Integrating Morphology and Genetics to Estimate Species Diversity, Host Specificity and Life Cycles of Echinostome Trematodes

by

Asma Sultana

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biological Sciences

University of Manitoba

Winnipeg, Manitoba

Abstract

Recognizing parasite diversity is important because treating several parasite species as one could be masking variation in infection dynamics, geographic distributions and host use. DNA sequencing aids in parasite identification, but only using genetic sequences limits our ability to understand parasite taxonomy and evolutionary ecology. Using integrative taxonomy to link genetics, morphology and host use provides insight into levels of biodiversity (e.g. populations and species) that can inform our knowledge of life cycles and host specificity. Echinostomes are a species-rich group of trematodes found as larvae in freshwater snails and as adults in vertebrates such as birds and mammals worldwide. Within this group, DNA sequencing has suggested misidentifications and cryptic species. This implies that our knowledge of life cycles and host specificity may be inaccurate. In Chapter 1, I sequenced larval and adult echinostomes collected from freshwater snails, birds, and muskrats in southern Manitoba at a mitochondrial (Nicotinamide Adenine Dinucleotide Hydrogenase subunit 1) and nuclear (28S rRNA) gene. To estimate genetic diversity and elucidate life cycles and host specificity in echinostomes, I performed phylogenetic, haplotype network and divergence analysis, which confirmed five known lineages/species and six new lineages/species that had not previously been genetically characterized. I genetically confirmed the first and final definitive hosts for five lineages/species and the host specificity of adults and larvae for three lineages/species each. In Chapter 2, I measured genetically-identified specimens and museum specimens to test the hypothesis of crypsis for genetic lineages within the nominal species (*Echinostoma trivolvis*) and between two nominal species (*E. trivolvis* and *Echinostoma revolutum* sensu lato) that co-occur within the same final bird and mammal hosts in North America. *Echinostoma trivolvis* genetic lineages were not morphologically distinguishable from each other and were also not

morphologically distinct from *E. revolutum* sensu lato. These results illustrate why misidentifications have occurred and suggest that species identification in past studies may not be reliable. Integrative taxonomy improved our ability to recognize echinostome species and revealed more diversity in North America than previously known. By using genetically-identified specimens, I elucidated previously unknown life cycles and found that some descriptions of life cycles and host specificity were inaccurate. Improving knowledge of life cycles and host specificity aids in understanding which species of echinostomes cause significant disease in wildlife and also in testing how intermediate and definitive hosts have influenced speciation.

Acknowledgements

First, I express sincere gratitude to my thesis advisor Dr. Jillian Detwiler. The door of her office was always open whenever I ran into trouble or had a question about my research or writing. I am really grateful for her excellent supervision and motivation throughout my MSc journey. I also thank Drs. Margaret Docker, Terry Galloway and Jason Gibbs for serving as my graduate committee members and for providing brilliant suggestions and comments about my research and writing. Also, special thanks to Dr. Norman Kenkel for his assistance with multivariate statistical analysis.

I acknowledge John Christie, Chris Benson and Wildlife Haven Rehabilitation Centre and Prairie Wildlife Rehabilitation Centre for donating birds and muskrats. I thank my fellow lab mates: Chenhua Li, Demi Gagnon, Jia Ming Tan, Laura Eliuk, Lauren Wiens, Osasumwen Osifo and Scott Malotka for helping me with necropsies and in snail collecting. Special thanks go to Parisa Selseleh and Rachel Donnelly for their assistance with genetic work.

Several scholarships and grants funded my research including an NSERC Discovery grant and Graduate Enhancement of Tri-Council Stipend awarded to Dr. Jillian Detwiler. I was awarded the Manitoba Graduate Scholarship, Alfred Rhea Tucker Memorial Scholarship and International Graduate Student Entrance Scholarship. Additionally, I received travel grants from Department of Biological Sciences and the Faculty of Science at the University of Manitoba as well as from the Canadian Society of Zoologists and American Society of Parasitologists to present at their annual conferences.

Finally, I must express my profound gratitude to my family and my loving husband for their continuous support and encouragement throughout my graduate career. This accomplishment would not have been possible without them.

Table of Contents

List of Tables

List of Figures

Thesis Introduction

Delimiting species boundaries is critical for taxonomic studies and biodiversity surveys to estimate species diversity accurately. Once the characteristics of species are clear, then the ecology and evolution can be investigated. Ecological studies of parasites help us understand the influence of parasites in natural ecosystems while evolutionary studies can give us insights into the factors that drive parasite speciation (Lafferty et al., 2006; Georgieva et al., 2013).

There are a variety of approaches that can be used to delimit species. Historically, most species were initially recognized because of morphological differences (Beaver, 1937; Kanev, 1994). In the last few decades, DNA sequencing and genetic analyses have been used to measure genetic differences and reconstruct phylogenies to delimit species (Brooks and McLennan, 1993; Morgan and Blair, 1998; Detwiler et al., 2010). For some taxonomic groups such as helminth parasites, the use of molecular tools began to replace morphological studies which ultimately led to a decline in our understanding of the morphology of parasite species (Blair and McManus, 1989; Luton et al., 1992; Adlard et al., 1993). This is problematic because for many parasites, there are few or no diagnostic gene sequences in public databases, which limits our ability to use only DNA sequences to identify species. Further, the DNA sequences were not often tied to a preserved specimen, meaning that genotype and phenotype from the same individual were unknown. Studying these groups from only a genetic perspective limits our ability to understand the evolutionary ecology of parasites leading to a call for integrative taxonomy (Shaw and Allen, 2000; Baker et al., 2003; Vieites et al., 2009; Herrmann et al., 2014). This approach delimits species with morphological characters in combination with other characteristics such as genetic, ecological, behavioural, etc (Dayrat et al., 2005). Integrative taxonomy will improve our ability

to identify species and enable us to address evolutionary ecology hypotheses relating to the genetics and morphology of parasites.

Before molecular-based approaches, trematode species were characterized by differences in morphology, host use and geographic distributions (Kostadinova and Gibson, 2000). However, this traditional approach to species delimitation has some limitations. When cryptic species are present (two or more species that were classified as one nominal species because they were at least superficially morphologically indistinguishable, Bickford et al., 2007), species diversity may be underestimated. DNA sequencing has revealed a growing number of parasites that are now considered cryptic species complexes (Pérez-Ponce de León and Poulin, 2018). In particular, trematodes (Platyhelminthes: Trematoda) have the highest number of cryptic species amongst platyhelminths, suggesting that many trematodes may have been misidentified (Pérez-Ponce de León and Poulin, 2018).

Another factor that can lead to inaccurate estimates of trematode species diversity is phenotypic plasticity. In contrast to crypsis, phenotypic plasticity can result in an overestimation of parasite species diversity because extensive morphological variation within species may be interpreted as evidence for several species rather than one species. Phenotypic plasticity is a widespread phenomenon in nature that occurs when the environment induces changes in a genotype's behaviour, morphology and physiology (Price et al., 2003). Several host-related factors such as species, age, physiological condition, and presence of previous infection may influence phenotypic plasticity of parasites (Stunkard, 1957). In the laboratory, Blankespoor (1974) exposed 51 definitive host species (host in which adult parasites reach sexual maturity) to *Plagiorchis noblei* Park, 1936 (Trematoda) and found that several morphological characters of the adult parasites differed in size, shape and position depending upon the host species in which

they developed. Evidence of host-induced phenotypic plasticity has also been suggested from field-collected trematodes. A high degree of morphological variation and lack of host specificity of specimens initially identified as *Isthmiophora melis* (Schrank, 1788) Lühe, 1909 were found in rodent and carnivore hosts, suggesting that more than one species was present. However, DNA sequencing revealed 1.4% divergence among cytochrome c oxidase (CO1) haplotypes suggesting that all the specimens belonged to one species, *I. melis* (Hildebrand et al., 2015).

Laboratory studies are ideal to test for phenotypic plasticity because the number of factors that can influence parasite phenotypes can be controlled. However, they also require animal care approval as well as expertise and facilities for proper maintenance of exposed vertebrate definitive hosts. Moreover, to perform adequate statistical analysis, relatively large sample sizes (e.g. >25 individuals) are necessary, which increases the demand on facilities and finances. In contrast, studying phenotypic plasticity of trematodes in naturally-infected host populations involves accounting for the effects of several factors that may each affect parasite phenotype including host age, host population density and exposure to previous infection (Blankespoor, 1974).

Given that morphology has proven problematic for delimitation in some parasite species, researchers turned to DNA sequencing to identify species because this approach is reliable, easier and faster than using morphology (Tautz et al., 2003; Blaxter, 2004; Gaston and O'Neill, 2004). For instance, in some cases, DNA barcoding can reveal cryptic species where morphology alone cannot (Hebert et al., 2003; Bickford et al., 2007). For parasites, DNA sequencing was only possible once genetic methods were adapted to working with small amounts of tissue (Nadler, 1990). Since the 1980s, an increasing number of cryptic helminth species have been found with the largest number being identified in trematodes (Poulin, 2011; Pérez-Ponce de León

and Poulin, 2018). It is not well understood why crypsis is so common within this group. Pérez-Ponce de León and Poulin (2018) presented three hypotheses that focused on aspects of trematode biology that distinguish them from other helminths. Cryptic trematodes may arise from somatic mutations that occur when the larvae asexually reproduce within a first intermediate mollusk host (Yin et al., 2008). If these mutations are common, then this could explain the abundance of cryptic trematode taxa. Another hypothesis is that there is a time lag between genetic speciation and morphological divergence because trematodes lack fast-evolving hard structures (Poulin, 2011). The final hypothesis was that the lack of hard structures may also mean that morphological differences may be subtle and require detailed morphological analysis. This last hypothesis speaks to the "at least superficially morphologically indistinguishable" part of Bickford et al.'s (2007) definition of cryptic species. Detailed investigations of morphology could be lacking especially relative to other groups of helminths such as cestodes, leading to an overestimate of the number of cryptic trematode species (Nadler and Pérez-Ponce de León, 2011).

To recognize cryptic species, DNA sequence analyses can be used to recognize intraspecific and interspecific variation. Selection of gene regions is key because genes differ in the rate at which they accumulate substitutions. Some gene regions have higher substitution rates while other genes slowly accumulate changes over time. Blasco-Costa et al. (2016) recommended using gene regions that show sufficient variation to differentiate taxa but not so much that substitution saturation occurs and phylogenetic inferences are impeded. The mitochondrial nicotinamide adenine dinucleotide dehydrogenase subunit 1 (ND1) and the cytochrome c oxidase subunit 1 (CO1) genes as well as the two internal transcribed spacers of the rRNA gene (ITS1 and ITS2) are the most frequently used gene regions for the identification

of cryptic species for subclass Digenea (Nolan and Cribb, 2005; Olson and Tkach 2005; Criscione et al., 2005). The relative divergence rates of nuclear and mitochondrial DNA are not well understood in flatworms. However, Vilas et al. (2005) demonstrated that mitochondrial DNA had a higher rate of nucleotide substitution than nuclear DNA for some trematodes and cestodes. Both sets of loci are used because they are independent (not linked, no recombination) and can detect intraspecific and interspecific genetic divergence of cryptic species (Pérez-Ponce de León and Nadler, 2010; Blasco-Costa et al., 2016).

Genetic differences can be assessed by comparing the number of differences in base pairs within and between species for each type of gene. Further, monophyletic groups can be compared between mitochondrial and nuclear phylogenies to determine if clades represent different putative species (agreement between mitochondrial and nuclear clades) or a variety of haplotypes at the population level (disagreement between mitochondrial and nuclear clades). However, these two approaches are not always reliable because genetic differences and relationships among species can differ depending upon the gene being used (Neigel and Avise, 1986; Nichols, 2001). Different factors contribute to the discordance between gene trees such as gene mutation, gene flow within populations, or migration which can make it difficult to recognize species based on single gene trees (Martin and Höhna, 2017). Although genetic sequences provide a measure of divergence among trematodes, they are just one type of information that can be used to answer the question "How do we recognize a species?"

There are several different species concepts and their use depends on the information that is known about a particular taxon of interest and on the philosophy of the researcher. The morphological species concept is commonly used to delimit trematode species. Using this model, individuals are characterized as species based on morphological distinctiveness (Mayden, 1997).

As discussed above, no or few morphological differences may lead to an underestimate of diversity if the species are cryptic. The evolutionary stability of some hard morphological structures such as collar spines and hooks make genetically different species appear similar. For example, several *Echinostoma* spp. share 37 collar spines (Fried and Toledo, 2004). Alternatively, overestimates of species diversity may occur if phenotypic plasticity within a parasite species is interpreted as evidence for several species. Because of these issues, the morphological species concept may be unreliable for delimiting trematode species.

The biological species concept (Mayer, 1942) defines species as a group of individuals which can interbreed and are reproductively isolated from other groups. This concept has rarely been tested in helminth parasites because of the challenges associated with studying mating behaviour (Detwiler et al., 2017). Given that helminths are endoparasites, direct observations of mating are very limited. Further, trematodes can self- and cross-fertilize, so even if two species are in contact, particular methods need to be used to verify parentage of offspring, such as analysing inter- and intra-isolate mating with isoenzymes or microsatellite genotypes (Trouvé et al., 1999; Detwiler et al., 2017).

One of the most commonly used species concepts for trematodes is the phylogenetic species concept (Nadler, 2002). According to this concept, a species is an irreducible monophyletic group derived from a common ancestor whose members share diagnostic/derived traits (Rosen, 1979; Donoghue, 1985). Molecular phylogenies are often used to differentiate putative species (Krone et al., 2007; Miranda et al., 2008). The phylogenetic approach is often used in concert with the genetic species concept which can determine differences in species and populations (Baker and Bradley, 2006). If pairwise genetic distances match or exceed the benchmark value of two sister species, then they are considered different species. For

trematodes, 5% and 1% are considered the benchmarks if using mitochondrial (CO1, ND1) and nuclear (ITS) genes, respectively (Vilas et al., 2005). Some researchers argue that differentiating species based on pairwise distance is arbitrary as the percent difference between species may differ depending upon the species being considered (DeSalle, 2006). Another issue is that different gene regions may vary in the nucleotide substitution rate among lineages. If the molecular substitution rate is higher for a certain gene region, the pairwise distance between two species will increase (Nadler, 2002). Because of the potential for this source of variation, it is important to understand the rates at several genes before choosing one for species delimitation.

Given the problems associated with relying on just one type of information (either morphological or genetic) for species delimitation, recent reviews have called for an integrative approach that synthetically combines several different features (phylogeography, comparative morphology, population genetics, ecology and behaviour) (Dayrat, 2005; Padial et al., 2010). For instance, Herrmann et al. (2014) revealed cryptic species complexes of the trematode, *Stegodexamene anguillae* Macfarlane, 1951, by combining molecular and ecological data. Although it may seem intuitive that all taxonomic studies should be integrative, this approach is not always used because it can be difficult and time-consuming to quantify morphology and additional expertise and expense are required to obtain other types of information about the parasites as well as to perform the genetic sequencing and analysis.

An integrative taxonomic approach for trematode taxonomy is particularly important because their small size and lack of species-specific morphological characteristics make it difficult to delineate these parasites using morphology alone (Luton et al., 1992; McManus and Bowles, 1996). Additionally, very few shared morphological features between larvae and adults make it challenging to link life cycle stages without experimental or genetic studies. This

cosmopolitan, diverse class of helminths has complex life cycles that include invertebrates (mostly mollusks) and vertebrate hosts such as birds and mammals to alternate their asexual and sexual generations. Accurate identification of trematode species can vastly improve our understanding of the epidemiology of wildlife and human disease. However, this task is challenging, considering that there are an estimated 18,000 nominal species (based on host species diversity) and this number may be higher based on preliminary investigations on cryptic species (Cribb et al., 2001; Pérez-Ponce de León and Poulin, 2018).

I will use integrative taxonomy to estimate the species diversity of echinostome trematodes in southern Manitoba. This group is species-rich with over 60 and 120 described species worldwide within just two genera (Kostadinova and Gibson, 2000). Several species of echinostomes have now been recognized as cryptic complexes in North America, Europe and Asia (Detwiler et al., 2010; Georgieva et al., 2014). Thus, estimates of diversity may be inaccurate due to cryptic species and phenotypic plasticity (Stillson and Platt, 2007; Detwiler et al., 2010). Characterizing the biodiversity of echinostomes is important to further their use as a model system for host-parasite interactions and to understand their role in human and wildlife disease. Various food-borne and zoonotic diseases have been reported in Southeast Asia associated with echinostomiasis. From 1980-1990, humans living in some areas within the Philippines, China and Thailand had prevalences ranging from 5-50% (Graczyk and Fried, 1998). Several echinostome species, such as *I. melis, Echinostoma revolutum* (Frolich, 1802) Looss, 1899 and *Echinoparyphium recurvatum* (Linstow 1873) Luehe, 1909, have been reported as causative agents of echinostomiasis in humans in Southeast Asia (Lu, 1982). In North America, echinostomiasis in humans is rare and more often recognized as playing a role in wildlife disease and especially in the decline of amphibians (Johnson and Mckenzie, 2009). In

particular, larval amphibians develop renal disorders after being infected by larval echinostomes, which may have severe consequences for amphibian conservation efforts (Holland et al., 2007).

Recent molecular-based studies have shown that echinostomes have been misidentified (Detwiler et al., 2010). For example, Holland et al. (2007) reported that frogs were infected by *E. revolutum* by matching their sequences to GenBank; subsequent DNA sequencing analysis confirmed they were infected by *Echinostoma trivolvis* Cort, 1914 because the GenBank sequences used for comparison had been misidentified (Detwiler et al., 2010). In Thailand, recent genetic studies show that the parasites identified as *E. revolutum* are likely a new species and thus whether *E. revolutum* can infect humans is not clear (Noikong et al., 2014; Nagataki et al., 2015). These misidentifications have been caused by mistakes or misinterpretations in morphological identification and can be partly attributed to the presence of cryptic species. Resolving the taxonomy of cryptic echinostome species is important because it will help to identify the causative agents of potential zoonotic and wildlife diseases. From an ecological perspective, it is important to identify echinostomes accurately because their trophic transmission can provide insights into host ecology (e.g. the habitat use, feeding and migratory behaviour of birds).

References

Adlard, R. D., S. C. Barker, D. Blair, and T. H. Cribb. 1993. Comparison of the second internal transcribed spacer (ribosomal DNA) from populations and species of Fasciolidae (Digenea). International Journal for Parasitology. 23: 423–425.

Baker, R. J. and R. D. Bradley. 2006. Speciation in mammals and the genetic species concept. Journal of Mammalogy. 87: 643–662.

Baker, A. M., C. Bartlett, S. E. Bunn, K. Goudkamp, F. Sheldon, and J. M. Hughes. 2003**.** Cryptic species and morphological plasticity in long-lived bivalves (Unionoida: Hyriidae) from inland Australia. Molecular Ecology. 12: 2707–2717.

Beaver, P. C. 1937. Experimental studies on *Echinostoma revolutum* (Frólich), a fluke from birds and mammals. Illinois Biological Monographs. p. 1–96.

Bickford, D., D. J. Lohman, N. S. Sodhi, P. K. L. Ng, R. Meier, K. Winker, K. K. Ingram, and I. Das. 2007. Cryptic diversity as a window on diversity and conservation. Trends in Ecology & Evolution. 22: 148–155.

Blair, D. and. P. McManus. 1989. Restriction enzyme mapping of ribosomal DNA can distinguish between fasciolid (liver fluke) species. Molecular and Biochemical Parasitology. 36: 201–208.

Blankespoor, H. D. 1974. Host-induced variation in *Plagiorchis noblei* Park, 1936 (Plagiorchiidae: Trematoda). American Midland Naturalist. 92: 415–433.

Blasco-Costa I, S. C. Cutmore, T. L. Miller, and M. J. Nolan. 2016. Molecular approaches to trematode systematics: 'Best practice' and implications for future study. Systematic Parasitology. 93: 295–306.

Blaxter, M. L. 2004. The promise of DNA taxonomy. Philosophical Transactions of the Royal Society of London. 359: 669–679.

Brooks, D. R. and D. A. McLennan. 1993. Parascript: Parasites and the language of evolution. Smithsonian Institution Press, Washington, DC. pp. 429.

Cribb, T.H., R. A. Bray, D. T. J. Littlewood, S. P. Pichelin, and E. A. Herniou. 2001. Interrelationships of the Platyhelminthes. Taylor and Francis, London. pp. 168–185.

Criscione, C. D., R. Poulin, and M. S. Blouin. 2005. Molecular ecology of parasites: Elucidating ecological and microevolutionary processes. Molecular Ecology. 14: 2247–2257.

Dayrat, B. 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society. 85: 407–415.

DeSalle, R. 2006. Species discovery versus species identification in DNA barcoding efforts: Response to Rubinoff. Conservation Biology. 20: 1545–1547.

Detwiler, J. T., D. H. Bos, and D. J. Minchella. 2010. Revealing the secret lives of cryptic species: Examining the phylogenetic relationships of echinostome parasites in North America. Molecular Phylogenetics and Evolution**.** 55: 611–620.

Detwiler, J. T., I. C. Caballero, and C. D. Criscione. 2017. Role of parasite transmission in promoting inbreeding: I. Infection intensities drive individual parasite selfing rates. Molecular Ecology. 26: 4391–4404.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. Bryologist. 88: 172–181.

Fried, B. and R. Toledo. 2004. Criteria for species determination in the '*revolutum*' group of *Echinostoma*. Journal of Parasitolgy. 90: 917.

Gaston, K. and M. A. O'Neill. 2004**.** Automated species identification: Why not? Philosophical Transactions of the Royal Society of London B. 359: 655–667.

Georgieva, S., A. Faltyńkova, R. Brown, I. Blasco-Costa, M. Solda´Nova, J. Sitko, T. Scholz, and A. Kostadinova. 2014. *Echinostoma 'revolutum'* (Digenea: Echinostomatidae) species complex revisited: Species delimitation based on novel molecular and morphological data gathered in Europe. Parasites & Vectors. 7: 520–537.

Georgieva S, C. Selbach, A. Faltýnková, M. Soldánová, B. Sures, K. Skírnisson, and

A, Kostadinova. 2013. New cryptic species of the 'revolutum' group of *Echinostoma* (Digenea: Echinostomatidae) revealed by molecular and morphological data. Parasites & Vectors. 6: 64–77.

Graczyk T. K. and B. Fried. 1998. Echinostomiasis: A common but forgotten food-borne disease. American Journal of Tropical Medicine and Hygiene. 58: 501–504.

Hebert P. D. N., A. Cywinska, S. L. Ball, and J. R. Dewaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. 270: 313–321.

Herrmann, K. K., R. Poulin, D. B. Keeney, and I. Blasco-Costa. 2014. Genetic structure in a progenetic trematode: Signs of cryptic species with contrasting reproductive strategies. International Journal for Parasitology. 44: 811–818.

Hildebrand, J., M. Adamczyk, Z. Laskowski, and G. Zaleśny. 2015. Host-dependent morphology of *Isthmiophora melis* (Schrank, 1788) Luhe, 1909 (Digenea, Echinostomatinae) morphological variation vs. molecular stability. Parasites & Vectors**.** 8: 481–488.

Holland, M.P., D. K. Skelly., M. Kashgarian, S. R. Bolden, L. M. Harrison, and M. Cappello. 2007. Echinostome infection in green frogs (*Rana clamitans*) is stage and age dependent. Journal of Zoology. 271: 455–462.

Johnson P. T. J. and V. J. Mckenzie. 2009. Effects of environmental change on helminth infections in amphibians: Exploring the emergence of *Ribeiroia* and *Echinostoma* infections in North America. In The Biology of Echinostomes. B. Fried and R. Toledo (eds.). Springer, NY, USA. p. 249- 280.

Kanev, I. 1994. Life-cycle, delimitation and redescription of *Echinostoma revolutum* (Frölich, 1802) (Trematoda: Echinostomatidae). Systematic Parasitology. 28: 125–144.

Kostadinova A. and Gibson D.I. 2000. The Systematics of the Echinostomes. In: Fried B., Graczyk T.K. (eds) Echinostomes as Experimental Models for Biological Research. Springer, Dordrecht. p. 31-57.

Krone, O., D. Friedrich, and M. Honisch. 2007. Specific status and pathogenicity of syngamid nematodes in bird species (Ciconiformes [sic], Falconiformes, Gruiformes) from Germany. Journal of Helminthology. 81: 67–73.

Lafferty, K. D., A. P. Dobson, and A. M. Kuris. 2006. Parasites dominate food web links. Proceedings of the National Academy of Sciences of the United States of America. 130: 211–216.

Lu, S. C. 1982. Echinostomiasis in Taiwan. International Journal of Zoonoses. 9: 33–38.

Luton, K., D. Walker, and D. Blair. 1992. Comparisons of ribosomal internal transcribed spacer from two congeneric species of flukes (Platyhelminthes: Trematoda: Digenea). Molecular and Biochemical Parasitology 56: 323–328.

Martin H. C. and S. Höhna. 2017**.** New evidence for the recent divergence of Devil's Hole pupfish and the plausibility of elevated mutation rates in endangered taxa. Molecular Ecology. 27: 831–838.

Mayden, R. L. 1997. A hierarchy of species concepts: The denouement in the saga of the species problem. In Claridge, M. F., Dawah, H. A., and Wilson, M. R. (eds.) Species: The units of biodiversity. London: Chapman and Hall, pp. 381–424.

Mayr, E. 1942. Systematics and the origin of species. New York: Columbia University Press pp. 372.

McManus, D.P. and J. Bowles. 1996. Molecular genetic approaches to parasite identification: Their value in diagnostic parasitology and systematics. International Journal for Parasitology. 26: 687–704.

Miranda, R. R., J. A. Tennessen, M. S. Blouin, and E. M. Rabelo. 2008. Mitochondrial DNA variation of the dog hookworm *Ancylostoma caninum* in Brazilian populations. Veterinary Parasitology. 151: 61–67.

Morgan, J. A. T., and D. Blair. 1998. Relative merits of nuclear ribosomal internal transcribed spacers and mitochondrial CO1 and ND1 genes for distinguishing among *Echinostoma* species (Trematoda). Parasitology. 116: 289–297.

Nadler, S. A. 1990. Molecular approaches to studying helminth population genetics and phylogeny. International Journal for Parasitology. 20: 11–29.

Nadler, S. A. 2002. Species delimitation and nematode biodiversity: Phylogenies rule. Nematology. 4: 615–625.

Nadler, S. A., and G. Pérez-Ponce De León. 2011. Integrating molecular and morphological approaches for characterizing parasite cryptic species: Implications for parasitology. Parasitology. 138: 1688–1709.

Nagataki, M., C. Tantrawatpan, T. Agatsuma, T. Sugiura, K. Duenngai, P. Sithithaworn, R. H. Andrews, T. N. Petney, and W. Saijuntha. 2015. Mitochondrial DNA sequences of 37 collar-spined echinostomes (Digenea: Echinostomatidae) in Thailand and Lao PDR reveals presence of two species: *Echinostoma revolutum* and *E. miyagawai*. Infection, Genetics and Evolution. 35: 56–62.

Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Evolutionary Processes and Theory: Academic Press. pp. 515–534.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology & Evolution. 60: 358–364.

Noikong, W., C. Wongsawad, J. Chai., S. Saenphetand, and A. Trudgett. 2014. Molecular analysis of echinostome metacercariae from their second intermediate host found in a localised geographic region reveals genetic heterogeneity and possible cryptic speciation. PLoS Neglected Tropical Diseases. 8: e2778.

Nolan, M. J, and T. H. Cribb. 2005. The use and implications of ribosomal DNA sequencing for the discrimination of digenean species. Advances in Parasitology.60: 101–163.

Olson, P. D., and V. V. Tkach. 2005. Advances and trends in the molecular systematics of the parasitic Platyhelminthes. Advances in Parasitology. 60: 165–243.

Padial, J. M., A. Miralles, I. De La Riva, and M. Vences. 2010. The integrative future of taxonomy. Frontiers in. Zoology. 7: 16–29.

Pérez -Ponce De León, G., and R. Poulin. 2018. An updated look at the uneven distribution of cryptic diversity among parasitic helminths. Journal of Helminthology. 92: 197–202.

Pérez-Ponce De León, G., and S. A. Nadler. 2010. What we don't recognize can hurt us: A plea for awareness about cryptic species. Journal of Parasitology. 96: 453–464.

Poulin, R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biology Letters. 7:241–244.

Price, T.D., A. Qvarnstrom, and D. E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society of London, Series B. 270: 1433–1440.

Rosen, D. E. 1979. Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. Bulletin of the American Museum of Natural History. 162: 269–375.

Shaw, A. J., and B. Allen. 2000. Phylogenetic relationships, morphological incongruence, and geographic speciation in the Fontinalaceae (Bryophyta). Molecular Phylogenetics and Evolution. 16: 225–237.

Stillson, L. L., and T. R. Platt. 2007. The crowding effect and morphometric variability in *Echinostoma caproni* (Digenea: Echinostomatidae) from ICR mice. Journal of Parasitology. 93: 242–246.

Stunkard, H. W. 1957. Intraspecific variation in parasitic flatworms. Systematic Biology. 6: 7–18.

Tautz, D. P., Arctander, A. Minelli., R. H. Thomas, and A. P. Vogler. 2003**.** A plea for DNA taxonomy. Trends in Ecology & Evolution. 18: 70–74.

Trouvé, S., F. Renaud, P. Durand, and J. Jourdane. 1999. Reproductive and mate choice strategies in the hermaphroditic flatworm *Echinostoma caproni*. Journal of Heredity. 90: 582–585.

Vieites, D. R., K. C. Walenberg, F. Andreone, J. Kölher, F. Glaw, and M. Vences. 2009. Vast underestimation of Madagascar biodiversity evidenced by an integrative amphibian inventory. Proceedings of the National Academy of Sciences of the United States of America. 106: 8267–8272.

Vilas, R., C. D. Criscione, and M. S. Blouin. 2005. A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. Parasitology. 131: 1–8.

Yin, M. B., W. Hu, X. J. Mo, S. Y. Wang, P. J. Brindley, D. P. McManus, G. M. Davis, Z. Feng, and D. Blair. 2008. Multiple near-identical genotypes of *Schistosoma japonicum* can occur in snails and have implications for population-genetic analyses. International Journal for Parasitology. 38: 1681–1691

Chapter 1: Elucidating life cycles with DNA sequencing reveals new insights into echinostome trematode biology

1.1 Introduction

For many trematodes, life cycles are inaccurately described or unknown. Elucidating life cycles is challenging because there are few shared morphological traits between larvae and adults, species descriptions include a few, but not all of the life stages in the life cycle, and some parasites have simply not been investigated sufficiently (Kostadinova and Gibson, 2000; Poulin, 2007). Life cycles may be initially described as the result of local or regional field surveys, but then as more field-infected hosts are evaluated and experimental infections are performed, they may require revision (Bolek et al., 2016). In addition, the presence of cryptic parasite species (morphologically similar, yet genetically distinct species) may affect our knowledge of life cycles because species within a cryptic complex may differ in their host use compared to what was described when that complex was considered one species. Elucidating life cycles is important because it helps us characterize biodiversity, host specificity and geographic distributions of parasites. These factors are fundamental to understanding the epidemiology, ecology and evolution and behaviour of trematodes.

Accurate identification of parasites is key to understanding life cycles. Historically, taxonomic studies used morphology and host specificity to define trematode species, but this classical approach has led to inaccurate estimates of species diversity. In many cases, trematode species diversity may be underestimated because genetic sequencing has revealed that many nominal species are comprised of cryptic species (Poulin, 2011; Pérez-Ponce de León and Poulin, 2018). Alternatively, species diversity can also be overestimated because phenotypic plasticity could be interpreted as evidence for several different species (Nolan and Cribb, 2005).

DNA sequencing allows us to recognize parasite diversity, link larval and adult stages to complete life cycles and determine if estimates of host specificity and geographic distributions are accurate.

The echinostomes (Digenea: Echinostomatidae) have long been the subject of taxonomic confusion (Roberts et al., 2009). Part of this confusion stems from the fact that this group is highly diverse relative to other trematode groups. Echinostomatidae has substantial species diversity relative to other trematode families with two particularly speciose genera with 120 and 60 described species (Kostadinova and Gibson, 2000; Jones, 2005). The true number of echinostome species may be higher because there are cryptic species complexes in several different genera (e.g. *Echinostoma,* Kanev et al., 1998; Detwiler et al., 2010). Especially for cryptic species, life cycles may be inaccurate if host specificity differs among the sister species in the cryptic complex. Echinostomes have a complex life cycle where larval stages develop in first and second intermediate hosts (invertebrates such as snails and vertebrates such as amphibians and freshwater fish) and the adult stage becomes sexually mature in final vertebrate hosts (birds and mammals).

Life cycles for echinostome species may be unclear because larval and adult stages are morphologically cryptic, making it difficult to identify the species that is infecting a particular host. For example, many species are distinguished based on the number of collar spines surrounding the oral sucker in adults and larvae. However, some congeners can have the same number of spines such as the 37-spined *Echinostoma* 'revolutum' complex, which means that more detailed morphological analysis is required to identify species accurately (Kostadinova and Gibson, 2000; Kostadinova, 2005). When more morphological characters are measured, it can result in splitting a species into several new species. Kostadinova (1999) re-examined the larval

sensory structure of the voucher materials described as *E. revolutum* by Kanev (1994) and found that it was a complex of at least two echinostome species- *E. revolutum* and *Echinostoma miyagawai* Ishii, 1932.

Another reason that echinostome life cycles are difficult to elucidate is because the larval and adult stages share few characteristics. Between larvae and adults, one of the only shared characteristics is the number of collar spines, but as pointed out, this trait can be shared by several species. Moreover, many researchers may not have the taxonomic expertise or guidance from a dichotomous key that promotes accurate identification based on morphology.

DNA sequencing can help resolve these two issues because each species can be distinguished based on genetic differences and larval and adult stages can be matched by genetic similarities. Morgan and Blair (1998a, b) delimited echinostome species based on the number of base pair differences at two gene regions. They found that mitochondrial DNA was more informative than nuclear DNA at differentiating closely related echinostome species. In particular, the Nicotinamide Adenine Dinucleotide Hydrogenase subunit 1 (ND1) gene was more variable (p-distance=12.3-30.8%) than the Internal Transcribed Spacer (ITS) rDNA (p-distance =1.1-3.7%), making it easier to distinguish sister *Echinostoma* species with mitochondrial DNA. In the initial studies by Morgan and Blair (1998a, b), there were few samples, so it appeared the genetic data largely supported the classical taxonomic approach. However, as more samples were collected from other geographic regions, researchers began to encounter cryptic species (Detwiler et al., 2010; Detwiler et al., 2012; Georgieva et al., 2014).

A wider sampling of species combined with phylogenetic analysis of ND1 and ITS sequences has clarified some aspects of echinostome taxonomy and estimates of diversity. In particular, sampling in Europe helped place genetic isolates into several genera (Kostadinova et

al., 2003). As a result of more comprehensive geographic sampling across Europe, several new species of echinostomes have been described using genetics and morphology (Georgieva et al., 2013, 2014). These studies also show that some species that were thought to be cosmopolitan are in fact complexes of cryptic species that each has smaller geographic distributions (e.g. *E. revolutum* sensu stricto in Europe vs *E. revolutum* sensu lato in USA). In North America, more geographic sampling has revealed the presence of several cryptic species complexes: *E. trivolvis*, *E. revolutum*, *Echinostoma robustum* Yamaguti, 1935 and *Echinoparyphium* lineage 1 (Detwiler et al., 2010). Specimens within these species complexes had the same number of collar spines, but were genetically distinct (>5% different at the partial ND1 gene). The hypothesis of crypsis has not been formally tested for most of these proposed cryptic species because only genetic analysis was conducted. Georgieva et al. (2014) called on researchers to use an integrative approach that includes genetics and morphology to clarify the taxonomic status of echinostomes in North America.

In addition to wider geographic sampling, more comprehensive sampling of hosts is required. Classical taxonomy often relied in part on host specificity for species identification. For example, it was thought that *E. revolutum* sensu stricto and sensu lato was only found in birds and not in mammals (Sorensen et al., 1997). Therefore, if a 37-spined adult echinostome was found in a mammal, it was not *E. revolutum*, but instead another species, *E. trivolvis*. However, DNA sequencing confirmed that *E. revolutum* sensu lato and *E. trivolvis* occur in mammals (Detwiler et al., 2012). Moreover, DNA sequencing has also shown that host specificity for first intermediate hosts is not always as expected. Detwiler et al. (2010) found that some echinostome species use more than one first intermediate host species, indicating that first intermediate host specificity is not always the best trait to use for accurate identification of larval parasites.

Increased host sampling in additional geographic locations will clarify host specificity and geographic distributions of species to promote a greater understanding of echinostome parasite ecology and evolution.

I used molecular phylogenetics to characterize echinostome biodiversity in southern Manitoba, which allowed me to determine if previous descriptions of life cycles, host specificity and geographic distributions were accurate. The abundance of wetlands and convergence of three migratory bird flyways increase the likelihood that Manitoba has higher echinostome diversity relative to other areas. Historically, echinostomes have been sampled in Manitoba (Bondar, 1950; McKenzie and Welch, 1979); however, most of these studies found species that are now hypothesized to be cryptic (e.g. *E. trivolvis*), so species diversity in this area remains unclear. I utilized the phylogenetic species concept that uses reciprocal monophyly to distinguish lineages/species (Rosen, 1979; Donoghue, 1985).

1.2. Materials and Methods

1.2.1. Sample collection

Six species of freshwater snails (*Helisoma trivolvis* Say, 1817, *Lymnaea elodes* Say, 1821, *Lymnaea stagnalis* Linnaeus, 1758 *Bulimnea megasoma* Say, 1824, *Physa gyrina*, Say, 1821 and *Gyralus* sp. Charpentier, 1837) were collected from 14 wetlands including ponds, marshes and rivers in southern Manitoba from May-August, 2013–2015 (Supplementary Table 1.1). Field-collected snails (n=3,786) were returned to the laboratory and placed in well plates under lights to stimulate cercarial (larval stage of trematodes within first intermediate host) emergence. After two hours, snails were examined under a stereomicroscope for emerging cercariae (Detwiler and Minchella, 2009). A total of 125 snails were infected with larval

echinostomes, which were distinguished from other larval trematode morphotypes based on the presence of collar spines (Schell, 1985).

Potential muskrat and bird hosts were obtained from hunters and the Wildlife Haven Rehabilitation Centre from 16 locations in southern Manitoba from April-October, 2014-2017 (Supplementary Table 1.2). A total of 113 muskrats were necropsied because they are known definitive hosts for echinostomes (e.g. Detwiler et al. 2012). A total of 114 birds from 20 species were necropsied. Seventeen bird species were chosen because they may eat second intermediate host snails or frogs (http://www.birds.cornell.edu) and thus act as definitive hosts if they consume echinostome metacercariae. In addition, ten of these seventeen species are known definitive hosts for echinostomes (Supplementary Table 1.2). Three additional bird species, the great grey owl (*Strix nebulosa* Forster, 1772), peregrine falcon (*Falco peregrinus* Tunstall, 1771) and snowy owl (*Bubo scandiacus* Linnaeus, 1758) are not known to consume snails or frogs, but were necropsied because they had previously been assessed for echinostomes (Taft, 1993).

The small and large intestines and caecae of vertebrate hosts were split longitudinally and placed in separate petri dishes. We searched for echinostome parasites (presence of collar spines) by systematically scraping debris and other contents from the tissues under a stereomicroscope. All larval and adult echinostomes were preserved in 70% ethanol and stored at 4°C for molecular and morphological analysis. From each host, 1-3 larval and adult individuals were haphazardly selected for DNA extraction.

1.2.2. Gene sequencing

For DNA extractions, a 1 mm^2 piece of the posterior end of the adult body was removed and soaked in water to rinse out the ethanol. For rediae, the whole body was soaked in water. The piece of adult tissue or redia was then placed into a 200 µl solution of 5% chelex containing

0.2 mg/ml of proteinase K, followed by incubation for 2 hours at 56°C and boiling for eight minutes at 99°C. Extracted DNA samples were stored at -20°C before running polymerase chain reaction (PCR). Genetic identification was based on amplification of partial fragments of the ND1 mitochondrial gene and the 28S rRNA gene regions. These genes were used because they are useful for delineating species (especially ND1) and allow comparison of sequences to those from past studies, especially when only one of the two gene regions was used (e.g. Detwiler et al., 2012; Pulis et al., 2011). PCR was performed using a 25 µl reaction mix containing 2 µl of extracted DNA, 1x buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.4 mM of each primer, and 0.05 unit/µl Taq polymerase to amplify partial fragment of ND1. JB11 (5′-AGA TTC GTA AGG GGC CTA ATA-3′) and JB12 (5′-ACC ACT AAC TAA TTC ACT TTC-3′) were the forward and reverse primers, respectively (Morgan and Blair, 1998b). The thermocycling profile began with an initial denaturation at 95°C for 3 min, followed by 35 cycles with 1-minute denaturation at 94°C, 45 s for primer annealing at 52°C, and 2 min for primer extension at 72°C, with a final extension step at 72°C for 7 min. To amplify the 28S region, the same PCR reaction mixture was used except for the primers: ZX-1F (5′-ACC CGC TGA ATT TAA GCA TAT-3′) (Bray et al., 2009) and LSU-5 (5′-TAG GTC GAC CCGCTG AAY TTA AGC A-3′) as forward and 1500-R (5′-GCT ATC CTG AGG GAA ACT TCG- 3′) as reverse (Tkach et al., 2003). The PCR profile was the same as used for ND1 except for 50°C annealing temperature.

All PCR products were visualized in a 2% agarose gel containing ethidium bromide and cleaned with the UltraClean PCR clean-up kit (MO BIO Laboratories, Inc). Sequencing was completed in the forward and reverse direction at the Hospital for Sick Children, Toronto, ON using an ABI 3730XL instrument.

1.2.3. Sequence alignment and analysis

ND1 and 28S contigs were formed in Sequencher™ software (GeneCodes Corp., ver. 4.1.4) and edited by eye. Our sequences were combined with those from GenBank (149 for ND1 and 24 for 28S) (Supplementary Table 1.3) which included five genera and 21 species for ND1 and four genera and 17 species for 28S sequences isolated from North America, Europe, and Australia. In addition, four sequences were included that originated from echinostomes collected from muskrats from New York (Detwiler, 2016 unpublished). Alignments were constructed with ClustalW in MEGA 7.0 (Kumar et al., 2016). I determined haplotype similarity using DNAsp v5 (Librado and Rozas, 2009) and constructed an alignment that included only unique haplotypes for phylogenetic analysis. The best fit model of nucleotide substitution was chosen based on Aikaike Information Criterion (Posada and Crandall, 1998) using jModeltest 2.1.6 (Darriba et al., 2012). The best models of molecular evolution were General Time Reversible model with invariant sites and gamma distribution for among site rate variation (GTR+I+G) for ND1 and General Time Reversible and gamma distribution (GTR+G) for the 28S dataset.

Bayesian and maximum likelihood methods were used to reconstruct phylogenetic trees. Bayesian analysis was performed in MrBayes v.3.1.1 (Ronquist et al., 2012). Two independent analyses were run with four heated chains for 10,000,000 generations with sampling frequency of 1000. The best consensus tree was found after a burnin of 1000. Maximum likelihood analysis was performed using RaxML 8.2.10 (Stamatakis, 2014) in the Cipres portal. I specified GTRCAT for the bootstrap model. Outgroups were *Fasciola hepatica* Linnaeus, 1758 (Morgan and Blair, 1998b) and *Isthmiophora melis* (Olson et al., 2003) for ND1 and 28S phylogenetic trees, respectively.

Following phylogenetic analysis, I estimated genetic distances between and within clades to confirm distinct species of echinostome isolates. Uncorrected p-distances were calculated with

MEGA 7.0. To visualize the genetic variation within monophyletic groups and to demonstrate gene flow between different geographic populations, I constructed haplotype networks of each species using TCS1.21 (Clement et al., 2000).

Putative species were monophyletic groups with >70 bootstrap and >.70 support values in Maximum Likelihood and Bayesian trees, respectively. A genetic benchmark of >5% difference at the ND1 gene was interpreted as evidence of distinct species (Vilas et al., 2005; Georgieva et al., 2014). Any base pair differences at the 28S gene were treated as notable as some species only differ by 1 base pair (Holland, 2010; Pulis et al., 2011). Unconnected haplotype networks (>9 base pair differences) corroborated the delineation of lineages/species.

1.3. Results

Five of six freshwater snail species were infected with echinostome larvae. Sampling effort varied per site and often targeted particular species, thus prevalence of echinostomes was not reported because it is likely not ecologically meaningful. Prevalence of echinostome infection was 52% (59/113) in muskrats and 25% (29/114) in birds. We do not report prevalence for each bird species because sample size per species was small (<10 individuals) and even smaller if reported by collection site. For the mitochondrial gene, I generated 198 ND1 sequences (each 474 base pairs) including 66 from 51 snail hosts, 36 from 18 bird hosts, and 96 from 46 muskrat hosts. The ND1 alignment had 231 unique haplotypes with 76 novel haplotypes from my study.

Seventeen ND1 sequences from the present study exactly matched GenBank sequences from nine hosts (eight snails and 1 muskrat) in two geographic locations (Indiana and Virginia in the USA). For the 28S nuclear gene region, 30 sequences (each 833 base pairs) were obtained: ten from snail hosts, 11 from bird hosts and nine from muskrat hosts. The 28S alignment

consisted of 26 unique haplotypes with 12 novel haplotypes from our study. Three sequences from my study exactly matched one GenBank *Ep. rubrum* sequence (JF820595) collected from metacercariae infecting wood frogs (*Lythobates sylvaticus* LeConte, 1825) in North Dakota.

The topology of Bayesian and Maximum likelihood ND1 and 28S consensus gene trees were largely similar but differed in the relationships among some non-North American taxa. However, since the goal of my study was not to elucidate phylogenetic relationships of non-North American taxa, I will not discuss these differences here. The Bayesian and Maximum likelihood trees for ND1 and 28S identified three distinct echinostome genera in Manitoba: *Echinostoma, Echinoparyphium* and *Hypoderaeum* (Figure 1.1; Figure 1.2).

The ND1 tree includes five groups that were previously genetically recognized lineages/species: *E. trivolvis* lineage a, *E. trivolvis* lineage c, *E. revolutum* sensu lato, *Ep.* lineage 2 and *Hypoderaeum* lineage 1 (Detwiler et al., 2010). I also found six novel lineages including a new *Echinostoma* sp., two new *Ep.* lineages 3a and 3b and three species of echinostomes whose generic designation are unclear (Figure 1.1). These lineages are considered novel because there were no matching GenBank sequences (as of 30 May 2018). Seven of eleven lineages/species recognized in the ND1 tree were monophyletic groups in the 28S tree including *E. trivolvis* lineage a, *E. trivolvis* lineage c, *E. revolutum* sensu lato, *Echinostoma* n. sp., and three new echinostome spp. (Figure 1.2). I did not successfully generate 28S sequences for several lineages that were identified based on the ND1 tree including *Ep*. lineage 2, *Ep.* lineages 3a and 3b, and *Hypoderaeum* lineage 1. Overall, ND1 sequencing suggested two cryptic species complexes (*E. trivolvis* and *Ep*. lineage 3) and 28S data further supported that one of these complexes, *E. trivolvis*, was cryptic.

For the hypothesized *E. trivolvis* cryptic species complex, phylogenetic analysis suggested two well-supported subclades for lineage a and lineages b and c combined (Figure 1.1). Similarly, haplotype network analysis showed lineage a as a distinct network and lineages b and c as one network with the lineages differing by nine mutational steps (Figure 1.3). Genetic p-distance was >5% between *E. trivolvis* lineages a and b, and a and c (Table 1.2). However, *E. trivolvis* b and c lineages differed by 2.9%, which is lower than the 5% p-distance cut-off value specified by Vilas et al. (2005) for species delimitation (Table 1.2).

Sequence matching and phylogenetic analysis elucidated the first and definitive hosts for 11 lineages/species and provided genetic confirmation of host use (Table 1.1). Six haplotypes of *E. trivolvis* lineage a collected from mallards (*Anas platyrhynchos* Linnaeus, 1758), sandhill crane (*Grus canadensis* Linnaeus, 1758) and muskrat (*Ondatra zibethicus* Linnaeus, 1766) from southern Manitoba exactly matched a larval sequence from GenBank that was collected from *H. trivolvis* in Indiana, USA. To my knowledge, the sandhill crane is a novel host for the *E. trivolvis* species complex. For *E. trivolvis* lineage c, five sequences from Manitoba muskrats exactly matched a haplotype collected from a muskrat in Virginia, USA. From all the infected snails, only one individual was infected with *E. trivolvis* lineage c. The host snail was *H. trivolvis*, which was interesting as it has only been genetically confirmed from *L. elodes* collected in Minnesota, USA (Detwiler et al., 2010). Definitive host use varied among the *E. trivolvis* lineages. *Echinostoma trivolvis* lineage a used both birds and muskrats as their final hosts, while *E. trivolvis* lineage c was found only in muskrat hosts.

I also identified the first and definitive hosts of another species from North America, *E. revolutum* sensu lato, which confirmed the life cycle description by Kanev (1994). *Echinostoma revolutum* cercariae infected *L. elodes*, while the adults used birds (i.e. mallards, sandhill crane

and Canada goose (*Branta canadensis* Linnaeus, 1758)) (Table 1. 1). Two sequences from mallards exactly matched two larval sequences of *E. revolutum* sensu lato collected from *L. elodes* in Indiana, USA. More samples from North America supported the conclusion that this group is a distinct species from *E revolutum* sensu stricto, originally described from Europe (Georgieva et al., 2014). The genetic distance between the two lineages was 6.5% and 0.4% according to ND1 and 28S, respectively (Table 1.2). These two well-supported species clades also formed two unconnected haplotype networks indicating no gene flow between Europe and North America.

In the *E. robustum* species complex, my sequences clustered with *E. robustum* lineages a, b (known as *E. robustum*, Morgan and Blair study, 1998b and now as *E. miyagawai*, Georgieva et al., 2014), c and d named by Detwiler et al. (2010). Genetic divergence (Table 1.2) and maximum parsimony haplotype network (Figure 1.4A) suggest that specimens from mallards represent a novel species 37-spined *Echinostoma* that is distinct from *E. miyagawai*. I identified this species to genus based on the key by Schell (1985). Clustering of larval sequences collected from *L. elodes* in Minnesota, USA (Detwiler et al., 2010) confirms a first intermediate host for this new species and suggests a geographic range that encompasses at least central North America.

In the *Ep.* lineage 2 complex identified by Detwiler et al. (2010), there were two strongly supported subclades. ND1 sequences from my study grouped with one of the subclades (Figure 1.1). Genetic divergence between these subclades was low (p-distance 3.4%) (Table 1.2), suggesting that they are not distinct species. This is the first genetically confirmed report of *Ep.* lineage 2 from a definitive host (broad-winged hawk, *Buteo platypterus* Vieillot, 1823) and

demonstrates that they at least infect birds and confirms that *L. elodes* is a first intermediate host (Table 1.1).

Echinoparyphium lineage 3 consisted of three monophyletic groups, suggesting that it is a complex of three distinct species (Figure 1.1). Additionally, haplotype analysis yielded three unconnected networks and genetic distance was >5% (Figure 1.4B and Table 1.2). Two of these lineages, *Ep.* lineage 3a and 3b, are identified for the first time in our study (Figure 1.1). *Echinoparyphium lineage* 3b infected *H. trivolvis* and was recorded from a novel first intermediate host, *P. gyrina*. Moreover, this is the first report of a definitive host (mallard) for the *Ep. lineage* 3 species complex (Table 1.1). I also report the first genetically-verified occurrence of *Hypoderaeum* lineage 1 in birds (mallard) in North America (Figure 1.1, Table 1.1). Larval isolates of *Hypoderaeum* lineages were previously identified by Detwiler et al. (2010) in *L. elodes*.

In some cases, larval sequences were not matched to any known adult sequences. I discovered a singleton and two well supported monophyletic groups of new echinostome spp. in Manitoba infecting *L. elodes* and *B. megasoma* as their first intermediate snail hosts, but their final host is unknown. In some cases, I was not able to sequence echinostomes successfully with ND1 sequences at the 28S gene and vice versa. For instance, three larval 28S sequences from our study exactly matched *Echinoparyphium rubrum* (Cort, 1914) metacercariae from frogs (JF820595, Pulis et al., 2011). This confirms that genetically identified *Ep. rubrum* uses *L. elodes* as a first host. This is a novel host record because only *Physa* spp. were identified as the first intermediate hosts in the original description. I did not obtain ND1 sequences for these samples, so it is unclear if our ND1 tree includes *Ep. rubrum*. It is possible that some of the lineages represented by only larval samples could be *Ep. rubrum*. The only lineage that shares

the same number of spines (43) as *Ep. rubrum* is *Echinoparyphium* lineage 1, which was not observed in Manitoba.

1.4. Discussion

In southern Manitoba, eleven lineages/species of echinostomes were identified from freshwater snails, birds and muskrats by integrating phylogenetic and network analyses and measuring inter and intraspecific genetic distance. Five of these previously recognized species were documented in central-midwestern USA (Detwiler et al., 2010). However, six were newly discovered, including *Ep*. lineages 3a and b, one *Echinostoma* sp. and three echinostome species. My results demonstrate that echinostome species diversity was underestimated due to the presence of crypsis and that wider host and geographic sampling revealed more diversity.

For the *E. trivolvis* cryptic complex, phylogenetic and network analysis confirmed the existence of three unique lineages that were characterized as cryptic lineages in Detwiler et al. (2010). Prior to this study, there were few haplotypes for *E. trivolvis* lineage c, so the taxonomic status of this lineage, especially relative to lineage b, was unclear. By sampling in Manitoba, I collected more *E. trivolvis* lineage c haplotypes, and subsequent genetic analysis showed that *E. trivolvis* lineage a is a distinct species from lineages b and c. The species status of lineages b and c is harder to interpret with only genetic data because the lineages were monophyletic in only one of the two gene trees and they differed by approximately 3% at the ND1 gene. Also, I was only able to generate 28S sequences successfully for lineages a and c. Nuclear sequences for lineage b and detailed morphological analysis are needed to determine whether these lineages are distinct species or populations.

I showed that crypsis can make current descriptions of life cycles inaccurate because host specificity for either the definitive or first intermediate host is incorrect. The *E. trivolvis* lineages

differed in definitive host specificity compared to the original description. According to Kanev et al. (1995), *E. trivolvis* infects birds and mammals as final hosts. In Manitoba *E. trivolvis* lineage c was not found in birds, suggesting that it may be restricted to muskrats. *Echinostoma trivolvis* lineage a infects mammals and birds in Manitoba, which matches the species description and suggests that perhaps lineage a represents the nominal species, *E. trivolvis*, as described by Kanev et al. (1995). I did not find any *E. trivolvis* lineage b in Manitoba, so to date, it has only been observed in muskrats in the eastern USA (Detwiler et al., 2012). More sampling of bird hosts are needed from additional geographic locations to confirm whether the lineages differ in their mammal and bird hosts use.

These cryptic lineages also differed based on their first intermediate host specificity. One paradigm about trematode taxonomy is that parasites are host specific to a first intermediate host species, meaning 1 parasite:1 host species (McCarthy, 1990). *Echinostoma trivolvis* lineage a was collected from mostly *H. trivolvis* snails although one *H. trivolvis* was infected with the *E. trivolvis* lineage c. Therefore, one trematode lineage may infect more than one first intermediate host species though it is primarily found in one first intermediate host species. In the USA, *E. trivolvis* lineage c was collected from *L. elodes* in Minnesota (Detwiler et al., 2010), yet despite sampling 1,641 *L. elodes* from several wetlands in Manitoba, I never recovered lineage c from this snail species.

Other echinostome species also infected more than one species of first intermediate host. In addition to *E. trivolvis* lineage c, *Ep*. lineage 3b infected two different snail host species, *H. trivolvis* and *P. gyrina* in Manitoba. In the USA, *E. revolutum* sensu lato was observed in *H. trivolivis* and *L. elodes*. In Manitoba, I only recovered this parasite from *L. elodes* suggesting that this is the primary first intermediate host. I found *Ep. rubrum* in *L. elodes*, while Kanev et al.

(1998) stated that it only uses *Physa* spp. as the first intermediate host. Either *Ep. rubrum* uses more than one first host, or the species description (for which there is no genetic voucher) is based on a different species than referred to in GenBank by Pulis et al. (2011).

In contrast, several lineages/species were found in one first host snail species as predicted. By linking larval and adult sequences for these echinostomes, I was able to elucidate the first and definitive hosts involved in their life cycle. The newly recognized *Echinostoma* n. sp., infects *L. elodes* as a first intermediate host and mallards as the final host. *Ep.* lineage 2 infects *L. elodes* and broad-winged hawk as a definitive host (first genetically confirmed definitive host). *Hypoderaeum* lineage 1 was found in *L. elodes* first intermediate hosts and mallards as the final host (first genetically confirmed definitive host). The other three new echinostome spp. such as Echinostome n. sp. 2 and 3 also used *L. elodes*, but their definitive hosts are unknown. The number of echinostome lineages/species that use *L. elodes* is large, suggesting that researchers need to be careful when identifying echinostomes from this species.

By sampling more hosts in an area predicted to have high echinostome biodiversity, I was able to confirm host specificity genetically for several echinostome species. Many of the previous DNA sequences of North American echinostomes come from first intermediate host snails and definitive host muskrats sampled in the USA. For the first time, I genetically confirmed that *E. trivolvis* lineage a was found in wild birds (mallards and sandhill cranes). The use of birds potentially explains why the same haplotypes are found throughout North America. In contrast, I did not find *E. trivolvis* lineage b or c in birds and only found the latter lineage in muskrats in Manitoba. All three *E. trivolvis* lineages were found in muskrats in Virginia, although most parasites in that study were *E. trivolvis* b (Detwiler et al., 2012). These two studies suggest that the distributions of the lineages may depend on geography. However, the role of

bird hosts behind gene flow is still unclear. Future studies require broader snail and definitive host sampling in North America to understand the role of first intermediate host specificity and host ecology on speciation.

Birds had greater echinostome species diversity than muskrats. Although a greater number of bird hosts and species were sampled compared to one mammal species, there were five lineages in one species of bird. Mallards were infected with *E. trivolvis* lineage a, *E. revolutum* sensu lato, *Ep*. lineage 3a, *Hypoderaeum* lineage 1 and one new *Echinostoma* sp., while muskrat hosts had only two lineages, *E. trivolvis* lineage a and *E. trivolvis* lineage c. Higher echinostome diversity in the mallard could be the result of acquiring infections by feeding in wetlands across their comparatively larger home range and during their continental-scaled migration. Additionally, differences in bird and mammal host physiology might play an important role in the differences in host specificity (Chowdhury et al., 1994).

My study suggests that several species may not have cosmopolitan distributions as previously described and should be described as new species. My analyses support the distinct species status of *E. revolutum* from Europe (sensu stricto) and North America (sensu lato) and means that *E. revolutum* sensu lato needs to be described formally. Further, I found that the European species, *E. miyagawai* (called *E. robustum* by Kostadinova et al., 2003 and Detwiler et al., 2010; 2012), is genetically distinct from specimens in North America that are part of the *E. robustum* cryptic species complex (Detwiler et al., 2010). More sequences from North and South America might help resolve the taxonomy of the remaining two singletons of the previously called *E. robustum* complex (Detwiler et al., 2010).

Although sampling in Manitoba has improved our knowledge of the biodiversity, host specificity and life cycles of echinostomes in North America, many questions remain. Several

species in North America were described without genetic vouchers, such as *Ep. rubrum* and *Echinoparyphium flexum* Linton, 1892. It would be useful to return to their type localities to obtain genetic sequences and confirm their life cycles.

1.5. Conclusion

My study is the first in North America to sequence echinostomes from a broad range of first intermediate and definitive hosts from the same geographic region. As predicted, I found that echinostome species diversity was previously underestimated due to the presence of cryptic complexes. ND1 sequences reliably inferred species boundaries and were corroborated by differences in the 28S gene region for eight species. This study improves our knowledge of the geographic distribution, host specificity and life cycles of echinostomes. Better estimates of life cycles will improve our understanding of the epidemiology of echinostomiasis in wildlife and help determine their potential role as causative agents of emerging disease.

References

Bondar, G. F. 1950. A biological survey of the muskrat, (*Ondatra zibethica alba*) (Sabine) Miller, 1912, in Manitoba. M.Sc. thesis, University of Manitoba, Winnipeg, Manitoba.

Bolek, M., H. Stigge, and K. Gustafson. 2016. The iron wheel of parasite life cycles: then and now! A century of parasitology: discoveries, ideas and lessons learned by scientists who published in The Journal of Parasitology, 1914–2014. 1st Edition. Wiley press. p 131–147.

Bray, R. A., A. Waeschenbach, T. H. Cribb, G. D. Weedall, P. Dyal, and D. T. Littlewood. 2009. The phylogeny of the Lepocreadioidea (Platyhelminthes, Digenea) inferred from nuclear and mitochondrial genes: implications for their systematics and evolution. Acta Parasitologica. 54: 310–329.

Broderson, D., A.G. Canaris and J.R Bristol. 1977. Parasites of waterfowl from Southwest Texas: II: The shoveler, *Anas clypeata*. Journal of Wildlife. Diseases. 13: 435-439.

Bush, A.O. and J.C. Holmes. 1986. Intestinal helminths of lesser scaup ducks: Patterns of association. Canadian Journal of Zoology. 64:132–141.

Chowdhury, N., M. L. Sood, and R. T. O'Grady. 1994 Evolution, parasitism and host specificity in helminths, in N. Chowdhury and I. Tada (eds.). Helminthology. Springer-Verlag, Narosa Publishing House, New York, pp. 1-19.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology. 9: 1657–1661.

Darriba, D, G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: More models, new heuristics and parallel computing. Nature Methods. 9: 772.

Davies, N. E. 2006. A survey of waterfowl for echinostomes and schistosomes from Lake Wanaka and the Waitaki River watershed, New Zealand. Journal of Helminthology. 80: 33-40.

Detwiler, J. T., D. H. Bos, and D. J. Minchella. 2010. Revealing the secret lives of cryptic species: Examining the phylogenetic relationships of echinostome parasites in North America. Molecular Phylogenetics and Evolution. 55: 611–620.

Detwiler, J. T. and D. J. Minchella. 2009. Intermediate host availability masks the strength of experimentally-derived colonisation patterns in echinostome trematodes. International Journal for Parasitology. 39: 585–590.

Detwiler, J. T., A. M. Zajac, D. J. Minchella, and L. K. Belden. 2012. Revealing cryptic parasite diversity in a definitive host: echinostomes in muskrats. Journal of Parasitology. 98: 1148–1155.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. Bryologist. 88: 172–181.

Georgieva, S., A. Faltýnková, R. Brown, I. Blasco-Costa, M.Soldánová, J. Sitko, and A. Kostadinova. 2014. *Echinostoma 'revolutum*'(Digenea: Echinostomatidae) species complex revisited: species delimitation based on novel molecular and morphological data gathered in Europe. Parasites & Vectors. 6: 520–537.

Georgieva, S., C. Selbach, A. Faltýnková, M. Soldánová, B. Sures, K. Skírnisson, and A. Kostadinova. 2013. New cryptic species of the 'revolutum' group of *Echinostoma* (Digenea: Echinostomatidae) revealed by molecular and morphological data. Parasites & Vectors. 7: 64–77.

Holland, M. P. 2010. Echinostome-induced mortality varies across amphibian species in the field. Journal of Parasitology. 96: 851–855.

Jones, A. 2005. Introduction and key to superfamilies. In: Jones, A., Bray, R.A., Gibson, D.I. (Eds.), Keys to Trematoda, vol. 2. CAB International and the Natural History Museum, London, p. 1–64.

Kanev, I. 1994. Life-cycle, delimitation and redescription of *Echinostoma revolutum* (Frölich, 1802) (Trematoda: Echinostomatidae). Systematic Parasitology. 28: 125–144.

Kanev, I., B. Fried, V. Dimitrov, and V. Radev. 1995. Redescription of *Echinostoma trivolvis* (Cort, 1914) (Trematoda: Echinostomatidae) with a discussion on its identity. Systematic Parasitology. 32: 61–70.

Kanev, I., R. Sorensen, and M. Sterner. 1998. The identification and characteristics of *Echinoparyphium rubrum* (Cort, 1914) comb. new (Trematoda, Echinostomatidae) based on experimental evidence of the life cycle. Acta Parasitologica. 43: 181–188.

Kinsella J. M., G. W. Foster and D. J. Forrester. 1995. Parasitic Helminths of Six Species of Hawks and Falcons in Florida. Journal of Raptor Research. 29: 117-122.

Kostadinova, A. 1999. Cercarial chaetotaxy of *Echinostoma miyagawai* Ishii, 1932 (Digenea: Echinostomatidae), with a review of the sensory patterns in the '*revolutum*' group. Systematic Parasitology. 44: 201–109.

Kostadinova, A. 2005. Family Echinostomatidae Looss, 1899. In Keys to the Trematoda. CABI Publishing and The Natural History Museum, London, U.K p. 9–64.

Kostadinova A. and Gibson D.I. 2000. The Systematics of the Echinostomes. In: Fried B., Graczyk T.K. (eds.) Echinostomes as Experimental Models for Biological Research. Springer, Dordrecht. p. 31-57.

Kostadinova, A., G. A. Herniou, J. Barrett, and D. T. L. Littlewood. 2003. Phylogenetic relationships of *Echinostoma* Rudolphi, 1809 (Digenea: Echinostomatidae) and related genera re-assessed via DNA and morphological analyses. Systematic Parasitology. 54: 159–176.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 33: 1870-1874.

Librado, P., & Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25: 1451-1452.

McKenzie, C. E. and H. E. Welch. 1979. Parasite fauna of the muskrat, *Ondatra zibethicus* (Linnaeus, 1766) in Manitoba, Canada. Canadian Journal of Zoology. 57: 640–646.

McCarthy, A.M. 1990. Speciation of echinostomes: evidence for the existence of two sympatric sibling species in the complex *Echinoparyphium recurvatum* (von Linstow 1873) (Digenea: Echinostomatidae). Parasitology. 101: 35–42.

McLaughlin, J.D. and Burt, M.D. 1979. A survey on the intestinal helminth of water fowl from New-Brunswick, Canada. Canadian Journal of Zoology. 57: 801-807.

Morgan, J. A. T. and D. Blair. 1998a. Relative merits of nuclear ribosomal internal transcribed spacers and mitochondrial CO1 and ND1 genes for distinguishing among *Echinostoma* species (Trematoda). Parasitology. 116: 289–297.

Morgan, J. A. T. and D. Blair. 1998b. Mitochondrial ND1 gene sequences used to identify echinostome isolates from Australia and New Zealand. International Journal for Parasitology. 28: 493–502.

Nolan, M. J. and T. H. Cribb. 2005. Use and implications of ribosomal DNA sequencing for the discrimination of digenean species. Advances in Parasitology. 60: 101–163.

Olson, P. D., T. H. Cribb, V. V. Tkach, R. A. Bray, and D. T. J. Littlewood. 2003. Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). International Journal for Parasitology. 33: 733–755.

Pérez-Ponce De León, G. and R. Poulin. 2018. An updated look at the uneven distribution of cryptic diversity among parasitic helminths. Journal of Helminthology. 92: 1–6.

Posada, D. and K. Crandall. 1998. Modeltest: testing the of DNA substitution. Bioinformatics. 14: 817–818.

Poulin, R. 2007. Evolutionary Ecology of Parasites. Chapman and Hall, London, UK. p. 360.

Poulin, R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biology Letters.7: 241–244.

Pulis, E., V. Tkach, and R. Newman. 2011. Helminth parasites of the wood frog, *Lithobates sylvaticus*, in prairie pothole wetlands of the northern Great Plains. Wetlands. 31: 675–685.

Roberts, L. S., J. Janovy, and G. D. Schimdt. 2009. Gerald D. Schimdt and Larry S Roberts' Foundations of Parasitology, Boston McGraw-Hill. pp. 701.

Ronquist, F., M. Teslenko, P. Van Der Mark, D. L. Ayres, A. Darling, S. Höhna, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology. 61: 539–542.

Rosen, D. E. 1979. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. Bulletin of the American Museum of Natural History. 162: 269–375.

Rothenburger, J. L., E. Hoberg, and B. Wagner. 2016. First Report of *Protechinostoma* mucronisertulatum (Echinostomatidae) in a Sandhill Crane (*Grus canadensis*) from Saskatchewan, Canada. Comparative Parasitology. 83: 111–116.

Schell, S. C. 1985. Trematodes of North America. University of Idaho Press, Moscow, Idaho, pp. 263.

Shaw M. G. and Alan, K. 1980. Helminth fauna of waterfowl in central Oklahoma. Journal of Wildlife Diseases. 16: 59-64.

Sorensen, R. E., I. Kanev, B. Fried, and D. J. Minchella. 1997. The occurrence and identification of *Echinostoma revolutum* from North American *Lymnaea elodes* snails. Journal of Parasitology. 83: 169–170.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30: 1312–1313.

Taft, S. J., K. Suchow, and M. Van Horn. 1993. Helminths from some Minnesota and Wisconsin raptors. Journal of Helminthology Society Washington. 60:260-263.

Tkach, V. V., D. T. J. Littlewood, P. D. Olson, J. M. Kinsella, and Z. Swiderski. 2003 Molecular phylogenetic analysis of the Microphalloidea Ward, 1901 (Trematoda: Digenea). Systematic Parasitology. 56: 1–15

Vilas, R., C. D. Criscione, and M. S. Blouin. 2005. A comparison between mitochondrial DNAand the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. Parasitology. 131: 1–8.

Table 1.1. First intermediate and final hosts of genetically-identified echinostomes in Manitoba wetlands compared to other locations in North America.

Lineages/species ID	**First intermediate hosts in Manitoba**	**First intermediate hosts elsewhere[1]**	**Final hosts in Manitoba**	**Final hosts elsewhere[1]**
E. trivolvis lineage a	*Helisoma trivolvis*	*Helisoma trivolvis*	*Anas platyrhynchos* *Grus canadensis* *Ondatra zibethicus*	*Ondatra zibethicus*
E. trivolvis lineage c	*Helisoma trivolvis*	*Lymnaea elodes*	*Ondatra zibethicus*	*Ondatra zibethicus*
E. revolutum (sensu lato)	*Lymnaea elodes*	*Lymnaea elodes* *Helisoma trivolvis*	*Anas platyrhynchos* *Branta canadensis* *Grus canadensis*	*Ondatra zibethicus*
Echinostoma n. sp.	*Lymnaea elodes*	unknown[2]	*Anas platyrhynchos*	unknown
Ep. lineage 2	*Helisoma trivolvis*	*Helisoma trivolvis*	*Buteo platypterus*	unknown
Ep. lineage 3a	unknown	unknown	*Anas platyrhynchos*	unknown
Ep. lineage 3b	*Helisoma trivolvis* *Physa gyrina*	*Helisoma trivolvis*	unknown	unknown
Hypoderaeum lineage 1	unknown	*Lymnaea elodes*	*Anas platyrhynchos*	unknown

Echinostome n. sp. 1	*Bulimnea megasoma*	unknown	unknown	unknown
Echinostome n. sp. 2	*Helisoma trivolvis*	unknown	unknown	unknown
Echinostome n. sp. 3	*Helisoma trivolvis*	unknown	unknown	unknown

[1] Collected from hosts in North America outside of Manitoba, Canada.

[2] "Unknown" indicates that, to our knowledge, these lineages are not reported from first intermediate hosts or final hosts in North America.

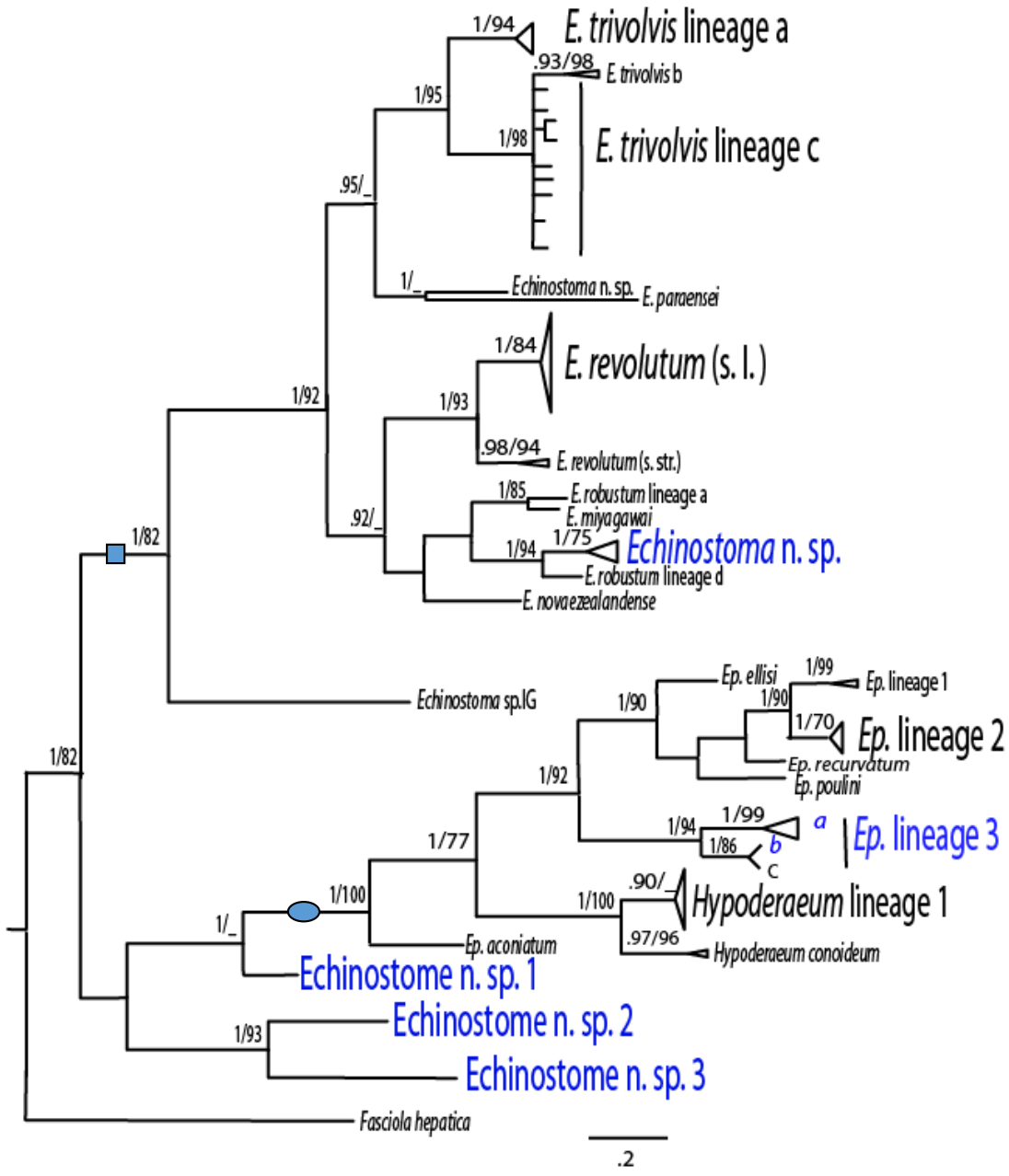
Table 1.2. Genetic distance (p-distance[1]) within and between echinostome lineages/species.

Species groups	**Gene Region**	
	ND1	**28S**
E. trivolvis		
Lineage a vs lineage b	8.1	
Lineage b vs lineage c	2.9	
Lineage a vs lineage c	8.3	0.90
E. revolutum		
E. revolutum sensu lato vs *E. revolutum* sensu stricto	6.5	0.40
E. robustum[2]		
E. robustum lineage a vs *E. miyagawai*	3.1	
*E. miyagawa*i vs *Echinostoma* n. sp.	6.8	0.12
E. miyagawai vs *E. robustum* lineage d	6.4	
Echinoparyphium lineage 2[3]	3.4	
Echinoparyphium lineage 3		
Ep. lineage 3a vs *Ep*. lineage 3b	22.8	
Ep. lineage 3b vs *Ep*. lineage 3c	15.8	
Echinostome n. sp.		
Echinostome n. sp. 1 vs Echinostome n. sp. 2	25.5	1.5
Echinostome n. sp. 2 vs Echinostome n. sp. 3	22.4	0.6
Echinostome n. sp. 1 vs Echinostome n. sp. 3	23.3	1.6
Hypoderaeum sp.		
Hypoderaeum lineage 1 vs *Hypoderaeum conoideum*	7.8	

[1]p-distance is the proportion of sites that differ between individuals/lineages/species. Based on 474 bp for ND1 and 833 for 28S rRNA.

[2]Nomenclature from Detwiler et al, 2010.

[3]Includes two subclades



.

Figure 1.1. Phylogenetic tree of echinostomes from Bayesian inference (BI) and Maximum likelihood (ML) analyses based on partial sequences of the nicotinamide adenine dinucleotide dehydrogenase subunit 1(ND1) gene region (474 base pairs). *Fasciola hepatica* was the outgroup. Bayesian support values and bootstrap values (shown as BI/ML) less than .70 and 70 are not shown and depicted with a dash. Black-coloured lineages were genetically identified in

previous studies and blue-coloured lineages represent newly recognized lineages in the present study. Square represents *Echinostoma* and circle represents *Echinoparyphium* and *Hypoderaeum*.

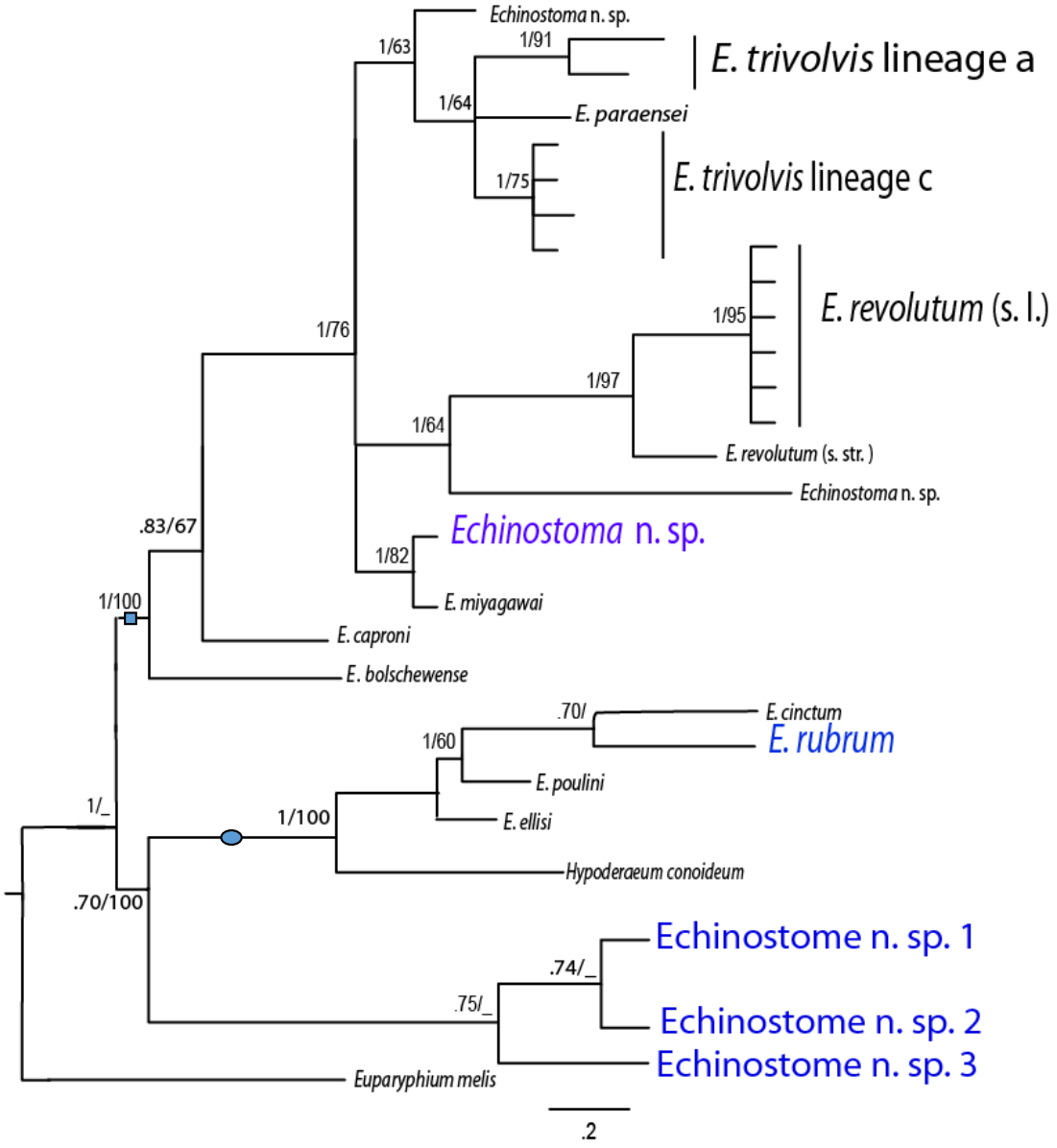


Figure 1.2. Phylogenetic tree of echinostomes from Bayesian inference (BI) and Maximum likelihood (ML) analyses based on partial sequences of the 28S gene region (833 base pairs). *Euparyphium melis* was the outgroup. Bayesian support values and bootstrap values (shown as BI/ML) less than .70 and 70 are not shown and depicted with a dash. Black-coloured lineages were genetically identified in previous studies and blue-coloured lineages represent newly

recognized lineages in the present study. Square represents *Echinostoma* and circle represents *Echinoparyphium* and *Hypoderaeum*.

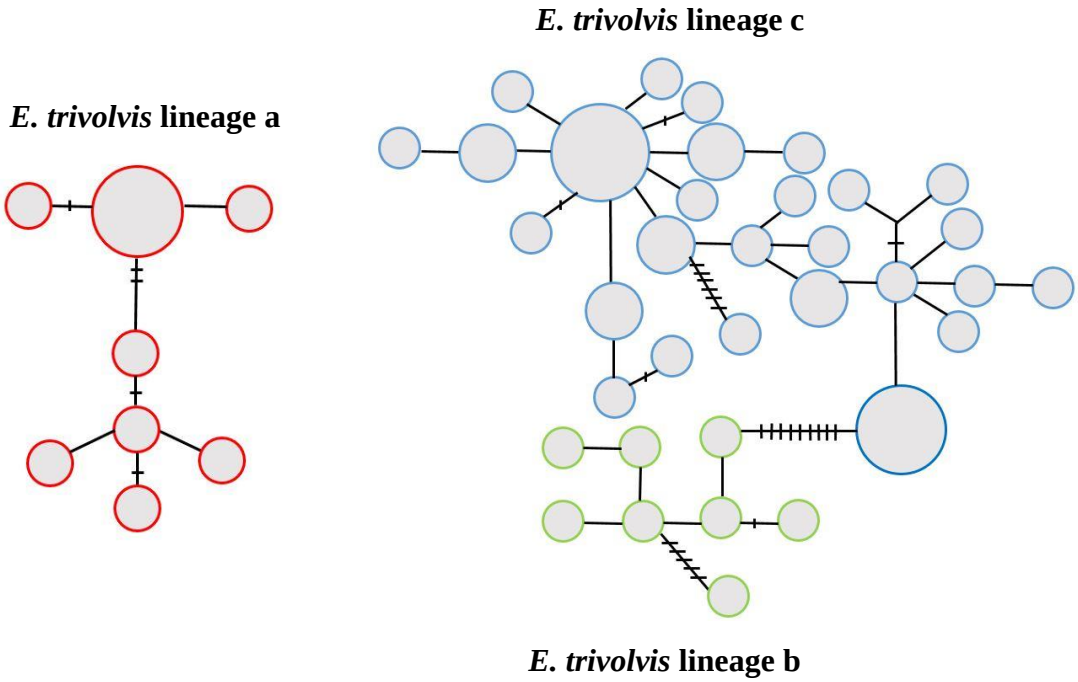


Figure 1.3. Haplotype networks of *Echinostoma trivolvis* lineages a, b and c based on partial nicotinamide adenine dinucleotide dehydrogenase subunit 1 (ND1) sequences. Haplotype frequency is represented by the diameter of the circle. Colours of the circles indicate lineages/species: red= *Echinostoma trivolvis* lineage a, green= *Echinostoma trivolvis* lineage b and blue= *Echinostoma trivolvis* lineage c.

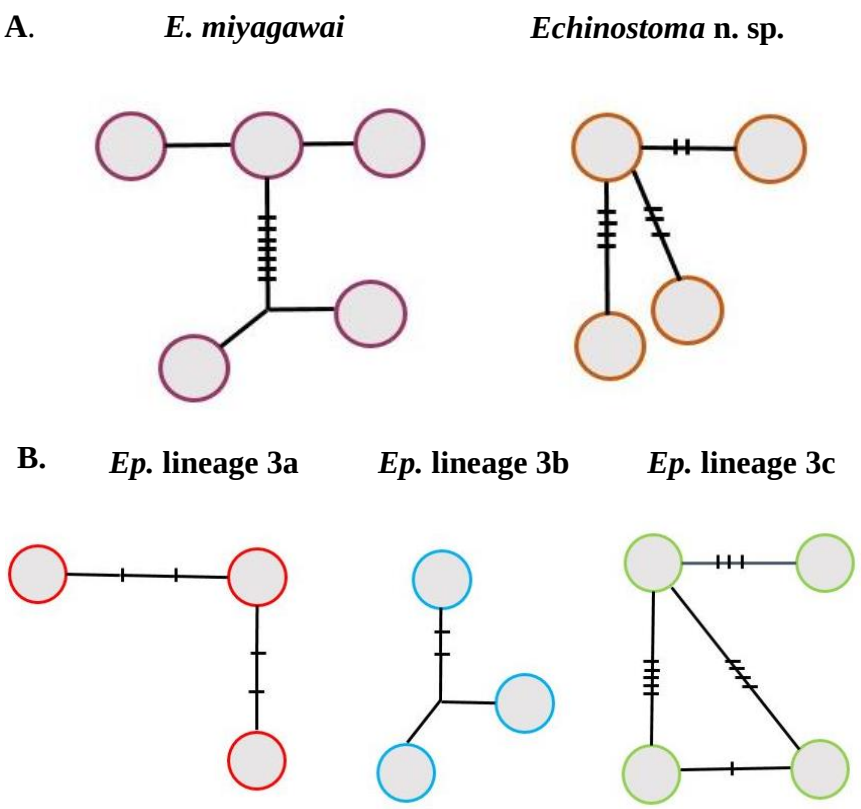


Figure 1.4. Haplotype networks of echinostomes based on partial nicotinamide adenine dinucleotide dehydrogenase subunit 1(ND1) sequences: A) *Echinostoma miyagawai* and *Echinostoma* n. sp. B) *Echinoparyphium* lineage 3 complex. *E miyagawai* haplotypes were identified from Europe (Morgan and Blair, 1998a) and *Echinostoma* n. sp. haplotypes were from my study. *Ep.* lineage 3a and *Ep.* lineage 3b were from my study and and 3c were collected in the USA (Detwiler et al. 2010). Haplotype frequency is represented by the diameter of the circle. Colours of the circles indicate lineages/species: purple= *E. miyagawai*, brown= *Echinostoma* n. sp., red= *Ep.* lineage 3a, blue= *Ep.* lineage 3b and green= *Ep.* lineage 3c.

Chapter 2: Using integrative taxonomy to test hypotheses about crypsis and host specificity of echinostome trematodes

2.1. Introduction

Host specificity is a fundamental natural life history characteristic of parasites because it influences many aspects of their biology. For instance, the range of hosts a parasite exploits influences the transmission success and patterns of co-evolution and speciation (Poulin and Mouillot, 2003). Host specificity varies widely among parasite taxa with a continuum extending from a specialist (using one host species) to a generalist (using several host species) (Poulin and Keeney, 2008). Further, host specificity can vary at each different life cycle stage so a parasite could specialize on one host species as a larva, but be a generalist as an adult (Poulin, 2011). Because most parasites are not maintained in the laboratory or easily culturable, our knowledge of host specificity originates primarily from field studies which can make it difficult to determine the factors that shape host specificity (Poulin and Keeney, 2008).

Phylogenetic, ecological and physiological factors can explain patterns of host specificity (Clark and Clegg, 2017). Some species of helminth parasites only infect closely related host species. For example, both trematodes and nematodes species such as *Crassicutis cichlasomae* Manter, 1936 and *Rhabdochona kidderi* Pearse, 1936 only infect closely related species of Cichlidae (Salgado-Maldonado et al., 2016). On the contrary, there are some parasites which exploit host species from less phylogenetically-related taxa. For instance, an acanthocephalan parasite, *Telosentis exiguous* von Linstaw, 1901 infects fish from different families (Kvach and Sasal, 2010). This pattern of infection could have arisen because the hosts shared the same ecological niche (Clark and Clegg, 2017). In other cases, neither host phylogeny nor ecology

explain infection patterns and instead, host specificity is mediated by physiological compatibility between hosts and parasite species (Locke et al., 2010).

Host specificity is a key life history trait of parasites because of its influence on parasite evolution (Dogiel et al., 1964). Opportunities for parasite dispersal and gene flow may be related to host mobility and geographic distribution (Criscione and Blouin, 2005). Generally, greater host vagility is predicted to increase parasite gene flow and reduce isolation, especially for hosts that migrate across large spatial scales. In addition, the range of hosts used may also affect transmission dynamics and parasite evolution. Infecting multiple hosts at a specific point in the life cycle may reduce the risk of extinction and increase colonization opportunities for parasites (Poulin and Keeney, 2008). However, being more of a generalist may not always increase the likelihood of parasite survival because parasites may not be equally successful in all potential hosts (i.e. achieve the same rates of growth, reproduction and survival) (Hayward, 2010). By infecting several hosts, a generalist parasite may not exploit any of the hosts at particular points in its life cycle very effectively and thus may have fewer opportunities to continue its life cycle. However, even if a parasite is better adapted to a few or a single host, being more host specific does not guarantee transmission especially if that particular host or its habitat is threatened or endangered (Koh et al., 2004).

Our ability to understand whether a species is a generalist or specialist depends on the context of the study. Typically, host specificity is estimated from field-collected hosts, which may result in inaccurate assessment of the true range of host use. Field-collections are limited in space and time so it is difficult to establish the true range of hosts a parasite uses (Poulin and Keeney, 2008). Due to economic and logistical constraints, field surveys are often spatially restricted, and usually samples are collected opportunistically rather than systematically, which

means that all potential hosts are not collected throughout the range of the parasite. This approach is problematic because there may be differences in parasite assemblages in a particular host based on geographic location or habitat type. For instance, parasite communities in muskrats from marshes were different than those inhabiting rivers (Anderson and Beaudoin, 1966). Often host sample sizes, especially for large vertebrates, is small (<50 individuals per species) and limited to a particular geographic region making it difficult to assess whether host phylogeny, ecology or physiology explains patterns of host specificity (Poulin and Keeney, 2008).

In addition to field-sampling issues, host specificity may not be well characterized because of inaccurate identification of parasites. Historically, parasites were recognized and classified by their morphological characteristics, but sole reliance on morphological traits has been considered unreliable for quite some time for some parasite taxa (Beaver, 1937). The inability to identify parasites confidently is problematic, especially when closely related and morphologically similar parasite species co-occur in the same hosts. More recently, genetic sequencing has confirmed that morphologically-identified helminths have been misidentified leading to confusion as published papers, genetic databases and museums often have incorrectly identified species (e.g. Terry and Gardener, 2008; Detwiler et al., 2010; Georgieva et al., 2014;). Misidentifications can also occur because cryptic species may have been sampled (morphologically indistinguishable, but genetically distinct) (Kostadinova et al., 2003). Recent genetic analyses have indicated that there are cryptic species in most helminth groups, but that trematodes have the highest number of cryptic taxa (Poulin, 2011; Pérez-Ponce de León and Poulin, 2018). Within a cryptic species complex, it is unclear if all species use the host range as described for the nominal species, or if cryptic species differ in the range of hosts they infect.

Distinct genetic lineages within an apparently generalist cryptic species complex can differ in their host specificity (Poulin and Keeney, 2008).

Echinostome trematodes can be used to test how cryptic species influence current estimates of host specificity. Several cryptic species complexes occur in Europe and have been proposed in North America (Detwiler et al., 2010, 2012; Georgieva et al., 2014). One of those proposed complexes, *Echinostoma trivolvis* Cort, 1914, consists of three genetically distinct lineages (a, b, c) (Detwiler et al. 2010; Detwiler et al. 2012). All three lineages have been genetically identified from muskrats in Virginia, USA. However, the nominal species, *E. trivolvis*, has been recorded from 29 field-infected host species which includes muskrats and birds (Fried and Graczyk, 2000).

Many field-collected echinostomes have not been deposited in museums (e.g. stained permanent slides or genetic specimens preserved in ethanol), so using integrative taxonomy to verify their identifications is impossible. This is problematic, as *E. trivolvis* is also a member of the 37-spined complex 'revolutum' group which includes a co-occurring, morphologically similar sister species *Echinostoma revolutum* (Frölich, 1802) Looss, 1899 sensu lato (Fried and Toledo, 2004). Although *E. revolutum* is described as parasitizing only birds (Kanev, 1994; Sorensen et al., 1997), it has been identified from muskrats in past field surveys using morphology (e.g. McKenzie and Welch, 1979). The only study to genetically confirm *E. revolutum* in muskrats was Detwiler et al. (2012). This genetic study makes it clear that these two species co-occur in muskrat hosts. However, the range of host specificity for birds has not been confirmed with DNA sequencing. Due to the morphological similarity of *E. trivolvis* and *E. revolutum* sensu lato, an integrative approach (using morphology and genetics) should be used to identify these two species. New collections will be required to test whether *E. trivolvis* lineages

are cryptic and to determine whether *E. trivolvis* and *E. revolutum* differ morphologically. Using this information will help to determine whether the cryptic echinostome lineages and species differ in their host specificity, particularly for bird definitive hosts.

To address these problems, I performed integrative taxonomy on adult echinostomes collected from naturally-infected bird and mammal hosts in Manitoba wetlands. After sequencing individuals at a mitochondrial and nuclear locus (see Section 1.2.2), I measured morphological features and used multivariate analysis to determine the intra lineage/species and inter lineage/species variation of *E. trivolvis* lineages a and c, and *E. revolutum* sensu lato. Based on these results, I recommend some morphological traits that can be used to distinguish between cryptic lineages and species.

2.2. Materials and Methods

2.2.1. Specimen collection

We collected 114 individuals of 20 bird species and 113 muskrats from 16 wetlands in southern Manitoba from licensed hunter and trappers. In most instances, the carcasses were frozen, so the parasites were dead upon necropsy. Host intestinal tracts were systematically necropsied as described in section 1.2.1. Initially, adults were identified as echinostomes by the presence of collar spines (Schell, 1985) and preserved in 70% ethanol and stored at -20 °C for molecular and morphological analysis.

2.2.2. Species delimitation by gene sequencing

DNA was extracted from haphazardly selected adult echinostomes (1-3 parasites per host individual) collected from naturally infected birds and muskrats, respectively. I attempted to sequence each individual at a mitochondrial and nuclear locus: Nicotinamide adenine dinucleotide dehydrogenase subunit 1 (ND1) mitochondrial and 28S rRNA gene regions.

Bayesian and Maximum likelihood analyses were performed to identify specimens. I used three lines of evidence to identify species: 1) clades with >70 bootstrap or nodal support values in gene trees, 2) haplotype networks to determine if each monophyletic group consisted of one network, 3) genetic distance between and within subclades to delimit echinostome species boundary (Section 1.2.3 for more details).

2.2.3. Permanent specimen preparation

Before a piece of the adult was removed for DNA extraction, a digital image was captured of each adult parasite to assess total body length using an Axio Cam ICc1 digital camera connected to a Stereo Discovery V8 microscope (Zeiss). From these images, certain traits such as Body length (BL), Body width (BW) and Hindbody length (HL) were measured (in micrometres) using Axio Vision 3.1 software.

A total of 48 genetically identified adult echinostomes from naturally infected birds and muskrats were permanently mounted on slides. I used Semichon's acetocarmine for 30 minutes to stain the specimens and then de-stained with concentrated hydrochloric acid (HCl) until the organs were clearly distinguishable from the surrounding tissue (1-2 minutes). Then the specimens were dehydrated using 70%, 85%, 90%, 95% and 100% ethanol, each for one hour. The adult worms were cleared in increasing concentrations of xylene (¾ 100% ethanol and ¼ xylene, ½ ethanol, ½ xylene; ¼ ethanol, ¾ xylene and finally 100% xylene), each for 30 minutes. Parasites were mounted using Canada balsam and glass cover slips.

2.2.4. Morphological measurement

Each stained worm was measured at 27 morphological features. I also measured 21 museum specimens from Kanev (1994) and Kanev et al. (1995) that were deposited in the Harold W. Manter Laboratory of Parasitology (HWML), Lincoln, Nebraska, USA. I could not

genetically identify the museum specimens as these were permanently mounted and had not been sequenced.

Images of the adult worms were taken with the Axio Cam ICc1 digital camera connected to Stereo Discovery. V8 microscope and Axio Imager M2 compound microscope. (Zeiss). From these images, I measured the following phenotypic traits (in micrometres) using Axio Vision 3.1 software: 1. Body length (BL); 2. Forebody length (FBL); 3. Head collar length (CL); 4. Collar spine number (CS); 5. Oral sucker length (OSL); 6. Oral sucker width (OSW); 7. Prepharynx length (PRL); 8. Pharynx length (PHL); 9. Pharynx width (PHW); 10. Oesophagus length (OL); 11. Acetabulum length (ACL); 12. Acetabulum width (ACW); 13. Ovary length (OVL); 14. Ovary width (OVW); 15. Ova maximum length (OMXL); 16. Ova maximum width (OMXW); 17. Ova minimum length (OMNL); 18. Ova minimum width (OMNW); 19. Cirrus sac length (CSL); 20. Cirrus sac width (CSW); 21. Anterior testis length (ATL); 22. Anterior testis width (ATW); 23. Posterior testis length (PTL); 24. Posterior testis width (PTW); 25. Inter-testis length (ITL); 26. Seminal vesicle length (SVL); 27. Space between anterior testis and ovary (SPATO). Example images demonstrating some of the measurements are shown in Figure 2.1.

2.2.5. Multivariate analysis

Morphological measurements included in the statistical analysis only came from the stained specimens rather than the digital images. In some preliminary analysis, I observed that Body length (BL), Body width (BW) and Hindbody length (HL), differed significantly between non-stained and stained adult worms (data not shown). To determine this, I imaged ten unstained adult worms and then stained them according to the protocol in Section 2.2.3. Then the three traits were measured from the pre- and post-staining images. Mean differences ± standard error for Body length, Body width and Hind body length were 145 ± .383.6, 126 ± 56 and 15 ± 280

(measurements in micrometres). These results demonstrate that these traits either expanded or shrank during the fixation process meaning pre-stained measurements should not be combined with post-stained measurements. Therefore, for the final analyses, I considered forebody length from the stained specimens as a proxy for body length because these two features were correlated (Pearson correlation coefficient =0.89, $P<0.05$).

To test the hypothesis of crypsis in *E. trivolvis*, I performed Principal Component Analysis (PCA) on genetically identified specimens from my study and museum specimens that were identified as *E. trivolvis* by Kanev et al. (1995) in the species description. To determine whether *E. trivolvis* and *E. revolutum* sensu lato were morphologically distinguishable, I performed a PCA that included genetically-identified specimens of *E. trivolvis* lineages and *E. revolutum* sensu lato as well as the museum specimens identified as *E. trivolvis* by Kanev et al. (1995). Additionally, I included definitive host use for each stained specimen to test for the influence of hosts on morphological variation. For all analyses, I used the package vegan in Rstudio 1.0153 (Oksanen et al., 2018). I tested whether the characters were normally distributed using a Shapiro-Wilk Normality test. If morphological characters were non-normally distributed, they were log-transformed.

2.3. Results

2.3.1 Genetic identification of echinostomes

Phylogenetic analyses with ND1 sequences confirmed six *E. trivolvis* lineage a specimens from Manitoba: four from mallards (three and one per two host individuals), one from a sandhill crane, and one from a muskrat (Figure 1.1). All of these sequences represented one unique haplotype because they were genetically identical. In addition, sequencing confirmed 30 *E. trivolvis* lineage c adult specimens that were all from muskrat hosts in Manitoba (Figure 1.1).

These sequences represented 12 unique haplotypes. No *E. trivolvis* lineage b was recovered from bird or muskrat hosts in Manitoba. Twelve *E. revolutum* sensu lato were found only in bird hosts including ten unique haplotypes.

2.3.2 Morphological variation within E. trivolvis *complex*

In total, I measured 22 characters for each of 42 adults including 31 genetically-identified specimens from my study and 11 museum specimens identified as *E. trivolvis* (Supplementary Table 2.1). A total of five genetically-identified and ten museum specimens were excluded from analysis because not all features could be measured. All morphological characters were normally distributed except inter-testis length, which was log-transformed prior to analysis.

Cumulatively, the first five PCs accounted for 90% of the total observed morphological variation of the specimens (Table 2.1). However, I only interpreted the first three PCs because together they encompassed >80% of the total variation. The first principal component (PC1) accounted for 61.1% of the total variation and was interpreted as describing body size because several morphological characters such as acetabulum length and width, forebody length, oral sucker length and width, pharynx length and width, and collar length contributed positively and significantly (considered cut-off value of factor scores as 0.8). The second principal component (PC2) explained 16% of the total variation and was interpreted as describing egg size because the main contributors were ova maximum and minimum length, and ova maximum and minimum width. The third principal component (PC3) accounted for 6.07% of the variation and was interpreted as describing reproductive organ size due to the contributions of anterior testis width, ovary length and inter-testis length.

The *E. trivolvis* lineages a and c specimens did not form separate clusters (Figure 2.2). There was one *E. trivolvis* lineage a specimen that was separate from lineage c and the other

three lineages a individuals. This specimen was gravid, but had smaller reproductive organs and the ovary and anterior testis overlapped, which was not seen in other individuals. Some of the museum specimens clustered with the genetically-identified specimens (HWML 23193-1, HWML 34804, HWML 699-2), but others were more distinct (HWML 35570-1, HWML 35571-2). There was no separate clustering based on definitive host use though specimens from waterfowl, muskrat, opossum and Canada goose were more morphologically similar than specimens from hawks and rodent (Figure 2.3). Two museum specimens from two hawk species (red-tailed hawk - *Buteo jamaicensis* Gmelin, 1788 and broad-winged hawk - *Buteo platypterus* Vieillot, 1823; Figure 2.3: diamond shape) and one museum specimen from a rodent (nutria, *Myocastor coypus* Molina, 1782 host (Figure 2.3: cross sign) were the most morphologically divergent specimens. The other outlier was one of the genetically-identified *E. trivolvis* museum specimens from a muskrat in USA.

2.3.3 Morphological variation between E. trivolvis *and* E. revolutum *sensu lato*

After excluding nine genetically identified and ten museum specimens because of missing data, I performed a PCA that included 22 characters and 50 adult worms, including 11 museum specimens and genetically-identified specimens, including four *E. trivolvis* lineage a, 27 *E. trivolvis* lineage c, and eight *E. revolutum* sensu lato. All morphological characters were normally distributed except inter-testis length, which was then log-transformed prior to PCA analysis. The first three PCs explained >80% of the variance (Table 2.2). PC1 explained 60.73% of the variation and best described body size as the main contributors were acetabulum size (length and width), collar length, oral sucker length and width, pharynx length and width, and ovary and cirrus sac length and width (cut off value for scores was considered as 1). Moreover,

PC2 and PC3 explained 15.27% and 5.82% variance and describe ova size and inter-testis length respectively.

Genetically identified specimens of *E. trivolvis* and *E. revolutum* sensu lato clustered together (Figure 2.4) and did not form distinct clusters based on host use (Figure 2.5). These specimens were found in an array of bird (mallard, sandhill crane, Canada goose) and mammal (muskrat) hosts. As in the PCA with only *E. trivolvis* specimens, some of the museum specimens from hawks and one rodent host (nutria) were the most divergent (HWML 35570-1, HWML 35571-2, HWML 22672-1).

2.4. Discussion

By genetically identifying parasites, I found that lineages of *E. trivolvis* differed in definitive host specificity in Manitoba. *Echinostoma trivolvis* lineage a used only birds (mallard and sandhill crane), while *E. trivolvis* lineage c used only muskrats. No lineage b specimens were collected from any of the sampled hosts, suggesting that it may not occur in Manitoba at least in the host species that were sampled. The definitive hosts for *E. revolutum* sensu lato were mallards, sandhill crane and Canada geese confirming that this echinostome species can be found in the same definitive hosts as *E. trivolvis* lineage a. The multivariate morphological analysis of 22 traits supports the hypothesis that *E. trivolvis* lineage a and c are cryptic. Moreover, *E. trivolvis* lineages are also morphologically indistinguishable with the sister species, *E. revolutum* sensu lato. Further, some of the museum specimens identified by experts (i.e. Kanev et al., 1994, 1995) may have been misidentified as some of them (e.g from the hawks and nutria) could be considered morphologically distinct.

Adult stages of echinostomes have few characters that can be used for reliable identification (i.e. *E. revolutum* sensu lato, Kanev, 1994). For these echinostomes, identification

should be based on larval morphology, host specificity and geographic distributions (Fried and Toledo, 2004). However, other experts argue that adult stages of echinostomes can be morphologically distinguished if detailed examinations are undertaken (Kostadinova et al., 2000). Morphological analysis of 22 features in two species, *E. trivolvis* and *E. revolutum* sensu lato within the 37-spined 'revolutum' complex showed that genetic lineages of *E. trivolvis* were cryptic compared to each other, and that *E. revolutum* sensu lato was cryptic compared to the *E. trivolvis* lineages. In this sense, they are morphologically indistinguishable because variation in their morphological characters was largely overlapping between the genetically distinct lineages/species. My dataset also provides preliminary information on how much morphological variation may occur within a mitochondrial haplotype, as all the *E. trivolvis* lineage a specimens had the same ND1 sequence. Additional sampling of this haplotype across different geographic regions in the same hosts could be used to test how geography and host use influences phenotype.

The morphological overlap between adult *E. trivolvis* genetic lineages a and c, and *E. revolutum* sensu lato contrasts with the finding that adult *E. trivolvis* and *E. revolutum* sensu lato of the same age from laboratory definitive chicken hosts (*Gallus gallus* (Linnaeus, 1758)) differ in body size and organ size (Fried et al., 1997). However, Sorensen et al. (1997) found that the adult body size of *E. trivolvis* (collected from *H. trivolvis* in Indiana and raised in chickens) overlapped with *E. revolutum* sensu stricto and sensu lato indicating that body size cannot always be used to distinguish among the three species. In my study, all the field-collected specimens were gravid, but the influence of age and host on phenotype is unknown. Because most researchers will encounter echinostomes in field-infected hosts, it is important to determine if these specimens can be accurately identified. Given the morphological overlap in the

genetically identified *E. revolutum* sensu lato and *E. trivolvis* specimens in my study, it is unclear how field-collected specimens could be accurately identified without genetic sequencing, especially considering that these species may be found in the same final host species. Thus, my results call into question the identifications of past studies. Unfortunately, many voucher specimens have not been deposited in museums either as permanent slides or in vials with formalin or ethanol, so it is impossible to re-evaluate their identifications. Even if permanent slides are available, no DNA sequencing from these individuals is possible as the chemicals used make it impossible to extract DNA (Detwiler, personal communication). In addition, it was common practice to preserve adult specimens in vials of formalin or denatured ethanol, which also makes it impossible/difficult to extract DNA. Yet, if a database of morphological measurements is constructed from genetically-identified specimens, then it could be possible to test whether echinostomes have been misidentified in past studies.

The museum specimens in this study were deposited under several different names (i.e. HWML 21801 as *E. revolutum*, HWML 21802 as *E. coalitum*). Kanev et al. (1995) later identified all of these specimens as *E. trivolvis* in the species description paper. These specimens were morphologically much more variable than the *E. trivolvis* lineages a and c collected from Manitoba. The greater morphological variance could be due to technical or biological reasons. Differences in slide preparation can influence the measurements of helminths (Criscione and Font, 2001). Biologically, the museum specimens could be exhibiting greater morphological variation from isolation by geography and host species as they were collected from different geographic locations (several states in the USA) and hosts (mallard, muskrat, Canada goose, broad-winged hawk, red-tailed hawk, nutria, and opossum) compared to my study. The specimens collected in Manitoba may exhibit less morphological variation overall because they

were collected from one geographic region and from a small subset of definitive host species. In the case of *E. trivolvis* lineage c, adults have now only been identified in muskrats in several geographic areas (Manitoba, this study; Wisconsin, Detwiler et al., 2010; Virginia, Detwiler et al., 2012). More samples from different geographic regions will help clarify the extent of the morphological variation in the *E. trivolvis* lineages. Ideally, the same definitive host species would be sampled as in this study, but samples from additional definitive host species would still be useful in determining the range of morphological variation of these lineages. In my dataset, some of the museum specimens collected from hawks and nutria could be considered outliers, suggesting that differences in geography and host species could be influencing morphology. However, because identification of these specimens cannot be verified with genetics, a more parsimonious explanation is that these specimens were misidentified by Kanev as *E. trivolvis*.

2.5. Conclusion

By integrating sequencing and morphology of adult echinostomes recovered from Manitoba wetlands, I provided the first test of crypsis, as no previous study had conducted detailed morphological analysis of the *E. trivolvis* genetic lineages. I supported the hypothesis that *E. trivolvis* lineage a and lineage c are cryptic, as the specimens did not form distinct morphological clusters in PCA analysis. By accurately identifying echinostomes with sequences, I confirmed that adult *E. trivolvis* lineage a infects birds and mammals while *E. trivolvis* lineage c appeared to be restricted to muskrats. In addition, genetically identified *E. revolutum* sensu lato were cryptic with *E. trivolvis*, and only found to infect bird hosts. This finding is important to ecologists as it confirms that *E. trivolvis* lineages and *E. revolutum* sensu lato can occur in the same host species and most likely cannot be accurately identified based on morphology. My detailed morphological analysis shows that there are no clear morphological characters that can

be used to distinguish these cryptic taxa in North America. Additional sampling in different geographic regions in more host species will be essential to understanding host specificity as well as the extent of crypsis in these groups.

References

Anderson, D. R. and R. L. Beaudoin. 1966. Host habitat and age as factors in the prevalence of intestinal parasites of the muskrat. Bulletin of the Wildlife Disease Association. 2: 70–77.

Beaver, P. C. 1937. Experimental studies on *Echinostoma revolutum* (Frólich), a fluke from birds and mammals. Illinois Biological Monographs. p. 1–96.

Clark, N. J. and S. M. Clegg. 2017. Integrating phylogenetic and ecological distances reveals new insights into parasite host specificity. Molecular Ecology. 26: 3074–3086.

Criscione, C. D. and M. S. Blouin. 2005. Effective sizes of macroparasite populations: a conceptual model. Trends in Parasitology. 21: 212–217.

Criscione, C. and W. Font. 2001. Development and specificity of *Oochoristica javaensis* (Eucestoda: Cyclophyllidea: Anoplocephalidae: Linstowiinae). Comparative Parasitology. 68. 149-155.

Detwiler, J. T., D. H. Bos, and D. J. Minchella. 2010. Revealing the secret lives of cryptic species: Examining the phylogenetic relationships of echinostome parasites in North America. Molecular Phylogenetics and Evolution. 55: 611–620.

Detwiler, J. T., A. M. Zajac, D. J. Minchella, and L. K. Belden. 2012. Revealing cryptic parasite diversity in a definitive host: echinostomes in muskrats. The Journal of Parasitology. 98: 1148–1155.

Dogiel, V. A., Y. I. Polyanski, and E. J. Kheisin.1964. General Parasitology. Oliver and Boyd, Edinburg, Scotland. pp. 516.

Fried, B. and T. K. Graczyk. 2000 Echinostomes as experimentalmodels for biological research. Dordrecht: Kluwer Academic Publishers, pp. 273.

Fried, B., Pane, P.L. and Reddy, A. 1997. Experimental infection of *Rana pipiens* tadpoles with *Echinostoma trivolvis* cercariae. Parasitology. 83, 666–669.

Fried, B. and R. Toledo. 2004. Criteria for species determination in the "Revolutum" group of Echinostoma. Journal of Parasitolgy. 90: 917.

Georgieva, S., A. Faltýnková, R. Brown, I. Blasco-Costa, M. Soldánová, J. Sitko,and A. Kostadinova. 2014. *Echinostoma 'revolutum*' (Digenea:Echinostomatidae) species complex revisited: species delimitation based on novel molecular and morphological data gathered in Europe. Parasites & Vectors. 7: 520–537.

Hayward, A. 2010. Cryptic diversity and patterns of host specificity in trematode flatworms. Molecular Ecology. 19: 2602–2604.

Kanev. I., 1994. Life-cycle, delimitation and redescription of *Echinostoma revolutum* (Frölich, 1802) (Trematoda: Echinostomatidae). Systematic Parasitology. 28: 125–144.

Kanev, I., B. Fried, V. Dimitrov, and V. Radev. 1995. Redescription of *Echinostoma trivolvis* (Cort, 1914) (Trematoda: Echinostomatidae) with a discussion on its identity. Systematic Parasitology. 32: 61–70.

Koh, L. P., R. R. Dunn, N. S. Sodhi, R.K. Colwell, H. C. Proctor, and V. S. Smith. 2004. Species co extinctions and the biodiversity crisis. Science. 305: 1632–1634.

Kostadinova, A., D. I. Gibson, V. Biserkov, and Ivanova, R. 2000. A quantitative approach to the evaluationof the morphological variability of two echinostomes, *Echinostoma miyagawai* Ishii, 1932 and *E. revolutum* (Frólich, 1802), from Europe. Systematic Parasitology. 45: 1–15.

Kostadinova, A., G. A. Herniou, J. Barrett, and D. T. L. Littlewood. 2003. Phylogenetic relationships of *Echinostoma* Rudolphi, 1809 (Digenea: Echinostomatidae) and related genera re-assessed via DNA and morphological analyses. Systematic Parasitology. 54: 159–176.

Kvach, Y. and P. Sasal. 2010. *Telosentis exiguus* (von Linstow, 1901) (Palaeacanthocephala: Illiosentidae), a generalist parasite of fishes in the Mediterranean basin. Systematic Parasitology. 76: 9–18.

Locke, S. A., J. D. McLaughlin, and D. J. Marcogliese. 2010. DNA barcodes show cryptic diversity and a potential physiological basis for host specificity among Diplostomoidea (Platyhelminthes: Digenea) parasitizing freshwater fishes in the St. Lawrence River, Canada. Molecular Ecology. 19: 2813–2827.

McKenzie, C. E. and H. E. Welch. 1979. Parasite fauna of the muskrat, *Ondatra zibethicus* (Linnaeus, 1766) in Manitoba, Canada. Canadian Journal of Zoology. 57: 640–646.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchi n, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, and H. Wagn er. 2018. community Ecology Package. R package version 2.5-2. https: // CRAN.R-project.org/ package.

Pérez -Ponce De León, G., and R. Poulin. 2018. An updated look at the uneven distribution of cryptic diversity among parasitic helminths. Journal of Helminthology. 92: 197–202.

Poulin, R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biology Letters. 7: 241–244.

Poulin, R. and D. Mouillot. 2003. Parasite specialization from a phylogenetic perspective: A new index of host specificity. Parasitology. 126: 473–480.

Poulin, R. and Keeney, D. B. 2008. Host specificity under molecular and experimental scrutiny. Trends in Parasitology. 24: 24–28.

Salgado-Maldonado, G. M. T. Novelo-Turcotte, J. M. Caspeta-Mandujano, G. Vazquez-Hurtado, B. Quiroz-Martínez, N. Mercado-Silva, and M. Favila. 2016. Host specificity and the structure of helminth parasite communities of fishes in a Neotropical river in Mexico. Parasite. 23: 61–71.

Schell, S. C. 1985. Trematodes of North America. University of Idaho Press, Moscow, Idaho, pp. 263.

Sorensen, R. E., I. Kanev, B. Fried, and D. J. Minchella. 1997. The occurrence and identification of *Echinostoma revolutum* from North American *Lymnaea elodes* snails. Journal of Parasitology. 83: 169–170.

Terry, R. H. and S. L. Gardner. 2008 A review of species in the genus *Rhopalias* (Rudolphi, 1819). Journal of Parasitology. 94: 716–726

Table 2.1. Principal component analysis (PCA) for the morphological measurements of *Echinostoma trivolvis* lineages a and c. Variance and loadings for the first five PCs of a PCA conducted on the correlation matrix of 22 morphological characters. Specimens were genetically identified in this study (n =31) or morphologically identified as *E. trivolvi*s by Kanev et al. (1995) (n =11).

	PC1	**PC2**	**PC3**	**PC4**	**PC5**
Variance (%)	61.1	16.1	6.070	5.29	2.94
FBL	0.800	-0.0682	0.426	0.542	0.142
CL	1.10	-0.112	0.183	-0.158	0.144
OSL	1.10	-0.100	-0.102	-0.142	0.151
OSW	1.10	-0.100	-0.102	-0.142	-0.151
PHL	1.08	-0.0221	-0.0700	-0.321	-0.0152
PHW	1.08	0.0210	-0.152	-0.309	0.0191
ACL	1.12	-0.0841	0.0822	-0.157	0.0107
ACW	1.12	0.0140	0.0616	-0.268	-0.100
OVL	1.0015	-0.120	0.024	-0.00607	-0.247
OVW	1.07	0.0421	0.800	0.0654	-0.177
SPATO	0.858	-0.256	0.482	-0.142	0.351
OVMXL	0.333	1.08	0.0162	-0.0419	-0.161
OVMXW	0.293	1.04	-0.151	0.197	-0.153
OVMNL	0.158	1.07	0.211	-0.140	-0.198
OVMNW	0.198	1.10	0.195	0.00362	-0.100
CSL	1.04	-0.0120	0.113	-0.236	0.0820
CSW	1.05	-0.117	-0.150	0.131	0.112
ATL	0.997	-0.0543	-0.160	0.456	-0.191
ATW	0.835	-0.0090	-0.600	0.260	0.316
PTL	0.984	0.0260	-0.180	0.491	-0.205
PTW	1.0087	0.0260	-0.181	-0.491	-0.205
ITL	0.427	-0.153	0.800	0.400	.539

Table 2.2. Principal component analysis (PCA) for the morphological measurements of *Echinostoma trivolvis* lineages a and c and *Echinostoma revolutum* sensu lato. Variance and loadings for the first five PCs of a PCA conducted on the correlation matrix of 22 morphological characters. Specimens were genetically identified (n= 39) in this study or were museum specimens that were identified as *E. trivolvis* by Kanev et al. (1995) (n=11).

	PC1	**PC1**	**PC3**	**PC4**	**PC5**
Variance (%)	60.7	15.2	5.82	4.92	3.00
FBL	0.843	-0.0900	0.400	-0.530	0.117
CL	1.14	-0.113	0.216	0.105	-0.150
OSL	1.14	-0.105	-0.0600	0.140	0.200
OSW	1.20	-0.0700	-0.0504	0.140	0.130
PHL	1.12	-0.0110	-0.0420	0.400	-0.100
PHW	1.12	0.300	-0.121	0.400	-0.043
ACL	1.16	-0.072	0.0840	0.181	-0.132
ACW	1.20	-0.00400	0.300	0.320	-0.0210
OVL	1.03	-0.100	-0.0230	0.0130	0.324
OVW	1.12	0.400	0.077	-0.100	-0.200
SPATO	0.900	-0.300	0.471	0.130	-0.400
OVMXL	0.274	1.11	0.031	0.082	-0.172
OVMXW	0.284	1.08	-0.193	-0.207	0.104
OVMNL	0.200	1.10	0.200	0.178	-0.200
OVMNW	0.261	1.10	0.200	-0.087	-0.088
CSL	1.07	-0.100	0.200	0.200	-0.105
CSW	1.09	-0.106	-0.242	0.150	-0.126
ATL	1.04	-0.0410	-0.232	0.421	0.233
ATW	0.900	-0.100	-0.600	-0.300	0.400
PTL	1.03	0.0333	-0.240	-0.500	0.230
PTW	1.06	-0.060	-0.414	-0.300	0.016
ITL	0.463	-0.129	0.831	-0.500	-0.467

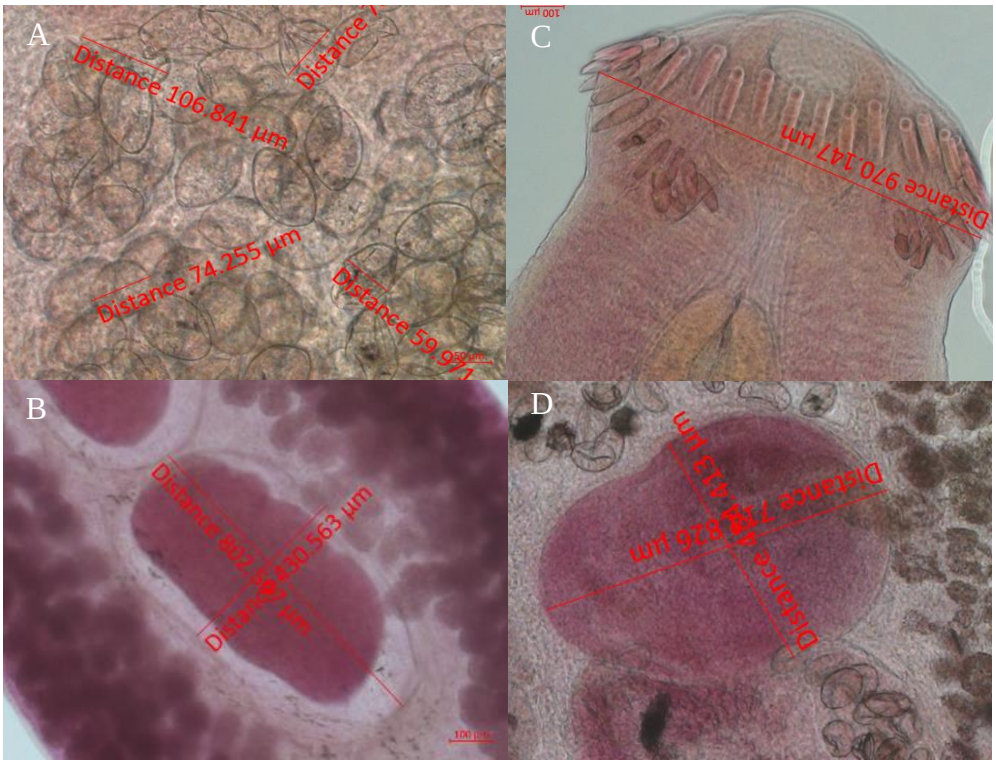


Figure 2.1. Digital images of four morphological characters of adult echinostomes with examples of measurements. A. Ova (length and width) B. Posterior testis (length and width) C. Collar (length) and D. Ovary (length and width).

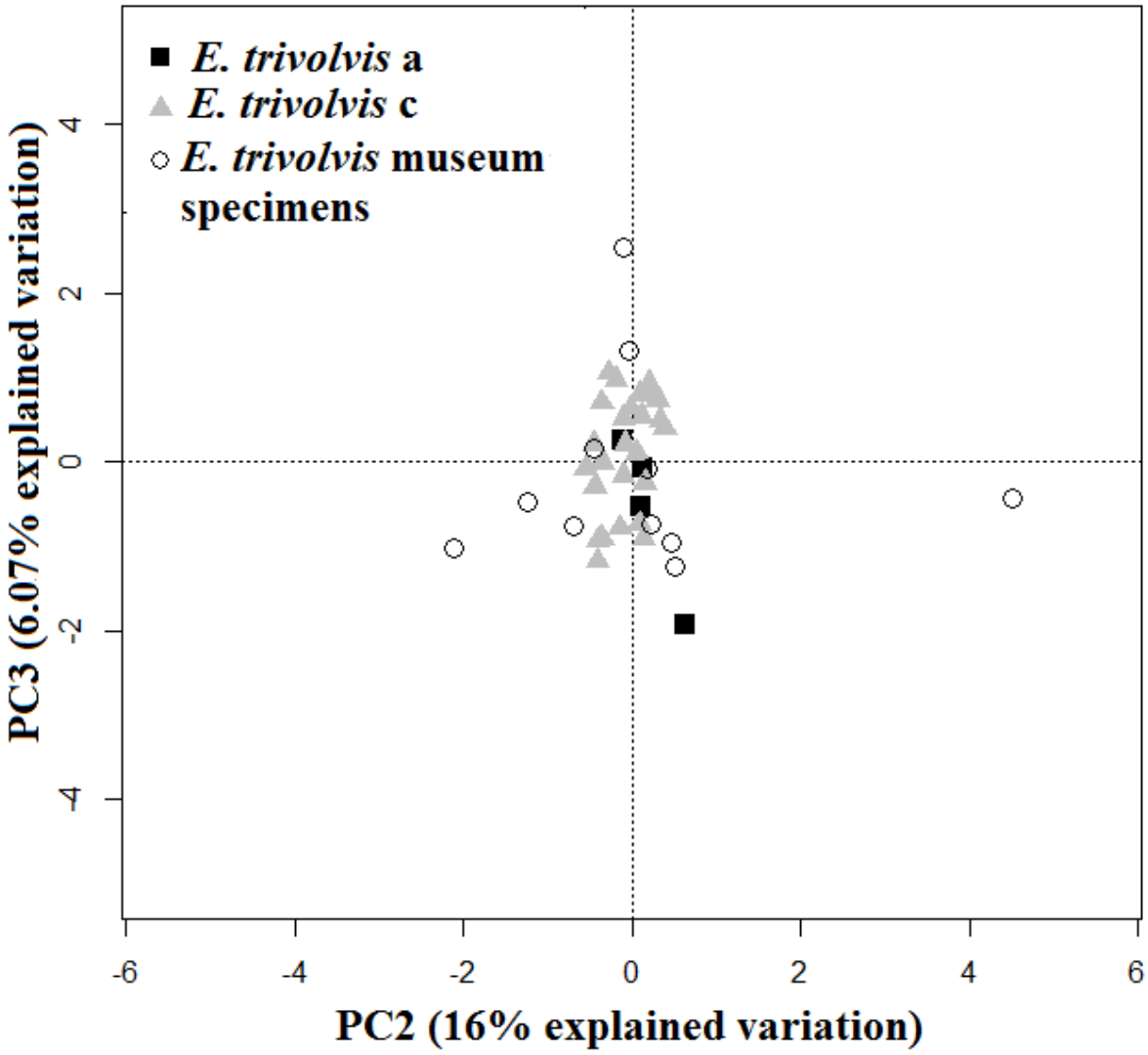


Figure 2.2. Principal Component Analysis (PCA) of the morphological features of genetically identified *Echinostoma trivolvis* lineage a and c, and *E. trivolvis* museum specimens (n=42).

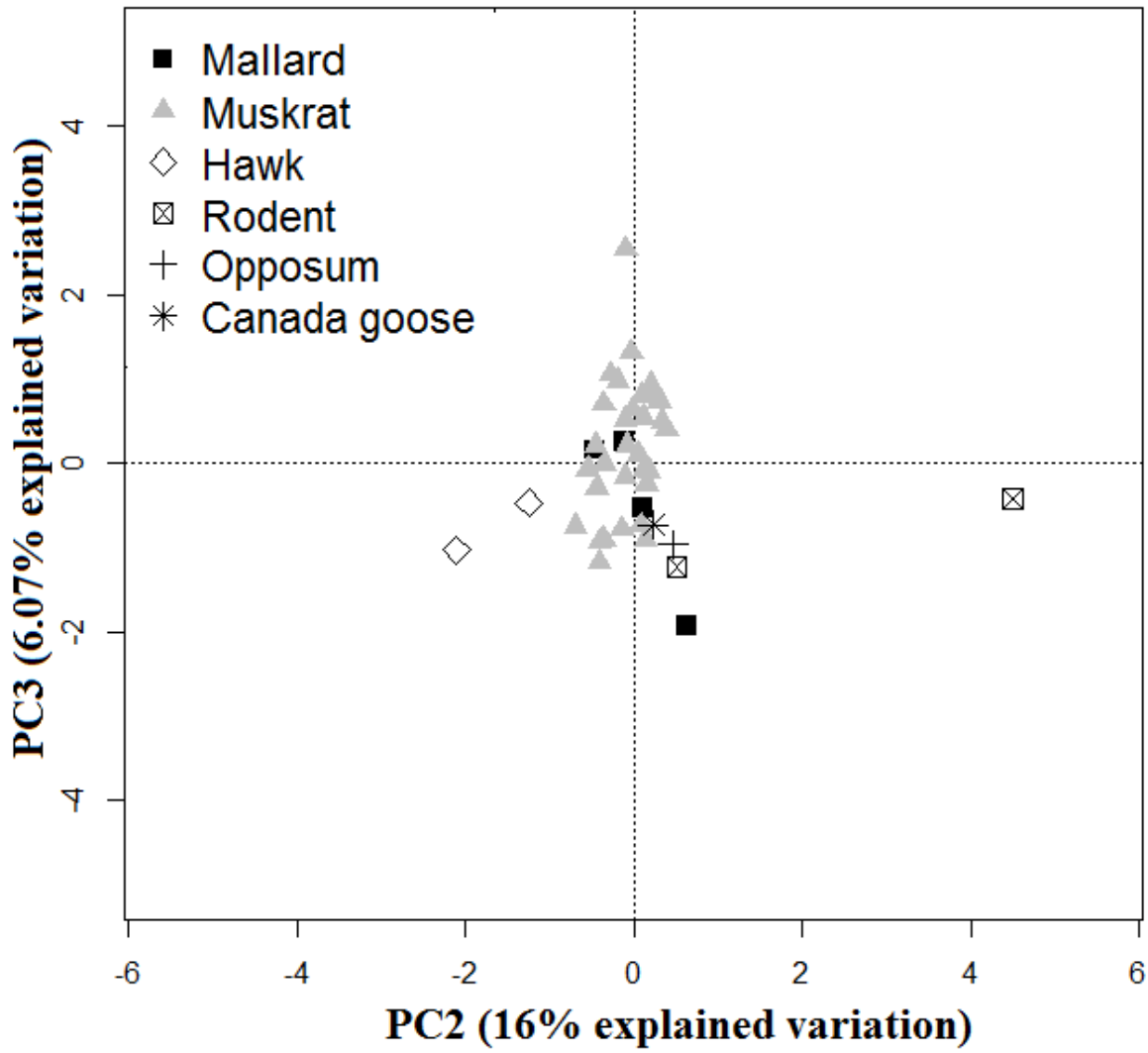


Figure 2.3. Principal Component Analysis (PCA) of the morphological features of *Echinostoma trivolvis* lineage a and c, and *E. trivolvis* museum specimens (n=42). Definitive hosts are shown with different coloured symbols

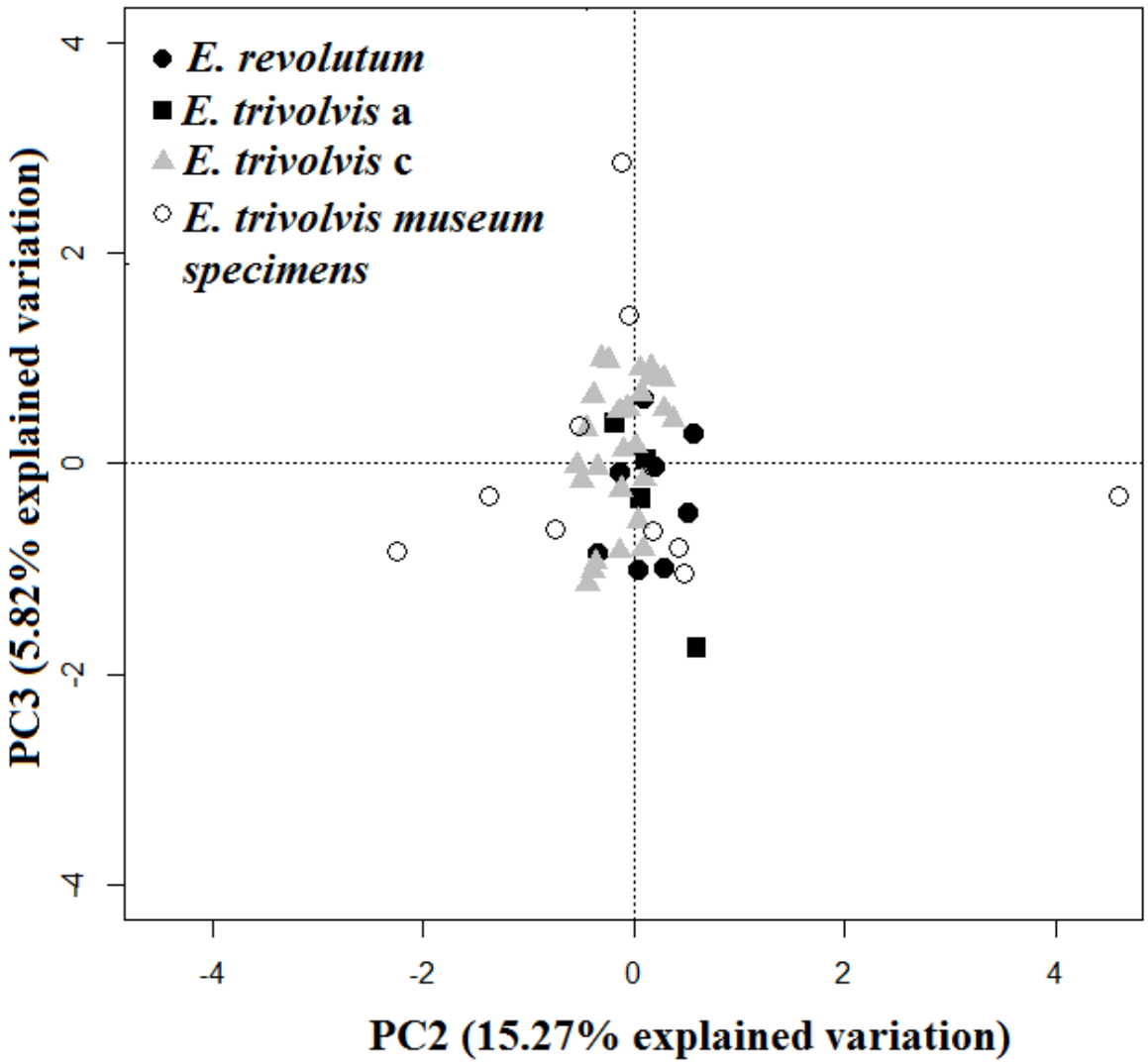


Figure 2.4. Principal Component Analysis (PCA) of the morphological features of genetically identified *Echinostoma trivolvis* lineage a and c, and *Echinostoma revolutum* as well as *E. trivolvis* museum specimens (n=50).

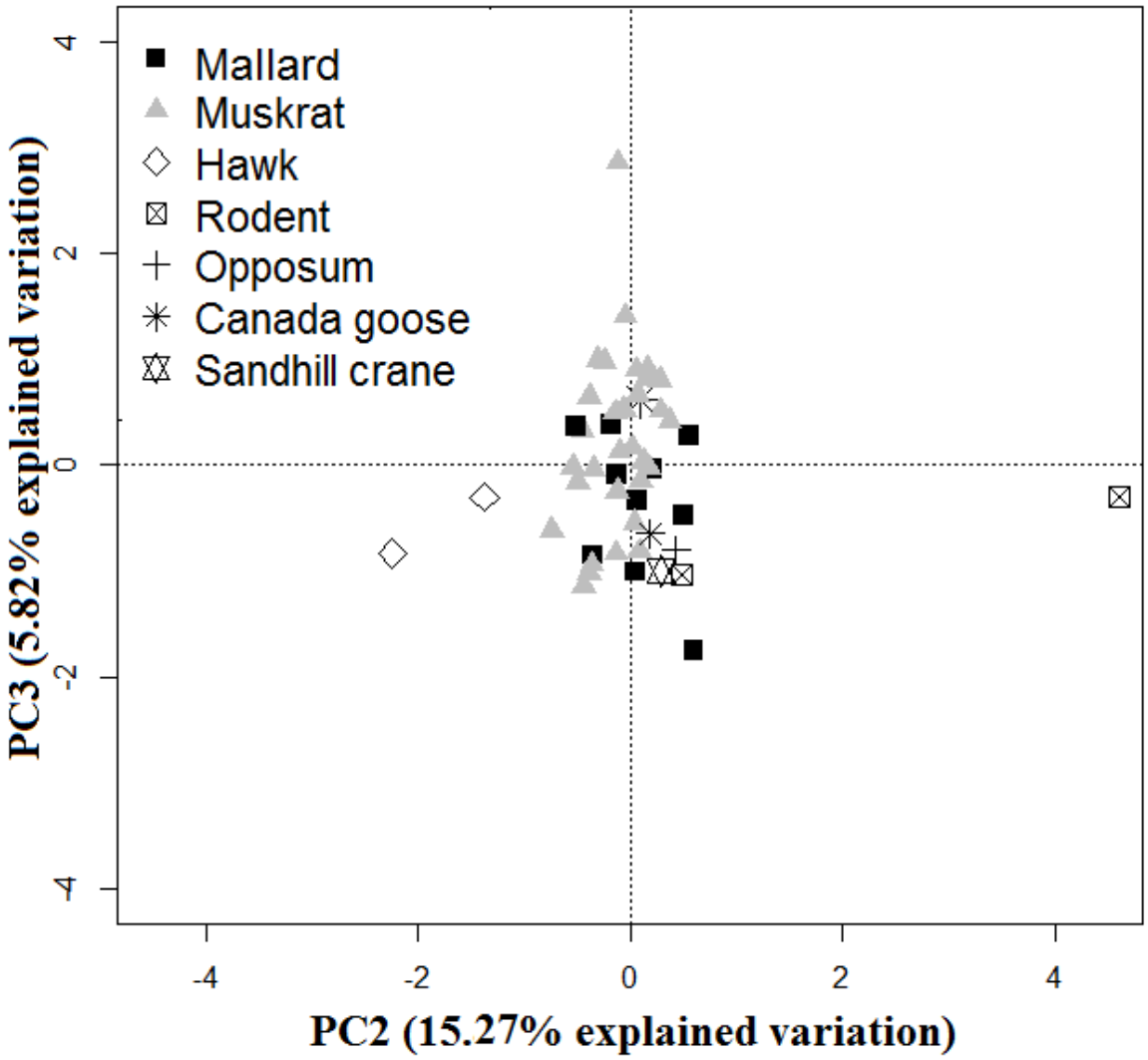


Figure 2.5. Principal Component Analysis (PCA) of the morphological features of *Echinostoma trivolvis* lineage a and c, and *Echinostoma revolutum* as well as *E. trivolvis* museum specimens (n=50). Definitive hosts are shown with the different coloured symbols.

Thesis Conclusion

I integrated morphology and genetics of echinostome trematodes collected from naturally-infected invertebrate and vertebrate hosts to estimate species diversity. Chapter 1 demonstrates that additional geographic sampling revealed greater species diversity than previously characterized in North America. To identify species, I used monophyletic groups from mitochondrial and nuclear gene trees, haplotype networks and genetic distance. Additionally, by linking DNA larval and adult echinostome sequences, I have demonstrated host specificity of 11 lineages/species and suggest that crypsis leads to inaccurate life cycle descriptions. In the present study, all echinostome larvae were not specific to a single first intermediate host species which was historically considered as an identifying characteristic for this trematode group. I also found that vagile bird hosts promote gene flow of cryptic lineages at a continental scale. In Chapter 2, I compared morphological and genetic diversity of echinostome species to test the hypothesis of crypsis in the *Echinostoma trivolvis* species complex and to determine morphological variation between two nominal species, *E. trivolvis* and *Echinostoma revolutum* sensu lato. Integrative taxonomy demonstrated crypsis between two lineages of the *E. trivolvis* species complex and confirmed that they differ in host use. Additionally, this study reveals *E. revolutum* sensu lato overlaps morphologically demonstrating that when it co-occurs in the same definitive host with *E. trivolvis* lineages, they cannot be morphologically distinguished. Further, my results suggest that some museum specimens have been misidentified, calling into question previous observations of host specificity.

Overall, the newly generated DNA sequences in this study will help to identify echinostome lineages in future studies. Accurate knowledge of echinostome biodiversity, life cycles, and host specificity will allow us to test how host specificity has influenced speciation.

Further, better estimates of geographic distributions for species of echinostomes improves our understanding of the epidemiology of wildlife disease and helps determine their potential role as causative agents of emerging disease.

Appendix A

Supplementary Table 1.1. First intermediate host snails infected with echinostome trematodes, collected from 14 wetlands in southern Manitoba.

Location (GPS coordinates)	Snail species	Number of examined hosts	Number of infected hosts
SmartPark Pond, University of Manitoba Campus	*Physa gyrina*	150	0
	Gyarulus sp.	52	0
Whiteshell Provincial Park (N 49° 54', W 95° 19')	*Helisoma trivolvis*	60	18
	Bulimnea megasoma	5	2
Lake Louis (N 49° 10', W 98° 52')	*Lymnaea stagnalis*	106	0
	Helisoma trivolvis	1	0
	Physa gyrina	4	0
Ditch 1, Netley-Libau Marsh (N 50° 26', W 96° 68')	*Lymnaea stagnalis*	92	0
	Lymnaea elodes	37	2
	Helisoma trivolvis	3	0
	Gyarulus sp.	6	0
	Physa gyrina	5	0
Ditch 2, Netley-Libau Marsh (N 50° 26', W 96° 68')	*Lymnaea stagnalis*	37	0
	Lymnaea elodes	296	37
	Helisoma trivolvis	107	0
	Gyarulus sp.	12	0
	Physa gyrina	29	0
Ditch 3, Netley-Libau Marsh (N 50° 26', W 96° 68')	*Lymnaea stagnalis*	1	0
	Lymnaea elodes	204	19
	Helisoma trivolvis	1	0
	Gyarulus sp.	4	0
	Physa gyrina	9	0
Marsh River (N 49° 56', W 97° 13')	*Lymnaea stagnalis*	47	4
	Lymnaea elodes	227	2
	Helisoma trivolvis	2	0
	Physa gyrina	52	0
West Shoal Lake (N 50° 43', W 100° 35')	*Lymnaea stagnalis*	94	1
	Lymnaea elodes	88	3
	Helisoma trivolvis	11	5

	Gyarulus sp.	28	0
	Physa gyrina	12	4
East Shoal Lake (N 50° 43', W 100° 35')	*Lymnaea stagnalis*	4	0
	Lymnaea elodes	166	9
North Shoal Lake (N 50° 43', W 100° 35')	*Lymnaea stagnalis*	80	1
	Lymnaea elodes	216	0
	Helisoma trivolvis	5	1
	Physa gyrina	6	0
Oak Hammock Marsh (N 50° 11' ,W 97° 07')	*Lymnaea stagnalis*	230	0
	Lymnaea elodes	333	12
	Helisoma trivolvis	81	0
	Physa gyrina	100	0
Lindalls Lake (N 58° 45', W 101° 48')	*Lymnaea stagnalis*	106	0
	Lymnaea elodes	21	1
	Helisoma trivolvis	95	0
	Physa gyrina	72	1
Clematis Wildlife Refuge (N 50° 61', W 97° 58')	*Lymnaea stagnalis*	66	0
	Lymnaea elodes	53	1
	Helisoma trivolvis	143	1
	Physa gyrina	2	1
Marais River (N 49° 08', W 97° 16')	*Lymnaea stagnalis*	33	0
	Lymnaea elodes	64	0
	Helisoma trivolvis	65	0
	Physa gyrina	128	0

Supplementary Table 1.2. Definitive host species collected from hunters and rehabilitation centres in southern Manitoba.

Location (With GPS)	Definitive host species	Common name	Number of host specimens infected (total examined)	Consume snails/frogs?	Confirmed host of echinostomes?
Rehabilititation Centre, Winnipeg*	*Aix sponsa*	Wood Duck	1 (2)	Yes	Yes (McLaughlin, 1979)
Rehabilititation Centre, Winnipeg*	*Anas platyrhynchos*	Mallard	1(1)	Yes	Yes (Shaw, 1980)
Inkster (N 45° 95', W - 97° 20')	*Branta canadensis*	Canada Goose	0(12)	Yes	Yes (Davis, 2006)
Pine Falls (N 50° 33', W 96° 11')	*Buteo platypterus*	Broad-winged Hawk	1(2)	Yes	Yes (Taft, 1993)
Rehabilititation Centre, Winnipeg*	*Buteo swainsoni*	Swainson's Hawk	0(2)	Yes	No
Brandon (N 49° 49', W - 99° 56')	*Strix nebulosa*	Great Grey Owl	0(3)	No	No
Tolstoi (N 49° 04' W - 96° 80')	*Strix variosa*	Barred Owl	0(1)	Yes	No
Rehabilititation Centre, Winnipeg*	*Falco peregrinus*	Peregrine Falcon	0(1)	No	No
Rehabilititation Centre, Winnipeg*	*Falco sparverius*	American Kestrel	0(1)	Yes	Yes (Kinsella, 1995)
Portage La Prairie (N 49° 97' W - 98° 17')	*Bubo scandiacus*	Snowy Owl	0(1)	No	No
Rehabilititation Centre, Winnipeg*	*Ondatra zibethicus*	Muskrat	1(1)	Yes	Yes (Detwiler et al., 2012)

Location (With GPS)	Definitive host species	Common name	Number of host specimens infected (total examined)	Consume snails/frogs?	Confirmed host of echinostomes?
Rehabilititation Centre, Winnipeg*	*Lophodytes cucullatus*	Hooded Merganser	0(1)	Yes	No
	Anas platyrhynchos	Mallard	10(24)	Yes	Yes
	Bucephala albeola	Bufflehead duck	0(1)	Yes	Yes (McLaughlin, 1979)
	Anas clypeata	Northern Shoveler	1(1)	Yes	Yes(Broderson, 1977)
Netley-Libau Marsh (N 50° 26', W 96° 68')	*Porzana carolina*	Sora	0(1)	Yes	No
	Aythya affinis	Lesser Scaup	0(1)	Yes	Yes (Bush, 1986)
	Fulica americana	American Coot	0(3)	Yes	No
	Anas caroliennsis	Green-winged Teal	0(3)	Yes	Yes (McLaughlin, 1979)
	Ondatra zibethicus	Muskrat	6(42)	Yes	Yes
	Anas strepera	Gadwall	1(1)	Yes	No
Coulter, MB (N 50° 26', W 96° 68')	*Anas carolinensis*	Green-winged Teal	0(3)	Yes	Yes
	Fulica americana	American Coot	2(2)	Yes	No
	Grus canadensis	Sandhill Crane	2(2)	Yes	Yes (Rothenburger et al., 2016)
Settler's Road, 207 Provincial Hwy (N 49° 49', W 97° 8')	*Branta canadensis*	Canada Goose	1(25)	Yes	Yes
	Anas platyrhynchos	Mallard	1(1)	Yes	Yes
South Winnipeg (N 50° 43', W100° 35')	*Branta canadensis*	Canada Goose	1(3)	Yes	Yes

Location (With GPS)	Definitive host species	Common name	Number of host specimens infected (total examined)	Consume snails/frogs?	Confirmed host of echinostomes?
	Anas platyrhynchos	Mallard	3(3)	Yes	Yes
Steinbach, MB (N 49° 31' W -96° 41')	*Bucephala clangula*	Common Goldeneye	0(1)	Yes	No
Morris, MB (N 49° 21' W97° 21')	*Branta canadensis*	Canada Goose	2(4)	Yes	Yes
Lyleton (N 49° 05', W 101° 14')	*Grus canadensis*	Sandhill Crane	1(3)	Yes	Yes
Whiteshell River (N 49° 54', W 95° 19)	*Ondatra zibethicus*	Muskrat	16(18)	Yes	Yes
Oak Hammock Marsh (N 50° 11 ,W -97° 07')	*Ondatra zibethicus*	Muskrat	8(17)	Yes	Yes
Marsh River (N 49° 30', W 97° 10')	*Ondatra zibethicus*	Muskrat	15(16)	Yes	Yes
Marsh River and Lake Louise	*Ondatra zibethicus*	Muskrat	13(19)	Yes	Yes
Unknown	*Branta canadensis*	Canada Goose	0(2)	Yes	Yes
		Mallard	1(3)	Yes	Yes

*Original location of collection is unknown; frozen carcasses donated from Wildlife Haven Rehabilitation Centre and Prairie Wildlife Rehabilitation Center in Winnipeg, Manitoba.

Supplementary Table 1.3. List of echinostome trematode sequences (ND1 and 28S genes) of in GenBank used for phylogenetic reconstruction. (*R= Redia, *A= Adult)

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinostoma trivolvis	A	*Physa heterostropha*	USA	GQ463049	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670850	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670852	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670854	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670861	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Wisconsin, USA	GQ463051	ND1
Echinostoma trivolvis	R	*Lymnaea elodes*	Minnesota, USA	GQ463113	ND1
Echinostoma trivolvis	R	*Helisoma trivolvis*	Indiana, USA	GQ463048	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Wisconsin, USA	JQ670857	ND1
Echinostoma trivolvis	Unknown	Unknown	North America	AF025831	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670853	ND1
Echinostoma trivolvis	R	*Lymnaea elodes*	Pond A, Indiana	GQ463050	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670851	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670855	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670856	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670858	ND1
Echinostoma trivolvis	Unknown	Unknown	NA	AF025831	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670859	ND1
Echinostoma trivolvis	A	*Ondatra zibethicus*	Virginia, USA	JQ670860	ND1
Echinostoma revolutum	A	*Ondatra zibethicus*	Virginia, USA	GQ463062	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463063	ND1
Echinostoma revolutum	R	*Lymnaea stagnalis*	Czech Republic	KP065594	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Australia	AF026288	ND1
Echinostoma revolutum	R	Planorbid snails	Australia	AF026287	ND1
Echinostoma revolutum	R	*Physa acuta*	Europe	AF025832	ND1
Echinostoma revolutum	R	Planorbid snails	Australia	AF026286	ND1

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463057	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463077	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463079	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Kake, Indiana	GQ463083	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463090	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463076	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463071	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463068	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463060	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Kake, Indiana	GQ463074	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463075	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463057	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463061	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463070	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463065	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463085	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463078	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463080	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463067	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463089	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463059	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Indiana	GQ463058	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Indiana	GQ463062	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Indiana	GQ463082	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Indiana	GQ463084	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463090	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463066	ND1
Echinostoma revolutum	R	*Lymnaea stagnalis*	Czech Republic	KP065650	ND1
Echinostoma revolutum	A	*Aythya fuligula*	Czech Republic	KP065653	ND1
Echinostoma revolutum	R	*Radix peregra*	Iceland	KC618453	ND1

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463072	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463077	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463056	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463073	ND1
Echinostoma revolutum	R	*Lymnaea stagnalis*	Finland	AY168936	ND1
Echinostoma revolutum	R	*Planorbid* snail	UK	AY168937	ND1
Echinostoma revolutum	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463069	ND1
Echinostoma robustum	A	*Anas platyrhynchos*	Bangladesh	LC224105	ND1
Echinostoma robustum	A	*Anas platyrhynchos*	Bangladesh	LC224086	ND1
Echinostoma robustum	R	*Lymnaea elodes*	Minnesota, USA	GQ463054	ND1
Echinostoma robustum	R	*Biomphalaria glabrata*	Brazil	GQ463055	ND1
Echinostoma robustum	R	*Lymnaea elodes*	Pond A, Indiana	GQ463053	ND1
Echinostoma friedi	A	*Mesocricetus auratus*	Valencia, Spain	AJ564379	ND1
Echinostoma miyagawai	A	*Anas platyrhynchos*	Poland	KP065624	ND1
Echinostoma miyagawai	A	*Anas platyrhynchos*	New Zealand	KY436400	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463091	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463092	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463093	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463094	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463095	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463096	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463097	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463098	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463099	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463100	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463101	ND1
Hypoderaeum linegae 1	R	*Lymnaea elodes*	Pond A, Indiana	GQ463102	ND1
Hypoderaeum conoideum	A	*Gallus gallus*	Valencia, Spain	AJ564381	ND1
Hypoderaeum conoideum	R	*Lymnaea peregra*	UK	AY168949	ND1

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Virginia, USA	JQ670863	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Wisconsin, USA	GQ463103	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Wisconsin, USA	GQ463104	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Wisconsin, USA	GQ463105	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Virginia, USA	JQ670864	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Virginia, USA	JQ670864	ND1
Echinoparyphium lineage 1	A	*Ondatra zibethicus*	Virginia, USA	JQ670865	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463105	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463106	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463107	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463108	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463109	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Shock lake, Indiana	GQ463110	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463111	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463112	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463114	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463115	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Shock Lake, Indiana	GQ463116	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463117	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463118	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463119	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463120	ND1
Echinoparyphium lineage 2	R	*Lymnaea elodes*	Pond A, Indiana	GQ463121	ND1
Echinoparyphium lineage 3	R	*Helisoma trivolvis*	Pond A, Indiana	GQ463122	ND1
Echinoparyphium lineage 3	R	*Helisoma trivolvis*	Pond A, Indiana	GQ463123	ND1
Echinoparyphium recurevatum	R	*Lymnea peregra*	Bulgaria	AY168938	ND1
Echinoparyphium recurevatum	R	*Lymnea peregra*	UK	AY168939	ND1
Echinoparyphium recurevatum	R	*Lymnea peregra*	UK	AY168940	ND1
Echinoparyphium recurevatum	R	*Lymnea peregra*	UK	AY168944	ND1
Echinostoma sp.	R	*Glyptophysa sp.*	Australia	AF026279	ND1

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinostoma sp.	R	*Glyptophysa sp.*	Australia	AF026280	ND1
Echinostoma sp.	A	*Glyptophysa sp.*	Australia	AF026281	ND1
Echinostoma sp.	R	lymnaeid snail	Australia	AF026284	ND1
Echinostoma sp.	unknown	unknown	New Zealand	AF026289	ND1
Echinostoma sp.	R	*Hydromys chrysogaster*	Australia	AF026290	ND1
Echinostoma sp. IG	R	*Radix peregra*	Iceland	KC618448	ND1
Echinostoma sp. IG	R	*Radix auricularia*	Germany	KC618449	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065659	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Czech Republic	KP065664	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Czech Republic	KP065666	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Czech Republic	KP065668	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065670	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065671	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065672	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065673	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065674	ND1
Echinostoma sp.	R	*Planorbarius corneus*	Slovakia	KP065675	ND1
Echinostoma sp.	unknown	unknown	Niger	AF025836	ND1
Echinostoma caproni	unknown	unknown	Madagascar	AF025837	ND1
Echinostoma caproni	unknown	unknown	Cameroon	AF025838	ND1
Petasiger sp.	R	*Gyraulus albus*	Germany	KM191810	ND1
Petasiger sp.	R	*Gyraulus albus*	Germany	KM191810	ND1
Petasiger sp.	R	*Planorbis planorbis*	Germany	KM191813	ND1
Petasiger sp.	R	*Planorbis planorbis*	Germany	KM191814	ND1
Petasiger sp.	R	*Gyraulus albus*	Germany	KM191815	ND1
Petasiger sp.	R	*Gyraulus albus*	Germany	KM191817	ND1
Petasiger sp.	R	*Helisoma trivolvis*	Canada	KT831343	ND1
Isthmiophora melis	R	*Helisoma trivolvis*	Canada	AY168948	ND1
Echinostoma paraensei	unknown	unknown	Brazil	AF025834	ND1
Echinoparyphium aconiatum	R	*Lymnaea stagnalis*	UK	AY168945	ND1

Species	Life cycle stages	Host species	Collection sites	Genbank accession number	Gene region
Echinoparyphium aconiatum	R	*Lymnaea stagnalis*	UK	AY168946	ND1
Euparyphium sp.	A	*Rattus norvegicus*	Spain	AJ564380	ND1
Echinostoma hortense	unknown	unknown	unknown	AF025835	ND1
Fasciola hepatica	A	bovine liver	USA	X15613	ND1
Echinostoma trivolvis	A	*Mesocricetus auratus*	UK	AY222246	28S
Echinostoma revolutum	R	*Lymnaea stagnalis*	Czech Republic	KP065594	28S
Echinostoma miyagawai	A	*Anas platyrhynchos*	New Zealand	KY436408	28S
Echinostoma paraulum	A	*Lymnaea stagnalis*	Germany	KP065604	28S
Echinostoma paraulum	A	*Aythya fuligula*	Czech Republic	KP065605	28S
Echinoparyphium cinctum	A	*Anas platyrhynchos*	Ukraine	AF184260	28S
Echinoparyphium cinctum	R	*Planorbella trivolvis*	USA	EF470906	28S
Echinoparyphium cinctum	R	*Planorbella trivolvis*	USA	EF470907	28S
Echinostoma rubrum	R	*Lythobates sylvaticus*	USA	JF820594	28S
Echinostoma rubrum	R	*Phasianus colchicus*	USA	JF820595	28S
Echinostoma bolschewense	R	*Viviparus acerosus*	Slovakia	KP065591	28S
Echinostoma caproni	A	*Mus musculus*	France	AF026104	28S
Echinostoma sp. n.	R	*Planorbarius corneus*	Czech Republic	KP065600	28S
Echinostoma sp. IG	R	*Radix auricularia*	Germany	KP065606	28S
Echinostoma paraensei	A	*Mesocricetus auratus*	USA	EU025867	28S
Hypoderaeum conoideum	R	*Lymnaea stagnalis*	Czech Republic	KP065607	28S
Echinoparyphium ellisi	A	*Anas platyrhynchos*	New Zealand	KY436410	28S
Echinoparyphium poulini	A	*Cygnus atratus*	New Zealand	KY436409	28S
Echinoparyphium recurvatum	R	*Radix ovata*	Slovakia	KT956913	28S
Echinoparyphium ellisi	A	*Anas platyrhynchos*	New Zealand	KY436406	28S
Echinostoma novaezealandense	A	*Anas platyrhynchos*	New Zealand	KY436398	28S
Echinostoma novaezealandense	A	*Anas platyrhynchos*	New Zealand	KY436399	28S
Echinoparyphium poulini	A	*Cygnus atratus*	New Zealand	KY436401	28S
Echinoparyphium poulini	A	*Cygnus atratus*	New Zealand	KY436404	28S
Euparyphium melis	A	*Apodemus agarius*	Poland	KT359583	28S

Supplementary Table 2.1. Final hosts and geographic locations of the 50 specimens included in Principal Component Analysis.

Species name	Number of specimens	Number of specimens from host species	Locations	Several individuals from the same host?
Echinostoma trivolvis lineage a	4	*Anas platyrhynchos* (3)	Manitoba, Canada	Yes, 1-2 per host
		Ondatra zibethicus (1)		NA
Echinostoma trivolvis lineage c	27	*Ondatra zibethicus* (27)	Manitoba, Canada	Yes, 1-2 per host
Echinostoma revolutum	8	*Grus canadensis* (1)	Manitoba, Canada	NA
		Anas platyrhynchos (6)		Yes, 1-2 per host
		Branta canadensis (1)		NA
Echinostoma trivolvis	11	*Buteo platypterus* (1)	Wisconsin, USA	NA
		Gallus gallus (1)	Pennsylvania, USA	NA
		Buteo jamaicensis (1)	Wisconsin, USA	NA
		Ondatra zibethicus (4)	Oregon, USA	Yes (1-2 per host)
		Branta canadensis (1)	Unknown[x]	NA
		Myocastor coypus (2)	Louisiana, USA	Yes (1 host)
		Didelphis virginiana (1)	Nebraska, USA	NA

Appendix B

Data manipulation in R:

```
#How to run Principle Component Analysis (PCA) in R
data <- read.csv(file.choose(), header=T)
#Before running PCA
Normality test of data:
Shapiro.test(data$Inter.testis)
#log 10 base transformation of the data
data <- cbind(data, log(1+data$Inter. testis)
#To remove not normally distributed features
data[,"Inter.testis"]<-NULL
#Removing rows that contain "NA", only used when removing worms that were missing data
data <- na.omit(data)
#Running PCA
library(vegan)
Echinos.pca. <- rda(data, scale = TRUE)
#summary of PCA
summary(Echinos. pca)
#Displaying biplot
biplot(Echinos.pca, choices = c(2,3), scaling = 0)
#For displaying individuals only
biplot(Echinos.pca, choices = c(2,3),display= “sites”)
#For displaying directions of characters only
biplot(Echinos.pca, choices = c(2,3),display= “species”)
# Plotting point data: symbols
Biplot (Echinos.pca, type="n", xlab="PC2 (X% explained var.)", ylab="PC2 (Y% explained var.)")
PCAscores=scores(Echinos.pca)
points(PCAscores$sites[1:15,1],PCAscores$sites[1:15,2],col="black", pch=15, cex=1)
# Adding a legend
Legend( “topleft”,c(“E. trivolvis”, pch= 21, col="black”)
```