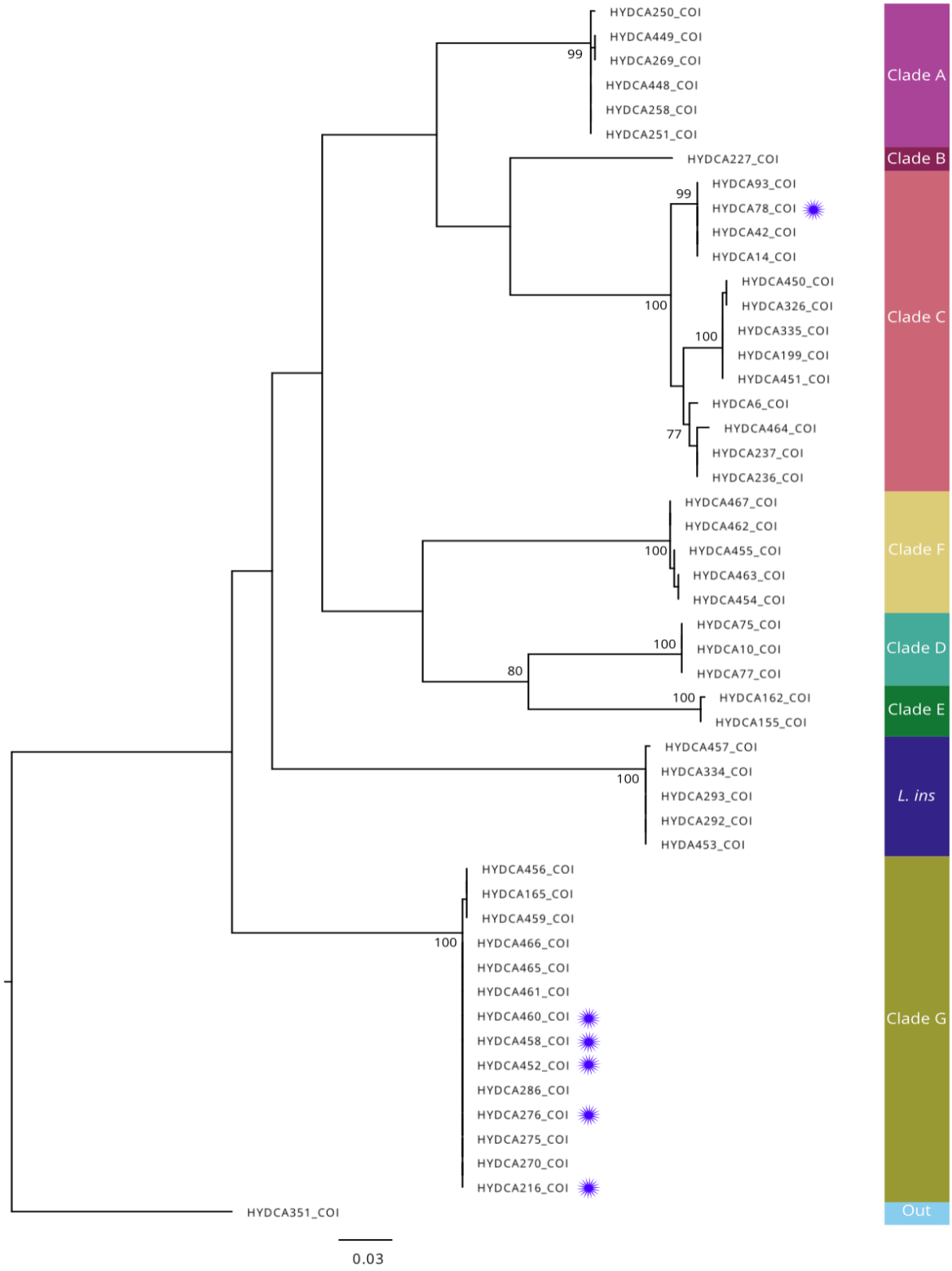

Figure 3: COI rooted maximum likelihood tree including *Lebertia insignis* and an outgroup, generated with complete gap deletion and 500 bootstrap replicates. Substitution model: Hasegawa-Kishino-Yano with Gamma distribution and invariant sites. Bootstrap values above 75 are shown. Scale bar represents genetic distance. Stars mark specimens infected with *Wolbachia*.

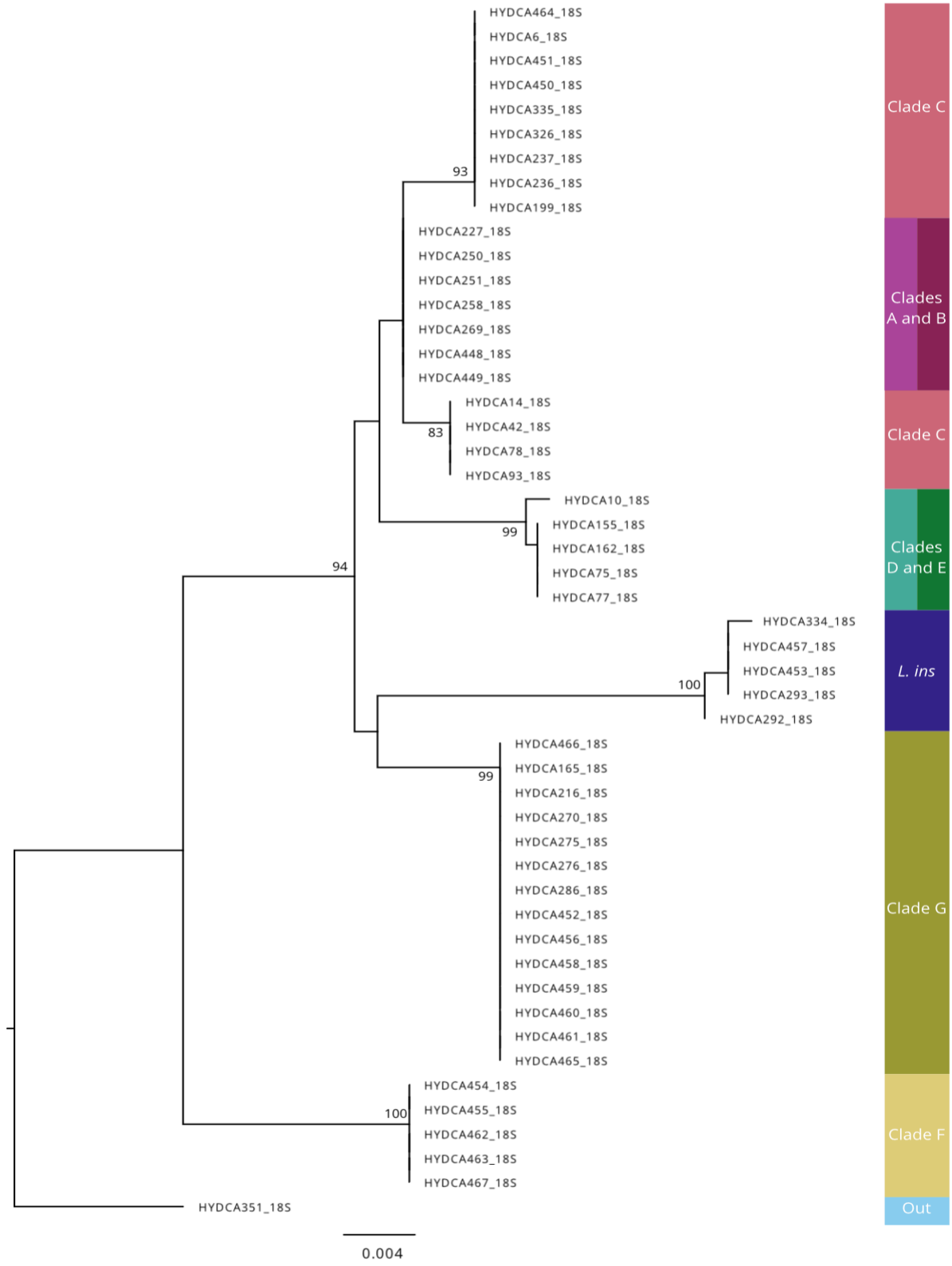

Figure 4: 18S rooted maximum likelihood tree including *Lebertia insignis* and an outgroup, generated with complete gap deletion and 500 bootstrap replicates. Substitution model: Kimura-2 parameter with Gamma distribution. Bootstrap values above 75 are shown. Scale bar represents genetic distance.

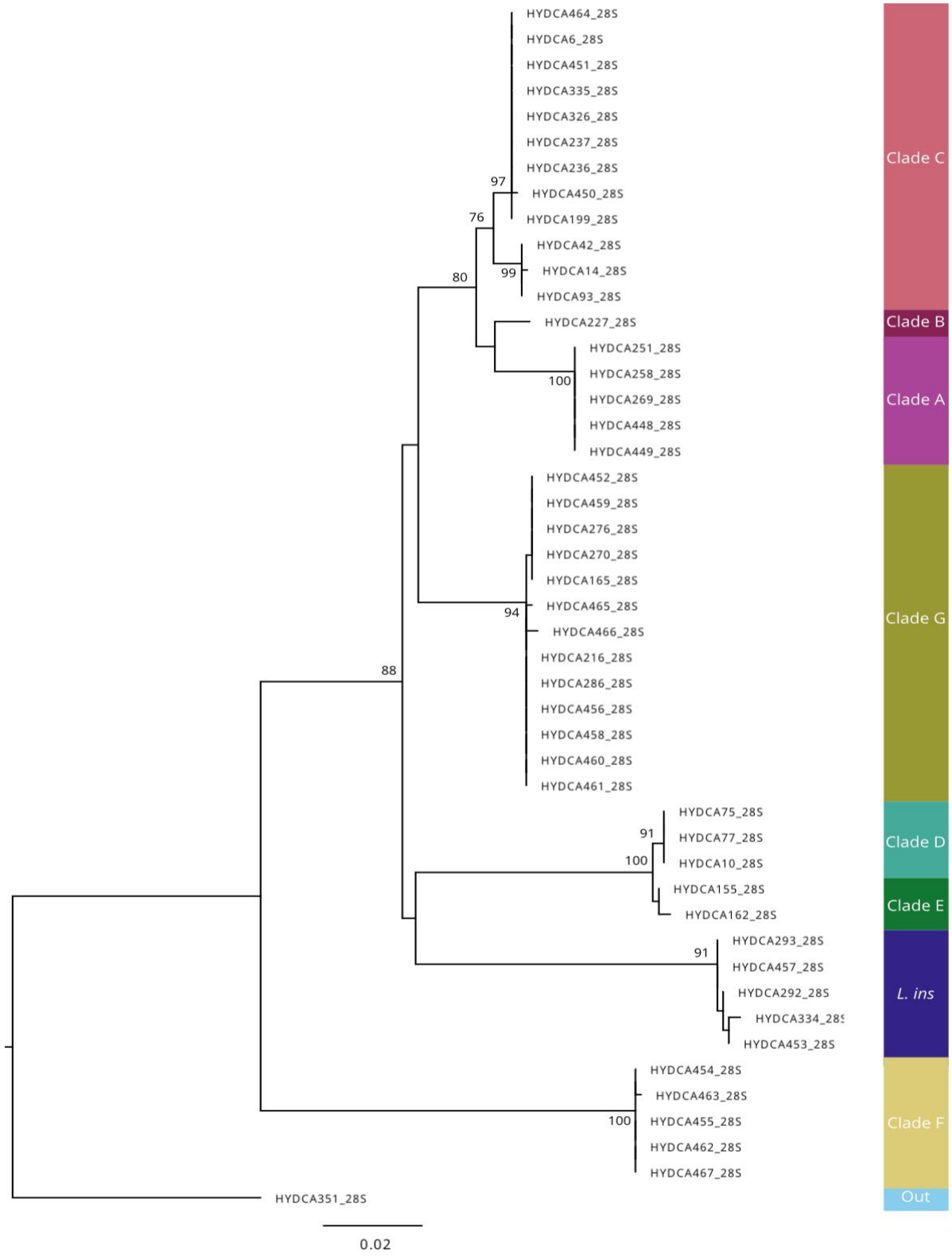

Figure 5: 28S rooted maximum likelihood tree including *Lebertia insignis* and an outgroup, generated with complete gap deletion and 500 bootstrap replicates. Substitution model: Kimura-2 parameter with Gamma distribution. Bootstrap values above 75 are shown. Scale bar represents genetic distance.

Delimitation

Lebertia porosa is divided into multiple OTUs by all single marker delimitation methods. (Table 3). For the 28S marker only GMYC delimits clades D and E as separate OTUs. None of the single marker methods delimit these clades as separate for 18S. GMYC, PTP, ASAP, and ABGD all delimit clade C as two OTUs for at least one marker. For the 18S marker, the sequences among clades A and B do not show any difference and these clades cannot be delimited by any single marker method.

Table 3: Single marker delimitation methods and the number of OTUs delimited by each. Both outgroups are excluded. All shown delimited OTUs belong to *L. porosa* clusters.

	GMYC	**PTP**	**ASAP**	**ABGD**
COI	19	7	7	7
18S	7	6	6	7
28S	19	6	7	7

BPP always delimits seven OTUs among the *L. porosa* clusters, with posterior probability for the entire model ranging from 0.74 to 0.95. The lower posterior probability in the model with more conservative values of θ and τ is mainly the result of lower posterior probability of the support for the node between clades D and E.

Network Analysis

The network of COI sequences (Figure 6) shows nine groups – seven *L. porosa* clades and two outgroups. The outgroups split the network into two halves, one of which contains the clades B and C, and the other where the rest of the clades are located. Aside from the clade C, the network has low reticulation with high number of mutations between nodes.

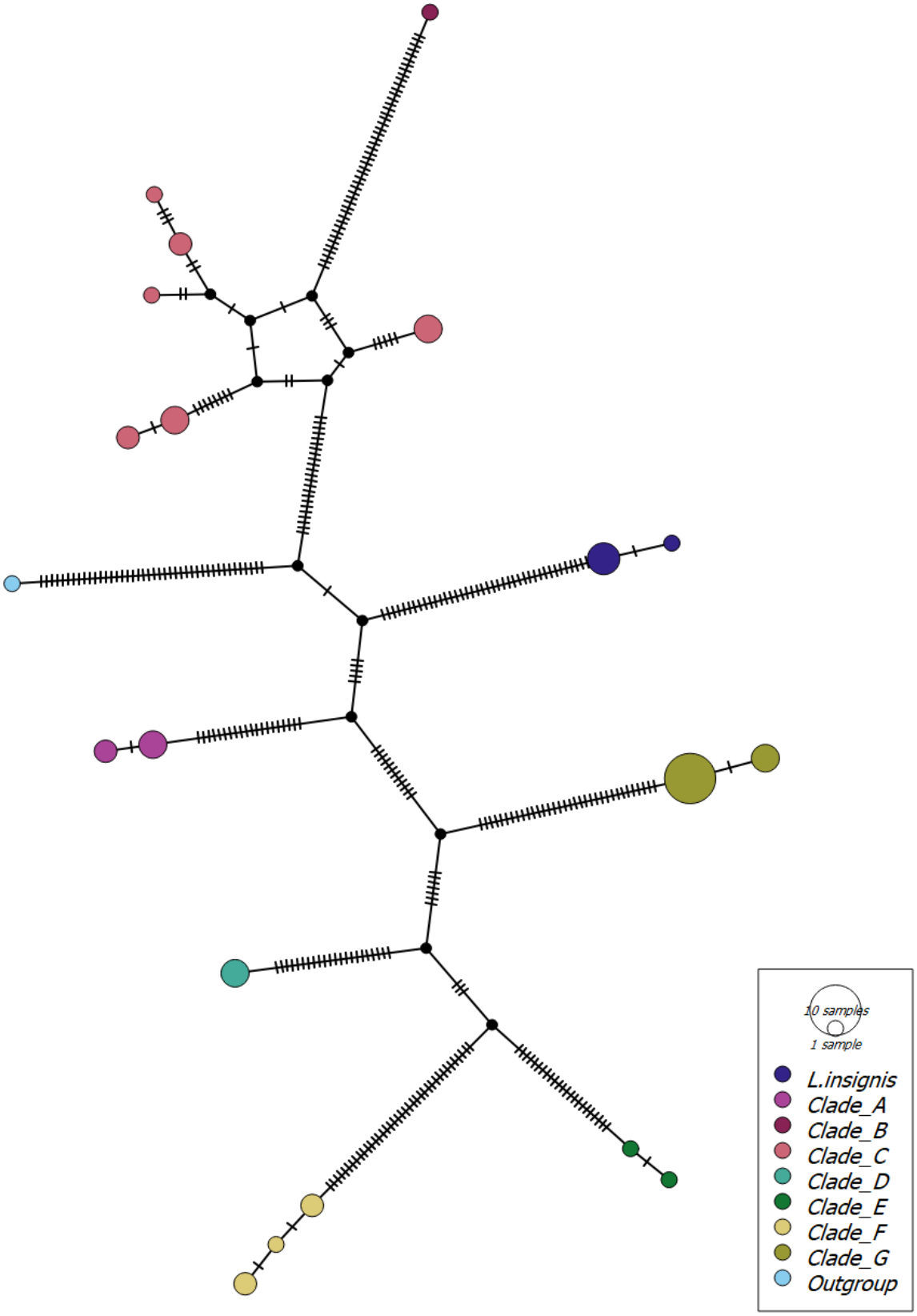

Figure 6: TCS haplotype network for COI sequences.

Wolbachia

All specimens were tested for the presence of *Wolbachia* using the 16S marker and *Wolbachia* specific primers and 17 showed positive results upon amplification. Three different fragment lengths were successfully sequenced: short (270 bp), medium (490 bp) and long (900 bp) (Supplement 3). Only the six long fragments were positively classified as *Wolbachia* in the Silva database. The highest identity match in Silva database was 95% and the 10 nearest neighbours included *Wolbachia* extracts from a quill mite (*Torotrogla cardueli* Bochkov & Mironov, 1999), a spider mite (*Bryobia sp*.), a nematode parasitizing ticks (*Cercopithifilaria japonica* Uni, 1983), giant yellow aphid (*Tuberolachnus salignus* Gmelin, 1790), and a springtail (*Neelus murinus* Folsom, 1896). The long sequences are highly similar, presenting only 3 variable sites and no parsimony informative sites. Of the six identifiable sequences, five originated from the specimens of clade G and one from clade C (Figure 3).

Morphology

Morphological analysis revealed that each preliminary clade possesses a unique combination of characteristics (Table 4, Figure 7). The diagnostic features of each clade (proportions of palp and leg segments, swimming setae numbers) concern character states which are different from those used for the current morphological identification of *L. porosa* (Gerecke, 2009). The measurements of claw dimensions and of the peg-like seta on palp segment three were not used due to low reliability.

Table 4: A presence/absence matrix of novel morphological characteristics observed in *L. porosa* clades. Characteristics marked with +/- were not present in all examined specimens. L indicates leg, palps are labelled P. The roman numerals refer to the order of legs starting from the front of the body. The final number stands for the order of segments starting from the most proximal. Acetabula are marked Ac. The exact position for measurement of P-3 distance D can be found in Figure 2vi.

Character\Clade	A	B	C	D	E	F	G
II-L-5 large gap in swimming setae	+/-	+	+				
III-L-5 >14 swimming setae		+		+			
IV-L-5 >14 swimming setae		+		+			
P-3 doubled most proximal seta				+/-			
IV-L-4 ≥9 swimming setae				+	+		
II-L-5 swimming setae stop before distal edge						+	
III-L-6 Stout (l/h<5)	+					+	+
IV-L-6 Stout (l/h<5)	+						+
P-3 distance D long (≥38µm)	+						
I-L slender (I-L-(4-6) l/h>3.2,4.2,4.2)		+					
IV-L slender (I-L-(5-6) l/h>6.0,4.2)			+				
Coxal field long (>1000 µm)				+			
Palp long and slender (P-3 l>140, P-4 l>195 µm, P-3 l/h>1.85)				+			
Legs long and slender (I-L-(4-6) l>450, 490, 430 µm, l/h>5…)				+			
P-2 relatively short (P-2/P-3 < 1.2)					+		
Genital field with large acetabula (Ac-1 l=85-95 µm)					+		
Leg segments short (I-L-(4-6) l=130-145, 138-150, 125-130 µm…)						+	
Palp short and slender (P-2 l<110, P-4 l<140 µm, P-3 l/h>2.2)						+	

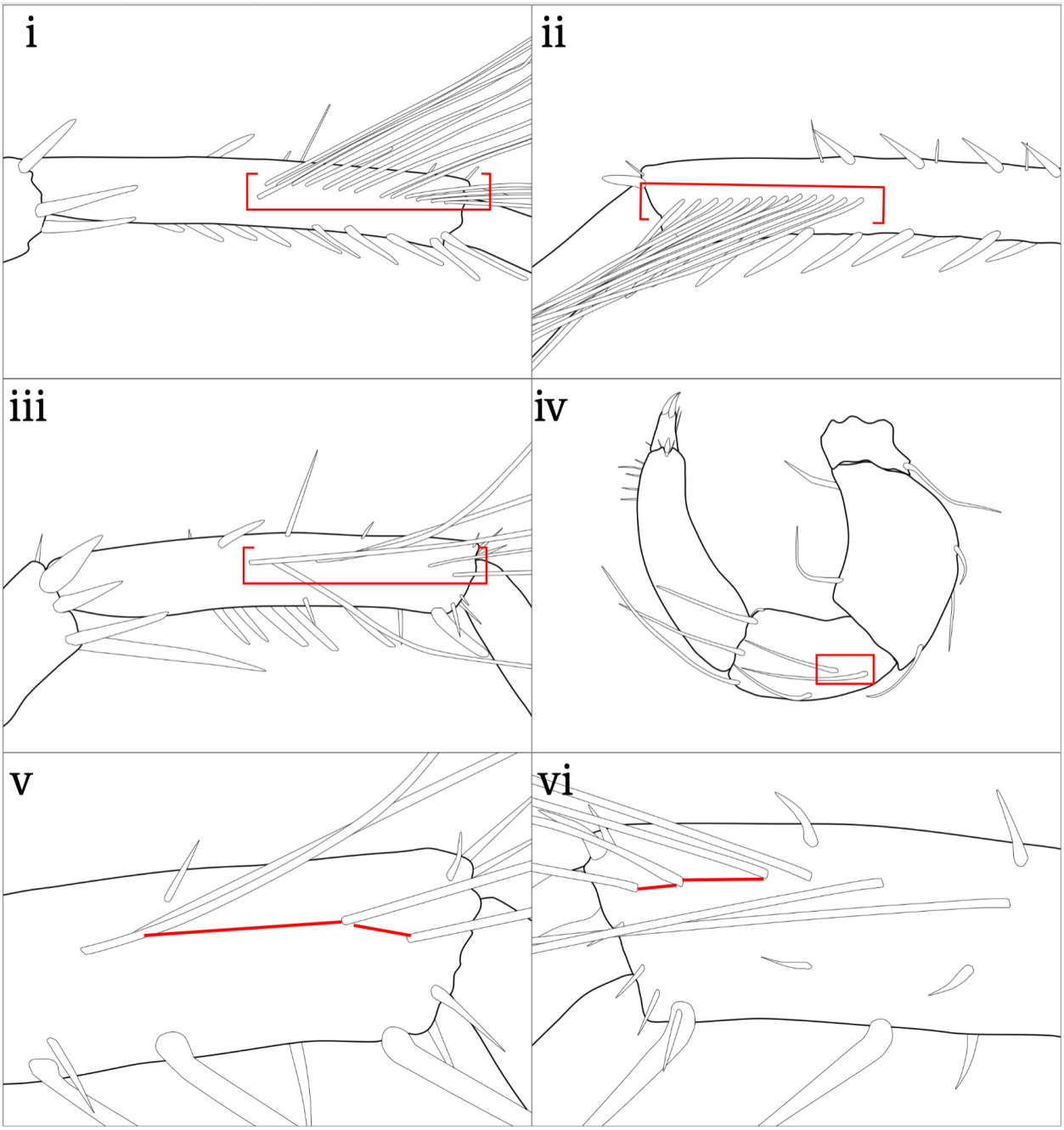

Figure 7: Examples of observed differences in setation. i) and ii) Examples of high number of setae on segment five of legs three and four (III-L-5, IV-L-5) in specimens from clades B and D. iii) The gap in swimming setae on the fifth segment of the second leg (II-L-5) in clades A, B, and C. iv) Segment three of palp (P-3) with a doubled proximal long seta. v) and vi) Segment five of the second leg (II-L-5) close-up; comparison of swimming setae with and without the large gap respectively.

Observations on habitat

From the specimens used in this thesis, the collection localities are often limited to a single clade, with some exceptions being shared across two or three clades. Clade C is reported from both lentic

and lotic environments. Some clades show clear distinction between the preferred habitats: lentic for clades B and D and lotic for clades A and G. Clades E and F are more often found in lotic habitats, but one specimen from each was collected in a lentic habitat.

Discussion

Species delimitation

Seven potential species level clades were hypothesized within the *L. porosa* complex based on BINs of the original COI data. This number is supported by the results of morphological and molecular analyses and is congruent with the main hypothesis: all three markers indicate similar divisions and differences in morphology are observable. This result adds this study to the ever-growing global list of literature resolving species complexes within Hydrachnidia and describing new species of water mites based on molecular methods and integrative taxonomy, (e.g., Blattner et al., 2019; Montes-Ortiz & Elías-Gutiérrez, 2020; Pešić et al., 2020; Pešić, & Smit, 2016; Stålstedt et al., 2013; Więcek et al., 2020) and provides further proof that the clade Hydrachnidia hides a great amount of undescribed diversity (Di Sabatino et al., 2008).

The single marker models do not always show the expected number of OTUs, but the differences occur in a manner consistent with the method's error patterns and with the structure of the phylogenies for individual markers, e.g., no single marker method could separate clades A and B in the 18S phylogeny, because there is no genetic difference in their sequences for this marker (Camargo et al., 2012; Dellicour & Flot, 2015; Pentinsaari & Vos, 2017; Puillandre et al., 2021).

GMYC obviously overestimates the number of OTUs in markers with higher ratio of variable sites (Table 3). It comes within a reasonable range only in the 18S marker. 18S has by far the lowest ratio of variation among sequences which probably puts it within the range of speciation rate values that GMYC can cope with (Dellicour & Flot, 2015; Esselstyn et al., 2012; Talavera et al., 2013). The COI marker presents a clear barcode gap in genetic distances (Supplement 4.1) and ABGD and ASAP delimit this marker according to the expected number of OTUs (Puillandre et al., 2021). They fare worse with the nuclear markers, where the genetic distances are shorter overall and no large difference between the intraspecific and interspecific distances is apparent (Supplement 4.2, 4.3). The numbers of OTUs are similar to the expected value, but in the 18S and 28S marker that is a result of these models both overestimating and underestimating the number of OTUs at the same time. This is a consequence of the large intraspecific distance among the clade C sequences and the small interspecific distance between clades D and E (Puillandre et al., 2012; 2021). PTP delimits the COI marker in a manner consistent with my initial assumption. The OTUs for the nuclear markers reflect the pattern of their individual phylogenies in a manner similar to the ASAP and ABGD results – underestimating in clades D and E and overestimating in clade C.

Studies with larger datasets, such as Pentinsaari & Vos (2017), might be forced to rely on single marker methods only and use further statistical analysis to resovle the discrepancies between the different results from each model, but the relatively small amount of data in my study allows for

the use of more sophisticated multilocus methods. BPP uses all markers and thus reduces the influence of possible incomplete lineage sorting in individual marker genealogies (Jacobs et al., 2018; Yang & Rannala, 2010). This model always delimits all seven clades of *L. porosa*. The posterior probability decreases with higher θ values. θ represents the average genetic difference between the members of the same population and runs using higher values tend to result in a more conservative assessment of the number of OTUs (Flouri et al., 2018; McKay et al., 2013).

Geography and Habitats

While the number of species in the *L. porosa* complex seems well supported at local geographical scale, globally the result is only preliminary. Specimens used in this study are all from Norway and do not represent the full Holarctic distribution of *L. porosa* (s.l.). Records in the BOLD and GBIF databases indicate that the species is also present in many localities across the entire Canada and the Western Palearctic. However, not all of these records included barcode data. The public barcode data also does not represent all of the delimited clades, as some BINs only appear within the dataset that served as a starting point for this study.

I did not analyse the geographic distribution deeply within this study, but, after sorting through the collection data in BOLD, clades B, D, and E only seem to occur in latitudes higher than 63°N and clade F has not been observed above 59°N latitude. The collection sites are not evenly distributed along the latitudinal gradient (Figure 1), so the true geographic limits of distribution of these clades cannot be determined from the current dataset. The clades B, D, and E have fewer barcoded specimens than some others (Table 2), possibly due to not occurring in the more southern latitudes which were the main target for sampling. The wider dataset also suggests habitat preference for either the lentic or the lotic environments across all clades except C, but this could merely reflect the low amount of available data. The two specimens, one in clade E and one in clade F, were collected at locations which present a habitat type different than the other members of their clades, which could indicate that clade C is not alone in inhabiting both habitats. To fully explore the possibility of using geographic distribution and habitat preference as an aid to identification, future sampling efforts should put an emphasis on noting down the exact habitat type present at the site of capture and including more sampling sites at intermediate latitudes.

Both the geographical and habitat patterns imply the possible presence of other *L. porosa* (s.l.) clades which might be limited to other latitudes or prefer habitats which are not common in Norway. It is not uncommon for a high number of new OTUs to emerge after deliberately sampling across a wide geographic range (Darwell & Cook, 2017; Stuart et al., 2006; Wattier et al.,2020). However, the public barcode data in BOLD also indicates that at least clade C truly has a very broad distribution. Publicly available data for one of the BINs in clade C (BOLD:AAM5138) shows multiple specimens collected across Canada. There is a possibility that such a wide distribution is a result of this clade being able to disperse more efficiently, perhaps with the aid of flying aquatic insects. Clade G provides an interesting contrast to clade C in this case. Both of these clades are distributed across the entire latitudinal range sampled in the dataset. Clade G has the highest number of sampled specimens of all the clades in this study, with clade C having the

second highest number. Despite these similarities Clade G presents no reticulation in the haplotype network and very low interspecific distance, reflected in its single BIN in comparison with the four BINs in Clade C.

Morphology

Despite being previously described as a single species based on morphology, the *L. porosa* complex presents a wide range of non-overlapping characteristics (Table 4). The characters used to describe the original species – equal distance of ventral setae on palp segment four and the equal distance among the three distal long setae on palp segment three – are quite variable within the clades (Gerecke, 2009). The new diagnostic features do not include these two characteristics. They instead rely on measurements of palps, legs and the coxal shield. Some of the originally taken measurements could not be used. For the claw dimensions it was mainly due to the mobility of the appendage. The claw is somewhat capable of both horizontal and vertical rotation. This, combined with the fact that the claws are fairly small, made it impossible to generate reliable measurements. The peg-like seta on palp segment three is also very small. Part of this seta is submerged within the palp and even a small variation in the positioning of the palp on the slide made a great difference in how much of this seta appeared to be protruding. Both the claws and the peg-like seta could still present distinguishing features, but these would need to be examined under larger magnification than the microscope used in this thesis was capable of.

The features that could be reliably measured show a pattern in defining characteristics that is somewhat congruent with the habitat preference of the clades. Clades B and D, which were thus far collected only in lentic environments have an increased number of swimming setae (Figure 7 i, ii) compared to the clades from purely lotic environments. There is no difference in the number of swimming setae for the specimens from different environment types across clade C. There is also a somewhat general tendency for long or slender legs in clades from lentic environments, while the ones from lotic environments tend towards short or stout legs and leg segments. Clade C does not show any observable differences among specimens here either. Clade D presents one specifically interesting feature – the occasional doubled seta on palp segment three (Figure 7 iv). This violates one of the defining characteristics of the subgenus *Pilolebertia* – five long setae on palp segment three (Gerecke, 2009). Considering that the sister group of this clade, clade E, does not present this characteristic at all, it could call to question the definition of the whole subgenus.

Wolbachia

A search across literature and databases has not revealed any published records of *Wolbachia* infections in water mites beyond speculation about their presence such as in Stryjecki et al., (2015). The identifiable *Wolbachia* sequences obtained from the water mites used in my study show very few genetic differences and are proven to be present in at least two clades – C and G. The unidentifiable sequences were isolated from more clades, but since their identity currently cannot be confirmed, they cannot be considered a proof of *Wolbachia* infection. The presence of *Wolbachia* within the other clades remains a possibility.

Since the closest identified *Wolbachia* sequences come from other, quite distant, groups of mites and one nematode parasite of mites, it is possible that the initial infection was a result of a horizontal transfer by a parasite. No common pattern of clustering can be seen between the bacteria and the water mites, but that could simply reflect a relatively recent infection. The 16S rDNA sequence has an estimated divergence rate of 1-2% per 50 million years and it is not useful in identifying newly diverged strains (Werren et al., 1995). An attempt was made to amplify more divergent sequences of *Wolbachia* genome, such as *wsp* and *ftsZ* (Werren et al., 1995). The PCR using the primers for these markers was not successful (Supplement 2). The fact that one of these primers had been recently successfully used to amplify *wsp* sequences of *Wolbachia* in aquatic insects could further indicate that the *Wolbachia* in these water mite specimens were horizontally transferred from other mite groups rather than originating from their prey and hosts.

The infected hosts are widely distributed across Norway and come from various habitats. Four out of six hosts were female, two were male. It is not possible to positively identify which *Wolbachia* phenotype is expressed in water mites from this data. Stryjecki et al., (2015) found intersex specimen in the genus *Arrenurus* Dugès, 1834, and proposed *Wolbachia* as one of possible causes of this phenomenon. These intersex specimens had outer morphological features of females but genital structure of males. *Arrenurus* presents striking sexual dimorphism so noticing feminized males could be much easier than in other water mite genera such as *Lebertia* (Więcek et al., 2021). Male killing phenotype does not seem likely since two infected specimens were male, but female parthenogenesis is a possibility. Female bias is common among water mites, but that does not necessarily prove the presence of sex ratio skewing *Wolbachia*. It could simply be a result of the species life cycle (Davids et al., 2006).

In species where the cytoplasmic incompatibility phenotype is expressed, the COI marker tends to give very different number of OTUs in molecular delimitation (Jiang et al., 2018; Sucháčková Bartoňová et al., 2021; Whitworth et al., 2007). Cytoplasmic incompatibility, therefore, does not seem very probable in most clades based on the nuclear markers and the mitochondrial marker showing very similar divergence patterns. Clade C could potentially be an exception. Its phylogenetic trees and single marker delimitations show some difference between the pattern in the mitochondrial and nuclear markers, but further investigation would be needed to form a conclusion. Lastly, it is also possible that the *Wolbachia* DNA was not extracted from the water mites at all, but rather from an internal parasite such as the previously mentioned nematode. The specimens were not checked for the presence of parasites during my study and the nearest neighbours of the identified sequences do not allow us to discard this possibility.

Taxonomy

The species *L. porosa* (s.l.) was originally described from southern Norway. Specimens from clades C, F and G were found coexisting at its type locality. Clade F is morphologically distinct from the original description of *L. porosa* and does not represent the species detected by Thor, but clades C and G were both possible candidates for the nominal species. Reinhard Gerecke conducted a revision of the original material available in museum collections. Based on

morphological comparison with syntypes found in the collections of the Natural History Museum in Oslo, clade C could be recognized as representing *L. obscura* Thor, 1900. This species was described as a variety of *L. porosa* (s.l.) in the same paper in which Thor introduced the name of *L. porosa*, both from the same type locality. Consequently, clade G would be considered true *L. porosa*. Specimens belonging to clade G agree with all features given in the original description (Thor, 1900) and in his by far the more detailed redescription (Thor, 1906). As the type material of *L. porosa* is lost, probably in consequence of the destruction of Thor's collection after his death (Lundblad, 1938; Viets, 1940), a neotype should be selected from the material collected at the type locality in 2019/2020, preferably one of the specimens sequenced in the course of the present study. The other junior synonyms of *L. porosa* might be applicable to the other clades, but an examination of the type specimens used to describe them will be needed for confirmation. If no matching type specimens are found among the junior synonyms, the clades will be described as species new to science.

Phylogeny

The trees constructed for this study were created for the purpose of species delimitation. Constructing the proper phylogeny of any part of the family Lebertiidae was not the aim of this research and the data is insufficient for full phylogenetic resolution. The *L. porosa* clades were not sampled evenly or broadly enough, and other potentially closely related *Lebertia* species are missing completely. A proper phylogenetic study would need to include not only *L. insignis*, but also at least *Lebertia fimbriata, Lebertia pusilla, Lebertia stigmatifera, Lebertia sefvei, Lebertia helvetica*, and *Lebertia inaequalis*, which cluster together with or between the various *L. porosa* clades in the COI marker Neighbour Joining tree created from the data in BOLD. The lack of data from these intermediate species could also explain the lack of reticulation in the haplotype network analysis (Figure 6).

A few observations on the general structure of the trees for all three markers can be made, however. Clades A, B and C will likely remain grouped together even after the insertion of the other species. The node connecting clades D and E as sister groups received strong support. The COI marker alone is not enough to make assumption about the true phylogeny, but *L. insignis* was placed between *L. porosa* clades for all three markers, with high bootstrap support values for the position for two of them. It is also not unlikely that other *Lebertia* species will be slotted between the clades as well. This could indicate that the *L. porosa* species cluster is not a monophyletic group and is more closely connected by morphological similarities than by phylogenetic relationships. Clades F and G occupy the most distant position closest to the outgroup depending on the marker. This could mean that these clades are the most distantly related among this species cluster.

Conclusion

The Norwegian populations of water mites belonging to the *L. porosa* species complex form seven clades which can be delimited by both molecular and morphological characters. One of the three delimited clades detected at the type of locality was given the nominal species name *Lebertia*

porosa. Another was assigned a junior synonym also described by Thor from the same locality. The remaining clades can be described as species new to science or be assigned names of other species previously regarded as junior synonyms of *L. porosa*. No pattern was detected in the *Wolbachia* sequences and the lack of discord between the nuclear and mitochondrial marker genealogies support the assumption that different bacterial strains are not responsible for the seven clades delimited in this study. There are questions regarding this species complex that are outside of the scope of my thesis, but the results presented here confirm my original hypothesis that *L. porosa*, as previously defined, is a complex of multiple species.

Acknowledgements

This thesis would not have been possible without Norwegian Taxonomy Initiative “Water mites and midges in southern Norway”, Knr 13-18 funding the collection of water mites. I would like to thank my supervisors, Torbjørn Ekrem, Elisabeth Stur, and Reinhard Gerecke, for guiding me through the practical and theoretical aspects of my research and for providing much advice during the writing of this thesis. Also thanks to Mohsen Falahati, for showing me the ropes around the labs and for taking on part of my sample preparation when I might not have been able to access the labs. Finally, thanks to Varsha Rani, for providing much sanity and an outside opinion to my writing.

References

Bilton, D. T., Freeland, J. R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics*, *32*(1), 159–181. https://doi.org/10.1146/annurev.ecolsys.32.081501.114016

Blattner, L., Gerecke, R., & Von Fumetti, S. (2019). Hidden biodiversity revealed by integrated morphology and genetic species delimitation of spring dwelling water mite species (Acari, Parasitengona: Hydrachnidia). *Parasites and Vectors*, *12*(1), 1–13. https://doi.org/10.1186/s13071-019-3750-y

Bohonak, A. J. (1999). Effect of insect-mediated dispersal on the genetic structure of postglacial water mite populations. *Heredity*, *82*(4), 451–461. https://doi.org/10.1038/sj.hdy.6885050

Bottger, K. (1976). The general life cycle of fresh water mites (Hydrachnelliae, Acari). *Acarologia*, *18*(3), 496–502. http://www1.montpellier.inra.fr/CBGP/acarologia/article.php?id=3063

Breeuwer, J. A. J., & Jacobs, G. (1996). *Wolbachia*: Intracellular manipulators of mite reproduction. *Experimental and Applied Acarology*, *20*(8), 421-434. https://doi.org/10.1007/BF00053306

Brown, A. N., & Lloyd, V. K. (2015). Evidence for horizontal transfer of *Wolbachia* by a Drosophila mite. *Experimental and Applied Acarology*, *66*(3), 301–311. https://doi.org/10.1007/s10493-015-9918-z

Callisto, M., Ribeiro, A. S., Santana, V. B., França, J. S., Ligeiro, R., Ferreira, W. R., Silva, D., Castro, D., Tupinambás, T. H., Santana, D., Souza, B., Gonçalves, F., Rodrigues, L., Andrade, C. B., Sales, S. C. M., & Souza, R. (2011). Rapid Ecological Assessment of benthic indicators of water quality: A successful capacity-building experience for Brazilian postgraduate students in ecology. *Brazilian Journal of Biology*, *71*(4), 937–947. https://doi.org/10.1590/s1519-69842011000500014

Camargo, A., Morando, M., Avila, L. J., & Sites Jr, J. W. (2012). Species delimitation with ABC and other coalescent-based methods: A test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). *Evolution: International Journal of Organic Evolution*, *66*(9), 2834-2849. https://doi.org/10.1111/j.1558-5646.2012.01640.x

Cantonati, M., Gerecke, R., & Bertuzzi, E. (2006). Springs of the Alps - Sensitive ecosystems to environmental change: From biodiversity assessments to long-term studies. *Hydrobiologia*, *562*(1), 59–96. https://doi.org/10.1007/s10750-005-1806-9

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology, (22)*17, 4369–4383. https://doi.org/10.1111/mec.12413

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*(10), 1657–1659. https://doi.org/10.1046/j.1365-294X.2000.01020.x

Clement, M., Snell, Q., Walke, P., Posada, D., & Crandall, K. (2002). TCS: Estimating gene genealogies. *Proceedings - International Parallel and Distributed Processing Symposium, IPDPS 2002*. https://doi.org/10.1109/IPDPS.2002.1016585

Cook, D. R. (1974). Water mite genera and subgenera. *Memoirs of the American Entomological Institute*. *21*, 1-860.

Cordaux, R., Michel-Salzat, A., & Bouchon, D. (2001). *Wolbachia* infection in crustaceans: Novel hosts and potential routes for horizontal transmission. *Journal of Evolutionary Biology*, *14*(2), 237–243. https://doi.org/10.1046/j.1420-9101.2001.00279.x

Dabert, M., Proctor, H., & Dabert, J. (2016). Higher-level molecular phylogeny of the water mites (Acariformes: Prostigmata: Parasitengonina: Hydrachnidiae). *Molecular Phylogenetics and Evolution*, *101*, 75–90. https://doi.org/10.1016/j.ympev.2016.05.004

Dabert, M., Witalinski, W., Kazmierski, A., Olszanowski, Z., Dabert, J. (2010). Molecular phylogeny of acariform mites (Acari, Arachnida): Strong conflict between phylogenetic signal and long-branch attraction artifacts. *Molecular Phylogenetics and Evolution*, *56*, 222-241. https://doi.org/10.1016/J.YMPEV.2009.12.020.

Darwell, C. T., & Cook, J. M. (2017). Cryptic diversity in a fig wasp community—morphologically differentiated species are sympatric but cryptic species are parapatric. *Molecular Ecology*, *26*(3), 937-950. https://doi.org/10.1111/mec.13985

Davids, K., Di Sabatino, A., Gerecke, R., Gledhill, T., Smit, H., & van der Hammen, H. (2006). Acari: Hydrachnidia. *Süßwasserfauna von Mitteleuropa*, *7*, 241-376.

De Queiroz, K. (1998). The General Lineage Concept of Species , Species Criteria , and the Process of Speciation and Terminological Recommendations. *Endless Forms: Species and Speciation*, 57–75. http://si-pddr.si.edu/jspui/handle/10088/4652

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Dellicour, S., & Flot, J. F. (2015). Delimiting species-poor data sets using single molecular markers: A study of barcode gaps, haplowebs and GMYC. *Systematic Biology*, *64*(6), 900–908. https://doi.org/10.1093/sysbio/syu130

Di Sabatino, A., Cicolani, B., & Gerecke, R. (2003). Biodiversity and distribution of water mites (Acari, Hydrachnidia) in spring habitats. *Freshwater Biology*, *48*(12), 2163–2173. https://doi.org/10.1046/j.1365-2427.2003.01151.x

Di Sabatino, A., Gerecke, R., & Martin, P. (2000). The biology and ecology of lotic water mites

(Hydrachnidia). *Freshwater Biology*, *44*(1), 47–62. https://doi.org/10.1046/j.1365-2427.2000.00591.x

Di Sabatino, A., Smit, H., Gerecke, R., Goldschmidt, T., Matsumoto, N., & Cicolani, B. (2008). Global diversity of water mites (Acari, Hydrachnidia; Arachnida) in freshwater. *Hydrobiologia*, *595*(1), 303-315. https://doi.org/10.1007/s10750-007-9025-1

Esselstyn, J. A., Evans, B. J., Sedlock, J. L., Khan, F. A. A., & Heaney, L. R. (2012). Single-locus species delimitation: A test of the mixed yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1743), 3678–3686. https://doi.org/10.1098/rspb.2012.0705

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. https://doi.org/10.1111/mec.14486

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution*, *35*(10), 2585–2593. https://doi.org/10.1093/molbev/msy147

Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2017). Biodiversity and the Species Concept - Lineages are not Enough. *Systematic Biology*, *66*(4), 644–656. https://doi.org/10.1093/sysbio/syw098

Gerecke, R. (2009). Revisional studies on the European species of the water mite genus *Lebertia* NEUMAN, 1880 (Acari: Hydrachnidia: Lebertiidae). *Abhandlungen Der Senckenberg Gesellschaft Fur Naturforschung*, *566*, 1–144.

Gerecke, R., Martin, P., & Gledhill, T. (2018). Water mites (Acari: Parasitengona: Hydrachnidia) as inhabitants of groundwater-influenced habitats - considerations following an update of Limnofauna Europaea. *Limnologica*, *69*, 81–93. https://doi.org/10.1016/j.limno.2017.11.008

Goldschmidt, T. (2016). Water mites (Acari, Hydrachnidia): powerful but widely neglected bioindicators–a review. *Neotropical Biodiversity*, *2*(1), 12–25. https://doi.org/10.1080/23766808.2016.1144359

Growns, J. E. (2001). Aquatic mites as bioindicators, with an australian example. *Acarology: Proceedings of the 10th International Congress*, 136–142. https://www.cabdirect.org/cabdirect/abstract/20013124092

Hall, B. G. (2013). Building phylogenetic trees from molecular data with MEGA. *Molecular Biology and Evolution*, *30*(5), 1229-1235. https://doi.org/10.1093/molbev/mst012

Hong, D. Y. (2020). Gen-morph species concept—A new and integrative species concept for outbreeding organisms. *Journal of Systematics and Evolution*, *58*(5), 725–742. https://doi.org/10.1111/jse.12660

Ilvonen, J. J., Kaunisto, K. M., & Suhonen, J. (2018). Odonates, gregarines and water mites: why are the same host species infected by both parasites?. *Ecological Entomology*, *43*(5), 591–600. https://doi.org/10.1111/een.12634

Jacobs, S. J., Kristofferson, C., Uribe-Convers, S., Latvis, M., & Tank, D. C. (2018). Incongruence in molecular species delimitation schemes: what to do when adding more data is difficult. *Molecular Ecology*, *27*(10), 2397-2413. https://doi.org/10.1111/mec.14590

Jiang, W., Zhu, J., Wu, Y., Li, L., Li, Y., Ge, C., Wang, Y., Endersby, N. M., Hoffmann, A. A., & Yu, W. (2018). Influence of *Wolbachia* infection on mitochondrial DNA variation in the genus *Polytremis* (Lepidoptera: Hesperiidae). *Molecular Phylogenetics and Evolution*, *129*, 158–170. https://doi.org/10.1016/j.ympev.2018.08.001

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lundblad, O. (1938): SIG THOR. - *Entomologisk Tidskrift*, 59 (3-4): 107-111.

Luo, A., Ling, C., Ho, S. Y., & Zhu, C. D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, *67*(5), 830-846. https://doi.org/10.1093/sysbio/syy011

Martin, P., & Gerecke, R. (2009). Diptera as hosts of water mite larvae - an interesting relationship with many open questions. *Lauterbornia*, *68*, 95–103.

Martin, P., Stur, E., & Wiedenbrug, S. (2010). Larval parasitism of spring-dwelling alpine water mites (Hydrachnidia, Acari): A study with particular reference to chironomid hosts. *Aquatic Ecology*, *44*(2), 431–448. https://doi.org/10.1007/s10452-009-9301-4

McKay, B. D., Mays Jr, H. L., Wu, Y., Li, H., Yao, C. T., Nishiumi, I., & Zou, F. (2013). An empirical comparison of character-based and coalescent-based approaches to species delimitation in a young avian complex. *Molecular Ecology*, *22*(19), 4943-4957. https://doi.org/10.1111/mec.12446

Montes-Ortiz, L., & Elías-Gutiérrez, M. (2020). Water Mite Diversity (Acariformes: Prostigmata: Parasitengonina: Hydrachnidiae) from Karst Ecosystems in Southern of Mexico: A Barcoding Approach. *Diversity*, *12*(9), 329. https://doi.org/10.3390/d12090329

Pentinsaari, M., & Vos, R. (2017). Algorithmic single-locus species delimitation: effects of sampling effort, variation and nonmonophyly in four methods and 1870 species of beetles. *Wiley Online Library*, *17*(3), 393–404. https://doi.org/10.1111/1755-0998.12557

R Core Team. (2020). R: A language and environment for statistical computing. https://www.R-project.org/

Pešić, V., Asadi, M., Cimpean, M., Dabert, M., Esen, Y., Gerecke, R., Martin, P., Savić, A., Smit, H., & Stur, E. (2017). Six species in one: Evidence of cryptic speciation in the *Hygrobates fluviatilis* complex (Acariformes, Hydrachnidia, Hygrobatidae). *Systematic and Applied Acarology, 22*(9), 1327–1337. https://doi.org/10.11158/saa.22.9.4

Pešić, V., Saboori, A., Jovanović, M., Manović, A., Bańkowska, A., & Zawal, A. (2020). *Torrenticola dowlingi* sp. nov. a new water mite from Iran based on morphometrical and molecular data (Acariformes, Hydrachnidia, Torrenticolidae). *International Journal of Acarology*, *46*(5), 298-303. https://doi.org/10.1080/01647954.2020.1802513

Pešić, V., & Smit, H. (2016). Evidence of cryptic and pseudocryptic speciation in *Brachypodopsis baumi* species complex (Acari, Hydrachnidia, Aturidae) from Borneo, with description of three new species. *Systematic and Applied Acarology*, *21*(8), 1092-1106. https://doi.org/10.11158/saa.21.8.10

Puillandre, N., Brouillet, S., & Achaz, G. (2021). ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources*, *21*(2), 609–620. https://doi.org/10.1111/1755-0998.13281

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Sazama, E. J., Bosch, M. J., Shouldis, C. S., Ouellette, S. P., & Wesner, J. S. (2017). Incidence of *Wolbachia* in aquatic insects. *Ecology and Evolution*, *7*(4), 1165–1169. https://doi.org/10.1002/ece3.2742

Sela, I., Ashkenazy, H., Katoh, K., & Pupko, T. (2015). GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Research*, *43*(1), 7–14. https://doi.org/10.1093/nar/gkv318

Simpson, G. G. (1951). The species concept. *Evolution*, *5*(4), 285-298. https://doi.org/10.2307/2405675

Smit, H. (2020). Water mites of the world, with keys to the families, subfamilies, genera and subgenera (Acari: Hydrachnidia). *Monogr. Nederl. Entomol. 12*, 1-774.

Stålstedt, J., Bergsten, J., & Ronquist, F. (2013). “Forms” of water mites (Acari: Hydrachnidia): intraspecific variation or valid species?. *Ecology and evolution*, *3*(10), 3415-3435. https://doi.org/10.1002/ece3.704

Stryjecki, R., Czepiel-Mil, K., Gryzinska, M., & Zawal, A. (2015). A very rare case of intersexuality in water mites of the genus *Arrenurus* Dugès, 1834 (Acari, Hydrachnidia). *Invertebrate Reproduction & Development*, *59*(3), 155-165.

Stuart, B. L., Inger, R. F., & Voris, H. K. (2006). High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology letters*, *2*(3), 470-474. https://doi.org/10.1098/rsbl.2006.0505

Sucháčková Bartoňová, A., Konvička, M., Marešová, J., Wiemers, M., Ignatev, N., Wahlberg, N., Schmitt, T., & Faltýnek Fric, Z. (2021). *Wolbachia* affects mitochondrial population structure in two systems of closely related Palaearctic blue butterflies. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-82433-8

Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: Insights from a butterfly survey. *Methods in Ecology and Evolution*, *4*(12), 1101–1110. https://doi.org/10.1111/2041-210X.12107

Viets, K. (1939). Sig Thor: Ein Nachruf: Mit 1 Bildnis. *Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen*, *9*(1), 356-357. https://doi.org/10.1080/03680770.1940.11898683

Walter, D. E., & Proctor, H. C. (2013). *Mites: Ecology, Evolution & Behaviour: Life at a Microscale*. Springer Science & Business Media.

Wattier, R., Mamos, T., Copilaş-Ciocianu, D., Jelić, M., Ollivier, A., Chaumot, A., Danger, M., Felten, V., Piscart, C., Žganec, K., Rewic, T., Wysocka, A., Rigaud, T., Grabowski, M. (2020). Continental-scale patterns of hyper-cryptic diversity within the freshwater model taxon *Gammarus fossarum* (Crustacea, Amphipoda). *Scientific reports*, *10*(1), 1-16. https://doi.org/10.1038/s41598-020-73739-0

Werren, J. H., Baldo, L., & Clark, M. E. (2008). *Wolbachia*: Master manipulators of invertebrate biology. *Nature Reviews Microbiology*, *6*(10), 741–751. https://doi.org/10.1038/nrmicro1969

Werren, J. H., & Windsor, D. M. (2000). *Wolbachia* infection frequencies in insects: evidence of a global equilibrium?. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1450), 1277-1285. https://doi.org/10.1098/rspb.2000.1139

Werren, J. H., Zhang, W., & Guo, L. R. (1995). Evolution and phylogeny of *Wolbachia*: Reproductive parasites of arthropods. *Proceedings of the Royal Society B: Biological Sciences*, *261*(1360), 55–63. https://doi.org/10.1098/rspb.1995.0117

Więcek, M., Broda, L., Proctor, H., Dabert, M., Smith, B. P., & Dabert, J. (2021). Species boundaries among extremely diverse and sexually dimorphic *Arrenurus* water mites (Acariformes: Hydrachnidiae: Arrenuridae). *bioRxiv*. https://doi.org/10.1101/2021.04.04.438411

Więcek, M., Szydło, W., Dabert, J., & Proctor, H. (2020). Delimiting species of water mites of the genus *Hydrodroma* (Acari: Hydrachnidiae: Hydrodromidae) from North America and Europe: Integrative evidence of species status from COI sequences and

morphology. *Zoologischer Anzeiger*, *284*, 16-29. https://doi.org/10.1016/j.jcz.2019.11.004

Whitworth, T. L., Dawson, R. D., Magalon, H., & Baudry, E. (2007). DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). *Proceedings of the Royal Society B: Biological Sciences*, *274*(1619), 1731-1739. https://doi.org/10.1098/rspb.2007.0062

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *107*(20), 9264-9269. https://doi.org/10.1073/pnas.0913022107

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology 61*(5), 854-865. https://doi.org/10.1093/czoolo/61.5.854

Zhang, C., Zhang, D. X., Zhu, T., & Yang, Z. (2011). Evaluation of a Bayesian coalescent method of species delimitation. *Systematic biology*, *60*(6), 747-761. https://doi.org/10.1093/sysbio/syr071

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

NTNU

Norwegian University of
Science and Technology