USE OF MOLECULAR MARKERS TO INVESTIGATE EPIDEMIOLOGY OF AND HOST MANIPULATION BY THE INVASIVE PARASITE *DICROCOELIUM DENDRITICUM*

BRADLEY VAN PARIDON
Bachelor of Science, University of Lethbridge, 2011

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

USE OF MOLECULAR MARKERS TO INVESTIGATE EPIDEMIOLOGY OF AND HOST MANIPULATION BY THE INVASIVE PARASITE *DICROCOELIUM DENDRITICUM*

BRADLEY VAN PARIDON

Date of defense: March 9th, 2017

	Rank	Highest Degree
Dr. Cameron Goater Co-Supervisor	Professor	PhD
Dr. John Gilleard Co-Supervisor	Professor	MRCVS, PhD
Dr. Douglas Colwell Thesis Examination Committee Member	Research Scientist	PhD
Dr. Theresa Burg Thesis Examination Committee Member	Associate Professor	PhD
Dr. Kevin Floate Internal Examiner	Research Scientist	PhD
Dr. Patrick Hannington External Examiner	Assistant Professor	PhD
Dr. Joe Rasmussen Chair, Thesis Examination Committee	Professor	PhD

ABSTRACT

Parasite distributions are rapidly expanding, in large part due to human mediated translocations of hosts. Thus, instances of parasite-spread to new hosts are common. Management of these emerging parasites depends on understanding the host-shifts occurring within invasive areas and the pathway of invasion; both of which can be investigated using molecular markers. *Dicrocoelium dendriticum* is one such emerging parasite, which has invaded Alberta. Here, I use molecular markers to identify two local Albertan hosts colonized by *D. dendriticum*; the snail *Oreohelix subrudis* and the ant *Formica aserva*. I also developed a panel of nine microsatellite markers that showed two separate introductions of the parasite into Canada. Furthermore, the markers also identified "clumped clonal transmission" throughout the life-cycle, which has implications for the potential evolution of cooperation via kin selection, in this behaviour altering parasite.

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor, Dr. Cameron Goater for introducing me to the world of parasites and research. Your enthusiasm for research and the bizarre world of parasites was infectious and I am grateful for the opportunity to work with you and for introducing me to my eventual co-supervisor Dr. John Gilleard. Thank you John for bringing me into your lab and for all the guidance and mentorship throughout the project. It has been my pleasure to work with and learn from both these great supervisors. Thank you also to members of the Gilleard lab who patiently trained me in the lab techniques I required especially, Drs. Umer Chaudry, Susan Stasiuk and Libby Redman. Thank you to my thesis committee members Dr. Douglas Colwell and Dr. Theresa Burg; your input on this project was invaluable. To my friends, thanks for the many needed breaks and stress relief activities- they were incredibly necessary. To my sister Tracy, thank you for your support and inspiration. To my brother Darrell, thanks for the place to live and all the laughs. Thank you to my fiancée Theresa- you are truly amazing and have put up with a lot during this time. I couldn't have done this without you, thank you for everything. Finally, I am incredibly fortunate to have the support (moral, financial and otherwise) of two amazing parents. Thank you Mom and Dad, you know how important you are to me and this chapter of my life and career are in large part thanks to you. I love you both.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BP	Base pairs
CHP	Cypress Hills Interprovincial Park
COX1	Cytochrome Oxidase 1 mitochondrial gene
DNA	Deoxyribonucleic acid
EID	Emerging infectious disease
GD	Genetic disequilibrium
ID	Identity disequilibrium
LD	Linkage disequilibrium
MLG	Multi locus genotype
mtDNA	Mitochondrial DNA
PCA	Principle component analysis
PCR	Polymerase chain reaction
STP	Staff Trailer Park (sampling site)
TC	Trans Canada Trail (sampling site)
SNP	Single nucleotide polymorphisms

CHAPTER 1: GENERAL INTRODUCTION

1.1 GENERAL OVERVIEW

Within the last two decades a noted increase in the prevalence and expansion in both geographical and host range of disease-causing parasites and pathogens has led to the broad categorization of such organisms as emerging infectious diseases (EIDs) (Daszak et al., 200; Tompkins & Poulin, 2006). EIDs are found globally and are represented by both microparasites (e.g., bacteria, virus, and myxozoan) and macroparasites (e.g., the 'worms' and parasitic arthropods) (Daszak et al., 2000). A number of factors have contributed to this temporal increase in EIDs. Climate change, for example, can increase the amount of suitable habitat for disease-carrying vectors and many other parasites, and alter their distributions such that new host populations are put at risk (Githeko et al., 2000). The same is true of human encroachment into wildlife habitat, which can increase contact rates between wild animals, their pathogens, with domestic animal and human populations, causing the potential for parasite spillover (Thompson et al., 2009). While each of these factors can work in combination to aid the expansion of EIDs, it has been the human-mediated movement of animals and invasive species that has been implicated most in disease emergence events (Daszak et al., 2000; Tompkins & Poulin, 2006). Regardless of the precise process of EID emergence within local host populations, these disease-causing agents can have detrimental impacts on the individuals, populations, and communities that they infect (Daszak et al., 2000; Tompkins & Poulin, 2006).

While it has been thought that introductions of EIDs into new habitats is less likely to occur with macroparasites such as helminths (Cleaveland et al., 2001; Torchin, et al., 2002), this is likely a case of under-reporting, perhaps because these kinds of parasites tend to have less dramatic effects on their hosts (Tompkins & Poulin, 2006). The translocation of fox species within the USA to expand hunting opportunities led to a significant public health concern due to the potential increase in the distribution of the zoonotic cestode *Echinococcus multilocularis* (Storandt et al., 2002). The movement of livestock within and between countries has aided the spread of the liver fluke *Fasciola hepatica*, a parasite that threatens agricultural production and human health globally (Mas-Coma et al., 2009). More dramatically, the human pathogen, *Schistosoma mansoni*, has expanded its range from Asia and Africa and now infects multiple species of snail hosts and threatens human populations in the South American continent (Crellen et al., 2016). With so many examples of the expansion of host and geographical ranges of macroparasites in animal and human populations, it is critical that we understand the processes that lead to successful invasion of parasite species (Araujo et al., 2015).

For any successful species invasion, whether involving parasites or free-living animals, there exists a number of traits that are considered to aid in successful establishment. When faced with a new environment, species with pre-existing traits that allow them to survive in a range of environments are more likely to become established via a process known as ecological fitting (Janzen, 1985). In the case of parasites, this concept is applied to host utilization within new environments and a parasite's ability to capture and exploit new hosts (Agosta et al., 2010). The trematode *Fascioloides magna* provides an example. This

species infects relatively few hosts within its endemic range in North America, but has overcome host species barriers to establish in continental Europe (Malcicka et al., 2015). A second major hurdle to invasive species is overcoming founder effects; i.e., a reduction in population size and loss of genetic diversity when a small founding population first colonizes a new area (Roman & Darling, 2007). A few ways, in which invaders can overcome this hurdle relies on the species having an initial invasive population made up of individuals from multiple sources, known as admixture, or by repeated introductions of new invaders that reinforce the new population (Roman & Darling, 2007). Thus, to understand the processes leading to the invasion of parasite EIDs and to understand their consequences within new host populations, we require the ability to simultaneously identify new hosts being captured by invaders, as well as identify the source and history of invasion; i.e., single vs. multiple introductions. Unfortunately, we often lack information regarding the historical origins of invasive populations and the continued connectivity of these populations with their sources (Rollins et al., 2009). Secondly, identifying the new hosts that invaders could potentially utilize can be difficult as parasites are often small and cryptic, lack easily recognizable morphological traits, and often use multiple hosts to complete their life cycles (Criscione et al., 2005). These difficulties are being increasingly addressed through the use of molecular markers to identify parasites and hosts and to infer historical phylogeography of parasite populations, especially those that have recently emerged in new host populations (Criscione et al., 2005).

1.2 MOLECULAR EPIDEMIOLOGY

The combination of analysis of molecular markers and phylogenetic methods with the field of epidemiology has been termed molecular epidemiology and it has led to an increased ability for researchers to study the transmission and epidemiological history of parasite infections, as well as the environmental or landscape features enhancing spread (Archie et al., 2009). The advent of DNA-based methods to study both parasite ecology (Criscione et al., 2005) and epidemiology (Archie et al., 2009) has led to the establishment of this field and to its increasing sophistication. The enormous variation in parasite life cycles and in parasite life-history strategies imposes a number of constraints on the researchers attempting to study them (Criscione et al., 2005). They are difficult to directly observe, living often within host tissues, and they can lack discernable morphological features that make them difficult to identify (Criscione et al., 2005). A molecular-based approach can resolve some of these issues by using genetic material sampled from an organism to evaluate the magnitude, and significance, of variation between individuals and populations (de Meeus et al., 2007).

The earliest and most widely adopted use for molecular markers in parasitology has been for species identification (McManus & Bowles, 1996). Molecular markers (e.g., ribosomal rDNA sequences) and mitochondrial DNA gene sequences (e.g., cytochrome oxidase one (COX1)) are especially useful for describing inter and intraspecific variation in parasites, which can then be used to discriminate between epidemiologically known species or strains (McManus & Bowles, 1996). DNA-based methods also can be applied to a wide variety of tissue, biological material (e.g., feces or saliva), life cycle stages (e.g., eggs or adults), and environmental samples, making sampling less invasive, easier, and less prone to ethical constraints (Lymbery & Thompson, 2012). By using molecular-based identification,

large surveys of host species can be efficiently screened for the presence of a parasite species of interest, thereby identifying the complete life cycle and host range of the study organism (Criscione et al., 2005).

Molecular methods can also allow the indirect inference of parasite life histories, evolution, and historical patterns of parasite and host dispersal (Criscione et al., 2005). Markers such as short tandemly repeating sections of DNA, known as microsatellites, are highly polymorphic, codominant, meaning both alleles at a single locus can be identified, and they are abundant in the genome (Lehmann & Rousset, 2010). They can be particularly effective in molecular epidemiology due to their ability to quantify variation in allele frequencies between individuals and populations of individuals (de Meeus et al., 2007). The quantification and distribution of this variation among and between populations is known as genetic structure and it is commonly quantified and analyzed using Wrights F-statistics (Wright, 1965). F-statistics are based on the infinite alleles model and assume that populations are all divided into sub-populations and mechanisms such as drift and migration act on allele frequencies within the total and sub-divided populations (de Meeus et al., 2007). In short, they provide a hierarchical measure of allele frequencies at three levels: F_{IS} measures the level of inbreeding of individuals due to the non-random union of gametes within each sub-population, F_{ST} is a measure of inbreeding resulting from subdivision of populations into smaller populations with no migration and is therefore a measure of genetic differentiation between sub-populations, and F_{IT} is a measure of inbreeding in the total population based on the previous two measures (de Meeus et al., 2007). The biological and epidemiological factors that influence allele frequencies can be inferred by quantifying the F-

statistics of targeted parasite populations. Since the application of F-statistics to molecular epidemiology, many other analytical tools such as assignment tests, phylogenetic analysis, and Bayesian approximation have been developed to study the epidemiology of parasite populations (Lymbery & Thompson, 2012).

1.2.1 Host identification and host range

The first challenge facing parasites emerging within new environments is the availability of suitable hosts. For parasites with multi-host life cycles, host availability can present a significant challenge, especially if the degree of host specificity is high, or the ability of parasites to infect multiple hosts is low. Early hypotheses proposed that parasites co-evolve with their hosts over evolutionary time in a manner that ultimately leads to specialization in, or on, a single host species (reviewed by Klassen, 1992). However, the co-evolutionary paradigm for the development of intimate host-parasite interactions leading to tight co-phylogenies between hosts and parasites has been challenged by the observation that host-switching events have occurred frequently in evolutionary and contemporary time (Hoberg & Brooks, 2008). One way in, which opportunities for host-switching can occur, is the alteration of host habitat or food webs and through ecological factors such as glaciation or range expansions that subsequently place new hosts in contact with novel parasites (Hoberg & Brooks, 2008). Human-mediated changes to environment will also provide opportunities for host-switching at shorter time scales (Patz et al., 2000; Hoberg et al., 2008). Thus, according to this framework, the rapid emergence of EIDs can be explained by anthropogenic

factors that lead to increased rates of host-switching that are not constrained by tight co-evolutionary interactions.

The use of molecular methods in parasite identification was one of the first molecular epidemiology practices to be widely adopted by parasitologists (McManus & Bowles, 1996) and their use on parasites has since been applied to epidemiological questions (Lymbery & Thompson, 2012). Molecular identification of parasites is an appropriate tool because it can distinguish stages that are morphologically indistinguishable (i.e., eggs, larvae) and these are often the most available stages and require non-invasive methods for sampling (Criscione et al., 2005). DNA-based identification was used to show that dogs were the source of zoonotic hookworm infections by identifying the shared species of hookworms from eggs shed in feces of humans and dogs (Traub et al., 2008). These tools are also invaluable in identifying cryptic species, those, which are morphologically identical, but genetically distinct (Grillo et al., 2007). Cryptic species can be found within the same host leading researchers to conclude a single infection when in reality there are multiple (Criscione & Blouin, 2004). The discovery of life cycles is another potential use for molecular markers (Criscione et al., 2005). This will play an important role when parasites have emerged into new areas and the hosts are unknown. Combinations of traditional and molecular methods were used to identify new hosts for pentastomid parasites introduced to Australia (Kelehear et al., 2014). Pentastomids represent a significant zoonotic threat and the identification of hosts being used in Australia was not previously understood making this finding critical to establishing the life cycle on the continent (Kelehear et al., 2014).

1.2.2 Population genetics of invasive parasites

Predictive population genetics models such as Hardy-Weinberg Equilibrium (HWE) are used to describe allele frequencies in natural populations in the absence of those forces known to influence allele frequencies (de Meeus et al., 2007). Population geneticists can then evaluate observed allele frequencies in populations and compare them with HWE predictions and hence infer the mechanisms behind any observed differences (de Meeus et al., 2007). Parasites, however, are a special case because they do not readily adhere to HWE assumptions due to unique biological and ecological factors of parasite life style (Gorton et al., 2012). Therefore, special considerations must be made when attempting to infer the mechanisms behind the observed distribution of alleles across parasite populations (Criscione et al., 2011).

Factors like parasite life cycles, transmission and life history will lead to a vast departure from the expectations of HWE and theoretical population models (Gorton et al., 2012). Parasitic trematodes provide an excellent example of these factors. Their life cycles are often complex (e.g., use multiple hosts) and are characterised by obligate alternation between asexual and sexual modes of reproduction. The asexual phase occurs in the first intermediate host, usually a mollusc, leading to the production of many genetic clones, which are then transmitted to subsequent hosts in the life cycle. Sexual reproduction then occurs in the final host. A further complication arises from the fact that most trematode species are hermaphroditic and can undergo selfing during the adult stage. The combination of clonal propagation, transmission between multiple hosts, and the potential for self-mating greatly

influences the distribution of alleles and genetic structure of a population (Prugnolle et al., 2005). Therefore, when assessing results from population genetic analyse in parasites, such as deviations from HWE, special consideration of life history must be accounted for (Prugnolle et al., 2005; Gorton et al., 2012). For example, the sampling of clonal individuals, created in the asexual phase and transmitted to the final host, will artificially inflate measures of F_{ST} (Prugnolle et al., 2005). In this case, researchers must identify the presence of clones and assess their effect on measures of F_{ST} before drawing conclusions (Criscione et al., 2011).

Factors related to the emergence of parasite populations into new hosts and environments will also affect genetic structure and therefore can be inferred from molecular epidemiological studies. First, when any population expands into novel locations, there is an expected drop in genetic diversity from founder effect or bottleneck because translocated individuals only comprise a fraction of the total genetic diversity of source populations (Dlugosch & Parker, 2008). However, one way, in which introduced populations can overcome bottlenecks, is by multiple introductions from a variety of sources (Dlugosch & Parker, 2008). Therefore, determining the level of genetic diversity between an invasive population and its putative source populations is one way to assess if the introduction is recent and of single origin. This information will be valuable to management strategies as continued introductions may help the invasive population succeed in establishing in its new environment (Estoup & Guillemaud, 2010). The presence of a genetic bottleneck was provided as evidence of a small founding population of the parasitic fly *Philornis downsi* to the Galapagos Islands (Dudaniec et al., 2008). The authors further showed little genetic

structure between islands and concluded the fly spread within the island chain; this helped inform a control strategy by showing flies on islands could be considered a single population (Dudaniec et al., 2008). This example shows that not only are the levels of genetic diversity affected by the nature of the introduction, but also the genetic structure.

Testing for genetic structure of parasites can provide important information about transmission that is not directly observable (Gorton et al., 2012). The presence of structure between parasite populations could be caused by a number of barriers to gene flow between them (Biek & Real, 2010). The separation of host populations by environmental factors like mountains or rivers could prevent the spread of parasites (Biek & Real, 2010) conversely, some presumed barriers such as host species may in fact not constrain dispersal (Gorton et al., 2012). A study of genetic structure between *Trichostrongylus axei* parasites infecting six species of sympatric wild ungulate hosts in North America revealed no structure, demonstrating that there was no barrier to transmission of *T. axei* amongst these hosts (Archie & Ezenwa, 2011). The presence or absence of barriers to gene flow then inform which populations are at risk and how host movements and sympatry may be exacerbating spread. Similar analysis undertaken to compare populations across much larger geographical distances can also identify source populations and indicate connectivity possibly by continued introductions from putative sources. The use of multiple markers, mitochondrial DNA and microsatellites, has identified the source of *Fascioloides magna* invasion into Europe and described its subsequent spread (Juhásová et al., 2016). Identifying the source populations can become more difficult, however, in the case of admixed invasive populations. Methods such as Bayesian clustering or multivariate analyses such as Principle

Components Analysis (PCA) can then be used to assign individuals to populations or groups based on patterns of genetic variation (Manel et al., 2003). In this way, a population of mixed origin can be separated out into genetic groups rather than grouping based on sampling location.

Molecular markers can also identify closely-related individuals and clonal individuals by identifying multi locus genotypes (MLG) for each individual. For microsatellites, a MLG is the complete set of alleles present at all loci tested. Statistical tests can then be applied to determine if identical MLGs in a population are likely to have arisen by chance or by asexual reproduction (Gregorius, 2005). This is important for assessing the impacts of life cycle type and life histories on genetic structure. The genotyping of the trematode parasites in the first intermediate snail host and second intermediate fish host showed that clonal parasites are less likely to be found co-infecting fish than snails (Rauch et al., 2005). This mixing of clonal genotypes in the second intermediate host reduces genetic structure in the parasite population and prevents inbreeding between clones in the next host (Criscione & Blouin, 2004). It can, however, also investigate unique aspects of parasite biology, e.g., the evolution of altruism among related parasites via kin selection. For example, larvae of some parasite species are known to manipulate the phenotypes of hosts in order to increase their rates of transmission to the next host (Hughes et al., 2012). Well known examples include the deformation of frog limbs by larvae of the trematode *Riboreia* sp., an alteration that increases rates of transmission to final avian hosts that prey on handicapped frogs (Johnson et al., 1999). Also, the conspicuous migration of larvae of the trematode, *Leucochloridium* spp. to the eye stalks of snails, creating a caterpillar mimic that encourages predation by final avian hosts

(Wesołowska & Wesołowski, 2014). In some cases of parasite-induced alteration in host phenotype, only a small number of individuals actively engage in the process of manipulation and can incur a cost for doing so (Poulin et al., 2005). Encysted larvae of the marine trematode, *Curtuteria australis*, reside within the foot of the cockle, *Austrovenus stutchburyi* (Poulin et al., 2005). Individual larvae located within the tip of the foot impede burrowing behaviour into the substrate and this exposes the host to predation by the definitive avian hosts. However, this impediment also exposes only the parasites at the tip of the foot, causing the manipulation, to predation by foot-cropping fish predators, in which the parasites cannot establish. This example demonstrates that the cost of behavioural manipulation is not shared by the other parasites encysted throughout the foot, but the benefit is. Analysis of genetic relatedness and the extreme case of this clonality on those incurring the cost of manipulation vs those benefiting would provide evidence for the presence of kin selection and thus altruistic-dependant transmission among manipulative parasites.

1.3 MODEL SYSTEM

The trematode *Dicrocoelium dendriticum* provides an excellent model, in which to study the processes and traits involved in the emergence of parasitic EIDs within new habitats and new hosts. It is also an iconic manipulator of host phenotype (Hohorst, 1962; Spindler et al., 1986), in which altruism among individuals within individual hosts has frequently been invoked (Mehlhorn, 2015). This fluke has expanded its range from areas in central Europe throughout Europe, Asia, and North Africa and most recently to isolated sites in North America (Soulsby, 1982; Otranto & Traversa, 2002). Adult *D. dendriticum* utilize a

broad range of mammalian hosts where it resides along the walls of the bile ducts in the liver and within the gall bladder. It is most often found in domestic and wild ruminants, but has also been documented in rabbits, pigs, dogs, horses and camelids (Otranto & Traversa, 2002). Pathology associated with infection of adult flukes occurs due to liver damage resulting in liver fibrosis, cholangitis, anaemia and reduced weight gain (Manga-Gonzalez et al., 2001). At high intensities (>4000 worms), liver dysfunction associated with alteration to bilirubin and albumen concentrations, and alteration of hepatic enzyme concentrations have also been documented (Manga-Gonzalez & Gonzalez-Lanza, 2005).

Larval *D. dendriticum* have been reported from approximately 90 species of terrestrial snails (Manga-Gonzalez et al., 2001). Snails ingest eggs that have been deposited on substrate in host feces. Following ingestion, repeated bouts of asexual reproduction lead to the production of numerous sporocysts, within which multiple free-swimming cercariae develop. Cercariae are then expelled via the snail respiratory pore encased within host mucous (slime balls). Slime balls are expelled onto the substrate, where they are available for ingestion by foraging ants (*Formica* spp.). Cercariae inside the ants lose their tails, penetrate the gut wall, and make their way to the host's hemocoel where they are enveloped by a host-derived cyst wall. However, one of the cercaria, following penetration of the gut wall, makes its way to the sub-esophageal ganglion at the anterior-ventral most region of the ant's brain (Hohorst, 1962). This so called "brain worm" is presumed to be responsible for the altered behaviour displayed by infected ants (Hohorst, 1962; Tarry, 1969; Spindler et al., 1986; Manga-Gonzalez & Gonzalez-Lanza, 2005). Infected ants ascend vegetation surrounding their nest and attach to it by their mandibles. Infected ants tend to spend the cool parts of the

day (dusk, evening, and dawn) attached to the plant, usually to the petals of a flower, but then detach when temperatures rise above approximately 18°C. It is presumed that attachment to plant material facilitates transmission into grazing mammals. Following ingestion of infected ants by grazing mammals, adults emerge in the duodenum and eventually settle in the bile ducts of the liver and gall bladder (reviewed by Manga-Gonzalez et al., 2001; Otranto & Traversa, 2002).

Dicrocoelium dendriticum is considered to have been imported to North America along with domestic animals after it was first found in the livers of sheep and white-tail deer collected from pastures in central and northern New York State in the 1950's (Mapes, 1951). Since this initial report, only sporadic reports have documented its spread west, first in cattle from mainland British Columbia (Lewis, 1974) and then in wild ungulates from southern Alberta (Pybus, 1990). Results from the Alberta survey of livers from >100 hunter-shot wild ungulates showed that the prevalence of *D. dendriticum* was <1% (Pybus, 1990). All of these infections were found in ungulates collected within Cypress Hills Interprovincial Park (CHP) in southeastern Alberta. By the mid-1990's, prevalence in CHP had increased to approximately 80% within samples of elk, white tail deer, mule deer and beef cattle (Goater & Colwell, 2007; Beck et al., 2014). This increase in prevalence within CHP demonstrates that this parasite is established within the region, where it presents a unique opportunity to study the dynamics of an emerging parasite.

The introduction of *D. dendriticum* into western Canada represents one of the few known sites in North America where the complete life cycle has been confirmed to have been

established. This means that the identification of local hosts is possible and that adult flukes are available via hunters and livestock producers for detail population genetics analyses. Secondly, Beck (2015) has identified multiple sites within the CHP where infected ants can routinely be found clinging to vegetation. The availability of larval *D. dendriticum* in ants allows, for the first time, an investigation into the genetic structure of metacercariae within terrestrial second intermediate hosts and more specifically, a test of the kin selection hypothesis for the evolution of behaviour manipulation by parasites. Finally, material from CHP in conjunction with the availability of *D. dendriticum* samples from colleagues across Canada and Europe will allow for the use of molecular tools to investigate the patterns of distribution and invasion of this EID into western Canada and specifically Alberta.

1.4 THESIS OBJECTIVES

Characterizing the life cycle of EIDs within sites of emergence is fundamental to mitigation efforts, and also to our general understanding of fundamental epidemiological rate parameters. Chapter 2 employs the use of molecular markers to identify the first and second intermediate hosts utilized by *D. dendriticum* in CHP. By doing so, I provide the first description of the complete life cycle of this emerging parasite in western North America. In this chapter, I also survey local populations of snails and ants to characterize patterns of host utilization and to determine spatial-temporal patterns of infection within intermediate hosts.

The selection and utilization of appropriate genetic markers is a fundamental requirement of studies in population genetics. In Chapter 3, I develop and validate a panel of microsatellite markers on adult *D. dendriticum* in CHP as a foundation for subsequent studies

in Chapters 4 and 5, and for future population genetic analyses involving this host/parasite interaction. Screening for potential loci and then tests for technical issues such as null alleles and linked loci were also undertaken. I then tested how life history traits of *D. dendriticum* such as self-mating and the non-random transmission of clones may impact the results of tests for population genetic structure. This provides a way forward for use of these markers and a protocol to minimize bias of results due to technical and/or biological factors as well as provide insight into the transmission of flukes within CHP.

In Chapter 4, a population genetics approach is used to investigate the distribution and invasion history of *D. dendriticum* into western Canada. Data collected from multiple populations of *D. dendriticum* in Canada and Europe by two molecular marker systems, nuclear microsatellite loci and mitochondrial cytochrome oxidase 1 gene sequences (COX1), were used to describe patterns of genetic diversity and population genetic structure. Fluke populations from 12 different regions (five from Canada and seven from Europe) were tested for genetic diversity and potential bottlenecks under the assumption of three alternative models: the infinite allele model (Kimura & Crow, 1964), the two-phase model (Di Rienzo et al., 1994), and the stepwise mutation model (Chakraborty & Nei, 1977). Tests for genetic structure between populations (F_{ST}) were also used to identify potential movement of flukes between populations. This was supported by clustering analysis to group populations based on genetic similarity, regardless of geographic origin. To strengthen the resolution of observed patterns, I developed a phylogenetic network (Bandelt et al., 1999) of the distribution of COX1 haplotypes and the presence of shared haplotypes between regions to infer potential sources and routes of invasion.

The same panel of microsatellite markers was then employed in Chapter 5 to characterize, for the first time, the genetic structure of metacercariae populations within individual ant hosts. The results of similar approaches involving metacercariae populations have been completed on aquatic hosts, but not on terrestrial hosts (Gorton et al., 2012). More specifically, in this chapter I test whether clonal metacercariae are infecting the same second intermediate host and whether the individuals located in the brain of the ant are clonemates with those in the haemocoel. These data are used to test the kin selection hypothesis for behaviour manipulation and transmission in this system. These analyses also open the door to comparisons on life history and genetic structuring with other trematode systems.

1.5 Literature Cited

Agosta, S. J., Janz, N., & Brooks, D. R. (2010). How specialists can be generalists: resolving the "parasite paradox" and implications for emerging infectious disease. *Zoologia, 27*, 151-162.

Araujo, S. B., Braga, M. P., Brooks, D. R., Agosta, S. J., Hoberg, E. P., von Hartenthal, F. W., & Boeger, W. A. (2015). Understanding host-switching by ecological fitting. *Plos One, 10*, e0139225.

Archie, E. A., & Ezenwa, V. O. (2011). Population genetic structure and history of a generalist parasite infecting multiple sympatric host species. *International Journal for Parasitology, 41*, 89-98.

Archie, E. A., Luikart, G., & Ezenwa, V. O. (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology & Evolution, 24*, 21-30.

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16*, 37-48.

Beck, M. A. (2015). Ecological epidemiology of an invasive host generalist parasite, *Dicrocoelium dendriticum*, in Cypress Hills Interprovincial Park, Alberta. (Ph.D.), University of Lethbridge, Lethbridge, Alberta.

Beck, M. A., Goater, C. P., Colwell, D. D., & van Paridon, B. J. (2014). Fluke abundance versus host age for an invasive trematode (*Dicrocoelium dendriticum*) of sympatric elk and beef cattle in southeastern Alberta, Canada. *International Journal for Parasitology: Parasites and Wildlife,3,* 263-268.

Biek, R., & Real, L. A. (2010). The landscape genetics of infectious disease emergence and spread. *Molecular Ecology, 19*, 3515-3531.

Chakraborty, R., & Nei, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution, 31*, 347-356.

Cleaveland, S., Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*, 991-999.

Crellen, T., Allan, F., David, S., Durrant, C., Huckvale, T., Holroyd, N., Emery, A.M., Rollinson, D., Aanensen, D.M., Berriman, M., Webster, J.P., Cotton, J.A. (2016). Whole genome resequencing of the human parasite Schistosoma mansoni reveals population history and effects of selection. *Scientific reports, 6*. 20954

Criscione, C. D., & Blouin, M. S. (2004). Life cycles shape parasite evolution: Comparative population genetics of salmon trematodes. *Evolution, 58*, 198-202.

Criscione, C. D., Poulin, R., & Blouin, M. S. (2005). Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology, 14*, 2247-2257.

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. *Molecular Ecology, 20*, 2510-2524.

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science, 287*, 443-449.

de Meeus, T., McCoy, K. D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Bousses, S., & Renaud, F. (2007). Population genetics and molecular epidemiology or how to "debusquer la bete". *Infection Genetics and Evolution, 7*, 308-332.

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., & Freimer, N. B. (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Science U S A, 91*, 3166-3170.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecolog,y, 17*, 431-449.

Dudaniec, R. Y., Gardner, M. G., Donnellan, S., & Kleindorfer, S. (2008). Genetic variation in the invasive avian parasite, *Philornis downsi* (Diptera, Muscidae) on the Galápagos archipelago. *BMC ecology, 8*, 13.

Estoup, A., & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology, 19*, 4113-4130.

Githeko, A. K., Lindsay, S. W., Confalonieri, U. E., & Patz, J. A. (2000). Climate change and vector-borne diseases: a regional analysis. *Bulletin of the World Health Organization, 78*, 1136-1147.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: *Dicrocoelium dendriticum* in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93*, 491-494.

Gorton, M. J., Kasl, E. L., Detwiler, J. T., & Criscione, C. D. (2012). Testing local-scale panmixia provides insights into the cryptic ecology, evolution, and epidemiology of metazoan animal parasites. *Parasitology, 139*, 981-997.

Gregorius, H. R. (2005). Testing for clonal propagation. *Heredity, 94*, 173-179.

Grillo, V., Jackson, F., Cabaret, J., & Gilleard, J. S. (2007). Population genetic analysis of the ovine parasitic nematode *Teladorsagia circumcincta* and evidence for a cryptic species. *International Journal for Parasitology, 37*, 435-447.

Hoberg, E. P., & Brooks, D. R. (2008). A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. *Journal of Biogeography, 35*, 1533-1550.

Hoberg, E. P., Polley, L., Jenkins, E. J., & Kutz, S. J. (2008). Pathogens of domestic and free-ranging ungulates: global climate change in temperate to boreal latitudes across North America. *Revue Scientifique Et Technique-Office International Des Epizooties, 27*, 511-528.

Hohorst, W. (1962). The role of ants in the developmental process of the lancet-shaped fluke (Dicrocoelium dendriticum). *Zeitschrift fur Parasitenkunde, 22*, 105-106.

Hughes, D. P., Brodeur, J., & Thomas, F. (2012). *Host manipulation by parasites*: Oxford, UK: Oxford University Press.

Janzen, D. H. (1985). On Ecological Fitting. *Oikos, 45*, 308-310.

Johnson, P. T., Lunde, K. B., Ritchie, E. G., & Launer, A. E. (1999). The effect of trematode infection on amphibian limb development and survivorship. *Science, 284*, 802-804.

Juhásová, L., Králová-Hromadová, I., Bazsalovicsová, E., Minárik, G., Štefka, J., Mikulíček, P., Pálková, L., Pybus, M. (2016). Population structure and dispersal routes of an invasive parasite, *Fascioloides magna*, in North America and Europe. *Parasites & Vectors, 9*, 547.

Kelehear, C., Spratt, D. M., O'Meally, D., & Shine, R. (2014). Pentastomids of wild snakes in the Australian tropics. *International Journal for Parasitology: Parasites and Wildlife, 3*, 20-31.

Kimura, M., & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics, 49*, 725-738.

Klassen, G. J. (1992). Coevolution: A history of the macroevolutionary approach to studying host-parasite associations. *The Journal of Parasitology, 78*, 573-587.

Lehmann, L., & Rousset, F. (2010). How life history and demography promote or inhibit the evolution of helping behaviours. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 365*, 2599-2617.

Lymbery, A. J., & Thompson, R. C. A. (2012). The molecular epidemiology of parasite infections: Tools and applications. *Molecular and Biochemical Parasitology, 181*, 102-116.

Malcicka, M., Agosta, S. J., & Harvey, J. A. (2015). Multi level ecological fitting: indirect life cycles are not a barrier to host switching and invasion. *Global change biology, 21*, 3210-3218.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution, 18*, 189-197.

Manga-Gonzalez, M. Y., & Gonzalez-Lanza, C. (2005). Field and experimental studies on *Dicrocoelium dendriticum* and dicrocoeliasis in northern Spain. *Journal of Helminthology, 79*, 291-302.

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123*, 91-114.

Mapes, C. R. (1951). Studies on the biology of Dicrocoelium-dendriticum (Rudolphi, 1819) Looss, 1899 (Trematoda, dicrocoeliidae), including its relation to the intermediate host, *Cionella-lubrica* (Muller) .1. A study of *Dicrocoelium dendriticum* and *Dicrocoelium* infection. *Cornell Veterinarian, 41*, 382-432.

Mas-Coma, S., Valero, M. A., & Bargues, M. D. (2009). *Fasciola*, Lymnaeids and human fascioliasis, with a global overview on disease transmission, epidemiology, evolutionary genetics, molecular epidemiology and control. *Advances in Parasitology, 69*, 41-146

McManus, D., & Bowles, J. (1996). Molecular genetic approaches to parasite identification: their value in diagnostic parasitology and systematics. *International Journal for Parasitology, 26*, 687-704.

Mehlhorn, H. (2015). The Brain Worm Story. In H. Mehlhorn (Ed.), *Host Manipulations by Parasites and Viruses* (pp. 101-108). Heidleberg, Germany: Springer International Publishing.

Otranto, D., & Traversa, D. (2002). A review of dicrocoeliosis of ruminants including recent advances in the diagnosis and treatment. *Veterinary Parasitology, 107*, 317-335.

Patz, J. A., Graczyk, T. K., Geller, N., & Vittor, A. Y. (2000). Effects of environmental change on emerging parasitic diseases. *International Journal for Parasitology, 30*, 1395-1405.

Poulin, R., Fredensborg, B. L., Hansen, E., & Leung, T. L. F. (2005). The true cost of host manipulation by parasites. *Behavioural Processes, 68*, 241-244.

Prugnolle, F., Liu, H., de Meeus, T., & Balloux, F. (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology, 35*, 255-263.

Pybus, M. J. (1990). Survey of hepatic and pulmonary helminths of wild cervid in Alberta, Canada. *Journal of Wildlife Diseases, 26*, 453-459.

Rauch, G., Kalbe, M., & Reusch, T. B. (2005). How a complex life cycle can improve a parasite's sex life. *Journal of Evolutionary Biology, 18*, 1069-1075.

Rollins, L. A., Woolnough, A. P., Wilton, A. N., Sinclair, R., & Sherwin, W. B. (2009). Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Molecular Ecology, 18*, 1560-1573.

Roman, J., & Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution, 22*, 454-464.

Soulsby, E. (1982). Helminths, arthropods and protozoa of domesticated animals. London, UK: Baillière Tindall & Cassell Ltd.

Spindler, E. M., Zahler, M., & Loosfrank, B. (1986). Behavioral-aspects of ants as 2nd intermediate hosts of *Dicrocoelium dendriticum*. *Zeitschrift Fur Parasitenkunde-Parasitology Research, 72*, 689-692.

Storandt, S., Virchow, D., Dryden, M., Hygnstrom, S. E., & Kazacos, K. (2002). Distribution and prevalence of *Echinococcus multilocularis* in wild predators in Nebraska, Kansas, and Wyoming. *Journal of Parasitology, 88*, 420-422.

Tarry, D. W. (1969). *Dicrocoelium dendriticum*: The life cycle in Britain. *Journal of Helminthology, 43*, 403-416.

Thompson, R. C. A., Kutz, S., & Smith, A. (2009). Parasite zoonoses and wildlife: Emerging issues. *International Journal of Environmental Research and Public Health, 6*, 678-693.

Tompkins, D. M., & Poulin, R. (2006). Parasites and Biological Invasions. In R. B. Allen & W. G. Lee (Eds.), *Biological Invasions in New Zealand* (pp. 67-84). Berlin, Heidelberg: Springer Berlin Heidelberg.

Torchin, M. E., Lafferty, K. D., & Kuris, A. M. (2002). Parasites and marine invasions. *Parasitology, 125*, 576-576.

Traub, R. J., Inpankaew, T., Sutthikornchai, C., Sukthana, Y., & Thompson, R. A. (2008). PCR-based coprodiagnostic tools reveal dogs as reservoirs of zoonotic ancylostomiasis caused by *Ancylostoma ceylanicum* in temple communities in Bangkok. *Veterinary parasitology, 155*, 67-73.

Wesołowska, W., & Wesołowski, T. (2014). Do *Leucochloridium* sporocysts manipulate the behaviour of their snail hosts? *Journal of Zoology, 292*, 151-155.

Wright, S. (1965). The interpretation of population structure by F-Statistics with special regard to systems of mating. *Evolution, 19*, 395-420.

CHAPTER 2

Life cycle, host utilization, and ecological fitting for invasive lancet liver fluke, *Dicrocoelium dendriticum*, emerging in southern Alberta, Canada

Bradley J. van Paridon[1], John S. Gilleard[2], Douglas D. Colwell[3], and Cameron P. Goater[1]

This chapter was published in the Journal of Parasitology (2017, Epub ahead of print)

[1]Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada. T1K 3M4.

[2]Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, Alberta, Canada. T2N 4Z6.

[3]Agriculture and Agri-Food Canada, Lethbridge Research and Development Station, Lethbridge, Alberta Canada, T1J 4B1

2.1 ABSTRACT

The expansion of parasite distributions outside of their native host and geographical ranges has occurred repeatedly over evolutionary time. Contemporary examples include emerging infectious diseases (EIDs), many of which pose threats to human, domestic animal, and wildlife populations. Theory predicts that parasites with complex life cycles will be rare as EIDs due to constraints imposed by host specialization at each life-cycle stage. In contrast to predictions of this theory, we report two new intermediate hosts in the 3-host life cycle of the liver fluke *Dicrocoelium dendriticum* in Cypress Hills Interprovincial Park, Alberta, Canada. Results of sequence analyses of the cytochrome oxidase 1 (COX1) mitochondrial gene identified the terrestrial snail *Oreohelix subrudis*, and the ant, *Formica aserva*, as first and second intermediate hosts, respectively, in the region. Neither of these intermediate hosts, nor their suite of domestic and wild mammalian grazers used in the life cycle, occurs within the native range of *D. dendriticum* in Europe. Our results from host surveys show that the prevalence of *D. dendriticum* in samples of *O. subrudis* varied between 4-10%, whereas mean metacercariae intensity in *F. aserva* varied in-between 33 to 41 (n = 163, mean ± s.d. = 38 ± 35). These results are the first to describe the complete life cycle of emerging lancet fluke in western North America. The process of multi-level ecological fitting, in which the lancet fluke possesses pre-existing traits to utilize host resources, rather than host species, at each life-cycle stage, provides a mechanism for the establishment of this complex life cycle in a novel habitat and in novel hosts.

2.2 INTRODUCTION

Emerging infectious diseases (EIDs) are pathogens and parasites that are currently undergoing geographical or host-species range extensions, or are currently increasing their patterns of occurrence within their host populations (Daszak et al.,2000; Tompkins & Poulin, 2006). There is general consensus among ecologists, parasitologists, and epidemiologists that parasite and pathogen range extensions have increased over the past several decades due to factors such as climate change, wildlife encroachment, and the anthropogenic movement of livestock and wildlife (Daszak et al., 2000). Emerging parasites can pose risks to human, domestic animal, and wildlife populations. One fundamental requirement for the mitigation of such risks is an understanding of the complete life-cycle and patterns of host utilization of EIDs within regions of emergence. In the absence of key life-cycle information, predictions regarding the potential for further expansion into new areas and into new hosts are exceedingly difficult to make.

Most instances of EIDs involve microparasites such as viruses, protozoa, and bacteria. Cleaveland et al. (2001) noted that of the > 200 species of emerging pathogens or parasites reported from humans and domestic animals, only 6-7% were macroparasites; i.e., helminths. These authors reasoned that helminths are unlikely to emerge because they are relatively complex organisms, often with complex life cycles and long generation times, compared to microparasites. One extension of this argument is that emergence will be rarer for helminths, because more life-cycle stages are incorporated into the life-cycle. For example, emergence of a typical trematode with an obligate three host life-cycle should be rare because conceivably, transmission would require establishment into new molluscan first intermediate hosts, new second intermediate hosts, and new definitive hosts. If we hold to the traditional

viewpoint that host-parasite associations are intimate, highly specialized interactions between two coevolving partners, then the host-switching events that would be required for establishment of this typical life cycle would be severely constraining, and presumably, rare.

Yet many parasites that have complex life-cycles have expanded their geographic and host-species distributions over evolutionary and ecological time scales (review by Hoberg & Brooks, 2015). Contemporary examples include the cestode *Echinococcus granulosis* that emerged in Australia from Europe (Thompson et al., 2010), the trematode *Schistosoma mansoni* that emerged from its homeland in Asia and Africa into South America (Crellen et al., 2016), the emergence of the North American lung fluke, *Haematolechus floedae*, in leopard frogs in Costa Rica (Brooks et al., 2006), and the emergence of the liver fluke *Fascioloides magna* into Europe from western North America (Malcicka et al., 2015). In each of these cases, establishment in new hosts was not constrained by the time required for intimate co-evolutionary interactions to occur within each life-cycle stage. This is not to say that co-evolutionary interactions are not important in the development of novel host-parasite associations, only that they are not required for their initial establishment.

Ecological fitting can explain the rapid establishment of parasites within new geographical locations and new hosts (Araujo et al., 2015). First invoked in the context of free-living organisms, Janzen (1985) proposed that certain defense traits of plants could retain their effectiveness against attacks by new species of insects, even if the plants no longer co-occurred with the insects that selected for the defense. In the case of parasites, host shifts can occur rapidly within new hosts if parasites track resources that are similar to those

within ancestral host, and not necessarily a particular host species. The probability that host shifts will occur will increase if the resource being tracked is widespread, or if a parasite is phenotypically plastic in its ability to utilize particular resources. For complex life-cycle parasites, shifts into new hosts can therefore occur via ecological fitting if larval and adult stages track ancestral resources when they become available.

Of the four known species of trematode in the genus *Dicrocoelium* (Trematoda: Dicrocoeliidae), the lancet fluke, *D. dendriticum*, has the widest geographical and host species range (review by Manga-Gonzalez et la., 2001). The obligate 3-host life cycle of *D. dendriticum* involving terrestrial snails as first intermediate host and formicid ants as second intermediate host, was first elucidated by Krull and Mapes (1952) for snails, ants and grazing mammals in the north eastern U.S.A. Since this initial life-cycle description, evidence from field surveys has indicated that putative *D. dendriticum* sporocysts are present in a wide range of land snails (Pulmonata, Stylommatophora), whereas metacercariae and adults have been reported from numerous species of formicid ant (Manga-Gonzalez et al., 2001), and domestic and wild mammalian grazers (Otranto & Traversa, 2002), respectively.

The lancet fluke has extended its geographical and host-species range beyond its native range in Central Europe. However, little is known regarding its life cycle, epidemiology, or patterns of host utilization within regions of emergence. Krull and colleagues reported adults from sheep and white-tail deer collected from pastures in central and northern New York State in the 1950s, reasoning that the worms originated from the importation of domestic sheep from central Europe (Mapes, 1951). Evidence regarding its

subsequent spread within the continent has been limited to sporadic reports of adult flukes in livers of sheep, cattle, and wildlife from British Columbia and Alberta (Lewis, 1974; Pybus, 1990). A survey of livers from >100 wild ungulates collected from western Canada in the 1990s showed that the prevalence of *D. dendriticum* was <1% (Pybus, 1990). The small numbers of infected animals reported in the latter study were found in samples of grazing mammals collected within Cypress Hills Interprovincial Park (CHP), in southeastern Alberta, Canada. Within this region, prevalence in final hosts increased to approximately 80% by the mid-1990's, with intensities frequently exceeding 1,000 worms/host (Goater & Colwell, 2007; Beck et al., 2014). This combination of results indicates that the introduction of *D. dendriticum* into CHP occurred prior to the 1980's, persisted within the region for an unknown period of time, and then emerged over the last 20-30 yr. The life-cycle of *D. dendriticum* at sites in North America is not known, other than at the sites originally studied by Krull and Mapes (1952) in New York State.

A full test of the ecological fitting verses coevolution models for the establishment of *D. dendriticum* in this region of emergence is outside the scope of the present report. Here, our focus is on the first description of the complete life-cycle of this emerging parasite to establish proof of principle that the requirements for ecological fitting exist within these host communities. First, we use molecular barcodes of parasite DNA to verify the occurrence of larval *D. dendriticum* in snail and ant intermediate hosts in CHP. We also analyze host DNA barcodes to identify the species of snail and ant that are utilized in this site of emergence in North America. Lastly, we report the results of field surveys of samples of snail and ant intermediate hosts on general patterns of host utilization within this region.

2.3 MATERIALS AND METHODS

2.3.1 Cypress Hills Interprovincial Park

CHP is situated on the southern Canadian plains (49°30'N, 110°W) covering an area of approximately 2,590 km^2 (Newsome & Dix, 1968). The hills have a maximum elevation of 1,420 m and comprise a plateau of pre-glacial bedrock that rises 430 m above the surrounding prairie (Kulig, 1996). The hills were surrounded by glaciers twice during the Quaternary period, but the plateau remained unglaciated (Westgate, 1968). Highland forest and grassland communities found in CHP are most similar to those of the Rocky Mountains located 300 km to the west and of the aspen parklands characteristic of the central regions of Canada's three Prairie Provinces (Newsome & Dix, 1968).

Forested areas comprise roughly 20% of the landscape of the hills, while 70% is grassland and 10% mixed or aquatic (Newsome & Dix, 1968). Tree species found in CHP are predominantly *Pinus contorta* (lodgepole pine) and *Populus tremuloides* (trembling aspen) followed by *Picea glauca* (white spruce) and *Populus balsamifera* (balsam poplar) that are common, but less abundant, with *Betula papyrifera* (white birch) and *Acer negundo* (Manitoba maple) being rare (Newsome & Dix, 1968). Grassland communities are dominated by *Festuca campestris* (rough fescue) and other grasses that characterize the mixed-grass natural sub-region (Downing & Pettapiece, 2006). Potential definitive hosts that graze within CHP include elk (*Cervus elaphus*), mule deer (*Odocoileus hemionus),* white-tail deer (*Odocoileus virginianus*) moose (*Alces alces*) and beef cattle (Hegel et al., 2009).

2.3.2 Sampling and host surveys

Our sampling efforts for potential snail and ant intermediate hosts focussed on a 50 X 50 m site in CHP, labelled the Ski Hill site, (12N 553430, 5500684 UTM) where results from our previous surveys indicated infected ants and terrestrial snails were common. In a field survey, from 2009-2013, of >100 randomly selected sites within CHP, Beck (2015), showed that ants at this site were consistently found clinging to vegetation, a behavior indicative of infection with metacercariae.

Samples of clinging ants and snails were collected each month between June-September, 2013. Samples of the unknown *Oreohelix* sp. present at the site were collected by hand by walking the perimeter of the plot and collecting the first 50 snails encountered and placing them in 90% ethanol. These samples of snails were returned to the laboratory, dissected and the presence of sporocysts and/or cercaria confirmed under a dissecting microscope. A sample of tissue and adjacent sporocysts/cercariae from a subset of the total infected samples was kept for use in molecular analysis to confirm the identity of the parasite as *D. dendriticum* as well as identify the unknown species of *Oreohelix*. Concurrently, clinging ants were sampled by entering the plot and collecting the first 50 ants that were observed clinging to the vegetation. Ants were fixed in 90% ethanol, after which they were returned to the laboratory and evaluated for metacercariae counts under a dissecting microscope. An additional five clinging ants were collected during the June sampling period for DNA extraction of metacercariae and host tissue.

2.3.3 Statistics

A two-tail Fisher's exact test was used to evaluate significant differences in prevalence of infected snails between months, critical p = 0.05. Kruskal-Wallis tests were used to test for significant differences in mean metacercariae intensity of infected ants between months, critical p = 0.05.

2.3.4 DNA lysate preparation, PCR and sequencing

Individual, 5 mg sections of snail/parasite tissue from the subset of total infected snails were rinsed twice in distilled water for 5 min and then were placed in lysis buffer and Proteinease K (10mg/ml, New England Biolabs). Lysis buffer contained 50 mM KCL, 10 mM Tris (pH 8.3), 2.5 mM $MgCl_2$, 0.045% Nonidet p-40, 0.45% Tween -20, 0.01% gelatin and dH_2O in 50 ml volumes. Samples were then lysed in 50 µl for 98 min at 60^{o}C followed by 15 min at 94^{o}C. Lysates were stored at -80^{o}C. The abdomens from the five additional ants collected in the June sampling period were removed, placed in lysis buffer and crushed with a sterile plastic pestle to increase lysis efficiency of the ant exoskeleton and parasite material after , which lysis was completed under the same conditions as the snail tissue.

Three separate PCR reactions were carried out; i.e., one to amplify a species-specific region of ITS-2 rDNA from *D. dendriticum* found in both hosts and then two additional reactions to amplify the cytochrome oxidase-1 (COX1) genes from either the snails or ants. The first reaction was carried out on both the snail and ant lysates using previously validated species-specific primers DD_ITS2_SPEC (Bazsalovicsova et al., 2010) that amplify a 176 bp section of ITS-2 ribosomal DNA from *D. dendriticum*. Reactions were carried out in 25 µl volumes containing 1X Thermopol reaction buffer (New England Biolabs), 2 mM $MgSO_4$,

100 uM dNTPs, 0.1 μM forward primer and reverse primers, and 1.25 U Taq DNA polymerase at 5,000 U/ml (New England Biolabs). All reactions were carried out alongside negative and positive controls containing no DNA or DNA from adult *D. dendriticum* collected from cattle in CHP and prepared in the same manner. Thermocycler conditions were 95^{o}C for 5 min followed by 35 cycles of 95^{o}C for 30 sec, 56^{o}C for 60 sec and 72^{o}C for 60 sec with a final extension of 72^{o}C for 5 min. Products were visualized by electrophoresis using a 2% agarose gel stained with SYBER Green loading dye.

The second PCR was carried out on snail lysates only and used universal primers HCO2198 and LCO1490 to amplify a 710 bp region of the snail COX1 genes (Folmer et al., 1994). Reactions were carried out in 25 μl volumes containing 1X Thermopol reaction buffer (New England Biolabs), 2 mM $MgSO_4$, 100 uM dNTPs, 0.1 μM forward primer and reverse primers, and 1.25 U Taq DNA polymerase at 5,000 U/ml (New England Biolabs). All reactions contained DNA negatives as control. Thermocycler conditions were 96^{o}C for 2 min, 28 cycles of 96^{0}C for 40 sec, 48^{o}C for 20 sec and 72^{o}C for 60 sec with a final extension of 72^{o}C for 7 min.

Finally, LepF1/R1 primers were used with the ant lysates only to amplify an 800 bp region of the ant cytochrome oxidase 1 gene (COX1) from the ant abdominal lysates (Hebert et al., 2004). Reactions were carried out in 25 μl volumes containing 1X Thermopol reaction buffer (New England Biolabs), 2 mM $MgSO_4$, 100 uM dNTPs, 0.1 μM forward primer and reverse primers, and 1.25 U Taq DNA polymerase at 5,000 U/ml (New England Biolabs). All reactions contained DNA negatives as control. Thermocycler conditions were 94^{o}C for 2 min

was followed by 5 cycles of 94^{o}C for 40 sec, 45^{o}C for 40 sec and 72^{o}C for 60 sec. This was followed by 36 cycles of 94^{o}C for 40 sec, 58^{o}C for 40 sec and 72^{o}C for 60 sec with a final extension of 72^{o}C for 5 min.

All PCR products for both snail and ant COX1 genes were visualized by electrophoresis using a 2% agarose gel stained with SYBER Green loading dye. PCR amplicons of the COX1 genes of both snails and ants were purified using Omega BioTek Micro Elute Cycle Pure Kit (D6293-02) and sequenced on both strands using the same amplification primers on an Applied Biosystems 3730xl genetic analyzer (Thermo-Fischer Scientific) through the University of Calgary Core DNA Services sequencing Lab (http://www.ucalgary.ca/dnalab/sequencing).

2.3.5 DNA sequence analyses

Chromatograms from both strands of the COX1 gene amplicons from snails or ants were imported into Geneious software (Kearse et al., 2012) and the primer sequences and poor quality 5' and 3' ends removed. Forward and reverse sequences for each product were then assembled into a single contig and the resulting consensus sequences used for subsequent analyses. Individual sequences that showed 100% sequence similarity were collapsed into one representative haplotype. The haplotypes were used in a BLAST search (http://blast.ncbi.nlm.nih.gov/) to find closely related sequences in the GENBANK database and confirm identification.

2.4 RESULTS

2.4.1 Molecular identification of D. dendriticum larval stages from snails and ants

All five snail lysates and all five ant lysates produced the 176 bp fragments of *D. dendriticum* ITS-2 rDNA using the DD_ITS2_SPEC (Bazsalovicsova et al., 2010) PCR primers. Positive controls of *D. dendriticum* DNA previously sequenced from adult worms collected from cattle in the Cypress Hills during pilot studies also showed the 176 bp fragment in each of these reactions.

2.4.2 Molecular identification of snail intermediate hosts

Five fragments of the snail COX1 gene sequences were obtained and sequenced from the infected snails (Accession no. KX923809 - KX923813). Sequences were imported into Geneious (Kearse et al., 2012), which produced a 640 bp alignment of five sequences that each showed 100% sequence identity. These five sequences were collapsed into a single haplotype that was used in a nucleotide BLAST search of GENBANK. The top hit was a COX1 sequence from *Oreohelix* sp. B (DQ858129) reported by Weaver et al. (2006). Query coverage was 100%, with an E value of 0.0. Sequence identity between the two sequences was 99% with only a single nucleotide change from G/A at position 321.

2.4.3 Molecular identification of ant intermediate hosts

A total of four ant COX1 gene sequences were obtained and sequenced from the infected ants; one individual failed to amplify (Accession no. KX923805 – KX923808). Sequences were imported into Geneious (Kearse et al., 2012), which produced a 639 bp alignment of one sequences all having 100% sequence identity. Results of the nucleotide

BLAST search between this haplotype and the GENBANK database produced a top hit with the COX1 sequence from *Formica aserva* (KR924623.1). Query coverage was 100%, with an E score of 0.0. Sequence identity between the two sequences was 99%, with only a single nucleotide change of G/A at position 316.

2.4.4 Patterns of infection in snails and ants

The prevalence of larval *D. dendriticum* within the total sample of 175 snails collected at the Ski Hill site in 2013 was 7 ± 0.04% (range = 4-10%; Table 2.1). There were no significant differences in prevalence between monthly samples ($p = 0.677$). Metacercariae intensity was highly variable between individual ants collected at this site, ranging between 1-182 metacercariae per ant (overall mean = 38 ± 35; Table 2.1). There were no significant differences in metacercariae intensity between monthly samples ($H = 1.533$, $p = 0.675$).

2.5 DISCUSSION

Species-level diagnosis of adult *D. dendriticum* is straightforward based upon standard morphological characters and their location within the bile ducts of grazing mammals. Identification of metacercariae is also straightforward based on morphological characteristics of the cyst and their location within the ant haemocoel and brain. In contrast, the identification of larval stages in terrestrial snails is challenging due to their morphological similarity to sympatric species of trematode. Our molecular-based diagnoses, based upon *D. dendriticum*-specific primers, confirmed that the sporocysts we recovered from snails, and the metacercariae that we recovered from ants, were *D. dendriticum*, and that these matched the adult worms we recovered from beef cattle and elk that had grazed previously within

Cypress Hills Park. These results are the first to confirm the sequence of intermediate and definitive hosts used in the life-cycle of this emerging parasite in western North America and they indicate that possession of a complex life-cycle, in this case involving 3 life-cycle stages, is not a barrier to establishment.

The distributions of at least five confirmed hosts of *D. dendriticum* – the snail, *O. subrudis*, the ant, *F. aserva*, and the ungulates, mule deer, white-tail deer, and elk - are confined to North America. In the case of the former two species, their distributions are centered in the western regions of the continent (Pilsbry, 1939; Gregg, 1972). Thus, the establishment of *D. dendriticum* within western North America had to involve colonization from outside the region, followed by establishment within new species of host at each of the three main life-cycle stages. Presumably, failure to establish within any one of these species during the colonization process would lead to local extirpation within just a few generations. The implication is that following the introduction of *D. dendriticum* into the Cypress Hills region, it rapidly became established, and persisted, within snail, ant, and mammalian hosts.

Evidence for host switching at each life-cycle stage is consistent with the process of multiple-level ecological fitting (Araujo et al., 2015). In another example involving a three host life-cycle, the North American lung fluke, *Haematolaechus floedae*, has been reported from Costa Rican leopard frogs, local species of snails, and local species of odonates in regions where none of its native hosts have been reported (Brooks et al., 2006). Similarly, the liver fluke *Fascioloides magna* has become established in Europe from its homeland in western North America by utilizing a new species of snail as first intermediate host and a

new definitive host (Malcicka et al., 2015). In these examples, rapid and simultaneous establishment within all life-cycle stage following range expansion is considered inconsistent with the requirement for intimate co-evolutionary interactions. Instead, host shifts were made possible into new hosts by certain pre-existing traits that were present at the time of colonization that enabled the recognition and utilization of ancestral resources. We do not know the nature of the pre-existing traits that may have been involved in the establishment of *D. dendriticum* within the sequence of new hosts in CHP, but these traits are likely to be associated with its plasticity, at each life-cycle stage, to utilize a wide range of host resources within their native ranges.

Snails in the genus *Oreohelix* are in the Order Stylommatophora, a taxa that includes all other known snail hosts of *D. dendriticum* (Manga-Gonzalez et al., 2001). These snails are endemic to mountainous regions in western North America, from southern Alberta, the eastern Sierra Nevadas, to northern Mexico, and as far east as the Black Hills of South Dakota (Pilsbry, 1939). Taxonomically, the classification of species within the genus is unresolved and complex (Turgeon et al., 1998). Numerous sub-species and alterations to species designations have been proposed (Weaver et al., 2006). The species of *Oreohelix* that we collected showed high sequence similarity to '*Oreohelix* sp. B' collected from various sites in Wyoming and Montana (Weaver et al., 2006). This species was denoted *O. subrudis* by Chak (2007) from collections in Montana, Wyoming, and South Dakota based upon the morphology of the radula and reproductive structures, and also DNA sequence similarities using numerous nuclear markers. Ongoing studies in our laboratory involving several genetic markers have documented the occurrence of this species at several sites

surrounding Elkwater Lake in the westernmost regions of CHP. Two other *Oreohelix* spp. occurs within CHP, both of which also contain *D. dendriticum* sporocysts (Dempsey, Burg and Goater, unpublished data). These results suggest that several species of *Oreohelix* land snail are compatible hosts for *D. dendriticum* within this region. Beck (2015) surveyed several other genera of land snail collected from numerous sites in CHP, including the Ski Hill site, but none of these were found to contain sporocysts. It is premature to conclude that no other species of snail is compatible in this region, given that so many other species of snail have been documented as first intermediate host in Europe. Considering that the distribution of snails in the genus *Oreohelix* is confined to the Rocky Mountains in western North America, yet adult *D. dendriticum* have been reported in Ontario and as far east as Newfoundland (B. J. van Paridon, unpublished observations), non-oreohelid terrestrial snails must be utilized in the life cycle within other regions of establishment within North America.

Our results indicate that in this region of emergence, *D. dendriticum* utilizes the North American ant *Formica aserva* as its second intermediate host. This result is consistent with the results of field surveys in Europe that report *D. dendriticum* metacercariae in over 20 species of ant belonging to the family Formicidae, mainly within the genus *Formica* (Manga-Gonzalez et al., 2001). According to the most recent survey of ant biodiversity in Alberta, 39 species of the genus *Formica* are found in the province (Glasier, 2013). Ants of the genus *Formica* are widespread and locally common across the northern hemisphere (Wheeler, 1913). Since multiple species of formicid ant are likely to be compatible intermediate hosts in

North America, including in Cypress Hills Park, it is unlikely that establishment within second intermediate hosts is a barrier to establishment of *D. dendriticum*.

Our limited host survey results involving larval *D. dendriticum* in the snail, *O. subrudis*, and in the ant, *F. aserva*, are consistent with the highly variable patterns reported at sites in Europe, Russia, and in several middle-eastern countries. Manga-Gonzalez et al. (2001) showed that the prevalence of *D. dendriticum* in samples of several species of terrestrial snail varied between approximately 0.5 – 5%, with rare extremes that exceeded 50%. The prevalence of larval *D. dendriticum* in *O. subrudis* at the Ski Hill site ranged between 4 – 10% and there was no evidence for seasonality between early summer and fall. Our survey results involving *F. aserva* sampled from the same site showed that 100% of ants found attached to vegetation harboured metacercariae in the hemocoel. Overall mean intensity (38 ± 35) was also highly variable, but it was consistent with reports involving samples of formicid ants collected from sites in Europe (review by Manga-Gonzalez et al., 2001). Following the results of this study that confirm the complete life-cycle of *D. dendriticum* in this region of emergence, more spatially and temporally-targeted sampling procedures can be developed to better understand spatio-temporal patterns of infection in these intermediate hosts.

Our results showed that intermediate hosts native to western North America are utilized by invasive *D. dendriticum*. If the process of multi-level ecological fitting is generalizable at other localities, then regions where the distributions of land snails, Formicid ants, and wild and/or mammalian grazers overlap creates the potential to support the

establishment of this life-cycle. Given the sporadic reports of *D. dendriticum* in Canada (Lewis, 1974; Pybus, 1990) and new reports from Vancouver and Salt Spring Islands on the west coast as well as Ontario and Newfoundland (B. J. van Paridon, unpublished data), this potential appears to be realized. We suggest that increased monitoring of livers from grazing mammals, as well as surveillance of snail and ant species for larval stages, will lead to the discovery of additional pockets of establishment and emergence.

2.6 ACKNOWLEDGMENTS

We would like to thank Melissa Beck, Stefano Catalano, Matthew Easton and John Acheson for help in sampling both snails and ants. Also the staff of Cypress Hills InterInterprovincialPark for their assistance during field seasons in Cypress Hills Park. Funding assistance was provided by a Grant in Biodiversity from the Alberta Conservation Association to Bradley van Paridon and a Research Grant from the National Centre for Veterinary Parasitology to Bradley van Paridon and John Gilleard. Funding by NSERC to Cameron Goater is gratefully acknowledged.

2.7 LITERATURE CITED

Araujo, S. B., Braga, M. P., Brooks, D. R., Agosta, S. J., Hoberg, E. P., von Hartenthal, F. W., & Boeger, W. A. (2015). Understanding host-switching by ecological fitting. *Plos One, 10*, e0139225.

Bazsalovicsova, E., Kralova-Hromadova, I., Spakulova, M., Reblanova, M., & Oberhauserova, K. (2010). Determination of ribosomal internal transcribed spacer 2 (ITS2) interspecific markers in *Fasciola hepatica*, *Fascioloides magna*, *Dicrocoelium dendriticum* and *Paramphistomum cervi* (Trematoda), parasites of wild and domestic ruminants. *Helminthologia, 47*, 76-82.

Beck, M. A. (2015). Ecological epidemiology of an invasive host generalist parasite, Dicrocoelium dendriticum, in Cypress Hills InterInterprovincialPark, Alberta. (Ph.D.), University of Lethbridge, Lethbridge, Alberta.

Beck, M. A., Goater, C. P., Colwell, D. D., & van Paridon, B. J. (2014). Fluke abundance versus host age for an invasive trematode (*Dicrocoelium dendriticum*) of sympatric elk and beef cattle in southeastern Alberta, Canada. *International Journal for Parasitology: Parasites and Wildlife*. *3*, 263-268

Brooks, D. R., León-Règagnon, V., McLennan, D. A., & Zelmer, D. (2006). Ecological fitting as a determinant of the community structure of platyhelminth parasites of anurans. *Ecology, 87*, 76-85.

Chak, T. C. S. (2007). Phylogenetic relationships and population differentiation of Oreohelix landsnails in Wyoming and adjacent South Dakota. (M.Sc.) University of Wyoming.

Cleaveland, S., Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*, 991-999.

Crellen, T., Allan, F., David, S., Durrant, C., Huckvale, T., Holroyd, N., Emery, A.M., Rollinson, D., Aanensen, D.M., Berriman, M., Webster, J.P., Cotton, J.A. (2016). Whole genome resequencing of the human parasite Schistosoma mansoni reveals population history and effects of selection. *Scientific reports, 6*. 20954

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science, 287*, 443-449.

Downing, D. J., & Pettapiece, W. (2006). Natural regions and subregions of Alberta: Natural Regions Committee. Edmonton, Canada: Government of Alberta.

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

Glasier, J. R. N., Acorn, J.H., Nielsen, S., Proctor, H. (2013). Ants (Hymenoptera: Formicidae) of Alberta: A key to species based primarily on the worker caste. *Canadian Journal of Arthropod Identification, 22*, 1-104.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: *Dicrocoelium dendriticum* in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93*, 491-494.

Gregg, R. E. (1972). The northward distribution of ants in North America. *The Canadian Entomologist, 104*, 1073-1091.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper , butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences USA, 101*, 14812-14817.

Hegel, T. M., Gates, C. C., & Eslinger, D. (2009). The geography of conflict between elk and agricultural values in the Cypress Hills, Canada. *Journal of Environmental Management, 90*, 222-235.

Hoberg, E. P., & Brooks, D. R. (2015). Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. *Philosphilcal Transactions of the Royal Society B, 370*, 20130553.

Janzen, D. H. (1985). On Ecological Fitting. *Oikos, 45*, 308-310.

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

Krull, W. H., & Mapes, C. R. (1952). Studies on the biology of *Dicrocoelium dendriticum* (Rudolphi, 1819) Looss, 1899 (Treamatoda Dicrocoeliidae), including its relation to the intermediate host, *Cionella-lubrica* (Muller) .7. The 2nd intermediate host of *Dicrocoelium dendriticum*. *Cornell Veterinarian, 42*, 603-604.

Kulig, J. J. (1996). The glaciation of the Cypress Hills of Alberta and Saskatchewan and its regional implications. *Quaternary International, 32*, 53-77.

Lewis, P. D. (1974). Presence of *Dicrocoelium denriticum* (Rudolphi, 1819) (Treamatoda Dicrocoeliidae) in western Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie, 52*, 662-663.

Malcicka, M., Agosta, S. J., & Harvey, J. A. (2015). Multi level ecological fitting: indirect life cycles are not a barrier to host switching and invasion. *Global change biology*. *21*, 3210-3218

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123*, 91-114.

Mapes, C. R. (1951). Studies on the biology of Dicrocoelium-dendriticum (Rudolphi, 1819) Looss, 1899 (Treamatoda Dicrocoeliidae), including its relation to the intermediate host, *Cionella-lubrica* (Muller) .1. A study of *Dicrocoelium dendriticum* and *Dicrocoelium* infection. *Cornell Veterinarian, 41*, 382-432.

Newsome, R. D., & Dix, R. L. (1968). The forests of the Cypress Hills, Alberta and Saskatchewan, Canada. *The American Midland Naturalist, 80*, 118-185.

Otranto, D., & Traversa, D. (2002). A review of dicrocoeliosis of ruminants including recent advances in the diagnosis and treatment. *Veterinary Parasitology, 107*, 317-335.

Pilsbry, H. A. (1939). *Land mollusca of North America:(north of Mexico)*: Philidelphia, USA: Academy of Natural Sciences.

Pybus, M. J. (1990). Survey of hepatic and pulmonary helminths of wild cervid in Alberta, Canada. *Journal of Wildlife Diseases, 26*, 453-459.

Thompson, R. C. A., Lymbery, A. J., & Smith, A. (2010). Parasites, emerging disease and wildlife conservation. *International Journal for Parasitology, 40*, 1163-1170.

Tompkins, D. M., & Poulin, R. (2006). Parasites and Biological Invasions. In R. B. Allen & W. G. Lee (Eds.), *Biological Invasions in New Zealand* (pp. 67-84). Berlin, Heidelberg: Springer Berlin Heidelberg.

Turgeon, D., Quinn Jr, J., Bogan, A., Coan, E., Hochberg, F., Lyons, W., Mikkelsen, P.M., Neves, R.J., Roper, C.F.E., Rosenberg, G. (1998). Common and scientific names of aquatic invertebrates from the United States and Canada: mollusks, 2nd edition. *American Fisheries Society Special Publication, 26*. 1-536.

Weaver, K. F., Anderson, T., & Guralnick, R. (2006). Combining phylogenetic and ecological niche modeling approaches to determine distri, bution and historical biogeography of Black Hills mountain snails (Oreohelicidae). *Diversity and Distributions, 12*, 756-766.

Westgate, J. (1968). Surficial geology of the Foremost-Cypress Hills area, Alberta. Research Council of Alberta. *RCA/AGS Bulletin, 22*, 1-137

Wheeler, W. M. (1913). A revision of the ants of the genus Formica (Linne) Mayr. *Bulletin of the Museum of Comparative Zoology, 53*, 379-565.

Table 2.1. Patterns of larval *Dicrocoelium dendriticum* infection in samples of the terrestrial snail *Oreohelix subrudis* and the ant *Formica aserva* in Cypress Hills Interprovincial Park, Alberta. Data are presented as prevalence in snails and mean intensity in ants

Month	Snails (N)	Ants (N)	Prevalence in Snails (±CI%)	Mean Intensity in Ants (±SD)
June	25	49	8 ± 0.11	41 ± 35
July	50	43	6 ± 0.06	39 ± 40
August	50	54	4 ± 0.04	33 ± 26
September	50	17	10 ± 0.08	38 ± 45
Total	175	163	7 ± 0.04	38 ± 35

CHAPTER 3

Characterization of nine microsatellite loci for *Dicrocoelium dendriticum*, an emerging liver fluke of ungulates in North America, and their use to detect clonemates and random mating

B.J van Paridon[1], C.P. Goater[1], J.S. Gilleard[2] and C.D. Criscione[3]

This chapter was published as a Short Communication in Molecular and Biochemical Parasitology (2016, 27(1): 19-22)

[1] Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada. T1K 3M4.

[2] Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, Alberta, Canada. T2N 4Z6.

[3] Department of Biology, Texas A&M University, College Station, Texas 77843, U.S.A.

3.1 ABSTRACT

This study characterizes polymorphic microsatellite loci from adults of the liver fluke *Dicrocoelium dendriticum* sampled from a population of sympatric beef cattle and elk in a region of emergence in southern Alberta, Canada. We also scrutinized the markers to validate their use in studying the population genetics of this complex life cycle parasite. Among the nine loci described, four deviated significantly from Hardy Weinberg Equilibrium (HWE) due to technical artifacts. The remaining five loci were in HWE. These five provided sufficient resolution to identify clonemates produced from the obligate asexual reproduction phase of the life cycle in snails and to assess the impact of non-random transmission of clonemates on measures of F_{IS}, F_{ST} and genotypic disequilibrium. Excluding clonemates, we show that the sub-population of worms was in HWE, that average F_{IS} within hosts was 0.003 (p = 0.4922) and that there was no population genetic structure among hosts F_{ST} = 0.001 (p = 0.3243). These markers will be useful for studies of *Dicrocoelium dendriticum* ecology, transmission, and evolution.

3.2 INTRODUCTION, METHODS, RESULTS AND DISCUSSION

Parasites are a diverse group of organisms with well-known negative impacts on the health of human, animal, and plant populations. They also present many challenges to researchers as questions concerning diagnosis, transmission, host specificity, host-parasite interactions, and the identification of cryptic species are often difficult to answer via egg counts and morphological analyses alone (Criscione et al., 2005; Lymbery & Thompson, 2012). Since their discovery, polymorphic genetic markers have been increasingly used to

indirectly and in some cases, directly, infer these aspects of parasite ecology and evolution (Criscione et al., 2005). In particular, due to their co-dominance and tendency for high polymorphism, microsatellite markers have a wide variety of applications in population genetic studies including investigating the genetic diversity and differentiation of populations, and distinguishing individuals via multilocus genotypes (MLGs) (Selkoe & Toonen, 2006). This utility is especially pertinent among trematode parasites where asexual reproduction in the first intermediate host and subsequent clonal transmission can influence the underlying pattern of parasite genetic structure among successive hosts in the life cycle (Gorton et al., 2012; Prugnolle et al., 2005). Identification of clonemates (individuals that are the product of an asexual reproductive event of a progenitor individual) is necessary to make correct inference on trematode transmission based on population genetics (Prugnolle et al., 2005). Microsatellites can be used to identify individuals with identical MLGs and subsequent statistical tests can be employed to test if these individuals are indeed clonemates (Arnaud-Haond & Belkhir, 2007; Gregorius, 2005). To date, studies examining fluke transmission via genetic markers have been conducted on aquatic or semi-aquatic systems (Gorton et al., 2012). Our understanding of trematode transmission (as inferred through population genetics) in a completely terrestrial environment is nonexistent, in part, because microsatellites have yet to be developed for a trematode with a completely terrestrial life cycle.

Here we present the development of microsatellite markers for the lancet liver fluke *Dicrocoelium dendriticum* (Trematoda: Dicrocoeliidae). This parasite is found throughout Europe, North Africa, Asia and North America (Manga-Gonzalez et al., 2001). *Dicrocoelium*

dendriticum has a three host life cycle that is completely terrestrial (Krull & Mapes, 1952). Snails (in which there is asexual reproduction) are first intermediate hosts, ants are second intermediate hosts, and several mammals can serve as the definitive host (where there is obligate sexual reproduction) (Manga-Gonzalez et al., 2001). The lancet liver fluke is famous for its ability to alter the behavior of ants. A single metacercaria encysted in the ant ganglia causes the ant to cling to vegetation, thus making them available for accidental ingestion by grazing mammals (Manga-Gonzalez et al., 2001). Once metacercariae are ingested by definitive hosts, juveniles emerge and migrate to the bile ducts where they develop to into adults. Infection in a range of ungulate definitive hosts is associated with anemia, edema emaciation, cirrhosis of the liver and in high-intensity infections, photosensitization (Otranto & Traversa, 2002; Sargison et al., 2012). *Dicrocoelium dendriticum* is an invasive species in North America, first described on the eastern coast of the continent in the 1950s (Mapes & Baker, 1950). Since then, sporadic reports have found it as far as the western coast of North America (Lewis, 1974; Pybus, 1990). In 1990 , it was first reported in Cypress Hills Interprovincial Park, Alberta, Canada where it can now be found in sympatric elk, cattle, and deer (Goater & Colwell, 2007).

All samples used in this study came from ongoing surveys involving sympatric elk and cattle in the Cypress Hills Interprovincial Park, Alberta, Canada (49^0 30'N 110^0 W). Adult worms were recovered via dissections of purchased cattle livers and hunter shot elk following Goater and Colwell (Goater & Colwell, 2007). An initial library enriched for microsatellites was constructed using Illumina MiSeq next generation sequencing to identify di-, tri-, tetra- and pentamer repeats throughout the genome of *D. dendriticum*. Library

construction and sequencing was conducted by S. Bogdanowicz at the Evolutionary Genetics Core Facility at Cornell University, USA using the methods described in Nali et al. (2014). Potential microsatallite loci with sizes of 150-450bp were selected for primer design. Testing and optimization of microsatellite loci was performed on a dataset of 66 adult worms collected from four elk livers and three cattle livers from Cypress Hills Interprovincial Park. Worms were lysed in lysis buffer and Protinease K (10mg/ml, New England Biolabs). Lysis buffer contained 50 mM KCL, 10 mM Tris (pH 8.3), 2.5 mM $MgCl_2$, 0.045% Nonidet p-40, 0.45% Tween -20, 0.01% gelatin and dH_2O in 50 ml volumes. Samples were lysed in 50 µl for 98 min at 60^{o}C followed by 15 min at 94^{o}C then stored at -20^{o}C. PCR reactions were carried out in reaction mixtures of 25 µl containing final concentrations of 1X Thermopol reaction buffer (New England Biolabs), 2 mM $MgSO_4$, 100 µM dNTPs, 0.1 µM 6-FAM labelled forward primer and reverse primers, 1.25U Taq DNA polymerase at 5000U/ml (New England Biolabs). Thermo cycling condition were 94^{o}C for 3 min followed by 35 cycles of 94^{o}C for 30 seconds, 54^{o}C for 30 seconds and 72^{o}C for 1 min with a final extension of 72^{o}C for 10 min. Products were then analysed on an Applied Biosystems 3500xl sequencer and sized using Genescan 500 (-250) LIZ size standard (Applied Biosystems) and scored manually using GeneMapper 4.1 (Applied Biosystems).

Gene diversity of each locus (H_s) as well as estimates and tests of F_{IS} and F_{ST} were conducted using FSTAT (Goudet, 1995). Significance of average within-host F_{IS} per locus and F_{ST} among hosts were determined using 10,000 randomizations of alleles among individual worms within hosts and of genotypes among hosts, respectively. Genotypic equilibrium (GD) between pairs of loci was tested in GENEPOP 4.2 (Markov chain

parameters: 1000 dememorizations; 100 batches; 1000 iterations) (Rousset, 2008). We also tested for identity disequilibrium (ID); i.e., correlations in heterozygosity among different loci that result from non-random mating such as partial selfing (David et al., 2007). ID is not affected by scoring artefacts such as null alleles. Thus, comparison of ID with F_{IS} is a useful means of determining if single locus estimates of F_{IS} are artificially elevated due to technical artefacts or are high due to real biological phenomena such as self-mating (David et al., 2007). ID was tested with 1000 iterations of resampling the single locus genotypes among individuals in the population using RMES (David et al., 2007). GENCLONE 2.0 was used to identify individuals with the same MLG and test whether or not identical MLGs arose via asexual reproduction (Arnaud-Haond & Belkhir, 2007). Clonemate testing is determined by calculating P_{SEX}, which is the probability of observing n copies of a MLG in a sample size N given sexual reproduction (significance determined with $P_{SEX} < 0.05$). If the P_{SEX} of a multicopy MLG at $n = 2$ is significant, it can be taken that all copies of this MLG are the result of asexual propagation (Gregorius, 2005).

Our initial library construction yielded 4517 potential microsatellite loci. Of these, 101 loci were tested on a random subset of the dataset ($n = 10$) using unlabeled primers and visualized on 2% agarose gel to confirm amplification and presence of only a single product. Of these, nine loci were found to be polymorphic, reliably amplified in all individuals, and provided unambiguous genotyping profiles.

The initial data analysis tested the average within-host F_{IS} per locus using all nine loci and all 66 individuals. A wide range of F_{IS} values was observed (-0.077 to 0.526) with four

loci testing significantly greater than 0 (p<0.05) and five loci not significant. Such a large variation in F_{IS} values among loci is atypical for a given mating system and, therefore, could be caused by a number of factors; e.g., clonal structure, null alleles, mating system, or cryptic structure (Criscione et al., 2011). To determine the potential factor(s) that may be causing this variation, we used a multistep approach described below.

First, individuals with missing data were removed resulting in a data set of n = 58 individuals. Removal was necessary as P_{SEX} cannot be calculated on individuals with missing data. Among these 58 individuals, there were 52 unique MLGs for which four MLGs had more than copy. At this point, we note that testing for the significance of clonemates (i.e., calculating P_{SEX}) relies on the estimation of allele frequencies from a given population. If there is population structuring, then within subpopulation allele frequencies need to be used to test for clonemates by subpopulation, otherwise P_{SEX} will be biased towards significance. However, clonemates themselves could drive among-host structure (Prugnolle et al., 2005). Thus, as noted by Criscione et al. (2011), the presence of clonemates creates a 'catch-22', i.e., does one test for clones if there is structure among hosts (not due to clonal transmission) while at the same time addressing if there is structure among hosts that is due to the presence of clonemates? We followed a similar approach used by Criscione et al. (2011) by determining the influence of having repeated MLGs in the data set compared to a data set where repeated MLGs were reduced to single copy within individual hosts. Thus, a second dataset was made where repeated MLGs were reduced to one copy within each host (n = 52). Genetic structuring among hosts was then tested on both datasets. In the complete dataset (n = 58), significant structure among hosts was found (multilocus F_{ST} = 0.02 (p = 0.0001)),

whereas when MLGs were reduced to only unique copies within hosts (n = 52), there was no significant among-host structure (multilocus F_{ST} = 0.003 (p = 0.8060)). This pattern indicates that the presence of clonemates is likely driving the significance of F_{ST}. Therefore, one can conclude that all worms are from the same population, as transmission prior to asexual reproduction is random (Prugnolle et al., 2005; Gorton et al., 2012).

Next, using the same two data sets above, we ignored the among-host division of worms and analyzed the datasets as a single population to test F_{IS} across loci. Whether repeated copies of MLGs were included or not, there remained extensive variation in F_{IS} among loci (-0.032 to 0.514 when n = 58, and -0.002 to 0.491 when n = 52; four loci still tested significantly greater than 0 in both data sets). Table 3.1 shows the F_{IS} estimates from the reduced data set. In contrast, GD was significant in 15 of 36 pairwise comparisons in the complete data set (n = 58), whereas this fell to 1 of 36 (a number expected by chance alone with α = 0.05) with repeated copies of MLGs reduced (n = 52). The lack of linkage disequilibrium (LD) in the reduced data sets suggests the presence of clonemates is driving the excess LD in the complete data set (de Meeûs et al., 2006). Moreover, the lack of LD in the reduced data set suggests that there is no underlying cryptic population structure as population admixture would cause LD.

To determine if technical issues were the cause of variation in F_{IS} values, we tested whether ID was significant in the reduced data set (n = 52). $\hat{g}_2$, the estimator of the two-locus heterozygosity disequilibrium (David et al., 2007), was not significantly greater than 0 (p = 0.774). Thus, there was no evidence of non-random mating such as might occur with self-

mating. Because ID is not biased by technical errors such as null amplification and because there was no evidence of cryptic structure (i.e., no LD), we were left with the conclusion that the significant F_{IS} values for the four loci DdMs43, DdMs44, DdMs89 and DdMs93 (Table 3.1) were caused by technical artifacts.

Now that we identified the variation in F_{IS} values was likely due to technical issues, we wanted to re-examine the data to interpret the biology of *D. dendriticum*. Thus, we repeated the above analyses using only the five loci (DdMs21, DdMs28, DdMs60, DdMs70 and DdMs95) that were in Hardy-Weinberg Equilibrium (HWE). Using only these five loci, we had a data set with no missing data of $n = 63$ individuals, among which there were 55 unique MLGs where all six MLGs present in more than one copy, had $P_{sex} < 0.05$. In re-examining structure among hosts, when all clonemates are included ($n = 63$), the average F_{IS} within hosts was 0 ($p = 0.6876$) and there was significant structure among hosts, $F_{ST} = 0.015$ ($p = 0.0025$). When only one individual of a set of clonemates was retained within a host ($n = 55$), average F_{IS} within hosts remained low at 0.003 ($p = 0.4922$), but structure among hosts disappeared, $F_{ST} = 0.001$ ($p = 0.3243$). In all cases where there were clonemates, the clonemates were in the same host individual. Thus, our results confirm the theoretical predictions of Prugnolle et al. (2005) that when there is little dispersion of clonemates after asexual reproduction in the first host (i.e., there is clumped clonal transmission), there will be significant genetic structure among subsequent hosts in the life cycle. In looking at the influence of clonemates on LD, we see that 8 out of 10 pairwise comparisons had significant GD ($p < 0.05$) when including clonemates ($n = 63$; treating the data set as one population). If

only one representative of each clonemate is retained (n = 55), then none of the pairwise comparisons were significant. Thus, again, sampling of clonemates will lead to LD.

Given the above results, we conclude that there is random mating among the adult population of flukes and random transmission of flukes prior to infection of the snail first host. After asexual reproduction in the snail host, however, there is some significant clonal transmission such that clonemates end up in the same definitive host more often than expected by chance alone. This pattern of clonal transmission implies that clonemates originating from snails would also be common in metacercariae subpopulations in ants. This prediction remains to be tested. Moreover, although we found that some loci may have technical issues, we ended up with five polymorphic loci that alone had enough power (likely due to their high polymorphism) to statistically detect the presence of clonemates. All nine polymorphic microsatellite loci are reported for *D. dendriticum*. These new markers will be valuable to study the population genetics, life cycle and global spread of this emerging trematode. In particular, these loci could be used in assignment tests to potential identify the source population(s) of the invasion in North America. From an ecological perspective, these markers will enable us to test if trematode clonal transmission varies in a terrestrial environment compared to aquatic environments. From an evolutionary perspective, moreover, these markers will enable us to address a long standing question of kin selection regarding the appearance of altruistic self-sacrificing behavior by the single worm that infects ant brains and alters their behavior. Lastly, whenever there is a large amount of variation in F_{IS} among loci, one should be suspect (Criscione et al., 2011; Detwiler & Criscione, 2011).

Our study highlights how new markers need to be thoroughly scrutinized as both technical (e.g., null alleles) and biological factors can cause deviations from HWE.

3.3 ACKNOWLEDGMENTS

We would like to acknowledge Steve Bogdanowicz at the Evolutionary Genetics Core Facility at Cornell University, USA for completing the initial library construction. This work was funded in part by a grant from the National Centre for Veterinary Parasitology. Charles Criscione is supported by a National Science Foundation grant (DEB 1145508), John Gilleard by the Natural Sciences and Engineering Research Council of Canada (NSERC) (Grant number RGPIN/371529-2209). B. van Paridon is a member of the NSERC-CREATE Host Pathogen Interactions (HPI) graduate training program at the University of Calgary.

3.4 LITERATURE CITED

Arnaud-Haond, S., & Belkhir, K. (2007). genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes, 7,* 15-17.

Criscione, C. D., Poulin, R., & Blouin, M. S. (2005). Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology, 14,* 2247-2257.

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. *Molecular Ecology, 20,* 2510-2524.

David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology, 16,* 2474-2487.

de Meeûs, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution, 6,* 163-170.

Detwiler, J. T., & Criscione, C. D. (2011). Testing Mendelian inheritance from field-collected parasites: revealing duplicated loci enables correct inference of reproductive mode and mating system. *International Journal for Parasitology, 41,* 1185-1195.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: Dicrocoelium dendriticum in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93,* 491-494.

Gorton, M. J., Kasl, E. L., Detwiler, J. T., & Criscione, C. D. (2012). Testing local-scale panmixia provides insights into the cryptic ecology, evolution, and epidemiology of metazoan animal parasites. *Parasitology, 139,* 981-997.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity, 86,* 485-486.

Gregorius, H. R. (2005). Testing for clonal propagation. *Heredity, 94,* 173-179.

Krull, W. H., & Mapes, C. R. (1952). Studies on the biology of *Dicrocoelium dendriticum* (Rudolphi, 1819) Looss, 1899 (Trematoda: Dicrocoeliidae), including its relation to the intermediate host, *Cionella lubrica* (Muller). VI. Observations on the life cycle and biology of *C. lubrica*. *The Cornell veterinarian, 42*(4), 465-489.

Lewis, P. D. (1974). Presence of *Dicrocoelium denriticum* (Rudolphi, 1819) (Trematoda: Dicrocoeliidae) in western Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie, 52*, 662-663.

Lymbery, A. J., & Thompson, R. C. A. (2012). The molecular epidemiology of parasite infections: Tools and applications. *Molecular and Biochemical Parasitology, 181*, 102-116.

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123*, 91-114.

Mapes, C. R., & Baker, D. W. (1950). The white-tail deer, a new host of *Dicrocoelium dendriticum* (Rudolphi, 1819) Looss, 1899 (Trematoda: Dicrocoeliidae). *Cornell Veterinarian, 40*, 211-212.

Nali, R. C., Zamudio, K. R., & Prado, C. P. (2014). Microsatellite markers for *Bokermannohyla* species (Anura, Hylidae) from the Brazilian Cerrado and Atlantic Forest domains. *Amphibia-Reptilia, 35*, 355-360.

Otranto, D., & Traversa, D. (2002). A review of dicrocoeliosis of ruminants including recent advances in the diagnosis and treatment. *Veterinary Parasitology, 107*, 317-335.

Prugnolle, F., Liu, H., de Meeus, T., & Balloux, F. (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology, 35*, 255-263.

Pybus, M. J. (1990). Survey of hepatic and pulmonary helminths of wild cervid in Alberta, Canada. *Journal of Wildlife Diseases, 26*, 453-459.

Rousset, F. (2008). GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources, 8*, 103-106.

Sargison, N. D., Baird, G. J., Sotiraki, S., Gilleard, J. S., & Busin, V. (2012). Hepatogenous photosensitisation in Scottish sheep casued by *Dicrocoelium dendriticum*. *Veterinary Parasitology, 189*, 233-237.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters, 9*, 615-629.

Table 3.1 Characteristics of nine polymorphic loci isolated from *D. dendriticum* based on dataset with clones reduced to one copy (n=52) $T_{a:}$ annealing temperature; Na: number of alleles; H_s: gene diversity; F_{IS}: inbreeding coefficient (values in bold are significantly greater than zero).

Locus	GenBank No.	Repeat Motif	Primer 5'-3'	T_a (^{0}C)	Na	H_s	Size Range (bp)	F_{IS}
DdMs21	KU094045	AAC	F:AACCTTGCCGTCTTCATTGATAC R:CGCTTTGCTCTGATTTCAAATGG	54	11	0.779	311-366	0.086
DdMs28	KU094044	ATCT	F:GATCACTCAGAGCGCTTTAAGTC R:CAACAATGCACTTTCGTCACTTC	54	9	0.811	344-366	-0.068
DdMs43	KU094043	AAC	F:TTGACTAGAGCCACCACAATTAG R:CCCACCCGTCCTTGTTAATAATC	54	7	0.621	463-481	**0.442**
DdMs44	KU094042	AAC	F:TGGAAGCCACGATTAAACTTAGC R:TATGGTGCCAGAATTGAAACGAG	54	20	0.919	248-346	**0.414**
DdMs60	KU094041	AAC	F:TGTGGAAACTTATCAAACACTGC R:AACTACTGACCGATTTGCTTGTC	54	11	0.805	243-288	-0.027
DdMs70	KU094040	AC	F:TAAAGCAACACCATGAGTAGCAC R:TAGCAACGTAACGAGAACCAATG	54	16	0.895	145-196	0.055
DdMs89	KU094039	AC	F:TTTCTGAACTCCGAACAACAGG R:TGCCATAGTGCCAACAGAATATC	54	9	0.793	169-185	**0.491**
DdMs93	KU094038	AC	F:CAAGTAAGGCGTTTCAGTTCCTC R:GCTCCAGTATCAAATTCGTGAC	54	13	0.686	180-237	**0.244**
DdMs95	KU094037	AG	F:CATCGCGAAGTAACTTGATTAGC R:AACTACACTGTTCGTACACTGGG	54	9	0.653	285-301	-0.002

CHAPTER 4

Population genetic analysis informs the invasion history of the emerging trematode *Dicrocoelium dendriticum* into Canada

B.J. van Paridon[1], D.D. Colwell[2], C.P. Goater[1] and J.S. Gilleard[3]

This chapter was prepared for and submitted to the International Journal for Parasitology

[1] Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada. T1K 3M4.

[2] Agriculture and Agri-Food Canada, Lethbridge Research and Development Station, Lethbridge, Alberta Canada, T1J 4B1

[3] Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, Alberta, Canada. T2N 4Z6.

4.1 ABSTRACT

Parasite distributions are constantly changing due to climate change, local and global movement of animals and humans, as well as land use and habitat change. The trematode *Dicrocoelium dendriticum* is a relatively recent invader into Canada being first reported in Eastern Canada in the 1930s and Western Canada in the 1970s. However, historical records are scarce and its emergence is poorly understood. The establishment of this parasite in Canada provides an interesting opportunity to explore the use of population genetic approaches to help elucidate the invasion history of a relatively recently established helminth parasite. In this study, we compare the genetic diversity and population structure of a number of *D. dendriticum* populations from Western and Eastern Canada and compare these with much longer established European populations. Two independent genetic marker systems were used; a microsatellite marker panel and a COX1 mtDNA sequence marker. We found distinct differences in both genetic diversity and population structure of the different Canadian populations that provide insights into their invasion histories when compared to the European populations. Two populations from British Columbia – Salt Spring Island and Vancouver Island – are of low diversity, show evidence of a population bottleneck, and are closely related to each other suggesting a shared recent history of establishment. These west coast populations are otherwise most closely related to those from Eastern Canada and Western Europe and in contrast are genetically divergent from those in Cypress Hills, Alberta. Although the Alberta parasite population is the most recently reported in Canada, being first identified there in the early 1990s, it was the most genetically diverse of those examined and showed a strong pattern of admixture of genotypes present in western and

Eastern Europe. Overall, our results are consistent with a model, in which Western Europe is likely the source of flukes on the east coast of Canada, which were then subsequently translocated to the west coast of Canada. The most recently reported *D. dendriticum* population in Canada appears to have a different history and likely has multiple origins.

4.2 INTRODUCTION

The occurrence of pathogens and/or parasites that have expanded beyond their host or geographic ranges, also known as emerging infectious diseases (EIDs), is an important issue for livestock, wildlife and human health (Jones et al., 2008). A number of factors such as climate change, wildlife encroachment, and the anthropogenic movement of livestock and wildlife have contributed to the increase of EIDs in recent decades (Daszak et al., 2000; Cleaveland et al., 2001). In the case of helminth parasites, movements are often human-mediated as parasites get transported to new locations inside of their hosts (Blouin et al., 1995). The spread and establishment of parasites can then threaten significant harm to human, domestic and wild animal populations (Rhyan & Spraker, 2010; Thompson et al., 2010).

The terrestrial liver fluke *Dicrocoelium dendriticum* is one such invasive parasite that has a broad geographic and host range, colonizing mammalian hosts in Europe and North America (Manga-Gonzalez et al., 2001). The three-host life cycle of this parasite involves terrestrial snails and ants as the first and second intermediate hosts respectively, and grazing mammals as the definitive host. The complete life-cycle was first elucidated in the 1950s by Krull and Mapes (1952) on the east coast of North America. It was believed that *D.*

dendriticum was imported around this time by way of domestic animals from Europe. However, very little is known about its invasion history into North America. The first published report of the parasite in Canada was in 1974 (Lewis, 1974). Observations of livers from cattle grazed in British Columbia found adult *D. dendriticum* (Lewis, 1974). The fluke was not reported in another western Canadian province until 1990 when it was found in a small number of elk originating in the Cypress Hills region of Alberta (Pybus, 1990). Since this initial report, the prevalence of *D. dendriticum* in Cypress Hills Interprovincial Park has increased and now at least four species of definitive host (elk, cattle, mule deer, and white-tailed deer) are routinely infected, indicating establishment in this region (Goater & Colwell, 2007; Beck et al., 2014). The apparently recent establishment of *D. dendriticum* in Canada, along with a paucity of historical reports makes this an interesting helminth system to explore the use of population genetic approaches to inform an invasion history.

In the absence of other data such as animal movements or parasite surveys, the tracking of invasive species and their subsequent spread can be achieved by assessing the population genetic characteristics that result from invasion events like bottlenecks, founder effects, or multiple introductions (Cristescu, 2015). In this study, we use a recently developed panel of microsatellite markers for *D. dendriticum* (van Paridon et al., 2016), along with cytochrome oxidase-1 (COX1) mitochondrial gene sequences, to compare the levels of genetic diversity and population structure in flukes from western Canada with those from eastern Canada and Europe to provide insights into its invasion history within Canada.

4.3 MATIERIALS AND METHODS

4.3.1 Sampling

Adult flukes were collected from Canada at the following sites. The Cypress Hills Interprovincial Park, Alberta (N 49^{o} 40’ 29” W 110 ^{o} 16’ 44”), two locations on the west coast of British Columbia (Salt Spring Island and Parksville, Vancouver Island), one location from southern Ontario and from Newfoundland (Table 4.1). Adult flukes from cattle and elk from Cypress Hills Park were made available by dissection of livers as part of ongoing studies of the liver fluke in this region (van Paridon et al., 2016). West coast flukes were obtained from the same abattoir on Vancouver Island, British Columbia where livers from sheep sent from various locations were inspected at time of slaughter. The Salt Spring Island flukes were obtained from sheep livers infected with *D. dendriticum* originating from Salt Spring Island (N 48° 48' 57" W 123° 30' 7"), one of many small islands located in the Gulf Islands chain off the coast of British Columbia. The second group of infected sheep livers originated from Parksville, Vancouver Island (N 49° 19' 9" W 124° 18' 49") a small town on the eastern coast of Vancouver Island. Flukes sampled from sheep in southern Ontario were provided from colleagues although the exact locations of farms where the animals originated were not known. Flukes from Newfoundland were obtained from hares trapped in Branch, Newfoundland (N 46 ^{o} 52’ 56” W 53 ^{o} 57’ 11”) in February 2015.

Additional samples of adult *D. dendriticum* from Europe were opportunistically obtained from colleagues. Detailed information was not always available on exact origins although the province and/or country of origin could always be verified (Table 4.1). Samples from three locations (Newfoundland, Ontario and Italy) were provided as pooled populations

of adult flukes from multiple hosts in the same region. As we could not distinguish flukes from separate hosts from these regions, these were treated as a single population from each region.

4.3.2 DNA extraction, PCR, genotyping and sequencing

All adult flukes were washed twice in dH_2O and lysed in lysis buffer and Proteinase K (10 mg/ml, New England Biolabs). Lysis buffer contained 50 mM KCL, 10 mM Tris (pH 8.3), 2.5 mM MgCl2, 0.045% Nonidet p-40, 0.45% Tween −20, 0.01% gelatin and dH2O in 50 ml volumes. Samples were lysed in 50 µl for 98 min at 60 ^{o} C followed by 15 min at 94 ^{o} C then stored at −80 ^{o} C.

PCR amplifications of the nine microsatellite loci (DdMs23, DdMs28, DdMs43, DdMs44, DdMs60, DdMs70, DdMs89, DdMs93 and DdMs95; van Paridon *et al*., 2016) were performed in reaction mixtures of 25 µl containing final concentrations of 1X Thermopol reaction buffer (New England Biolabs), 2 mM MgSO4, 100 µM dNTPs, 0.1 µM 6-FAM labelled forward primer and reverse primers, 1.25U Taq DNA polymerase at 5000U/ml (New England Biolabs). Thermocycling conditions were 94^{o}C for 3 min followed by 35 cycles of 94^{o}C for 30 s, 54^{o}C for 30 s and 72^{o}C for 1 min with a final extension of 72^{o}C for 10 min. Products were then analysed on an Applied Biosystems 3500xl sequencer and sized using Genescan 500 (-250) LIZ size standard (Applied Biosystems) and scored manually using GeneMapper 4.1 (Applied Biosystems).

PCR amplification of the *D. dendriticum* cytochrome oxidase 1(COX1) mitochondrial gene was carried out using primers *cox1*^{int1} 5'- GGGCTGGTTATTCTGGTTGGGG - 3'and *cox*^{int2} 5'- GCCATTCCTAAATAATGCATAGGG – 3'originally designed for the trematode *Fascioloides magna* (Kralova-Hromadova et al., 2008). PCR reactions were carried out in 25 μl containing final concentrations of 1X Thermopol reaction buffer (New England Biolabs), 2 mM MgSO4, 100 μM dNTPs, 0.1 μM of forward and reverse primers, 1.25U Taq DNA polymerase at 5000U/ml (New England Biolabs). Thermocycling conditions were 94^{o}C for 5 min followed by 35 cycles of 94^{o}C for 30 s, 50^{o}C for 1 min and 72^{o}C for 1 min with a final extension of 72^{o}C for 5 min. Products were visualized on 2% agarose gel stained with SYBER Green loading dye (New England Biolabs). *Dicrocoelium dendriticum* COX1 amplicons were purified using Omega BioTek Micro Elute Cycle Pure Kit (D6293-02) and both strands sequenced using the same amplification primers on an Applied Biosystems 3730xl genetic analyzer through the University of Calgary Core DNA Services sequencing Lab (http://www.ucalgary.ca/dnalab/sequencing).

4.3.3 Microsatellite validation

An approach similar to van Paridon et al. (2016) and Criscione et al. (2011) was used to assess the utility of a panel of nine microsatellite loci that were originally isolated from adult *D. dendriticum* collected from the Cypress Hills Interprovincial Park, Alberta (van Paridon et al., 2016). This approach was necessary to confirm the amplification of the microsatellites with flukes sampled from across Canada and Europe. With this approach, we could also assess the presence of null alleles and the effects of repeated multilocus genotypes

(MLG) (a common phenomenon with asexually multiplying trematodes). Null alleles are known to cause underestimations of true heterozygosity (Falush et al., 2003) and the non-random transmission of clones during the trematode life cycle can bias estimates of F_{ST} and F_{IS} towards significance (Prugnolle et al., 2005; Criscione et al., 2011). First, for all analyses, individuals that scored zero alleles (failed to amplify) at any locus were removed from the dataset. This reduced the original dataset of n=189 to n=166. Estimates and tests of F_{IS} and F_{ST} were performed using FSTAT (Goudet, 1995). Significance of average within-host F_{IS} per locus and F_{ST} among hosts were determined using 10,000 randomizations of alleles among individual worms within hosts and of genotypes among hosts, respectively. Loci appearing to depart from HWE according to significant deviations of F_{IS} were tested for identity disequilibrium (ID). Identity disequilibrium (ID) is a measure in correlations in heterozygosity among different loci that result from non-random mating such as partial selfing. It was tested as it is a method that is not influenced by null alleles, which may be artificially inflating estimates of F_{IS} (David et al., 2007). ID was tested with 1000 iterations of resampling the single locus genotypes among individuals in the population using the software RMES (David et al., 2007). Genotypic disequilibrium (GD) between pairs of loci was tested on both datasets using GENEPOP 4.2 (Markov chain parameters: 1000 dememorizations; 100 batches; 1000 iterations) (Rousset, 2008). GENCLONE 2.0 was used to identify individuals with the same MLG and to test whether or not identical MLGs arose via asexual reproduction (Arnaud-Haond & Belkhir, 2007). The identities of individuals with identical MLGs were also used to create a second dataset in, which multiple copies of MLGs were reduced to one. This is done in order to determine the influence of multiple MLGs on

genetic structure, which will in turn bias clonemate testing (Criscione et al., 2011). Clonemate testing then proceeds by calculating P_{sex} in GENCLONE 2.0, which is the probability of observing n copies of a MLG in a sample size N, given sexual reproduction (significance determined with $P_{sex} < 0.05$). If the P_{sex} of a multicopy MLG at n = 2 is significant, it can be taken that all copies of this MLG are the result of asexual propagation.

4.3.4 Genetic diversity and population genetic analysis of microsatellite data

We used a number of analyses to assess the levels of genetic diversity in Canadian populations of *D. dendriticum* and compared them with populations from the other locations. Genetic diversity (H_s) of all populations was assessed using FSTAT to provide unbiased estimates of genetic diversity per sample and per locus (Goudet, 1995). To ensure the differences in sample size did not influence results, we compared this measure with values of genetic diversity adjusted for sample size from POPTREE using the Wilcoxon rank sum test. Allelic richness (A_c) based on minimum sample size (accounting for sample size differences) was calculated using FSTAT (Goudet, 1995). Measures of diversity were carried out on two different groupings of the flukes: 1) within individual hosts 2) groups of hosts sampled from the same country or region. Wilcoxon sign tests in BOTTLENECK 1.2.02 (Cornuet & Luikart, 1996) were run to detect recent bottlenecking in populations. P values for the Wilcoxon tests were determined using 10,000 replicate estimations of heterozygosity under three models of selection - the infinite allele model (IAM; Kimura & Crow, 1964), the two-phase model (TPM; Di Rienzo, 1994), and the stepwise mutation model (SMM; Chakraborty & Nei, 1977).

To test for population structure between regions, analysis of molecular variance (AMOVA) for the partitioning of genetic diversity among and between populations pairwise F_{ST} values between the regions and countries were calculated in GenAlEx 6.5 (Peakall & Smouse, 2012). Significance of F_{ST} was tested with 999 permutations of the data and the False Discovery Rate (FDR) method used for correction of multiple comparisons (Benjamini, 2010). Pairwise F_{ST} values were represented visually with a PCA plot created in GenAlEx 6.5. The Bayesian method of clustering implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000) was also applied to assign population assignment probability estimates. STRUCTURE was run with minimum/maximum cluster settings of K = 1 to K = 30, with 5 iterations for each K value in order to test whether flukes within hosts from similar geographic areas grouped together. We also grouped flukes from all hosts representing a single country or location and tested minimum/maximum cluster settings of K=1 to K=12, again with 5 iterations for each K value. Each Markov chain consisted of 500,000 steps preceded by a burn-in of 100,000 steps under an admixture model for correlated frequencies (Falush et al., 2003). STRUCTURE HARVESTER version 0.6.93 (Earl, 2012) was used to process the STRUCTURE output, and Evanno's delta K method (Evanno et al., 2005) identified the appropriate number of clusters for the dataset. CLUMPP version 1.1 (Jakobsson & Rosenberg, 2007) was used to permute multimodal solutions for the appropriate number of clusters (K) deciphered with STRUCTURE HARVESTER. Using the K-specific output files of STRUCTURE HARVESTER, CLUMPP was run using the Full Search method on individual and population output files. DISTRUCT 1.1 (Rosenberg, 2004) was used to visualize the clusters to, which individuals belong and the degrees of admixture.

4.3.5 Mitochondrial COX1 sequence analysis

Mitochondrial sequences were also used to assess the genetic diversity and structuring of the parasite populations. Firstly, chromatograms of sequenced COX1 genes were aligned in Geneious 7.1.9 (Kearse et al., 2012). Poor quality ends and primer regions were trimmed to the shortest obtained sequence containing all polymorphic sites. The alignment was then imported into GenAlEx 6.5 (Peakall & Smouse, 2012) to calculate the number of haplotypes and unbiased haplotype diversity (uH). Alignments were also imported into MEGA 6 (Tamura et al., 2013) to calculate the number of segregating sites (S) and nucleotide diversity (∏). Relationships between haplotypes shared across the dataset were explored via a maximum parsimony network created using Network v4.611, which is designed to construct the shortest, least complex network (Bandelt et al., 1999). A median joining network of haplotypes was created using the settings: weights = 10 for each character and a 1:1 weighting of transitions/transversions and an elipson value of 0.

4.4. RESULTS

4.4.1 Assessment of a microsatellite marker panel on a geographically diverse set of D. dendriticum populations

We first assessed the full panel of all nine microsatellite markers (van Paridon et al., 2016) on the geographically diverse set of *D. dendriticum* populations obtained from Canada and Europe. After removal of missing data, two datasets were created - one containing all individuals (n = 166) and one with identical MLGs reduced to one copy each (n=152). Per

locus F_{IS}, F_{ST} and genetic disequilibrium (GD) were tested on both. A wide range of F_{IS} per locus was found in both; -0.012 to 0.392 with all MLGs included and 0.003 to 0.402 when MLGs were reduced to one copy. Five of nine loci (DdMs43, DdMs44, DdMs70, DdMs89, and DdMs93) showed significant deviation from HWE in each dataset (Supplementary Table S4.1). Global F_{ST} values were significant whether MLGs were included or reduced to one copy, which was expected as these flukes were sampled from distant geographical regions, but the value decreased when MLGs were reduced (F_{ST} = 0.082 p = 0.0001 and F_{ST} =0.065 p =0.0001). Tests for GD were significant in 16 of 36 pairwise comparisons with MLGs present and dropped to 2/36 when MLGs were reduced. We conclude that the presence of clones (identical MLGs) contributes to structure and linkage disequilibrium and those technical issues such as null alleles increased F_{IS} at five loci (DdMs43, DdMs44, DdMs70, DdMs89, DdMs93). To test the potential effects of null alleles, we investigated whether identity disequilibrium was significant in the reduced dataset (n = 152): ˆg2, the estimator of the two-locus heterozygosity disequilibrium, was not significantly greater than 0 (p = 0.758). From this analysis, we conclude that null alleles were likely present in five of nine loci (DdMs43, DdMs44, DdMs70, DdMs89, and DdMs93).

We then reduced the dataset to the four remaining loci (DdMs21, DdMs28, DdMs60 and DdMs95) and repeated the above analysis. The dataset with all identical MLGs included was n = 180 where 159 MLGs were unique and 21 MLGs appeared more than once making the reduced dataset n = 159. F_{ST} was again significant in both datasets, although slightly lower than expected when MLGs were reduced to a single copy (F_{ST} = 0.068, p = 0.001 and F_{ST} = 0.063 p =0.001). F_{IS} per locus ranged from 0.013 to 0.042 with all MLGs included, and

from -0.015 to 0.052 without, and no loci deviated significantly from HWE. Global F_{IS} was not significantly greater than zero in either dataset (F_{IS} = 0.023 p= 0.19040) and (F_{IS} = 0.028 p = 0.14510). P_{SEX} values for these 21 repeated MLGs were 0.0479. Since the P_{SEX} values for the repeated MLGs were extremely close to 0.05, suggesting the repeated MLGs in the dataset may have arisen by chance rather than asexual multiplication, we retained any MLGs appearing in two different final hosts, but reduced those found within the same final host to one copy. We did this because it is more likely that MLGs found within a single host are the result of clones that were transmitted together throughout the life cycle, while those occurring across geographic distance are likely not clones.

We concluded that five of the loci (DdMs43, DdMs44, DdMs70, DdMs89, and DdMs93) have null alleles in some populations and that four of the loci (DdMs21, DdMs28, DdMs60 and DdMs95) do not. Consequently, although we undertook subsequent population genetic analyses using all nine loci, we also undertook a separate analysis using the subset of four markers without null alleles (results of the analysis using only markers DdMs21, DdMs28, DdMs60 and DdMs95 is included as supplementary material in Appendix I; Supplementary Fig. S4.1-S4.5). The effects of null alleles on the dataset were investigated by running the analysis with and without suspected loci, as their presence would result in heterozygous individuals being scored as homozygous. This would then decrease measures of diversity and heterozygosity. If these null alleles affected populations' disproportionately, then exaggerated structure may be observed, leading to false conclusions. Our results showed no differences in the conclusions drawn from the two analyses, leading us to conclude that their effects on the analysis of structure would be limited. Also, when repeated MLGs were

included there were increases in F_{ST} and genetic disequilibrium suggesting the presence of clonal genotypes in hosts was responsible for this increase. Consequently, we reduced repeated MLGs to one copy per final host to eliminate significant structure appearing for this reason alone.

4.4.2 Salt Spring Island and Vancouver Island D. dendriticum populations show low genetic diversity, whereas the Alberta population shows high genetic diversity relative to global populations

We first assessed the overall levels of genetic diversity of the microsatellite marker panel in each parasite population taken from individual hosts. Measures of genetic diversity (H_s) were not significantly different in any grouping of hosts and so H_s values from FSTAT are reported here. H_s within individual hosts varied from 0.278 in Spain-2 to 0.861 in Greece-1 (Fig. 4.1). Allelic richness (A_c) adjusted for the smallest sample size ranged from 1.62 in Salt Spring-1 to 2.94 in Alberta-3 (Fig 4.1). Of the Western Canadian populations, all three populations from Salt Spring Island hosts were amongst the lowest by both measures of genetic diversity and the populations from Vancouver Island hosts had low levels of genetic diversity compared with most populations from other regions. In contrast, the parasite populations from Alberta were amongst the most genetically diverse (Fig. 4.1). When individual hosts were grouped into populations based on geographical location, values of H_s ranged from 0.396 (Salt Spring Island) to 0.809 (Greece) (Fig 4.2). A_c in hosts grouped into populations ranged from 2.62 (Salt Spring Island) to 3.89 (Greece) (Fig 4.2). Salt Spring Island was the lowest by both measures and Alberta was second highest. The Vancouver

Island population was ranked third lowest according to values of H_s and fourth lowest according to A_c. The results of the BOTTLENECK analysis showed potential bottlenecking events in three populations, Salt Spring Island, Vancouver Island and Norway. Wilcoxon sign tests showed a significant heterozygote excess based on all three models for Salt Spring (P value for each model tested was 0.0039) and based on only the infinite alleles model for Vancouver Island (P = 0.0137) and Norway (P = 0.0371).

4.4.3 Microsatellite genotyping reveals significant geographical population structure of D. dendriticum

AMOVA analysis revealed that 6% of the overall genetic variation, based on genotypes of the nine microsatellite markers, was explained by differences between parasites from the twelve geographical regions sampled, 17% between hosts, and 77% among parasites within an individual host. To next test for genetic structure between geographic regions, we calculated pairwise F_{ST} between all individual hosts (n = 30) and plotted the results with PCA (Fig. 4.3). A PCA plot of F_{ST} values between parasites isolated from individual hosts showed a trend for grouping according to geographical region, suggesting strong population structure (Fig. 4.3). We then further tested if parasite populations from individual hosts clustered by geographical region using the Bayesian assignment protocol of STRUCTURE. When this analysis was run with each host as a separate population, the optimal value of K was 2 with the individual host populations from Salt Spring Island and Vancouver Island hosts forming one cluster and the Alberta, Greek, Italian and Norwegian host forming the second cluster. All remaining populations showed admixture between the two clusters (Supplementary Data

Fig. S4.4). As K was increased, hosts from the same geographical region continued to remain within the same cluster showing that hosts sampled from the same region are more similar to each other than to those sampled from other regions (Supplementary Data Fig. S4.4).

4.4.4 Salt Spring Island and Vancouver Island D. dendriticum populations share a close genetic relationship with each other but not with the Alberta population

Pairwise F_{ST} values were calculated between the *D. dendriticum* populations in each of the 12 geographical regions to determine the relationships of the Western Canadian populations with each other and with the other global populations (Table 4.2). The most highly divergent of the 12 geographical regions overall was the Salt Spring Island population with all comparisons being significant after correction. The lowest pairwise F_{ST} value for the Salt Spring Island population was with the Vancouver Island population (F_{ST} = 0.054). The next most genetically divergent of the twelve regions overall, based on pairwise F_{ST}, was the Vancouver Island population with 9 of 11 significant pairwise comparisons; only pairwise comparisons with Albania and Germany were not significant (Table 4.2). The Alberta population showed significant F_{ST} values with eight other populations. Notably, in the case of the Alberta population, the geographically closest populations were the most genetically divergent. The pairwise F_{ST} values, which were found to be non-significant for Alberta, were the comparisons with Germany (0.00), Italy (0.007) and Albania (0.037). Visualisation of pairwise F_{ST} using PCA showed that the Salt Spring Island and Vancouver Island populations were separate from the main cluster, with the Salt Spring Island population being the most

isolated. The Alberta population was grouped with the main cluster and is closest to Norway and Germany (Fig 4.4).

Bayesian assignment in STRUCTURE performed to compare the *D. dendriticum* populations from each of the 12 geographical regions determined that K = 2 clusters was the optimal number of clusters. In this analysis, the Salt Spring Island and Vancouver Island populations formed a distinct cluster with little admixture. Alberta, Greece, Italy and Norway again formed a separate second cluster with little admixture and the remaining populations split between the two clusters (Fig. 4.5). To further investigate potentially admixed populations and genetic relationships not shown at K = 2, we investigated an increasing number of K values. At K = 3, Salt Spring Island and Vancouver Island remained a group, Alberta continued to cluster with the Eastern Europe countries Greece and Italy, and a third Western European cluster formed with Spain and the UK. Newfoundland, Ontario, Germany and Norway most resemble the Western Europe group with some admixture with the Eastern Europe group. The Albanian population was split between the three groups. As K was increased to K = 4 and K = 5 this pattern of three distinct clusters remained; i.e., the west coast Canadian populations, Western Europe, and Eastern Europe. The Canadian populations of Ontario and Newfoundland resembled the Western European cluster, while the Alberta population formed a separate cluster (Fig. 4.5). A random sample of nine worms from Alberta was taken and the analysis re-run to investigate potential bias caused by the larger numbers of individuals sampled from this location. This process was repeated three times, each with a new sample of nine worms. The results of this analysis were consistent with the broad trends and groupings, but showed some additional structure within some of the groups

described in the original run at K = 3 to K = 5 (Supplementary Data Fig. S4.6), specifically, the smaller groupings that included Alberta flukes show admixture, and clustered mainly with the Eastern European group. The West Coast populations also showed more admixture, with the Salt Spring Island group splitting into two groupings shared with only Vancouver Island. The latter also shows some clustering with the UK group.

4.4.5 COX1 mtDNA sequence analysis is consistent with the microsatellite genotyping results and suggests multiple geographical origins of the Alberta D. dendriticum population.

A 794 bp fragment of the *D. dendriticum* COX1 gene was sequenced from 213 individual adult parasites across the 12 geographical regions. Sequence alignments showed 42 segregating sites across the 794 bp fragment. Measures of genetic diversity at both the nucleotide and haplotype level indicated the Alberta population was the most diverse with values of 0.0156 and 0.294, respectively (Table 4.3). Salt Spring Island was the lowest by both measures, nucleotide diversity was 0.0002, and haplotype diversity was 0.004 (Table 4.3). The Vancouver Island population also had low diversity for all measures; i.e., 0.0014 and 0.037 for nucleotide and haplotype diversity, respectively (Table 4.3).

Haplotype network analysis of all 213 individuals revealed 30 unique haplotypes overall with 20 haplotypes being present in European populations and 16 in the Canadian populations (Fig 4.7). Of these, 20 haplotypes were unique to a single country or region and 10 were shared between at least two regions. Of the 10 shared haplotypes, H1 was the most common being found in all but four regions; UK, Norway, Albania and Greece (Fig 4.6). Considering the European regions first, the shared haplotypes were distributed into two broad

groups in Europe. A group of four haplotypes, (H1-H4) were found predominantly in the UK and Spain forming a Western European group while haplotypes H7-H10 were found predominantly in Norway, Albania and Greece forming the Eastern European group. Germany and Italy contained haplotypes from both groups. Considering the Canadian populations in relation to these shared haplotypes, Salt Spring Island contained only a single shared haplotype (H1) and Vancouver Island just two shared haplotypes (H1 and H2), which are from the Western European group. Newfoundland and Ontario populations also contained those shared haplotypes predominantly found in Western European haplotypes (H1 H2 and H4) and, like Salt Spring Island and Vancouver Islands, no shared haplotypes from the Eastern European group. Of the Canadian populations, Alberta was the sole population to contain haplotypes from both the western and eastern European groups including two haplotypes otherwise unique the Eastern European group (H5 and H10).

All 30 COX1 haplotypes (both shared and unique) were placed in a median joining network to examine their relatedness. The network revealed two distinct groups of haplotypes separated by 17 single nucleotide polymorphisms; one containing 16 and the other 14 haplotypes (Fig 4.7). These two groupings broadly correspond to the eastern and western European split of shared haplotypes with the UK and Spain appearing solely on the left group and Greece, Albania and Norway solely in the right (as orientated in Figure 4.7). The Alberta population, along with Germany and Italy, were the only populations that contained haplotypes from both sides of the network. These three populations contained nine, five and four haplotypes respectively. Alberta had three unique haplotypes, two of which varied by only one base pair from H9 (an Eastern European shared haplotype) and the third differing by

one base pair from H1 (a Western European shared haplotype). Germany also had three unique haplotypes, one which was shared with only Alberta and Italy had two unique haplotypes. Three of the five haplotypes found in the Vancouver Island population were unique to this area and differed by only one base pair from H1 and H2. The Salt Spring Island population had the lowest number of haplotypes with only two, one of which was unique to Salt Spring Island and only varied by one base pair from the H1 haplotype.

4.5 DISCUSSION

The distribution of *D. dendriticum* is considered to be long standing in Europe and a relatively recent invader into North America, presumably having been translocated by anthropogenic animal movement (Otranto & Traversa, 2003). We have undertaken a population genetic study on a number of populations from Canada and compared them with longer established European populations, the likely region from where the Canadian parasite populations originated.

Analyses with microsatellite and mtDNA markers were congruent overall and revealed significant population structure for both Canadian and European populations. There is a marked east-west divide in Europe according to both the microsatellite markers and COX1 gene sequence analysis. Worms from the UK and Spain clustered as a western European group and worms from Greece and Albania as an eastern European group on F_{ST} PCA plots and STRUCTURE analysis of microsatellite data, and COX1 mtDNA haplotype network analysis. The two central populations of Germany and Italy appear to be a mix of these two broad regional populations, based on the presence of haplotypes from both groups

in the haplotype network and their central position within the PCA plots and evidence of admixture in the clustering analysis. The Norwegian flukes were also central on the PCA plots but only had haplotypes from the eastern European group. These results suggest that two lineages of *D. dendriticum* co-occur in Europe, with a zone of mixing in the middle of the continent. This pattern is consistent with the mixed regions representing the remnants of the original population that has since spread east and west and become genetically isolated, similar to the pattern reported for *Fascioloides magna* on the endemic continent of North America (Juhásová et al., 2016). Further sampling of European worms is required to test the robustness of these patterns.

The observed population structure in Europe provides an opportunity to infer the likely geographical origins of *D. dendriticum* in Canada. The eastern Canadian populations (Newfoundland and Ontario) are genetically more similar to western European than to eastern European populations based on both the microsatellite and COX1 mtDNA sequence data. The Ontario and Newfoundland populations cluster most closely with the UK, Spain and Germany populations based on F_{ST} PCA and STRUCTURE microsatellite analysis and are most distant to Albania and Greece of the European populations. Similarly, the Newfoundland and Ontario populations share two COX1 mtDNA haplotypes with the western or central European populations and not with the Eastern European populations; H1 is shared with UK, Spain, Germany and Italy and H2 is shared with UK and Spain with no haplotypes being shared with the eastern European populations. When we re-ran STRUCTURE with the resampled Alberta flukes, the Ontario flukes showed more admixture with the Eastern European group, which is inconsistent with rest of the data. One explanation

for this is the non-recombining nature of the mtDNA marker, meaning mtDNA identity can resemble one group while the microsatellite markers could show admixture. Further sampling of flukes from this region may reveal more mtDNA haplotypes from the Eastern European group are present in Ontario. The use of a more robust marker system such as genome wide single nucleotide polymorphisms (SNPs) might better capture deeper levels of sub-structuring occurring in our sampled populations. Given that the earliest Canadian reports of *D. dendriticum* are from Eastern Canada, this suggests the initial invasion of Canada by *D. dendriticum* was into Eastern Canada from a Western European source, most likely by anthropogenic movement of domestic livestock.

In the case of the Canadian West Coast *D. dendriticum* populations, the genetic diversity of both the microsatellite and COX1 mtDNA markers suggests that these populations – from Salt Spring Island and Vancouver Island – are the most recently established of all the populations examined. They showed the lowest genetic diversity by measures of diversity with both nuclear and mitochondrial markers. Decreased genetic diversity is often a sign of a recent bottleneck events associated with introductions (Bryan et al., 2005). We tested these lower-diversity populations for the presence of heterozygote excess to determine if a bottleneck had occurred. Heterozygote excess is a useful test for bottlenecking because low-frequency alleles present in source populations are less likely to be sampled and transported with the invasive population compared to more common alleles (Cornuet & Luikart, 1996). The Salt Spring Island population had a significant heterozygous excess regardless, of which of the three mutation models were evaluated, providing evidence of a recent bottleneck (Cornuet & Luikart, 1996). The most realistic model of mutation for

microsatellites is the TPM model (Di Rienzo et al., 1994). The Vancouver Island population only showed a significant heterozygote excess when the infinite alleles model was evaluated, suggesting the effect was less pronounced than that on Salt Spring Island.

This more recent invasion of the Canadian west coast, suggested by the genetic data, is generally consistent with the limited historical information. *Dicrocoelium dendriticum* was first anecdotally reported on the Canadian west coast, as late as ~1965, with the first published report being in cattle from a ranch in Sechelt, British Columbia in 1974 and anecdotal reports in sheep around the same time (Lewis, 1974; M. Bond, personal communication). The genetic data also support a common origin of these west coast island populations. The pairwise F_{ST}, PCA and STRUCTURE analyses of the microsatellite and COX1 mtDNA data indicates these two populations are much more similar to each other than to any of the other populations examined. The resampled STRUCTURE runs do indicate that Vancouver Island is admixed, and is split with clusters resembling the UK flukes and clusters found only on the Salt Spring Island, indicating Vancouver Island as a likely source of the Salt Spring worms. This may also explain the stronger evidence for bottlenecking seen in the latter population. Further, they are amongst the most isolated populations overall, which is probably a reflection of their low genetic diversity as a result of population bottlenecking. The only two COX1 mtDNA haplotypes present in these western Canadian populations (H1 and H2) are also present in the eastern Canadian populations and shared with Western but not the Eastern European populations. This is consistent with translocation from Canadian east coast populations to the west with some loss of genetic diversity due to population bottlenecking. Similar patterns of isolation of invasive parasite populations have been

observed in Europe. In a population genetic analysis of adult *Fascioloides magna*, Juhásová et al. (2016) showed evidence for two separate introduction events into continental Europe from western North America with serial dilution along the dispersal route from a presumed second invasion site located in Europe. This general pattern was confirmed through analyses of mtDNA sequences of the same populations (Králová-Hromadová et al., 2011).

Albertan *D. dendriticum* have a markedly different genetic structure compared to the other Canadian populations, suggesting more complex origins. The Alberta population is remarkably diverse considering its apparent recent introduction. *Dicrocoelium dendriticum* was not identified in Alberta until 1990, despite extensive host surveys of wild and domestic hosts (Pybus, 1990). Since the initial reports of infection in this region, prevalence has increased with at least four species of sympatric final hosts being infected (Goater & Colwell, 2007). Despite being the most recently reported region in Canada to harbour *D. dendriticum*, the Alberta population was the most genetically diverse of any Canadian population. There was also no evidence for bottlenecking in the Alberta population. In previous population genetic studies on this population, there was no evidence of genetic structure existing between worm populations in elk versus cattle (van Paridon et al., 2016). Thus, it is unlikely that the overall pattern of high genetic diversity of the Alberta population is due to any genetic partitioning between host species. The most likely explanations for the unexpectedly high genetic diversity in the Alberta *D. dendriticum* population is that there have been multiple introductions from different sources or a separate introduction from a single highly admixed source. The clearest indication of this is from the COX1 mtDNA data. Alberta is the only Canadian population that contains haplotypes shared with both Western and Eastern

European populations: H1 and H2 haplotypes are shared with all other Canadian populations and with Western European but not Eastern European populations. In contrast, H9 and H10 haplotypes are found nowhere else in Canada and are unique to Eastern European populations. This suggests the origin of the Alberta populations is different to the other Canadian populations perhaps with at least two separate introductions; likely one from elsewhere in Canada and one from outside Canada, perhaps Eastern Europe. This hypothesis is supported by the STRUCTURE analysis of microsatellite data, which in the original analysis that included all the Alberta flukes, clusters them with Greece and Italy until K is increased such that Alberta is forced into its own group at K = 4.

However, after adjusting the sample size to better resemble the other locations, a more distinct pattern of admixture emerged. The resampled STRUCTURE runs show Alberta grouping consistently with the Eastern Europe group with some admixture with the UK, which is supported by the shared mtDNA haplotypes from both regions. In these runs of STRUCTURE, Ontario and Alberta also show similar patterns of structure, which may be a signal that Ontario has a similar invasion history or may have contributed to the Albertan population. Increased sampling from all regions of Canada may support multiple introductions into Eastern Canada from Europe and multiple routes of spread west. The incorporation of additional sampling and a more robust marker system would be the next step to providing further evidence of the patterns of movement of the fluke throughout North America.

Ideally, parasite invasion scenarios that are indirectly inferred from population genetic data should be matched with information on host movements. The invasion pathway

of the nematode *Nematodirus battus* into western North America and northern Europe could be tied to known importation routes of domestic sheep from central Europe (Nadler et al., 2000). The timing of the establishment of the liver fluke, *F. magna* in ungulate populations in Eastern Europe coincided with the importation of elk from western North America to support local hunting economies (Králová-Hromadová et al., 2011). Unfortunately, information on the historical movement of potential hosts of *D. dendriticum* in Canada, particularly wildlife, is very limited. The historical roots of livestock production in eastern Canada can be traced to the importation of cattle and sheep to support trading posts (I. MacLachlan, personal communication.). The earliest cattle were imported in the 16th and 17th centuries from the Iberian Peninsula in Spain, and the Brittany region of France, into the Acadia Valley and the St. Lawrence River Valley into what is now Nova Scotia, Ontario, and Quebec, respectively. Each of these source populations of cattle include regions, in which livestock are currently infected with *D. dendriticum*. The importation of livestock into the western Canadian prairies did not occur until the 1880s (Evans, 1978). Initially, the main flow of beef cattle that supplied early homesteaders in southern Alberta was from large ranches in the US, especially Texas. But catastrophic winter-kill events in 1887 and 1907 encouraged ranchers to diversify their stock to focus on the Hereford and Aberdeen Angus breeds that dominate contemporary ranching operations in this region. This original stock was imported onto the southern Canadian prairies starting in the late 1800's from sources in eastern Canada, and also directly from Britain (Evans, 1978). These limited anecdotal observations are consistent with two central conclusions based upon our population genetics data. First, the opportunities for introduction of *D. dendriticum* from Europe into Canada via the importation of domestic

stock were present hundreds of years earlier in eastern Canada compared to opportunities available in the west. Second, opportunities for multiple introductions of *D. dendriticum* into the Cypress Hills region exist through animal translocations from known regions of infection in eastern Canada and Western Europe.

In summary, our population genetic data provide the first insights into the invasion history of *D. dendriticum* into Canada. The combined historical and genetic evidence suggests that the parasite arrived in eastern Canada via animal trade with countries in Western Europe with the populations found on the coastal islands of BC being the most recent introductions. The invasion history of the Alberta population is likely to be quite different with at least two independent origins; one likely being from elsewhere in Canada and the other from outside the country, perhaps from Eastern Europe. Of course, these two origins need not be direct introductions and many more parasite populations will need to be sampled to produce a more detailed view of the invasion history. Overall, this work provides a proof of concept of how population genetic data can be used to help elucidate the invasion history of a very recently established non-indigenous parasite into a region.

4.6 ACKNOWLEDGMENTS

The authors thank Tanya Dunlop, Tim Goater, Kyran Power and Danielle Broaders for providing flukes from British Columbia and Newfoundland. We also thank Stefania Weigl for providing flukes from Italy, Neil Sargison for the flukes from the UK and Norway, and Yolanda Manga for the samples from Spain. Assistance with microsatellite genotyping was provided by Sean Rogers, Elizabeth Redman and Jocelyn Poissant. This work was

funded by a grant from the National Centre for Veterinary Parasitology (http://www.ncvetp.org/) to B. van Paridon and grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to CPG (#40130) and JSG (#371529-2209). B. van Paridon is a member of the NSERC-CREATE Host Pathogen Interactions (HPI) graduate training program at the University of Calgary.

4.7 LITERATURE CITED

Arnaud-Haond, S., & Belkhir, K. (2007). genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes, 7,* 15-17.

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16,* 37-48.

Beck, M. A., Goater, C. P., Colwell, D. D., & van Paridon, B. J. (2014). Fluke abundance versus host age for an invasive trematode (*Dicrocoelium dendriticum*) of sympatric elk and beef cattle in southeastern Alberta, Canada. *International Journal for Parasitology: Parasites and Wildlife,3,* 263-268.

Benjamini, Y. (2010). Discovering the false discovery rate. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 72*, 405-416.

Blouin, M. S., Yowell, C. A., Courtney, C. H., & Dame, J. B. (1995). Host movement and the genetic-structure of populations of parasitic nematodes. *Genetics, 141*, 1007-1014.

Bryan, M. B., Zalinski, D., Filcek, K. B., Libants, S., Li, W., & Scribner, K. T. (2005). Patterns of invasion and colonization of the sea lamprey (*Petromyzon marinus*) in North America as revealed by microsatellite genotypes. *Molecular Ecology, 14*, 3757-3773.

Chakraborty, R., & Nei, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution, 31*, 347-356.

Cleaveland, S., Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356,* 991-999.

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics, 144,* 2001-2014.

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. *Molecular Ecology, 20*, 2510-2524.

Cristescu, M. E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology, 24,* 2212-2225.

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science, 287,* 443-449.

David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology, 16,* 2474-2487.

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., & Freimer, N. B. (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Science U S A, 91,* 3166-3170.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources, 4,* 359-361.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14,* 2611-2620.

Evans, S. M. (1978). Stocking the Canadian range. *Alberta History, 26,* 1-8.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics, 164,* 1567-1587.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: *Dicrocoelium dendriticum* in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93,* 491-494.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity, 86,* 485-486.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23,* 1801-1806.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature, 451,* 990-993.

Juhásová, Ludmila, Ivica Królová-Hromadová, Eva Bazsalovicsová, Gabriel Minárik, Jan Štefka, Peter Mikulíček, Lenka Pálková, & Pybus, M. (2016) Population structure and dispersal routes of an invasive parasite, *Fascioloides magna,* in North America and Europe. *Parasites & Vectors, 9,* 547.

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

Kimura, M., & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics, 49*, 725-738.

Králová-Hromadová, I., Bazsalovicsová, E., Štefka, J., Špakulová, M., Vávrová, S., & Szemes, T. (2011). Multiple origins of European populations of the giant liver fluke *Fascioloides magna* (Trematoda: Fasciolidae), a liver parasite of ruminants. *International Journal for Parasitology, 41*, 373-383.

Krull, W. H., & Mapes, C. R. (1952). Studies on the biology of *Dicrocoelium dendriticum* (Rudolphi, 1819) Looss, 1899 (Trematoda Dicrocoeliidae), including its relation to the intermediate host, *Cionella-lubrica* (MULLER) .7. The 2nd intermediate host of *Dicrocoelium dendriticum*. *Cornell Veterinarian, 42*, 603-604.

Lewis, P. D. (1974). Presence of Dicrocoelium-denriticum (Rudolphi, 1819) (Trematoda Dicrocoeliidae) in western Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie, 52*, 662-663.

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123*, 91-114.

Nadler, S. A., Hoberg, E. P., Hudspeth, D. S., & Rickard, L. G. (2000). Relationships of *Nematodirus* species and *Nematodirus battus* isolates (Nematoda: Trichostrongyloidea) based on nuclear ribosomal DNA sequences. *Journal of Parasitology, 86*, 588-601.

Otranto, D., & Traversa, D. (2003). Dicrocoeliosis of ruminants: a little known fluke disease. *Trends in Parasitology, 19*, 12-15.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics, 28*, 2537-2539.

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

Prugnolle, F., Liu, H., de Meeus, T., & Balloux, F. (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology, 35*, 255-263.

Pybus, M. J. (1990). Survey of hepatic and pulmonary helminths of wild cervid in Alberta, Canada. *Journal of Wildlife Diseases, 26*, 453-459.

Rhyan, J. C., & Spraker, T. R. (2010). Emergence of diseases from wildlife reservoirs. *Veterinary Pathology, 47*, 34-39.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes, 4*, 137-138.

Rousset, F. (2008). GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources, 8*, 103-106.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution, 30*, 2725-2729.

Thompson, R. C. A., Lymbery, A. J., & Smith, A. (2010). Parasites, emerging disease and wildlife conservation. *International Journal for Parasitology, 40*, 1163-1170.

van Paridon, B. J., Goater, C. P., Gilleard, J. S., & Criscione, C. D. (2016). Characterization of nine microsatellite loci for *Dicrocoelium dendriticum*, an emerging liver fluke of ungulates in North America, and their use to detect clonemates and random mating. *Molecular and Biochemical Parasitology, 207*, 19-22.

Table 4.1. Host identifications (IDs), geographical locations and host species of *D. dendriticum* hosts sampled and number of flukes (N) genotyped from each host using either microsatellite genotyping or COX1 gene sequencing.

			Microsatellites	COX1
Host ID	Location	Host Species	N	N
Albania	Albania	Sheep (single)	5	12
Alberta1	Cypress Hills, Alberta	Cow (single)	10	10
Alberta2	Cypress Hills, Alberta	Cow (single)	10	10
Alberta3	Cypress Hills, Alberta	Cow (single)	10	10
Alberta4	Cypress Hills, Alberta	Elk (single)	10	10
Alberta5	Cypress Hills, Alberta	Elk (single)	10	10
Alberta6	Cypress Hills, Alberta	Elk (single)	10	10
Alberta7	Cypress Hills, Alberta	Elk (single)	6	6
Germany	Swabian Alps, Germany	Sheep (single)	5	6
Glasgow	Glasgow, United Kingdom	???	5	5
Greece1	Lagadas,Greece	Sheep (single)	5	10
Greece2	Lagadas,Greece	Sheep (single)	5	10
Greece3	Lagadas,Greece	Sheep (single)	5	10
Ile Coll	Ile of Coll, United Kingdom	Sheep (single)	5	5
Italy	Apulia, Italy	Sheep (multiple)	5	5
Newfoundland	Branch, Newfoundland	Hares (multiple)	12	11
Norway1	Hordland County, Norway	???	5	8
Norway2	Oppdal, Norway	???	5	8
Ontario	Ontario	Sheep (multiple)	10	10
SaltSpring1	Salt Spring Island	Sheep (single)	3	3
SaltSpring2	Salt Spring Island	Sheep (single)	4	4
SaltSpring3	Salt Spring Island	Sheep (single)	4	4
Spain1	Salamanca,Spain	Sheep (single)	5	5
Spain2	Salamanca,Spain	Sheep (single)	5	5
Spain3	Leon, Spain	Ox (single)	5	5
Spain4	Villa de Soto, Leon, Spain	Cow (single)	5	5
VanIsl1	Parksville, Vancouver Island	Sheep (single)	5	4
VanIsl2	Parksville, Vancouver Island	Sheep (single)	5	4
VanIsl3	Parksville, Vancouver Island	Sheep (single)	5	5
VanIsl4	Parksville, Vancouver Island	Sheep (single)	5	3
Total			189	213

Table 4.2. Pairwise F_{ST} comparisons between *D. dendriticum* flukes from hosts grouped into 12 geographical regions based on nine microsatellite loci. Values in bold are significant after FDR correction of P values for multiple comparisons.

	Alberta	Albania	Ontario	Salt Spring Island	Vancouver Island	Newfoundland	Germany	Spain	UK	Greece	Italy	Norway
Alberta	0.000											
Albania	0.037	0.000										
Ontario	**0.042**	0.064	0.000									
Saltspring	**0.121**	**0.157**	**0.147**	0.000								
Vancouver Island	**0.059**	0.038	**0.081**	**0.054**	0.000							
Newfoundland	**0.038**	**0.063**	**0.033**	**0.094**	**0.029**	0.000						
Germany	0.000	0.003	0.005	**0.110**	0.015	0.000	0.000					
Spain	**0.055**	**0.086**	**0.051**	**0.148**	**0.075**	**0.034**	0.022	0.000				
UK	**0.049**	**0.071**	0.016	**0.158**	**0.053**	**0.042**	0.000	0.006	0.000			
Greece	**0.048**	0.000	**0.058**	**0.156**	**0.097**	**0.082**	0.032	**0.118**	**0.101**	0.000		
Italy	0.007	0.029	**0.064**	**0.181**	**0.062**	0.036	0.006	**0.057**	**0.046**	0.039	0.000	
Norway	**0.025**	0.019	**0.061**	**0.109**	**0.035**	**0.037**	0.013	**0.036**	**0.036**	**0.054**	0.017	0.000

Table 4.3. Summary of diversity statistics for COX1 gene sequences of *D. dendriticum* sampled from hosts grouped into 12 geographical regions and for Canada and Europe separately. m = the number of sequences, S = segregating sites, π = nucleotide diversity, uH = unbiased haplotype diversity.

Region	m	S	π	uH
Saltspring Island	11	1	0.000229	0.004
UK	10	2	0.000701	0.013
Ontario	10	5	0.001595	0.030
Vancouver Island	16	4	0.001370	0.037
Spain	20	6	0.002250	0.042
Albania	12	6	0.002465	0.043
Norway	16	6	0.002480	0.047
Greece	30	8	0.002595	0.047
Newfoundland	11	9	0.003856	0.066
Italy	5	23	0.012091	0.229
Germany	6	29	0.013854	0.254
Alberta	66	33	0.015552	0.294
Canada	114	31	0.010235	0.223
Europe	99	37	0.015670	0.296

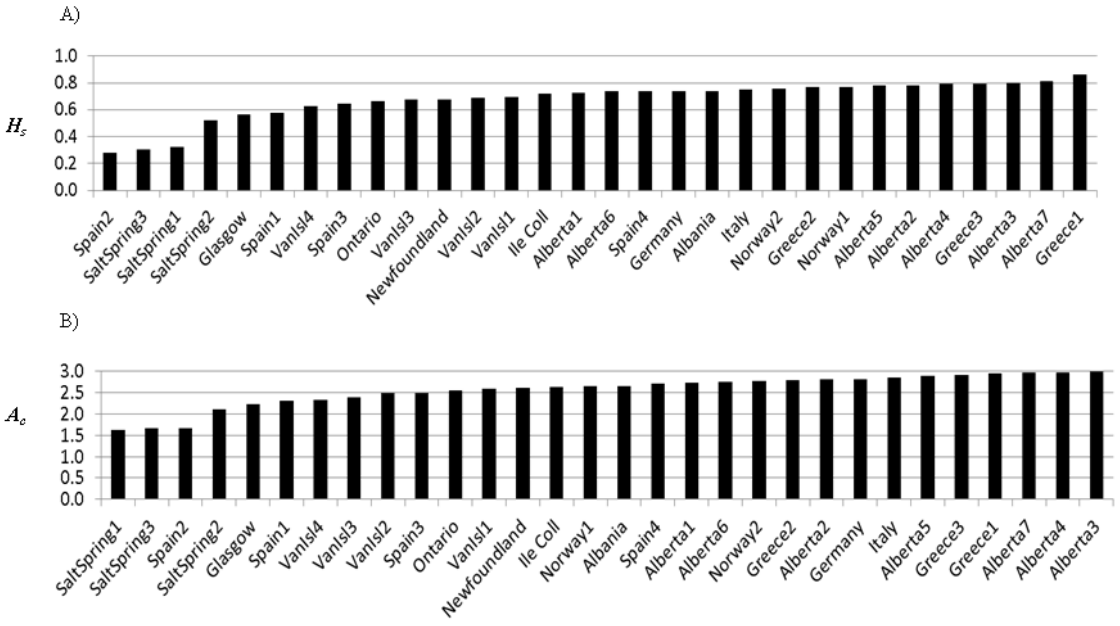


Figure 4.1. Genetic diversity (H_s) (A) and allelic richness (A_c) (B) of *D. dendriticum* within each individual host based on nine microsatellite loci.

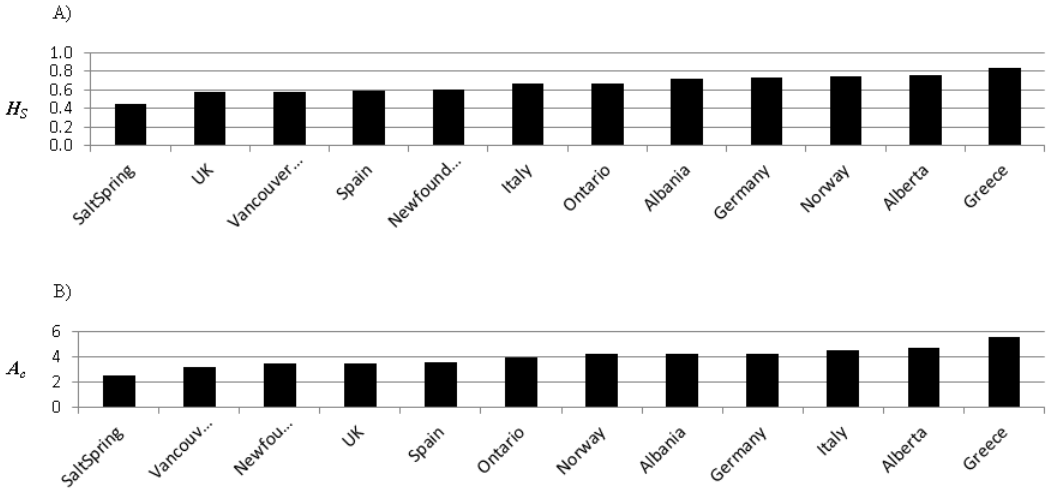


Figure 4.2. Genetic diversity (H_s) (A) and allelic richness (A_c) (B) of *D. dendriticum* within hosts grouped into regions based on nine microsatellite loci.

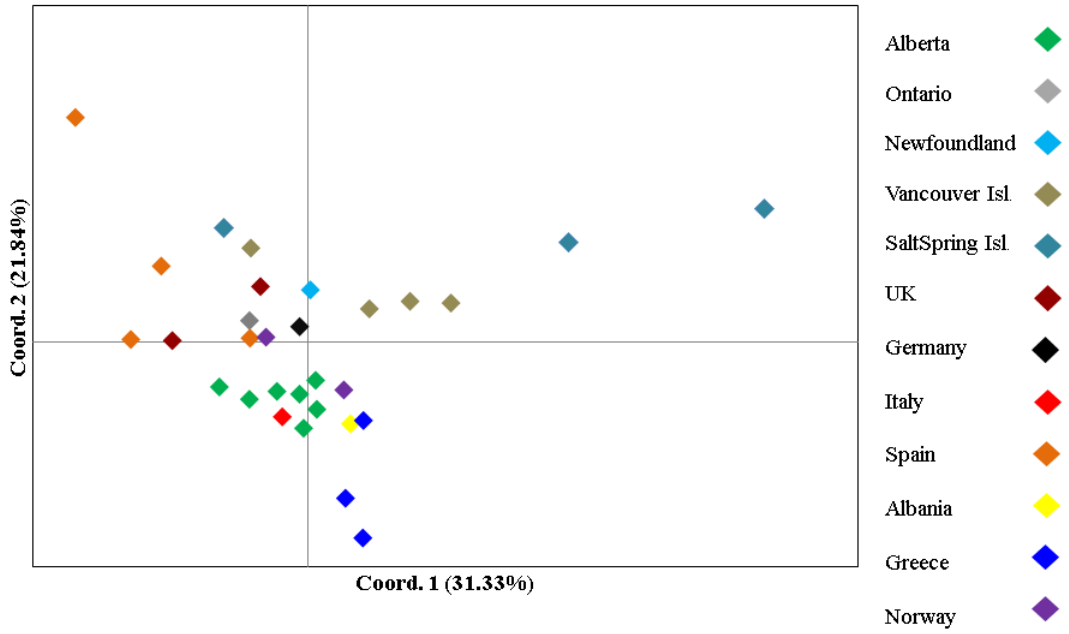


Figure 4.3. PCA plot of pairwise F_{ST} values between *D. dendriticum* within individual hosts based on nine microsatellite loci.

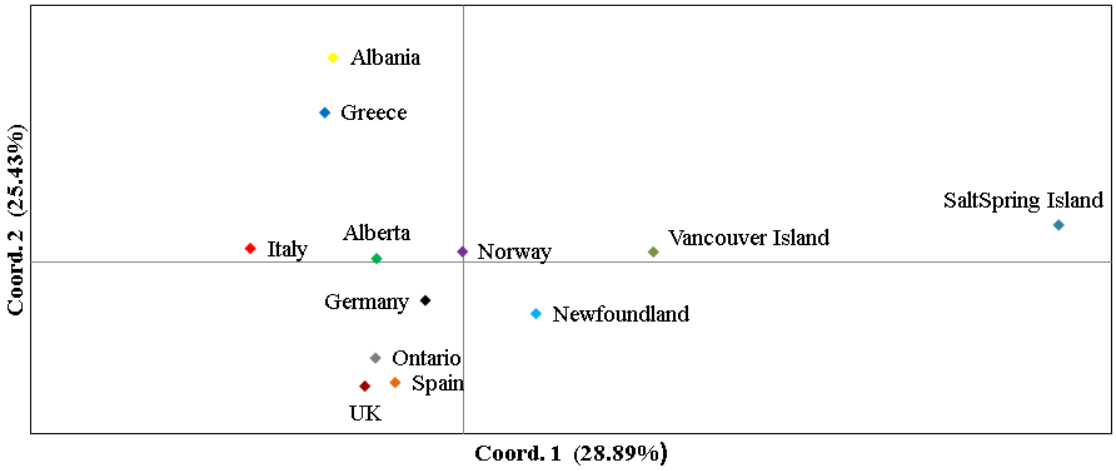


Figure 4.4. PCA plot of pairwise F_{ST} values between *D. dendriticum* in hosts grouped as regions based on nine microsatellite loci.

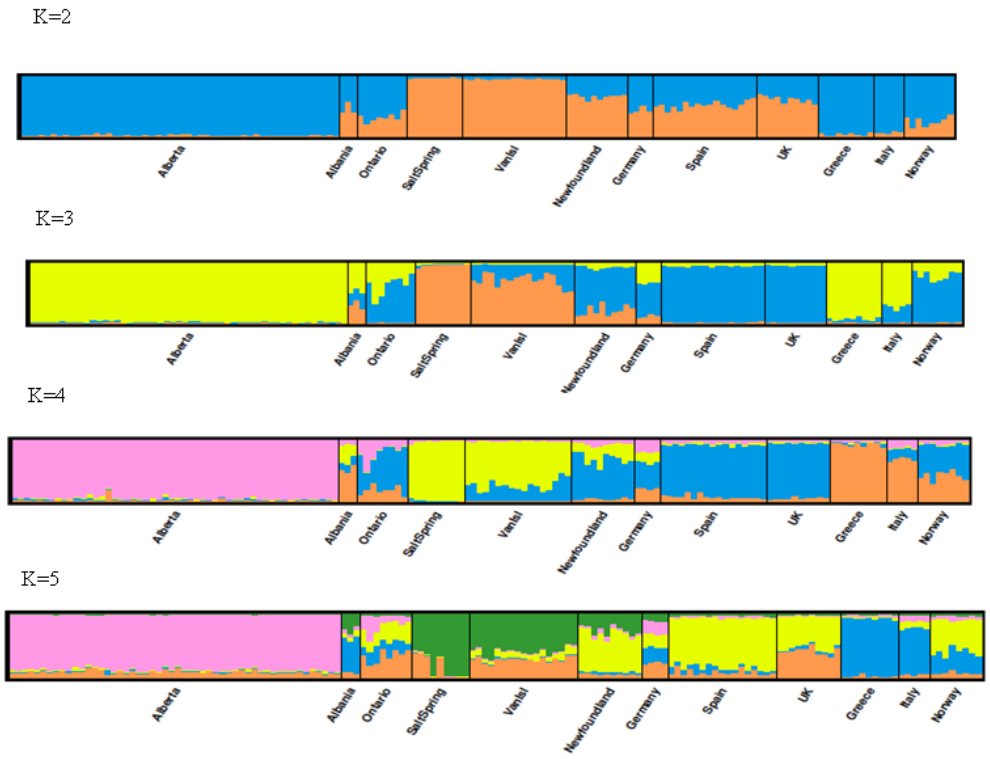


Figure 4.5. Results of STRUCTURE clustering analysis of *D. dendriticum* from each of the 12 geographical regions based on nine microsatellite loci. Values of K=2-5. Region IDs are shown below each plot.

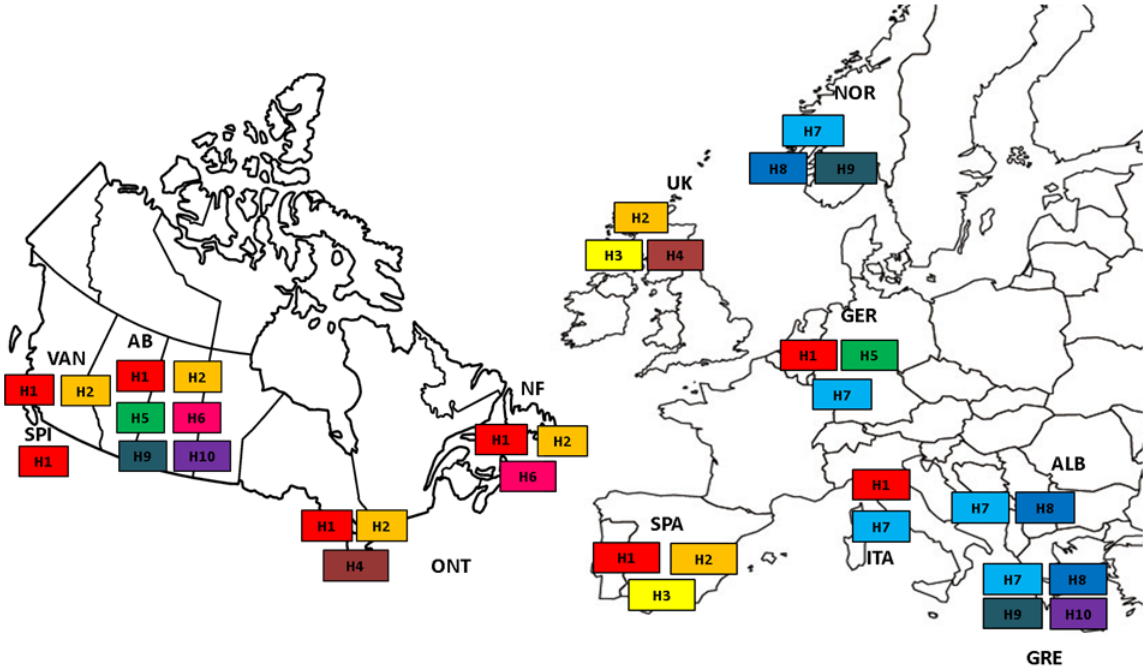


Figure 4.6. Maps of Canada and Europe showing shared COX1 haplotypes of *D. dendriticum* present in each geographical region.

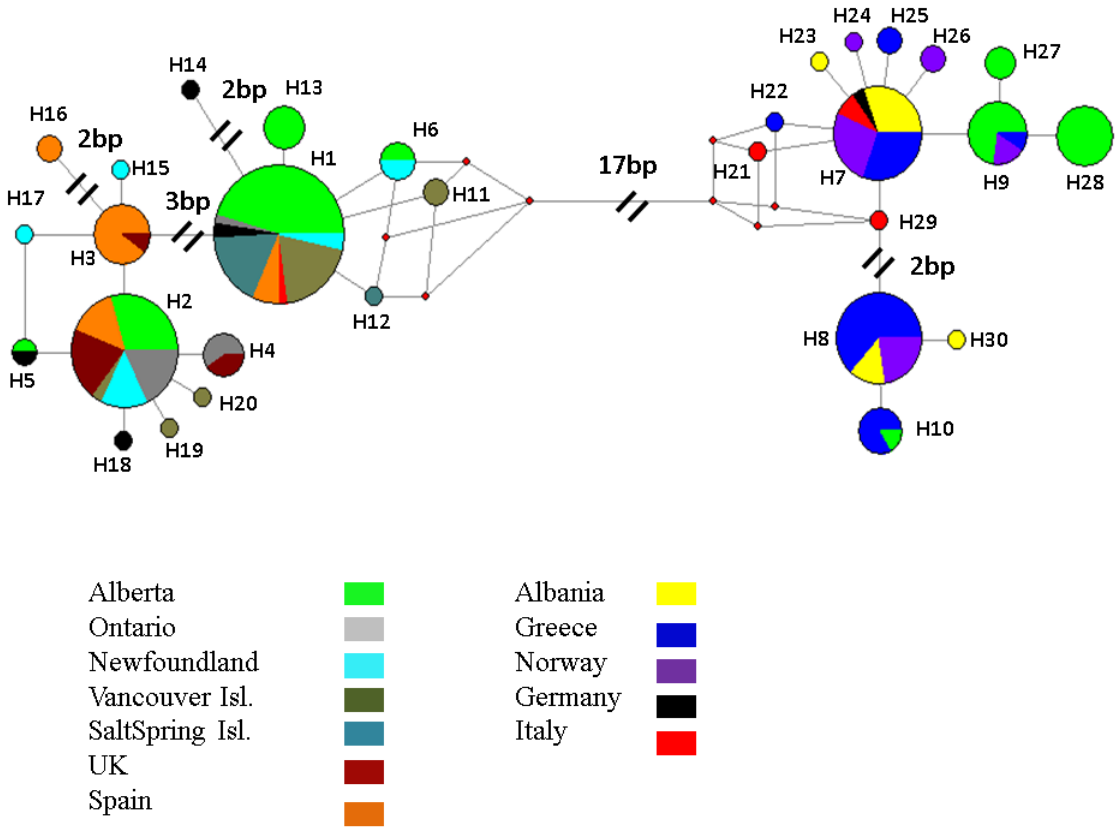


Figure 4.7. Maximum parsimony network of all 30 COX1 haplotypes. Circles represent haplotypes with the size of the circle representing the overall frequency and the colored slices indicating the proportion of that haplotype contributed by different regions. All branches haplotypes differ by 1 bp unless denoted with a dashed line. Small red circles represent inferred or missing haplotypes.

CHAPTER 5

Clonal parasite diversity in an iconic host manipulator: the trematode *Dicrocoelium dendriticum* in populations of zombie ants, *Formica aserva*

5.1 ABSTRACT

The evolution and maintenance of cooperation in animals can arise from a combination of factors. Ecological challenges, in which cooperation increases the fitness of a group, coupled with a high degree of relatedness between individuals of the group (kin selection) can lead to social behaviour. The parasite *Dicrocoelium dendriticum* engages in highly altruistic behaviour whereby some individuals sacrifice themselves to manipulate their ant hosts to facilitate transmission into definitive hosts. To test for a role of kin selection in the evolution of this behaviour, we investigate for the first time the multilocus genotype (MLG) identities of metacercarial populations within ant intermediate hosts. Using a panel of microsatellite markers, we genotyped individual metacercariae within both the body cavity of ants and also the single "brain worm" that drives the behavioural manipulation. Our results showed that 11 of the 18 ants sampled contained multiple batches of metacercariae with identical MLGs. We further determined that 80% of these batches arose from asexual reproduction that would have occurred within snail intermediate hosts. These results indicate that clumped clonal transmission of cercariae to ants occurs and that final hosts ingesting ants are exposed to clonal individuals. Results from 11 successfully genotyped brain worms showed that in most hosts (9 of 11), the majority of individuals found in the abdomen of ants were not clonemates of the manipulator found in the brain. These result, therefore provide no evidence for kin selection in this host-parasite system.

5.2 INTRODUCTION

The evolution of cooperation in animals has been a debated topic among biologists since the 1960s (Sachs et al., 2004). Over the years, several hypotheses have been developed and tested that aim to describe how cooperation within animal groups might evolve and be maintained over time. One of the first and most enduring hypotheses is that cooperation arises under conditions when genes are shared among individuals in a group (Hamilton, 1964; Smith, 1964). Thus, according to the kin selection hypothesis, an individual in a group can benefit from the behaviour of another in a group, in a fitness sense, when it shares alleles by descent from a common ancestor. Through kin selection, alleles can be passed between generations, regardless of which individual actually reproduces. In Hamilton's (1964) original model, therefore, evolved cooperative behaviour is directed at individuals depending on their average degree of relatedness within a group.

The best empirical evidence for the role of kin selection in the evolution of cooperation comes from the eusocial insects (Queller & Strassmann, 1998). The division of labour in social insect societies represents an extreme form of cooperation, whereby many related individuals forgo reproduction for the benefit of relatives (Queller & Strassmann, 1998). This type of sociality has also been described in some species of shrimp (*Synalpheus*) (Duffy, 2003) and even rodents (Jarvis & Bennett, 1993). In each case, individuals are highly related and follow the predicted patterns of kin selection. However, recent work indicates that relatedness is not the sole factor in the evolution of cooperation and that ecological factors such as a challenging environment conditions interact with relatedness to determine observed patterns (Duffy & Macdonald, 2010). The snapping shrimp (*Synalpheus*) that inhabit marine

sponges for example are faced with the challenge of competing for the extremely valuable resources provided by their host sponge (Duffy, 2003). The increased benefits of cooperation in the face of this challenge coupled with the accumulation of related individuals inside a host due to non-dispersing larvae leads to both kin selection and the ecological challenge resulting in cooperation (Duffy & Macdonald, 2010).

The life history of trematode parasites represents a similar scenario whereby ecological factors and kin selection may interact to promote cooperation. The life cycles of all digenean trematodes require multiple species of hosts. In the first intermediate host, usually a gastropod mollusc, an obligate asexual stage occurs that ultimately involves the production of large numbers of clonal cercariae that are infective to the subsequent host (Goater et al., 2013). These highly related larvae face the challenges of finding and successfully infecting subsequent hosts and also competing for host resources with the host and with co-infecting parasites. Currently, the only evidence for cooperation in trematodes comes from recent work showing the development of castes for the division of labour within snail hosts (Hechinger et al., 2011; Leung & Poulin, 2011). For two species of trematode, two functional types of rediae are asexually produced - a non-reproducing defensive type and a smaller type that disperses to the next host. These results show that the evolution of cooperation can occur within highly-related clones of trematode larvae in the face of interspecific competition for host resources.

It is conceivable that cooperation among clones might also provide a solution to the problem of transmission between resting stages and definitive hosts. One well-known

strategy involves the manipulation of host phenotypes, including behaviours, by encysted larval stages. Well known examples include the extra limbs of *Ribeiroia*-infected froglets (Johnson et al., 1999), trematode-induced cataracts in rainbow trout (*Oncorhynchus mykiss*) caused by larvae of *Diplostomum pseudospathaceum*, and the highly conspicuous banded colouration of *Leucochloridium*-infected tentacles of some terrestrial snails (review by Hughes et al., 2012). In each of these cases, the alterations in host phenotype facilitate the transmission of resting larval stages into definitive hosts. Not surprisingly, adaptive manipulations such these are costly, either energetically through the production of neurotransmitters that alter host behaviour (Thomas et al., 2005) or in fitness by the potential foregoing of reproduction by those individuals that manipulate host behaviour (Poulin et al., 2005). Faced with these costs and the significant ecological challenge of transmission one would predict that a high degree of relatedness in manipulative systems. The kin selection hypothesis for cooperation in the resting stage of manipulative parasites has not been tested.

Adults of the lancet liver fluke, *Dicrocoelium dendriticum,* reside along the bile ducts of grazing mammals. Its three host life cycle begins when eggs released in the feces of infected mammals are ingested by terrestrial snails (review by Manga-Gonzalez et al., 2001). A miracidium within the egg hatches and then develops asexually into sporocysts. These undergo repeated bouts of asexual reproduction to produce hundreds of clonal cercariae. Cercariae are then packaged within host-derived mucus and egested into the environment within ‘slime balls’ (Manga-Gonzalez et al., 2001). Slime balls are consumed by a wide range of species formicid ants, either by the individual that encountered it, or by fellow workers within the nest. Following ingestion of a slime ball, a single individual, known as the

"brain worm" penetrates the stomach and travels to the sub-esophageal ganglion. The remaining larva within the slime ball, if present, penetrate the stomach, and then travel to the hemocoel where they are enveloped by a thick, host-derived, cyst wall (Hohorst, 1962). These metacercariae enter a characteristic resting phase, reaching infectivity to mammals approximately 60 days following ingestion (Manga-Gonzalez et al., 2001; Otranto & Traversa, 2002). In contrast, the brain worm is approximately 30% smaller than the larvae within cysts, remains unencysted, and is never infective to mammalian grazers (Hohorst, 1962; Hohorst & Lammler, 1962). The brain worm causes the radically altered behaviour characteristic of infected ants. This alteration involves leaving the nest each day, then climbing up adjacent vegetation, onto which they firmly attach with their mandibles. Infected ants remain attached overnight, then detach the next day if temperatures exceed approximately 18C (Hohorst, 1962; Spindler et al., 1986; Botnevik et al., 2016). Infected ants repeat this cycle of attachment and detachment each day until they are ingested by an appropriate grazing mammal, or they die of other causes. Presumably, this radically altered behaviour caused by the brain worm facilitates the accidental ingestion of infective larvae that reside within the hemocoel.

The altruistic behaviour of the brain worm has perplexed evolutionary biologists and parasitologists for many years (Wilson, 1977; Mehlhorn, 2015). On the one hand, cooperation between the non-infective individual that resides in the brain and the infective ones in the hemocoel can be expected, via kin selection, because the ingestion of a single slime ball is likely to contain clonemates that are genetically identical. On the other hand, over the course of a worker ants' lifetime of 2-3 years (Hölldobler & Wilson, 1990) there is

little to prevent an individual ant from being exposed to cercariae within multiple slime balls that originate from different infected snails. In this case, a previously-infected ant is likely to be exposed to additional clones of cercariae that are genetically distinct. Under this scenario, the role of kin selection in determining cooperation between non-infective brain worm and infective larvae in the hemocoel would be harder to invoke. Resolution of this perplexing problem requires an understanding of the clonal structure of *D. dendriticum* larvae within individual infected ants. More specifically, a full test of the kin selection hypothesis for cooperation among larvae in ants requires an understanding of the genetic relationship between the larva that resides within the brain of an individual ant, and those that reside within the hemocoel.

The recent development of a panel of microsatellite markers for adult *D. dendriticum* (van Paridon et al., 2016) provides the opportunity to genotype individual larvae located with the brain and hemocoel of infected ants. The purpose of this study is to characterize the clonal structure of larval *D. dendriticum* within samples of infected ants collected from two populations in a region where the parasite has been introduced in southern Alberta, Canada (Chapter 4). We also test the kin selection hypothesis by evaluating genotype differences between the uninfective larvae that resides within the brain and infective ones that reside within the hemocoel of individual ants.

5.3 MATERIALS AND METHODS

5.3.1 Sampling

Samples of the ant, *Formica aserva* were collected during the early morning of June 8th 2013 between 6:00 and 9:00 am from two sites located within Cypress Hills Interprovincial Park. General habitat characteristics of CHP are described in van Paridon et al. (2016). Both sites contained *F. aserva* nests, in which workers were known to be infected and to cling to adjacent vegetation (Beck, 2015). The first site, known as Staff Trailer Park (STP; 12N 551975, 5500920) is located within the town of Elkwater, AB. Beef cattle are excluded from the town site, but other definitive hosts (e.g., elk, white-tail deer, and mule deer; Goater & Colwell, 2007) are commonly observed foraging at this site. At STP, *F. aserva* nests are common in the mixed-wood forests adjacent to an area that has been cleared for accommodation for Park personnel. The second site, known as Trans-Canada Trail (TC; 12N 564452, 5501246) is approximately 12 km southeast of the STP site, located in a mixed coniferous/deciduous forest stand on a gently south-facing slope (Beck, 2015). At each site, an ant nest was first identified and ants clinging in closest proximity to it were sampled. Nine ants were collected clinging to either a single plant or from adjacent plants within a 1m^2 area. This sampling protocol was used to increase the probability that the ants sampled were from the same nest. Ants were placed directly into 90% ethanol and stored at -80^{o}C until DNA extraction.

5.3.2 DNA extraction

Individual metacercariae were isolated from ants by opening the abdomen into a Petri dish containing distilled water (dH_2O). Metacercariae intensity was assessed in each ant following methods in van Paridon et al. (2016). If present, up to 21 fully-encysted

metacercariae were selected and fixed in 200 µl PCR tubes containing 50 µl of lysis buffer and Protinease K (10 mg/ml, New England BioLabs, USA). In one sample, TC_B2, greater than 20 metacercariae were fixed and lysed for a preliminary test of the effectiveness of genotyping individual metacercariae. In cases where metacercariae intensity was < 21 (n = 4), all metacercaria were lysed. The heads of each ant were also removed from each body and placed in a 1.5 ml Eppendorf tube containing 50 µl of lysis buffer and Protinease K. There exists a single metacercaria in the heads of infected ants, which we confirmed, as ongoing work in our lab dissections of thousands of ants heads always revealed a single metacercaria. The heads were crushed using a sterile plastic pestle within the tube to increase exposure of the metacercariae inside to the lysis solution. Lysis buffer contained 50 mM KCl, 10 mM Tris (pH 8.3), 2.5 mM MgCl2, 0.045% Nonidet p-40, 0.45% Tween-20, 0.01% gelatin and 129 dH2O in 50- µl volumes. Samples were lysed in 50 µl for 98 min at 60^{o}C followed by 15 min at 94^{o}C, then stored at -20^{o}C.

5.3.3 PCR Reactions and Genotyping

All metacercariae were genotyped with nine microsatellite loci (DdMs21, DdMs28, DdMs43, DdMs44, DdMs60, DdMs70, DdMs89, DdMs93 DdMs95) described in van Paridon et al. (2016). PCR reactions were carried out in reaction mixtures of 25 µl containing final concentrations of 1X Thermopol reaction buffer (New England Biolabs), 2mM $MgSO_4$, 100 uM dNTPs, 0.1 uM 6-FAM labelled forward primer and reverse primers, 1.25U Taq DNA polymerase at 5000 U/ml (New England Biolabs). Thermocycling conditions were 94^{o}C for 3 minutes followed by 35 cycles of 94^{o}C for 30 seconds, 54^{o}C for

30 seconds and 72°C for 1 minute with a final extension of 72°C for 10 minutes. Products were then analyzed on an Applied Biosystems 3500xl sequencer and sized using Genescan 500 (-250) LIZ size standard (Applied Biosystems) and scored manually using GeneMapper 4.1 (Applied Biosystems).

5.3.4 Microsatellite Validation

After genotyping, individual metacercariae that failed to amplify at a locus were removed to allow for appropriate tests for clonality (Arnaud-Haond & Belkhir, 2007). Gene diversity of each locus (H_S) as well as estimates and tests of F_{IS} and F_{ST} were conducted using FSTAT (Goudet, 1995). Significance of average within-host F_{IS} per locus and F_{ST} among hosts were determined using 10,000 randomizations of alleles among individual worms within hosts and of genotypes among hosts, respectively. Genotypic equilibrium (GD) between pairs of loci was tested in GENEPOP 4.2 (Markov chain parameters: 1000 dememorizations; 100 batches; 1000 iterations) (Rousset, 2008).

5.3.5 Clonality and diversity of metacercaria in ants

GENCLONE 2.0 was used to identify individuals with identical multi locus genotypes (MLG) and test whether they arose from asexual reproduction (Arnaud-Haond & Belkhir, 2007). The clonemate testing is done by calculating P_{SEX}, which is the probability of observing n copies of a MLG in a sample size of N given sexual reproduction (significance determined as $P_{SEX} < 0.05$). Asexual reproduction is assumed if a MLG has a significant P_{SEX} value at $n = 2$ (Gregorius, 2005). To test for the significance of clonemates we must first note that structure among subpopulations can bias estimates of P_{sex} towards significance and also

that the presence of clones can drive structuring (Criscione et al., 2011; van Paridon et al., 2016). In order to test for structuring, we used the approach of Criscione et al. (2011), which was previously demonstrated, with these same microsatellites using adult *D. dendriticum* (van Paridon et al., 2016). In short, two datasets were created, one including all individuals and a second, in which identical MLGs were reduced to one copy each. Both datasets were tested for F_{ST} and GD between loci to determine if the presence of identical MLGs is associated with an increase in genetic structure and GD.

5.3.6 Identity of head MLG vs. body MLG

The MLG identity of metacercariae located in the head of an individual ant was identified and compared with those found in the corresponding body. We quantified the number of times an identified head MLG appeared in the body and compared with all other MLGs discovered. We also assessed whether MLGs found in the heads of individual ants were also found in other hosts, either in the head or body.

5.4 RESULTS

5.4.1 Microsatellite validation

The total number of metacercaria found in the abdomens of the 18 ants collected was 775 with the numbers of metacercaria per ant ranging from 17 to 176 (Table 5.1). A total of 393 were lysed and run through the PCR and genotyping procedure (Table 5.1). Of the 393 attempted, 261 were successfully genotyped at all nine loci (Table 5.1). Genotyping the 18

metacercariae from the heads of ants using all nine loci was unsuccessful because four loci consistently did not amplify (DdMs43, DdMs44, DdMs89 and DdMs93). These four loci were also shown to suffer from null alleles in a previous analysis completed with adult *D. dendriticum* from CHP (van Paridon et al., 2016). Given the consistent inability to amplify these loci, especially in the head metacercariae, and the previous demonstration of null alleles in the sample of adult worms, we proceeded by using only the five remaining loci (DdMs21, DdMs28, DdMs60, DdMs70 and DdMs95). This allows us to compare the MLG identity of metacercariae in both the head and body as well as investigate genetic structure and compare with results from the previous work on adults by van Paridon et al. (2016).

In total 272 metacercaria, were successfully genotyped at all five loci, including 11 of the 18 sampled from the heads of the ants (Table 5.1). Average F_{IS} per locus ranged from -0.304 to -0.548 with no values significantly differing from zero. The numbers of alleles sampled per locus ranged from 9-13 and H_S per locus ranged from 0.485 to 0.640. Global F_{ST} was 0.272 and significant ($p = 0.0001$). Analyzing the dataset with only these five loci showed all loci were in HWE, no deviations from F_{IS} per locus. Significant structure did exist between hosts as demonstrated by the significant F_{ST} value.

We then assessed the effect of clonality on HWE and genetic structure as the presence of clones can influence these measures (Prugnolle et al., 2005; Criscione et al., 2011). We created two datasets; i.e., one containing all copies of MLGs determined using five loci ($n = 272$) and, one in which identical copies of MLGs determined using five loci were reduced to one copy each ($n = 66$). We then calculated F_{IS}, the number of alleles sampled and H_S per

locus. F_{IS} per locus ranged from -0.118 to 0.069 again with no values differing significantly from zero (Table 5.2). The number of alleles sampled ranged from 9-13 and H_S ranged from 0.570 to 0.807 (Table 5.2). When clones were included, average F_{IS} was 0 (p = 0.9999) and there was significant structure between hosts F_{ST} 0.272 (p = 0.0001). When identical genotypes were reduced to one copy each, F_{IS} was 0.003 (p = 0.5226) and structure remained between hosts, although the value of F_{ST} was reduced, F_{ST} 0.043 (p = 0.0011). We next looked at the effects of identical MLGs on GD by testing for linkage disequilibrium (LD) in both datasets. When all MLGS were included, all 10 comparisons were significant for LD. This dropped to 0 of 10 significant comparisons when identical copies were reduced to one. The absence of LD when clones were reduced indicates that clones were driving up the excess LD when identical copies were included (de Meeus et al., 2007). We also concluded that there was no cryptic population structure as admixture would cause LD. However, significant structure between hosts still remained.

With the reduced dataset with clones removed (n = 66), we split the sample of hosts into the two sites and tested for structure between ant hosts within each site. There was a total of 34 metacercariae sampled at the STP site and 32 at the TC site. F_{IS} in STP was 0 (p = 0.6138) and F_{ST} was 0.06 (p = 0.0083); F_{IS} in TC was 0.003 (p = 0.5255) and F_{ST} was 0.043 (p = 0.0015). The inclusion of clones appeared to inflate the value of F_{ST}, but significant structure remained between and within sites. We conclude that the presence of clones is not solely responsible for the presence of structure and this structure can influence the calculations of P_{SEX}, biasing them towards significance.

5.4.2 Clonality and diversity of metacercaria in ants

To test for $P_{SEX,}$ the unique and repeated MLGs were first identified. A total of 59 unique MLGs were identified among the 27 metacercariae that were successfully genotyped. Of the 59 MLGs, 36 were shared by at least 2 or more metacercaria and 23 were observed only once (Table 5.3). The numbers of MLGs per ant ranged from 1 to 9 (Table 5.3). Three pairs of ant hosts shared MLGs: three of four MLGs found in host STP_C3 (F, G and H) were also found in STP_C8. Three of four MLGs from STP_C5 (P, Q and S) were found in STP_C10 and a single genotype from TC_B2 (EE) was identified in TC_B4 (Table 5.3).

Due to the presence of structure between hosts, the allele frequencies of the sub-populations; i.e., individual hosts, should be used to test for P_{SEX} (Criscione et al., 2011). Therefore, we tested for P_{SEX} on three groupings of the data to determine how structuring may bias our estimates. First, we tested repeated MLGs for significance at $n = 2$ on the entire dataset as a whole ($n = 272$). Then, we tested each site separately STP ($n = 134$) and TC ($n = 138$) and finally within each host individually. When treated as a whole, all 34 MLGs present in more than one copy had $P_{SEX} < 0.05$. When each site was tested separately, all 19 MLGs present more than once in Site STP had $P_{SEX} < 0.05$ and all 15 MLGs present in Site TC at more than one copy had $P_{SEX} < 0.05$. When each host was considered individually, 27 of the 34 MLGs found in multi copies within hosts had $P_{SEX} < 0.05$ while 7 of 34 were not significant at $P_{SEX} < 0.05$ (Table 5.4).

5.4.3 Identity of head MLG vs. body MLG

The MLG identities of 11 of 18 metacercariae sampled from ant heads were successfully identified from hosts: STP_C2, STP_C3, STP_C4, STP_C6, STP_C8, STP_C9, STP_C10, TC_B3, TC_B6, TC_B7 and TC_B9. The relationship between the genotype identities of the head metacercariae to those found in the body varied between individual ants (Figure 5.1). The MLG of five head metacercariae were found only once in the entire dataset (STP_C3, STP_C6, STP_C9, TC_B6 and TC_B9). Two head metacercariae were part of MLGs shared between two hosts. The MLG of head metacercariae from STP_C8 was G and found in both STP_C3 and STP_C8 while the MLG of the head metacercaria from STP_C10 was Q and found in both STP_C5 and STP_C10 (Table 5.3). In only three hosts was the MLG found in the head also the most frequently found MLG in the body (Figure 5.1).

5.5 DISCUSSION

In this study we provide data showing the genetic diversity of metacercariae within the 2nd intermediate ant host of *D. dendriticum*. This is the first study to investigate clonal transmission or co-sibling transmission in a completely terrestrial system. Our data reveal that multiple groups of clones infect a single ant and therefore these individuals are not always clones of the 'brain worm' that controls altered behaviour. Our analyses of MLG diversity of larval *D. dendriticum* in ants showed two important findings.

First, in 15 of the 18 ants sampled, more than one MLG genotype was found within the metacercariae population. In three individual ants, only a single MLG was found. In these three hosts, the percentage of the infrapopulation sampled was 71%, 53% and 31%. The range of MLGs found within 18 ants was 1-9. These data indicate that although clonal

diversity in this completely terrestrial system was lower than for metacercariae populations in aquatic hosts (Rauch et al., 2005; Keeney et al., 2007a, 2007b; Karvonen et al., 2012) single clone infections appear to be rare. Second, the MLG of the 'brain worm' was found to be present among >50% of the metacercariae sampled in the body cavity in only two individuals. In one of these, 20 of 22 metacercariae in the body cavity of ant STP_C2 shared the same MLG with that found in the head. In the second, 20 of the 21 metacercariae in the body cavity of ant TC_B3 shared the same MLG as the individual in the head. In a third ant, TC_B7, the most abundant MLG in the body cavity was the same as the individual in the head, but these seven metacercaria only represented 35% of the total metacercariae genotyped. Our results show the dominance of multiple clone genotypes in infected ants and incongruence in genotype structure between the head and body cavity metacercariae.

The presence of multiple clonal lineages in ants can be explained by repeated exposure of ants to slime balls, each containing different sets of clones. The rates of exposure of ants to slime balls will depend on factors such as snail density and the prevalence of *D. dendriticum* infection within this snail population, slime ball shedding rates, slime ball size and density, numbers of cercariae/slime ball, ant feeding rates, and ant lifespan. Unfortunately, estimates of these key rate parameters are unknown. In boreal regions in Canada, *F. aserva* workers are thought to live for approximately two years (Hölldobler & Wilson, 1990). The density of infected snails at the sites where we collected the infected ants is as high as 20%, depending on the time of year (Dempsey, Goater & Burg, unpublished data). It would thus seem likely that ants would be exposed and infected multiple times over their lifespan. However, the genetic diversity of cercariae in snails has not yet been

investigated in this system. Previous studies on aquatic trematodes have shown that snails can harbor multiple genotypes (Rauch et al., 2005; Keeney et al., 2007a; Karvonen et al., 2012). The sampling of MLGs among *D. dendriticum* cercariae in snails is a critical next step.

The presence of high numbers of MLGs in second intermediate aquatic hosts is considered to be a product of the ability of free living cercariae to disperse in water and to be advantageous as mixing reduces the chances of inbreeding when metacercariae are transmitted to the next host (Criscione & Blouin, 2006). Admixture of this type is not possible in the terrestrial environment and, in the case of *D. dendriticum;* cercariae from a single snail are packaged together into slime balls. Not surprisingly then, the extent of clonal mixing found in the terrestrial 2nd intermediate hosts of *D. dendriticum* was much less than those reported in aquatic systems. The range of MLGs found within ants was one to nine. In contrast there was an average of 22 MLGs of *Diplostomum pseudospathaceum* metacercariae in freshwater stickleback intermediate hosts (Rauch et al., 2005). In a marine system involving the trematode *Maritrema novaezealandensis*, the range of MLGs found within crab intermediate hosts was 16 to 25 (Keeney et al., 2007a). The marine amphipod *Paracalliope novizealandiae*, which also acts as second intermediate to *Maritrema novaezealandensis* also showed mixing as only 11% of amphipods contained identical trematode clones (Keeney et al., 2007b). These studies have all confirmed the “aquatic mixing” theory proposed by Criscione & Blouin (2006) and our data show that this same level of mixing does not occur in terrestrial systems.

Regardless of the processes leading to multiple clone infections, the common occurrence of mixed clones in ants has important implications for the role of kin selection in determining and maintaining the altruistic behaviour of the brain worm. Each of the ants whose metacercariae we genotyped were collected while they were attached firmly to plants. Thus, the radical behavioural manipulation occurred in each ant, despite the wide variation in genotype diversity in metacercaria subpopulations and despite the incongruence in genotype between brain and body metacercariae. All of the ants were being manipulated at time of collection, even though in only one of 11 ants were the sole metacercariae found in the body clones of the metacercaria in the head, and in only two of 11 was the brain MLG represented by > 50 % of metacercaria sampled. These results indicate that all metacercariae in the haemocoel need not be clones of the brain worm in order for manipulation to occur.

Despite this mixing of clonal lineages in ants, the evolution and maintenance of altruistic behaviour involving *D. dendriticum* metacercercariae may still be possible. Wilson (1977) used mathematical models to predict the probability of selection for altruistic individuals in this system. His model showed that the altruistic behaviour could evolve in scenarios where the parasites within hosts are derived from as many as five different parents. Methods assessing the sibship structure within ants (Jones & Wang, 2010) could show that metacercariae co-infecting ants may in fact be siblings albeit not clones. The evolution of eusociality in shrimp also did not require clonal individuals, but merely related individuals (Duffy & Macdonald, 2010). Co-sibling transmission has been assessed in order to inform transmission patterns of ticks and schistosomes (Guzinski et al., 2009; Lu et al., 2010), but not, to our knowledge, in a system involving parasite altered behaviour. A test of sibship

structure in *D. dendriticum* metacercariae in all three hosts sampled from these sites would provide data to test this hypothesis.

Before a full test of the kin selection hypothesis can be completed, there are several assumptions that should to be tested. First, it is generally assumed that only a single brain worm enters the sub-esophageal ganglion and that it is typically the first one to penetrate the host gut (Hohorst, 1962). Others have reported rare cases when 2-3 metacercariae have been found to reside within various regions of the brain of a single ant (Manga-Gonzalez et al., 2001). Genotyping those additional metacercariae would demonstrate whether they all came from a single exposure or if individuals from subsequent exposures also attempted to reach the brain. Furthermore, we and others assume that the individual in the sub-esophageal ganglia is permanent and long-lived. Although it seems unlikely that a large metacercaria could re-locate to other sites within the host, where it could then possibly encyst, this intriguing idea should be tested.

The research on genetic structure of parasites in intermediate hosts has primarily been concerned with mixing of clones within intermediate hosts as mechanism to avoid inbreeding within definitive hosts (Criscione & Blouin, 2006). The life cycle of *D. dendriticum*, however, appears ideally suited to promote clonal or co-sibling transmission of individuals from host to host. The terrestrial environment is not suitable for free-living dispersal stages and the manipulative effort required to reach final hosts is dependent on altruistic behaviour. The extent of the level of relatedness required for this altruism though is not fully understood. A new avenue of research has been opened here that can shed light onto the

effects of genetic structure on the extent of manipulative behaviour in intermediate hosts. The altered behaviour of ants has been shown to be primarily affected by temperature, with most ants only clinging below 20^{o}C (Badie et al., 1973; Spindler et al., 1986; Botnevik et al., 2016). However, there is variation in this pattern. In studies involving marked ants at these same sites (Goater, van Paridon & Colwell, unpublished observation), we have observed that some infected ants remain in their nests at low temperatures, and that others do not detach from vegetation at temperatures > 20°C. It is intriguing to consider that variation in manipulative effort in individual ants might be associated with variation in either genetic diversity in the overall metacercariae population, or in the magnitude of incongruence between the genotypes of head vs body cavity larvae. A more detail sampling regime and further analysis of relatedness between larval stages of *D. dendriticum* is needed to shed light on to this bizarre phenomenon.

5.6 Litereture Cited

Arnaud-Haond, S., & Belkhir, K. (2007). genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes, 7,* 15-17.

Badie, A., Vincent, M., Morel-Vareille, C., & Rondelaud, D. (1973). Cycle de *Dicrocoelium dendriticum* (Rudolphi, 1819) en Limousin: Ethologie des fourmis parasitees par les metacercaires. *Comptes rendus des Séances de la Société de Biologie, 167,* 725-727.

Beck, M. A. (2015). Ecological epidemiology of an invasive host generalist parasite, *Dicrocoelium dendriticum*, in Cypress Hills Interprovincial Park, Alberta. (Ph.D.) University of Lethbridge, Lethbridge, Alberta.

Botnevik, C. F., Malagocka, J., Jensen, A. B., & Fredensborg, B. L. (2016). Relative effects of temperature, light, and humidity on clinging behavior of metacercariae-infected ants. *Journal of Parasitology, 102,* 495-500.

Criscione, C. D., & Blouin, M. S. (2006). Minimal selfing, few clones, and no among-host genetic structure in a hermaphroditic parasite with asexual larval propagation. *Evolution, 60,* 553-562.

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. *Molecular Ecology, 20,* 2510-2524.

de Meeus, T., McCoy, K. D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Bousses, S., & Renaud, F. (2007). Population genetics and molecular epidemiology or how to "debusquer la bete". *Infection Genetics and Evolution, 7,* 308-332.

Duffy, J. E. (2003). The ecology and evolution of eusociality in sponge-dwelling shrimp. *Genes, Behavior, and Evolution in Social Insects.* Sapporo, Japan: University of Hokkaido Press.

Duffy, J. E., & Macdonald, K. S. (2010). Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis. *Proceedings of the Royal Society B: Biological Sciences, 277,* 575-584.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: *Dicrocoelium dendriticum* in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93,* 491-494.

Goater, T. M., Goater, C. P., & Esch, G. W. (2013). *Parasitism: the diversity and ecology of animal parasite.* Cambridge, UK: Cambridge University Press.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity, 86*, 485-486.

Gregorius, H. R. (2005). Testing for clonal propagation. *Heredity, 94*, 173-179.

Guzinski, J., Bull, C. M., Donnellan, S. C., & Gardner, M. G. (2009). Molecular genetic data provide support for a model of transmission dynamics in an Australian reptile tick, *Bothriocroton hydrosauri*. *Molecular ecology, 18*, 227-234.

Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology, 7*, l-52.

Hechinger, R. F., Wood, A. C., & Kuris, A. M. (2011). Social organization in a flatworm: trematode parasites form soldier and reproductive castes. *Proceedings of the Royal Society of London B: Biological Sciences, 278*, 656-665.

Hohorst, W. (1962). The role of ants in the developmental process of the lancet-shaped fluke (*Dicrocoelium dendriticum*). *Zeitschrift fur Parasitenkunde, 22*, 105-106.

Hohorst, W., & Lammler, G. (1962). Experimental dicrocoeliasis studies. *Zeitschrift fur Tropenmed und Parasitologie, 13*, 377-396.

Hölldobler, B., & Wilson, E. O. (1990). *The ants*: Cambridge, USA: Harvard University Press.

Hughes, D. P., Brodeur, J., & Thomas, F. (2012). *Host manipulation by parasites*: Oxford, UK: Oxford University Press.

Jarvis, J. U. M., & Bennett, N. C. (1993). Eusociality has evolved independently in two genera of bathyergid mole-rats, but occurs in no other subterranean mammal. *Behavioral Ecology and Sociobiology, 33*, 253-260.

Johnson, P. T., Lunde, K. B., Ritchie, E. G., & Launer, A. E. (1999). The effect of trematode infection on amphibian limb development and survivorship. *Science, 284*, 802-804.

Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular ecology resources, 10*, 551-555.

Karvonen, A., Rellstab, C., Louhi, K. R., & Jokela, J. (2012). Synchronous attack is advantageous: mixed genotype infections lead to higher infection success in

trematode parasites. *Proceedings of the Royal Society B-Biological Sciences, 279*, 171-176.

Keeney, D. B., Waters, J. M., & Poulin, R. (2007a). Clonal diversity of the marine trematode Maritrema novaezealandensis within intermediate hosts: the molecular ecology of parasite life cycles. *Molecular Ecology, 16*, 431-439.

Keeney, D. B., Waters, J. M., & Poulin, R. (2007b). Diversity of trematode genetic clones within amphipods and the timing of same-clone infections. *International Journal for Parasitology, 37*, 351-357.

Leung, T. L., & Poulin, R. (2011). Small worms, big appetites: ratios of different functional morphs in relation to interspecific competition in trematode parasites. *International Journal for Parasitology, 41*, 1063-1068.

Lu, D.-B., Rudge, J. W., Wang, T.-P., Donnelly, C. A., Fang, G.-R., & Webster, J. P. (2010). Transmission of *Schistosoma japonicum* in marshland and hilly regions of China: parasite population genetic and sibship structure. *Plos Neglected Tropical Diseases, 4*, e781.

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123*, 91-114.

Mehlhorn, H. (2015). The Brain Worm Story. In H. Mehlhorn (Ed.), *Host Manipulations by Parasites and Viruses*: Heidelberg, Germany: Springer International Publishing.

Otranto, D., & Traversa, D. (2002). A review of dicrocoeliosis of ruminants including recent advances in the diagnosis and treatment. *Veterinary Parasitology, 107*, 317-335.

Poulin, R., Fredensborg, B. L., Hansen, E., & Leung, T. L. F. (2005). The true cost of host manipulation by parasites. *Behavioural Processes, 68*, 241-244.

Prugnolle, F., Liu, H., de Meeus, T., & Balloux, F. (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology, 35*, 255-263.

Queller, D. C., & Strassmann, J. E. (1998). Kin selection and social insects social insects provide the most surprising predictions and satisfying tests of kin selection. *Bioscience, 48*, 165-175.

Rauch, G., Kalbe, M., & Reusch, T. B. (2005). How a complex life cycle can improve a parasite's sex life. *Journal of Evolutionary Biology, 18*, 1069-1075.

Rousset, F. (2008). GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources, 8*, 103-106.

Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. (2004). The evolution of cooperation. *The Quarterly Review of Biology, 79*, 135-160.

Smith, J. M. (1964). Group selection and kin selection. *Nature, 201*, 1145-1147.

Spindler, E. M., Zahler, M., & Loosfrank, B. (1986). Behavioural-aspects of ants as 2nd intermediate hosts of *Dicrocoelium dendriticum*. *Zeitschrift Fur Parasitenkunde-Parasitology Research, 72*, 689-692.

Thomas, F., Adamo, S., & Moore, J. (2005). Parasitic manipulation: where are we and where should we go? *Behavioural Processes, 68*, 185-199.

van Paridon, B. J., Goater, C. P., Gilleard, J. S., & Criscione, C. D. (2016). Characterization of nine microsatellite loci for *Dicrocoelium dendriticum*, an emerging liver fluke of ungulates in North America, and their use to detect clonemates and random mating. *Molecular and Biochemical Parasitology, 207*, 19-22.

Wilson, D. S. (1977). How nepotistic is the brain worm? *Behavioral Ecology and Sociobiology, 2*, 421-425.

Table 5.1 Numbers of *D. dendriticum* metacercariae sampled, lysed and the numbers of successful genotypes obtained from 18 ants at two sites (STP and TC) within Cypress Hills Interprovincial Park.

Host ID	Total Metacercariae	Metacercariae Lysed	Complete Genotypes 9 Loci	Complete Genotypes 5 Loci
STP_C1	17	17	11	11
STP_C2	47	21	20	21
STP_C3	25	21	14	15
STP_C4	49	21	21	22
STP_C5	108	21	16	16
STP_C6	40	21	17	18
STP_C8	45	21	19	20
STP_C9	71	21	2	3
STP_C10	28	21	7	8
TC_B1	17	17	12	12
TC_B2	51	51	28	28
TC_B3	64	21	19	20
TC_B4	17	17	4	4
TC_B5	36	21	19	19
TC_B6	44	21	8	9
TC_B7	21	21	19	20
TC_B8	18	18	4	4
TC_B9	176	21	21	22
Total	775	393	261	272

Table 5.2. F_{IS}, genetic diversity (H_s) and numbers of alleles sampled per locus from *D. dendriticum* metacercariae based on five microsatellite loci.

Locus	F_{IS} per Locus (P value)	H_S	Alleles Sampled
DdMs23	0.061 (0.2835)	0.714	10
DdMs28	-0.036 (0.7638)	0.778	10
DdMs60	0.008 (0.5748)	0.715	11
DdMs70	0.069 (0.1371)	0.807	13
DdMs95	-0.118 (0.9507)	0.570	9
Overall	0.003 (0.5226)	0.717	53

Table 5.3. Numbers of *D. dendriticum* metacercariae successfully genotyped at five loci and the numbers and identities of multi locus genotypes (MLG) within each ant host. Symbols indicate hosts, which share genotypes.

Host	Metacercariae	#MLG	Metacercariae per MLG
STP_C1	11	2	**A**(10) **B**(1)
STP_C2	21	2	**C**(20) **D**(1)
STP_C3*	15	4	**E**(1) **F** (1) **G**(3) **H**(10)
STP_C4	22	7	**I**(1) **J**(1) **K**(1) **L**(3) **M**(3) **N**(6) **O**(7)
STP_C5$^{\Omega}$	16	4	**P**(1) **Q**(3) **R**(4) **S**(8)
STP_C6	18	6	**T**(1) **U**(1) **V**(3) **W**(3) **X**(4) **Y**(6)
STP_C8*	20	3	**F**(2) **G**(6) **H**(12)
STP_C9	3	3	**Z**(1) **AA**(1) **BB**(1)
STP_C10$^{\Omega}$	8	3	**P**(4) **Q**(1) **S**(3)
TC_B1	12	1	**CC**(12)
TC_B2$^{\Pi}$	28	3	**DD**(1) **EE**(5) **FF**(22)
TC_B3	20	1	**GG**(20)
TC_B4$^{\Pi}$	4	2	**HH**(1) **EE**(3)
TC_B5	19	1	**II**(19)
TC_B6	9	4	**JJ**(1) **KK** (1) **LL** (1) **MM**(6)
TC_B7	20	8	**NN**(1) **OO**(1) **PP**(1) **QQ**(1) **RR**(2) **SS**(2) **TT**(5) **UU**(7)
TC_B8	4	3	**VV**(1) **WW**(1) **XX**(2)
TC_B9	22	9	**YY**(1) **ZZ**(1) **AB**(1) **AC**(1) **AD**(1) **AE**(3) **AF**(4) **AG**(4) **AH**(6)

Table 5.4. List of *D. dendriticum* metacercarial genotypes present at more than one copy in all 18 ant hosts from two sites (STP and TC) in Cypress Hills Park. Genotypes in red had P_{SEX} values not significant at P > 0.05.

Host	Genotypes present at n >1 copies
STP_C1	A (10)
STP_C2	C (20)
STP_C3	G (3) H (10)
STP_C4	L (3) M (3) N (6) O (7)
STP_C5	Q (3) R (4) S (8)
STP_C6	V (3) W (3) X (4) Y (6)
STP_C8	F (2) G (6) H (12)
STP_C10	P (4) S (3)
TC_B1	CC (12)
TC_B2	EE (5) FF (22)
TC_B3	GG (20)
TC_B4	EE (3)
TC_B5	II (19)
TC_B6	MM (6)
TC_B7	RR (2) SS (2) TT (5) UU (7)
TC_B8	XX (2)
TC_B9	AE (3) AF (4) AG (4) AH (6)

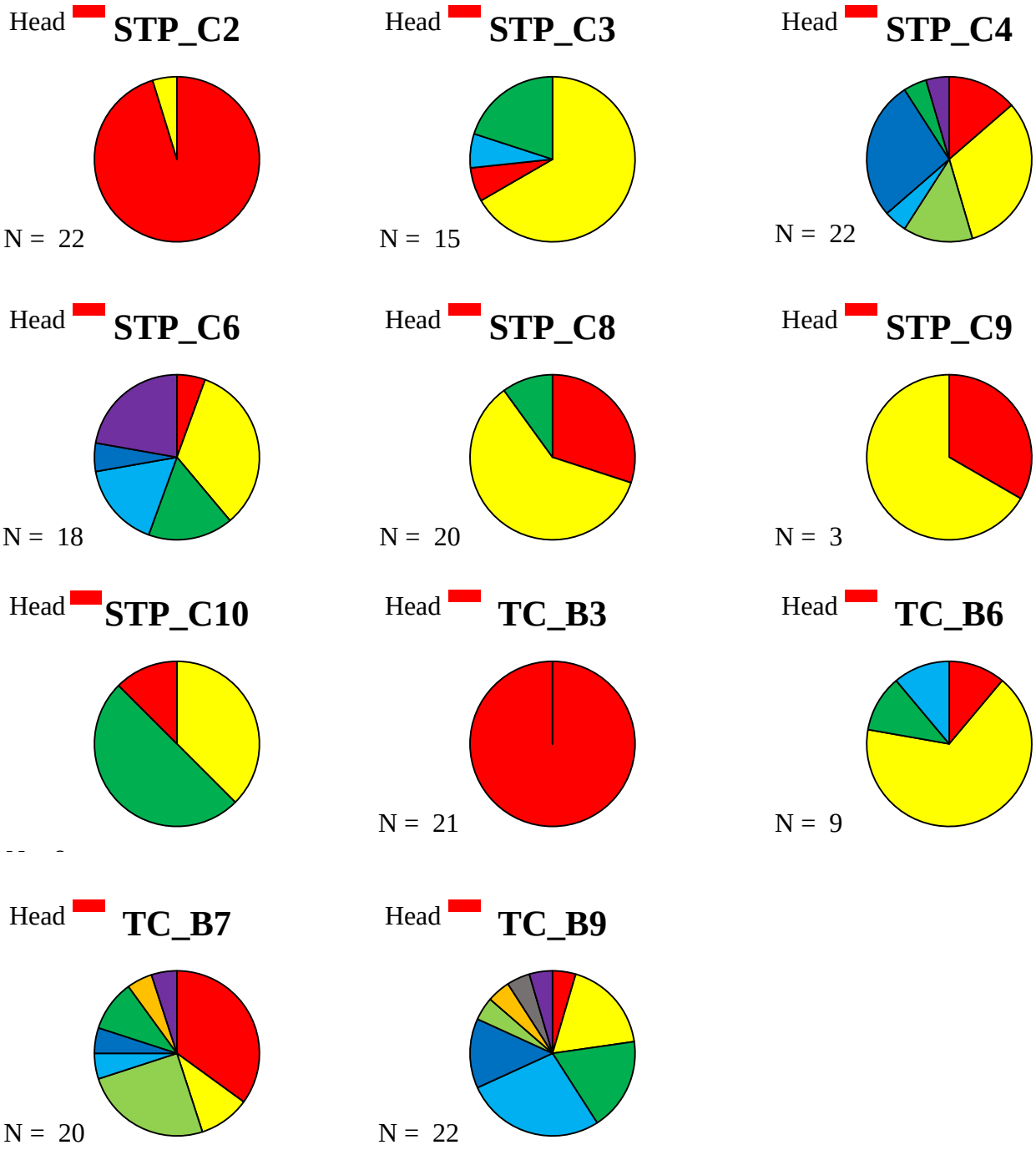


Figure 5.1. Proportion of different *D. dendriticum* metacercarial multilocus genotypes (MLG) within the 11 ant hosts, in which the head metacercariae was successfully genotyped.

CHAPTER 6: GENERAL DISCUSSION

6.1 CHAPTER SYNTHESIS

The increase in EIDs globally has led to enhanced focus on the factors leading to the emergence of pathogens and parasites. While attention is often placed on microparasites, the emergence of macroparasites is of equal concern (Tompkins & Poulin, 2006). The emergence of macroparasites such as helminths involves a combination of ecological and biological interactions (Hoberg & Brooks, 2015). The study of these factors thus draws from multiple fields such as invasion biology, molecular epidemiology and landscape genetics (Tompkins & Poulin, 2006; Biek & Real, 2010; Lymbery & Thompson, 2012). An understanding of how and where parasites have travelled prior to emergence in new environments is necessary and often impeded by the inability to directly track parasites and a lack of information regarding historical animal movement. The ability of parasites to infect new hosts will also be vital to their establishment in new areas and therefore understanding the host range of invasive parasites is critical to describing their emergence and assessing the risk of spread (Hoberg & Brooks, 2015). The emergence of *Dicrocoelium dendriticum* within Cypress Hills Interprovincial Park has provided me an excellent opportunity to develop molecular methods for use on a complex life cycle EID and investigate some of the factors leading to establishment in this region, such as host use in a new range, its pattern of distribution within its intermediate hosts, its invasion history, as well as some unique features of trematode biology.

In Chapter 2, I have used the established method of molecular identification of potential host species (McManus & Bowles, 1996) to confirm the identity of the intermediate hosts used by *D. dendriticum* in CHP. By isolating *D. dendriticum* DNA from within suspected first intermediate host snail species and second intermediate host ant species, I clarified host use in this region of emergence. The DNA barcoding of COX1 mitochondrial gene sequences (Savolainen et al., 2005) to identify the host species filled a fundamental gap in our understanding of the potential for this invader to spread into other areas of North America. As expected, it utilizes local snail and ant hosts to complete its life cycle in the Cypress Hills - *Oreohelix subrudis* and *Formica aserva*, respectively. This result also informs the growing host range of *D. dendriticum* and speaks to its ability to infect a range of related snail and ant species (Manga-Gonzalez et al., 2001). This range extends to multiple families of snail in the Order Stylommatophora and multiple species of ants in the genus *Formica*. The exact traits responsible for this parasite's ability to use such a wide range of hosts are unknown, but the framework of ecological fitting provides an explanation in that these pre-adapted traits allow *D. dendriticum* to track host resources within these groups of related hosts (Araujo et al., 2015). Both species identified here are endemic to high elevation habitats in western North America (Wheeler, 1913; Pilsbry, 1939) demonstrating that suitable hosts exist for this parasite throughout the region. The growing data on host range further informs which potential species are suitable hosts for this parasite.

Understanding host range is a critical piece of knowledge, but to further understand the distribution and transmission of this invader the development of a reliable set of molecular markers is required. In Chapter 3, the protocol for identifying, validating and

applying a panel of microsatellite markers for *D. dendriticum* is developed for the first time. Using a whole genome sequence approach, a panel of nine polymorphic microsatellite loci were identified within the population of *D. dendriticum* in Cypress Hills Interprovincial Park (CHP). To validate the markers for use in population genetics studies, the potential biological and technical reasons results that may show deviations from population genetics models must be identified, as these can bias results and lead to misinterpretations of the data (Gorton et al., 2012). The observation of large deviations from HWE in multiple loci, which could not be explained by population structure or inbreeding, was concluded to be the result of null alleles. This finding validated our screening methods and identified potential biases to be aware of in subsequent chapters. I next was able to identify the role of asexually replicating sporocysts and cercariae on estimates of genetic structure, another crucial piece of information when studying complex life cycle trematodes (Prugnolle et al., 2005). After accounting for the role of asexual reproduction, I showed that adult flukes sampled from two different species of final host, elk and cattle are a panmictic population. This demonstrates that there are no barriers to transmission between different host species. Beck (2015) showed that the majority of contribution of *D. dendriticum* eggs into pasture within CHP is by the population of beef cattle and I now show that the population of parasites is being shared by these two species and likely other suitable hosts in the area such as mule and white-tail deer (Goater & Colwell, 2007). The development of these microsatellite markers now provides a useful tool to study patterns of transmission within CHP, between geographic locations, as well as the invasion history of *D. dendriticum*.

The markers developed in the previous chapter were then put to use on a larger group of *D. dendriticum* flukes from across Canada and Europe in Chapter 4. First, the validity of these markers on the larger dataset were shown, using the methods from Chapter 3, for use on the geographically diverse set of flukes. With this panel of microsatellites and the use of COX1 mitochondrial markers (Králová-Hromadová et al., 2008), I investigated the distribution and potential invasion route of *D. dendriticum* into Canada. Data from measures of genetic diversity using both markers showed that the populations located on the west coast of British Columbia were less diverse than the next closest population, the flukes from CHP in Alberta. This disparity in diversity combined with patterns of genetic structure and shared haplotypes indicated that the British Columbia fluke populations were of single origin, likely eastern Canada, which was itself founded by flukes from Western Europe. The Alberta flukes though appear to be of mixed origin as they had higher levels of diversity and were the only Canadian populations to show genetic similarity in the form of COX1 shared haplotypes, PCA and cluster analysis with both western and eastern European populations. These results are similar to the analysis of the introduction of the liver fluke *Fascioloides magna* to the European continent in that two separate introductions were identified, one in isolation and another that appears to have spread further into the continent (Králová-Hromadová et al., 2011; Bazsalovicsová et al., 2015). The data in this thesis are not consistent with the hypothesis of a single introduction into Canada, followed by subsequent spread. It also shows that in the absence of human-mediated animal movement between the two closest sampled populations (i.e., British Columbia and Alberta) there may be physical barriers, such as the Rocky Mountains, that restrict colonization.

Chapter 5 saw the microsatellite markers used for a study on the life cycle of *D. dendriticum*. This life cycle is unique in that it is terrestrial and altered behaviour is observed in infected ants. For the first time multiple clonal lineages of metacercaria were identified inside ant intermediate hosts. These data have implications regarding the evolution of cooperation via kin selection in this system. I discuss how the unique terrestrial life cycle of *D. dendriticum* may favour the co-transmission of clones and possibly lead to the evolution of cooperation between the individual manipulating behaviour in the heads of ants with those encysted in the haemocoel. Understanding more regarding the mechanisms of evolution and transmission of this parasite will also further aid the understanding of why it is such a successful invader.

As complete analyses, the data presented in these chapters identify some of the factors that have allowed this parasite to arrive and establish in western Canada and specifically Alberta. The first identification of this parasite in CHP in the 1990s (Pybus, 1990) and its subsequent increase in prevalence among hosts, up to 80% in some host species (Goater & Colwell, 2007), showed that *D. dendriticum* had emerged in this region. Until now, though, the identification of the intermediate hosts used by the parasite in CHP and its relation to those used in endemic regions was unknown (Chapter 2). The origins of the fluke and its relation to other populations in western Canada were also not described until now (Chapter 4). I show that the dynamics of *D. dendriticum* invasion are different in the two provinces and the fluke populations are of different origin. Taken together with our data on host range, I hypothesize that local species of intermediate hosts will differ in British Columbia, and other regions, and this has not been a barrier to establishment. This pattern of

broad host range coupled with multiple introductions has likely increased the parasite's ability to emerge in these provinces. Low host specificity increases the available options for incoming parasites (Hoberg & Brooks, 2008) while multiple introductions and human mediated transport can aid transmission and reinforce founding populations (Tompkins & Poulin, 2006; Estoup & Guillemaud, 2010).

The conclusions provided here would not be possible without the development of the microsatellite markers in Chapter 3. The uses of molecular markers to infer patterns of transmission and establishment that are not otherwise directly observable are invaluable to the study of parasites (Criscione et al., 2005). I was also able to investigate the genetic structure of this parasite within the CHP, showing that host species are not a barrier to transmission (Chapter 2) and the results from Chapter 5 are the first to show the potential for genetic structure of a complex life cycle parasite to influence transmission and host manipulation.

6.2 METHODOLOGY/ADVANCES

The use of molecular markers has greatly informed the field of epidemiology and parasitology by providing the ability to distinguish genetic variation between individuals, populations and species (Criscione et al., 2005; Lymbery & Thompson, 2012). Without these tools, many questions concerning parasite transmission and evolution would be difficult if not impossible to answer (Criscione et al., 2005). However, careful consideration of marker choice and the potential technical and biological sources of bias are sometimes ignored

(Bonin et al., 2004; Selkoe & Toonen, 2006). One of the most common types of errors in microsatellite genotyping is that of null alleles (Chapuis & Estoup, 2007). Our use of microsatellites thus required a careful screening for these, which was further complicated by the nature of our study system. The population genetics of trematodes warrants specific attention, as the non-random transmission of clones can influence estimates of genetic structure (Prugnolle et al., 2005). A number of studies have highlighted this and have proposed methods for identifying and reducing bias (Criscione et al., 2011; Yin et al., 2015; Huo et al., 2016). Here, I followed the method proposed by Criscione et al. (2011) that addresses the "catch 22" of testing for the influence of clones when the identification of clones will itself be influenced by their presence. The demonstration of the effectiveness of this method paves the way for future studies on this and other parasites. I also showed in Chapter 3 and Chapter 4 that after consideration of these issues, reliable data and conclusions can still be obtained.

6.3 STUDY LIMITATIONS

6.3.1 Opportunistic sampling

Issues of sample size, sampling accuracy and precision on the quality of parasitological data are not new (Gregory & Blackburn, 1991; Jovani & Tella, 2006). The use of molecular markers is not free of this constraint either (Cornuet et al., 1999; Traub et al., 2005). Population genetic techniques are based on comparisons of allele frequencies, and greater sampling of individuals and populations will increase the amount of information

available for context and comparison. For example, the accurate assignment of individuals to populations of origin requires that the true population of origin was sampled (Cornuet et al., 1999). However, when studying parasites, researchers must often rely on opportunistic sampling if and when infected hosts become available. The availability of hosts within CHP allowed for accurate sampling of this population, but we could not guarantee this for locations outside of CHP. The identification of more foci of infection between the east and west coast of Canada would clarify the invasion route proposed here. Increased sampling of populations of adult worms in Europe would also aid in describing the global movement patterns of *D. dendriticum*. I often relied on pooled samples from some locations, meaning we could not assess between-host differences at all locations. Increased surveillance of potential hosts within Canada would therefore, greatly benefit future research.

6.3.2 Marker resolution

The selection of appropriate markers for a select purpose is of utmost importance in population genetic studies, as each have separate modes of inheritance and provide varying degrees of resolution (Lymbery & Thompson, 2012). The microsatellites developed here are suitable for our study purpose (Selkoe & Toonen, 2006), and our use of a second mitochondrial marker provides more resolution, and corroborates results. However, advances in whole genome sequencing have led to a wealth of new genetic information that has become available for parasite research (Wasmuth, 2014). Specifically, the discovery of thousands of single nucleotide polymorphisms (SNP) by way of restriction site associated

DNA sequencing (Baird et al., 2008) can provide a panel of thousands of SNP markers for in-depth population genetic studies. These SNP markers have been used to recreate phylogenies and identify regions of the genome under selection (Hohenlohe et al., 2010), which one could apply to invasive *D. dendriticum* populations and see which areas may have undergone selection during an invasion. Furthermore, these markers can generate large amounts of genomic data in organisms for which there is little to no prior genetic information, making them useful for linkage mapping and identification of quantitative trait loci (QTL) (Baxter et al., 2011). Identifying potential QTLs associated with either invasion or host manipulation would be useful in understanding exactly which traits and genes may be involved in both of these processes.

6.4 PROPSPECTS OF FUTURE RESEARCH

6.4.1 Increased investigation of D. dendriticum invasion from Europe to Canada

The identification of adult *Dicrocoelium dendriticum* flukes in final hosts across Canada (Chapter 4) and Europe, where it is known to be a common parasite of ruminants (Otranto & Traversa, 2002), coupled with the development of the microsatellites (Chapter 2), sets the stage for a more in-depth analysis of this parasite between the two continents. With increased surveillance and sampling of potential intermediate hosts in Canada, it is possible to add to the growing list of suitable hosts for *D. dendriticum*. The phylogenetic relationships of these hosts could, further be explored to test ideas concerning the relative roles of ecological fitting and coevolution in related hosts (Araujo et al., 2015). Differences in environmental characteristics at the sites, in which *D. dendriticum* is found to be emerging would further provide information about the ability of this parasite to invade and to emerge at other sites. A more in-depth sampling strategy would provide greater resolution for future molecular epidemiological studies (Constantine, 2003). Are there potential populations from eastern Canada that are also of mixed origin similar to Alberta, or are they solely of western European origin? These types of data may support additional evidence for multiple introductions to the east coast of Canada that have subsequently coalesced in Alberta. This would, in turn provide useful information for the monitoring of wild and domestic host movements within Canada, and not just between the two continents. The application of the microsatellite and COX1 markers to a broader set of samples representing more populations

is the easiest next step for population genetics research of this system without, the need for, development of new molecular markers.

6.4.2 Genetic structure of larval stages within intermediate hosts

The data stemming from Chapter 5 on the MLG identities of metacercariae in ants leads to a number of interesting questions. The next step with these microsatellite markers will be to apply them to cercariae within slime balls. This would fill an important gap in our understanding, especially if it was matched with field studies on rates of encounter and ingestion of slime balls by foraging workers. Similar studies identified the average number of *Schistosoma mansoni* genotypes within host snails compared with the number in rat definitive host to estimate how many snails contributed to infection (Theron et al., 2004). This general approach could be extended to the snails and ants within patches of infection within CHP. Beck (2015) has created a GIS risk map that predicts the presence of infected ants within CHP and found a patchy distribution of infection. Understanding how many snails contribute to infection in ants within a patch, and then extending this to the mobile final hosts would provide an in depth map of transmission showing how host movement may be responsible for the mixing of clonal genotypes within the final hosts. My data showed that ants contained multiple clonal individuals and there was significant structure between ant hosts, yet this structure was not seen in the population of adult worms (van Paridon et al., 2016). These results indicate that mixing of clones occurs in definitive hosts.

Another fascinating avenue for future research would be to identify sibling relationships within metacercariae populations in individual ants. The methods for sibship

assignment can now be determined with analyses of microsatellite markers (Jones & Wang, 2010). The presence of clonal individuals is known to increase structure within hosts, but siblings within a host can do the same (Criscione et al., 2005), and it is likely that this may be occurring within ants. Data from Chapter 5 showed that individuals within an ant were not all clonemates and that the individuals in the body were not necessarily clonemates of the brain worm. The kin selection hypothesis of altruism by the brain worm, though, may not require clonality (Wilson, 1977), and the identification of sibship among these individuals may reveal a high level of relatedness between non-clonemates co-infecting ants. Also, greater sampling of metacercariae within ants, and more sites would provide a more accurate picture of the level of relatedness of parasites being co-transmitted within individuals and within sites.

6.5 CONCLUSIONS

As human-mediated climate change, encroachment, and animal translocations increase, it is predicted that the incidence of emerging parasite will also increase (Daszak et al., 2000; Cleaveland et al., 2001; Tompkins & Poulin, 2006). A parasite's ability to infect multiple hosts will also breakdown further barriers to its spread (Malcicka et al., 2015). The data presented here show how *D. dendriticum*, as a notorious host generalist within each life-cycle stage, is one of these parasites and begins to describe the factors that led to its establishment. It can infect local hosts in western Canada and has been introduced multiple times, likely by human sources, which have led to its emergence in western Canada.

Increased surveillance for this parasite in adjacent regions of North America will likely reveal further sites where this parasite has become established.

Dicrocoelium dendriticum also represents one of the classic examples of host manipulation, and is one of the few trematodes with an entirely terrestrial life cycle. These two factors make it a fascinating example of how life history impacts genetic structure. Investigations into the interplay of these two factors will highlight how the evolution of altered behaviours in hosts is able to occur. An increased focus on the population genetics of this parasite system will lead to a better understanding of its emergence around the globe and the evolution of parasite life cycles.

6.6 Literature Cited

Araujo, S. B., Braga, M. P., Brooks, D. R., Agosta, S. J., Hoberg, E. P., von Hartenthal, F. W., & Boeger, W. A. (2015). Understanding host-switching by ecological fitting. *Plos One, 10*, e0139225.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. & Johnson, E.A., (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *Plos One*, 3, e3376.

Baxter, S. W., Davey, J. W., Johnston, J. S., Shelton, A. M., Heckel, D. G., Jiggins, C. D., & Blaxter, M. L. (2011). Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *Plos One, 6*, e19315.

Bazsalovicsová, E., Králová-Hromadová, I., Štefka, J., Minárik, G., Bokorová, S., & Pybus, M. (2015). Genetic interrelationships of North American populations of giant liver fluke *Fascioloides magna*. *Parasites & Vectors, 8*, 288.

Beck, M. A. (2015). Ecological epidemiology of an invasive host generalist parasite, *Dicrocoelium dendriticum*, in Cypress Hills Interprovincial Park, Alberta. (Ph.D.), University of Lethbridge, Lethbridge, Alberta.

Biek, R., & Real, L. A. (2010). The landscape genetics of infectious disease emergence and spread. *Molecular Ecology, 19*, 3515-3531.

Bonin, A., Bellemain, E., Eidesen, P. B., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology, 13*, 3261-3273.

Chapuis, M.-P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*, 621-631.

Cleaveland, S., Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*, 991-999.

Constantine, C. C. (2003). Importance and pitfalls of molecular analysis to parasite epidemiology. *Trends in Parasitology, 19*, 346-348.

Cornuet, J.-M., Piry, S., Luikart, G., Estoup, A., & Solignac, M. (1999). New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics, 153*, 1989-2000.

Criscione, C. D., Poulin, R., & Blouin, M. S. (2005). Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology, 14*, 2247-2257.

Criscione, C. D., Vilas, R., Paniagua, E., & Blouin, M. S. (2011). More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. *Molecular Ecology, 20*, 2510-2524.

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science, 287*, 443-449.

Estoup, A., & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology, 19*, 4113-4130.

Goater, C. P., & Colwell, D. D. (2007). Epidemiological characteristics of an invading parasite: *Dicrocoelium dendriticum* in sympatric wapiti and beef cattle in southern Alberta, Canada. *Journal of Parasitology, 93*, 491-494.

Gorton, M. J., Kasl, E. L., Detwiler, J. T., & Criscione, C. D. (2012). Testing local-scale panmixia provides insights into the cryptic ecology, evolution, and epidemiology of metazoan animal parasites. *Parasitology, 139*, 981-997.

Gregory, R. D., & Blackburn, T. M. (1991). Parasite prevalence and host sample size. *Parasitology Today, 7*, 316-318.

Hoberg, E. P., & Brooks, D. R. (2008). A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. *Journal of Biogeography, 35*, 1533-1550.

Hoberg, E. P., & Brooks, D. R. (2015). Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. *Philosophical Transactions of the Royal Society B, 370*, 20130553.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *Plos Genetics, 6*. e1000862

Huo, G.-N., Liu, L., He, H.-B., & Attwood, S. W. (2016). An investigation into the potential effects of infrapopulation structure and other sources of sampling error, on population genetic studies of the transmission of *Schistosoma japonicum* (Trematoda: Digenea). *Parasites & Vectors, 9*, 165.

Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources, 10,* 551-555.

Jovani, R., & Tella, J. L. (2006). Parasite prevalence and sample size: misconceptions and solutions. *Trends in Parasitology, 22,* 214-218.

Králová-Hromadová, I., Bazsalovicsová, E., Štefka, J., Špakulová, M., Vávrová, S., & Szemes, T. (2011). Multiple origins of European populations of the giant liver fluke *Fascioloides magna* (Trematoda: Fasciolidae), a liver parasite of ruminants. *International Journal for Parasitology, 41,* 373-383.

Králová-Hromadová, I., Špakulová, M., Horáčková, E., Turčeková, L., Novobilský, A., & Beck, R. (2008). Sequence analysis of ribosomal and mitochondrial genes of the giant liver fluke *Fascioloides magna* (Trematoda: Fasciolidae): intraspecific variation and differentiation from *Fasciola hepatica*. *Journal of Parasitology, 94,* 58-67.

Lymbery, A. J., & Thompson, R. C. A. (2012). The molecular epidemiology of parasite infections: Tools and applications. *Molecular and Biochemical Parasitology, 181,* 102-116.

Malcicka, M., Agosta, S. J., & Harvey, J. A. (2015). Multi level ecological fitting: indirect life cycles are not a barrier to host switching and invasion. *Global change biology, 21,* 3210-3218.

Manga-Gonzalez, M. Y., Gonzalez-Lanza, C., Cabanas, E., & Campo, R. (2001). Contributions to and review of dicrocoeliosis, with special reference to the intermediate hosts of *Dicrocoelium dendriticum*. *Parasitology, 123,* 91-114.

McManus, D., & Bowles, J. (1996). Molecular genetic approaches to parasite identification: their value in diagnostic parasitology and systematics. *International Journal for Parasitology, 26,* 687-704.

Otranto, D., & Traversa, D. (2002). A review of dicrocoeliosis of ruminants including recent advances in the diagnosis and treatment. *Veterinary Parasitology, 107,* 317-335.

Pilsbry, H. A. (1939). *Land mollusca of North America:(north of Mexico)*: Philidelphia, USA: Academy of Natural Sciences.

Prugnolle, F., Liu, H., de Meeus, T., & Balloux, F. (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology, 35,* 255-263.

Pybus, M. J. (1990). Survey of hepatic and pulmonary helminths of wild cervid in Alberta, Canada. *Journal of Wildlife Diseases, 26*, 453-459.

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K., & Lane, R. (2005). Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360*, 1805-1811.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters, 9*, 615-629.

Theron, A., Sire, C., Rognon, A., Prugnolle, F., & Durand, P. (2004). Molecular ecology of *Schistosoma mansoni* transmission inferred from the genetic composition of larval and adult infrapopulations within intermediate and definitive hosts. *Parasitology, 129*, 571-585.

Tompkins, D. M., & Poulin, R. (2006). Parasites and Biological Invasions. In R. B. Allen & W. G. Lee (Eds.), *Biological Invasions in New Zealand*: Berlin, Heidelberg: Springer Berlin Heidelberg.

Traub, R., Monis, P., & Robertson, I. (2005). Molecular epidemiology: a multidisciplinary approach to understanding parasitic zoonoses. *International Journal for Parasitology, 35*, 1295-1307.

van Paridon, B. J., Goater, C. P., Gilleard, J. S., & Criscione, C. D. (2016). Characterization of nine microsatellite loci for *Dicrocoelium dendriticum*, an emerging liver fluke of ungulates in North America, and their use to detect clonemates and random mating. *Molecular and Biochemical Parasitology, 207*, 19-22.

Wasmuth, J. D. (2014). Realizing the promise of parasite genomics. *Trends in Parasitology, 30*, 321-323.

Wheeler, W. M. (1913). A revision of the ants of the genus Formica (Linne) Mayr. *Bulletin of the Museum of Comparative Zoology, 53*, 379-565.

Wilson, D. S. (1977). How nepotistic is the brain worm? *Behavioral Ecology and Sociobiology, 2*, 421-425.

Yin, M., Li, H., McManus, D. P., Blair, D., Su, J., Yang, Z., . . . Hu, W. (2015). Geographical genetic structure of *Schistosoma japonicum* revealed by analysis of mitochondrial DNA and microsatellite markers. *Parasites & Vectors, 8*, 150.

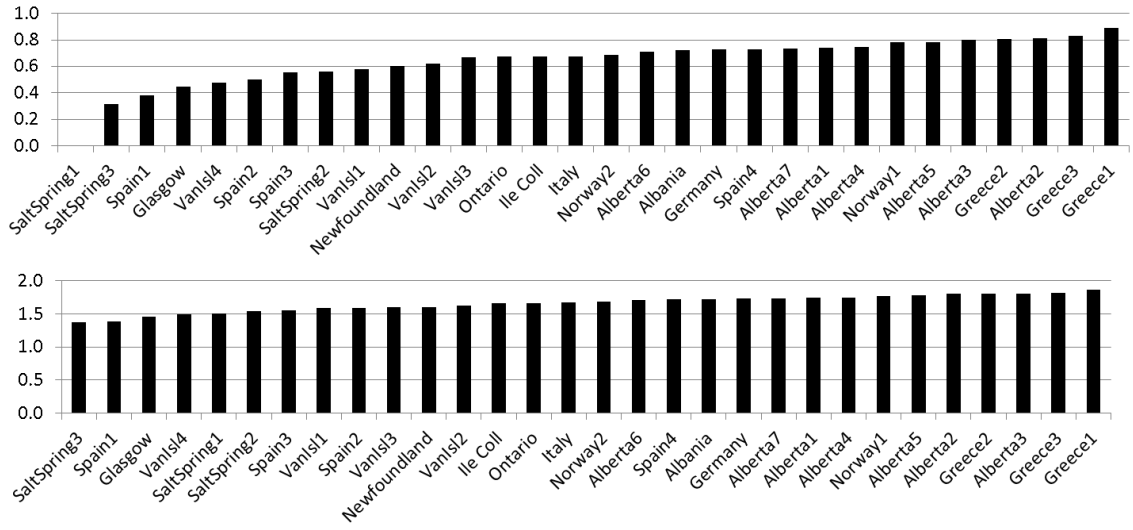
APPENDIX I

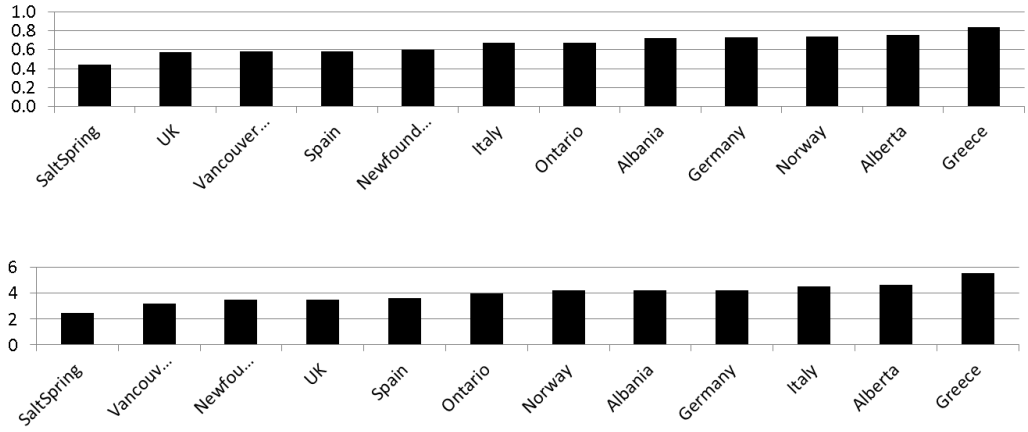
Supplementary Data Chapter 4

Supplementary Table S1. Results of microsatellite validation using all nine microsatellite loci. Results of calculations for F_{IS} overall and by locus, F_{ST}, genetic equilibrium (GD) and estimates of selfing are shown. Values in bold are significant. $P < 0.05$

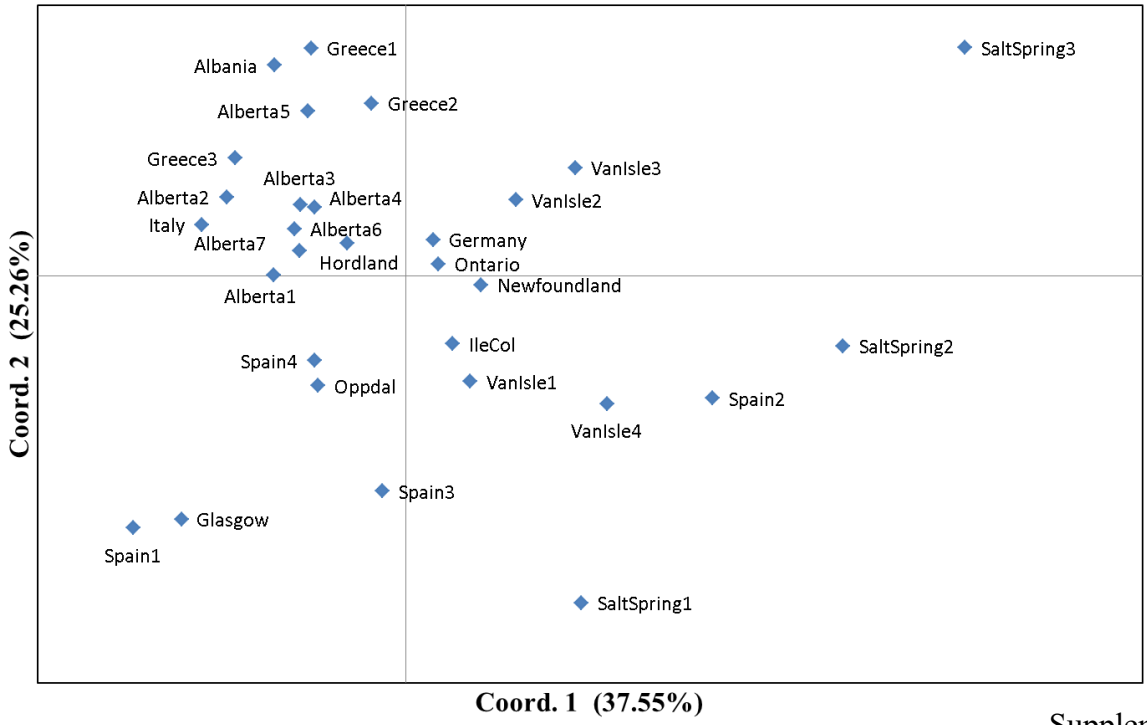
	All Nine Loci	
	With Clones (n = 166)	Without Clones (n = 152)
F_{IS} (overall)	0.168 (0.00010)	0.178 (0.00010)
F_{IS} (by Locus)		
DdMs21	0.059	0.068
DdMs28	-0.012	0.003
DdMs43	**0.262**	**0.259**
DdMs44	**0.392**	**0.402**
DdMs60	0.041	0.071
DdMs70	**0.122**	**0.129**
DdMs89	**0.319**	**0.324**
DdMs93	**0.276**	**0.264**
DdMs95	0.002	0.008
F_{ST}	**0.082 (0.00010)**	**0.065 (0.00010)**
GD	16/36 significant	2/36 significant
estimate of selfing(s = $\hat{g} = 0$)	N.A	none

Supplementary Figure S1. Genetic diversity (H_s) (top) and allelic richness (A_c) (bottom) of *D. dendriticum* within each individual host based on four microsatellite loci.

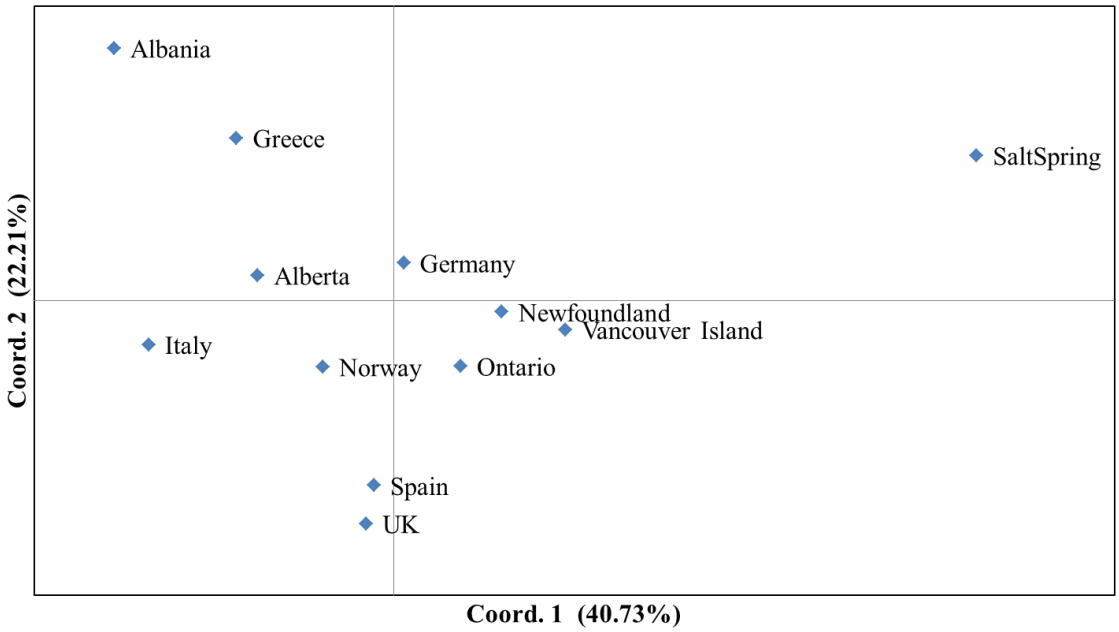




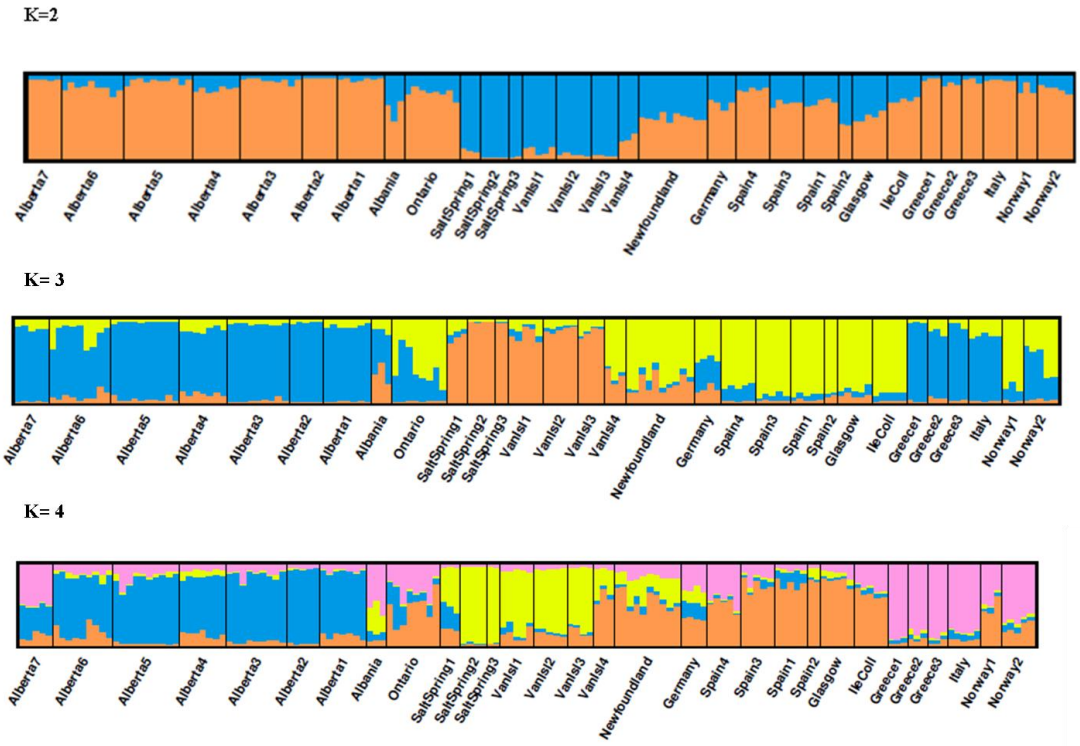
Supplementary Figure S2. Genetic diversity (H_s) (top) and allelic richness (A_c) (bottom) of *D. dendriticum* within each host grouped as regions based on four microsatellite loci.



Supplementary Figure S3. PCA plot of pairwise F_{ST} values between *D. dendriticum* within individual hosts based on four microsatellite loci.



Supplementary Figure S4. PCA plot of pairwise F_{ST} values between *D. dendriticum* in hosts grouped as regions based on four microsatellite loci.



Supplementary Figure S5. Results of STRUCTURE clustering analysis of *D. dendriticum* from individual hosts based on four microsatellite loci. Values of K = 2-4. Host IDs are shown below each plot.