Population structure and associated larval host variation of the forest tent caterpillar, *Malacosoma disstria*

by

Kyle Lawrence Snape

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Systematics and Evolution

Department of Biological Sciences

University of Alberta

Abstract

The forest tent caterpillar, *Malacosoma disstria* (*M. disstria*) Hübner, is a major forest defoliator with regional differences in host association across its range, but the factors shaping its population structure are poorly understood. In eastern Canada, *M. disstria* primarily feeds on maple (*Acer saccharum*) or aspen (*Populus tremuloides*), and earlier studies have documented functional differences between populations on different larval hosts. However, it is not known whether these populations differ genetically. Clarification of the link between host races, genetic variation and geographic distribution can help to inform our understanding of *M. disstria* population dynamics. I collected 130 *M. disstria* larvae from eastern Canada, Alberta and Saskatchewan to characterize their population genetic structure, using a reduced representation library to genotype 9,284 SNPs (single nucleotide polymorphisms) across their genome. I found no meaningful genetic differences between *M. disstria* sampled on different larval hosts. However, I did detect regional genetic variation between populations sampled from different ecozones within eastern Canada. On a broader geographic scale, I also found strong divergence between eastern and central populations. Mitochondrial sequences (new and previously published) loosely supported this east-west division. *M. disstria* population structure therefore appears to be shaped by geography and regional forest structure, rather than larval host.

Acknowledgements

I have many to thank for their support that helped me through this degree, but most importantly my supervisors, Dr. Amanda Roe and Dr. Felix Sperling. Without your guidance and constant motivation, I am not sure I would have made it to where I am today. Thank you for believing in me and your mentorship during these past two years.

For the many members and associates of the Sperling Lab, I would like to thank you all as well. Tyler Nelson, Erin Campbell, Brittany Wingert, Victor Shegelski, Rowan French, Zac MacDonald, Janet Sperling, Oksana Vernygora, Marnie Wright, Shawn Abraham, Lisa Lumley and Stephen Trevoy. Thank you for all your contributions and support.

In addition, there were many others who contributed to my success. For the lab assistance at Natural Resources Canada, I would like to thank Alice Yuehong, Eric Lemieux, Kevin Ong and Meng Zhang. Thank you also to the many research scientists, professors, field rangers and technicians that helped to further my work, notably Caroline Whitehouse, John Acorn, Emma Despland, Tyler Wist, Joshua Jarry, Julian Dupuis, Anne-Sophie Caron, Vanessa Chaimbrone, Lia Fricano, Mike Francis, Ariel Ilic, Lia Fricano, Kristin Hicks, Christi Jaeger, Chris McVeety and Reshma Jose. Your support greatly expanded my research, and without your support, would have not been possible.

I would also like to thank my other committee members, Maya Evenden and Carol Frost. Thank you for your feedback and advice, which improved and enriched this project.

Last, my family that provided the support to help me through this degree. Mackenzie, Shannon, Terra-Lynn and Sam, thank you for everything. Mom and Mike, Dad and Andrea, thank you for helping me pursue this dream and passion of mine.

Funding was provided by Natural Resources Canada and Engineering Research Council of Canada Discovery Grant to Felix Sperling, and with support from Natural Resources Canada through Amanda Roe.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

bp: base pair

BOLD: Barcode of Life Data System

COI: cytochrome c oxidase subunit I

DAPC: discriminant analysis of principal components

ddRAD: double-digest restriction-site-associated

dNTP: deoxynucleoside triphosphate

EXO: exonuclease

EXOSAP: exonuclease-shrimp alkaline phosphatase

GLFC: Great Lakes Forestry Centre

IPQL: Insect Production and Quarantine Laboratory

MBSU: Molecular Biology Service Unit

MCMC: Markov Chain Monte Carlo

mtDNA: mitochondrial deoxyribonucleic acid

PC1: principal component axis one

PC2: principal component axis two

PCA: principal component analysis

PCR: polymerase chain reaction

RNase: ribonuclease

SAP: shrimp alkaline phosphatase

SNP: single nucleotide polymorphism

TCS: Templeton, Crandall and Sing; also known as statistical parsimony

List of Appendices

Chapter 1

Introduction and Thesis Objective

1.1 General introduction

Insects are a highly diverse lineage of organisms, comprising half of all described species (Nakadai 2017). This diversity is a result of barriers to mating and reproduction arising from intraspecific variation within a species, which is influenced by many factors interacting on a population's fitness such as habitat specialization (Loxdale *et al.* 2011). These habitats possess many novel factors acting on a diverging population, with interactions between insects and novel host plants often attributed to leading to population divergence and speciation (Nymen *et al.* 2006). Other factors can also influence divergence, such as predation (Nosil and Crespi 2006), regional abiotic conditions (Dillon and Lozier 2019) and historical biogeography (Braby *et al.* 2007). Thus, intraspecific diversification is a result of interacting factors, and establishing how their relationship affects population structure can be critical in understanding evolutionary processes (Bankhead-Dronnet *et al.* 2015). It is therefore necessary to assess the impact that these distinct processes have in shaping population structure.

Insect-plant coevolution has largely driven intraspecific diversification, having a strong effect on host plants and their associated insects (Endara *et al.* 2015). Through strong selective pressures on a population, a subset of the population that is feeding on a novel host may adapt by accumulating beneficial alleles (Fricke and Arnqvist 2007; Ohshima 2007). In addition, functional differences in variable life history traits enable population subsets to better adapt to a novel host (Carroll *et al.* 2003). These alter population fitness, allowing future generations a survival advantage on that novel host (Nosil 2012). Over time, divergence in an insect population will occur, associated with the different hosts, as mate choice (Ferrari *et al.* 2006) and

oviposition preference (Downey and Nice 2013) affect gene flow between the populations. As a result, genetic divergence increases between host race populations (Peccoud *et al.* 2009).

Even with the prominence of studies examining host races, intraspecific diversity is not merely the result of one factor influencing a population, but rather a consequence of many interacting factors that drive divergence. Tri-trophic interactions can shape a population through ovipositional preference. Herbivores feeding on a potential host plant can trigger an indirect plant defence through chemical release that attracts predators, thus pressuring females to select an alternative host (Shiojiri and Takabayashi 2003). Environmental conditions can structure populations based on a thermal response to temperature fluctuation and tolerance, with a thermal gradient driving distribution and polymorphisms (Chu *et al.* 2014). These types of environmental pressures will lead to different population structures and can cause ecological and, eventually, genetic divergence from the parent population (Rundle and Nosil 2005; Schluter 2009).

Historical biogeography can also shape the population structure of insects. Historic host dispersal, with increasing ranges due to changing geological conditions, permitted many species to spread into new areas. These dispersal events were responsible for the origins of new biodiversity across North America (Meseguer *et al.* 2015). In addition, historical geological and climate events would have affected multiple species in the same region, leading to similarities among different species' phylogeographic trees (Lapointe and Rissler 2005). For example, in northern temperate species, the post-glaciation recolonization of temperate habitats would have left distinct genetic signatures within populations (Gratton *et al.* 2008). Therefore, historic climate and geological barriers could have also heavily influenced the distribution of many species and contributed to intraspecific variation.

However, none of these factors work in isolation. Populations are always changing with time and represent a brief moment in an ever-changing evolutionary spectrum (de Quieroz 2007). Insect herbivores are particularly suited to studying these factors through a nuanced examination of host plant differences, environmental factors, and historical biogeography.

1.2 *Malacosoma*

Multiple factors influencing diversification are seen in the genus *Malacosoma*, the tent caterpillars, that form a well-known group of Lepidoptera in the family Lasiocampidae. Consisting of several well-known forest pest species, *Malacosoma* species occur along a large geographic range spanning multiple continents and forest ecozones suitable for teasing apart these different factors. Tent caterpillars are typically active from early spring until mid-summer. Neonate larvae overwinter in an egg mass laid the previous summer and emerge in the following spring (Gray and Ostaff 2012). As bud break varies across the landscape, warm conditions promote synchrony between various hosts and larvae (Parry *et al.* 1998). Newly emerged larvae are dependent on new foliage, both due to its relative availability and its quality (Despland and Noseworthy 2006). Most *Malacosoma* are well known for their tent building and gregarious behaviour during this larval stage (Progar *et al.* 2010; McClure and Despland 2010; Franklin *et al.* 2012), with *M. disstria* instead spinning silken mats (Despland and Hamzeh 2004). Sociality improves survival (Despland and Huu 2006), and these group dynamics may increase fitness by decreasing predation, increasing foraging efficiency and improving thermoregulation (Despland and Hamzeh 2004). *Malacosoma* species remain in colonies throughout their early larval stage with siblings, becoming more solitary as they progress to later instars.

Globally, 26 species are recognized in *Malacosoma* by Fitzgerald (1995), all in the northern hemisphere, in North America, Europe, north Africa and Asia (Fitzgerald 1995). North America has six species and several subspecies: *M. disstria* (Hübner), *M. constrictum* (Edwards), *M. tigris* (Dyar), *M. americanum* (Fabricius), *M. californicum* (Packard) *and M. incurvum* (Edwards) (Franclemont 1973). Larvae of these species vary considerably in body colours and patterns (Figure 1.1). In addition to colouration, host associations, geographic distributions, tent type and use, egg mass shape and ovipositor shape are key taxonomic features.

Malacosoma species have received several population genetic studies. Costa and Ross (1994) examined population structuring of *M. americanum* along the Atlantic coast of the United States and found only minor differences between populations over moderate to large geographic distances. Franklin *et al.* (2014) found little genetic differentiation among populations of *M. californicum pluviale* (Dyar) from sequenced mitochondrial COI samples of populations from south western British Columbia. Lait and Hebert (2018) examined BOLD (Barcode of Life Data System) COI sequences of *M. americana*, *M. californica* and *M. disstria.* They found strong population structure among *M. californica* and *M. disstria* populations but *M. americana* had the weakest structuring (Lait and Hebert 2018). While Lait and Hebert examined several *Malacosoma* species across North America, I focused on forest tent caterpillars (*M. disstria)* populations, to see which factors are more strongly influencing population structure and intraspecific diversity.

1.3 *Malacosoma disstria*

Studies examining *M. disstria* have found intraspecific differences among the *M. disstria* populations, with geographic variation in several life history traits such as outbreak dynamics,

larval size, and egg size. In eastern populations, *M. disstria* outbreaks have been documented to occur approximately every 10 years and are more common in fragmented forests (Cooke and Lorenzetti 2006). Western populations, however, experience more variation in outbreaks, being more unpredictable and asynchronous than in eastern populations (Cooke and Roland 2018). Larval size also varies with latitude. Larvae from northern populations tend to be larger than those from southern populations. Parry *et al.* (2001) suggested that a larger body size (and by extension larger eggs) in northern latitudes is advantageous as it equips emerging larvae with greater energy stores if larvae emerge prior to bud burst. The trade off is that females lay fewer eggs per clutch than those in southern populations, which have smaller eggs but more per clutch and smaller larvae when they emerge. Similar patterns are observed in other widely distributed forest pests, such as *Choristoneura fumiferana* (Clemens) (Harvey 1983).

In addition to observed geographic variation, *M. disstria* shows functional differences in life history traits that relate to the different host plants found along its Canadian range. *M. disstria* has been documented on a variety of larval host families (Fagaceae, Hamamelidaceae, Rosceae and Salicaceae: Futuyma and Saks 1981; Aceraceae, Cornaceae, Nyssaceae: Fitzgerald 1995); however, as *M. disstria* oviposits on select regional species, *M. disstria* may be regionally oligophagous (Parry and Goyer 2004). In the northeastern boreal forests of Ontario and Quebec, *M. disstria* are primarily found on trembling aspen (*Populus tremuloides* Michx). Further south, in the temperate deciduous forests, *M. disstria* feeds and oviposits on several deciduous trees: trembling aspen (*Populus tremuloides* Michaux), sugar maple (*Acer saccharum* Marsh) and northern red oak (*Quercus rubra* L. synonym *Q. borealis* Michx). Between these two forest types, the eastern temperate mixed forest transition zone separates northern conifer and aspen forests from southern deciduous forests. This area is defined by hardwood covered hills, with

conifer dominated valleys (Goldblum and Rigg 2005). Forests dominated by sugar maple give way to balsam fir (*Abies balsamea* (L.) Mill), white spruce (*Picea glauca* (Moench) Voss), trembling aspen and red maple (*Acer rubrum* L.) forests.

Elsewhere in its range, in the deltas of the southern United States, *M. disstria* populations in Alabama and Louisiana prefer water tupelo (*Nyssa aquatica* L.), blackgum (*Nyssa sylvactia var biflora* Walt), and sweetgum (*Liquidambar styraciflua* L.) (Stark and Harper 1982). In western Canada, *M. disstria* populations are located in aspen parklands, spanning Alberta and western Saskatchewan (Williams and Langor 2011). Along the Pacific coast, *M. disstria* can be found in sub-boreal spruce forests intermixed between white spruce, balsam poplar (*Populus balsamifera* L.) and trembling aspen (Schmidt *et al.* 2003). Among all these different tree hosts, adults mate in the forest canopy at low densities, and during outbreaks, mating shifts toward the forest understory (Miller 2006). In captivity, adult males have been shown to fly up to about 3.3 km in a single sustained flight (Evenden *et al.* 2015). Wind currents may lead to greater dispersal distances (~500 km) (Fitzgerald 1995). *M. disstria* dispersal is driven mainly by males seeking out females through female-produced sex-pheromones (Evenden *et al.* 2015). However, females do show limited flight propensity (Fitzgerald 1995), with a brief pre-oviposition flight. After mating, females select a host for egg laying, and it remains unclear if female host preference is influenced by the host on which they matured, and whether they exhibit diversified risk spreading (Gray and Ostaff 2012). A single egg mass is laid per female, with all eggs laid together to promote larval aggregation, a survival mechanism (Fitzgerald 1995).

Larval performance is strongly affected by host plant phytochemical variation and phenology. Some closely related tree species, such as red maple, are resistant to *M. disstria*. Red maple foliage is toxic to *M. disstria* larvae, resulting in high mortality for individuals that feed

upon its foliage. Survival rates vary between larvae that feed on aspen versus maple (Nicol *et al.* 1997), with larvae on sugar maple showing higher mortality than on aspen. However, this study did not account for potential underlying regional differences in host preference. Host plants also impact pupal mass and developmental time (Parry and Goyer 2004). Therefore in *M. disstria,* larval weight and development times are affected differently by different hosts. Aspen-fed *M. disstria* develop faster and have greater biomass than maple-fed *M. disstria*. Among northern populations, female *M. disstria* show regional ovipositional preference for trembling aspen but can also be found alternatively on sugar maple (Trudeau *et al.* 2010). Host phenology can also impact larval performance. Host phytochemistry varies over time, becoming less nutritious and therefore less beneficial to the development of recently emerged larvae than later foliage (Fuentealba *et al.* 2017). Larval performance is also affected by climate change, through changes in plant phytochemistry (Jamieson *et al.* 2015). With temperature increases due to climate change, different leaf nutrients can be affected, as well as defensive chemicals, that negatively impact larval development. Therefore, synchrony between emerging larvae and budbreak is critical for the survival of the young larvae, as starvation or harmful phytochemical intake are possible when larvae emerge outside the appropriate time window.

The economic impact of *M. disstria* varies regionally. In eastern Canada, *M. disstria* is commonly found feeding on sugar maple. Sugar maple, in addition to Ontario's syrup production, also happens to be one of the dominant hardwood species used in hardwood lumber (Niese and Strong 1992). Trembling aspen, the primary host of *M. disstria* in western Canada, is another widespread harvest tree that is commonly devastated during outbreaks (Brandt *et al.* 2003). During these outbreak years, *M. disstria* caterpillars can cause severe defoliation (Arimura *et al.* 2004) and larvae can disperse into nearby forest stands. While dispersing, these

populations have caused traffic and rail delays, and been responsible for power disruption as far back as the 1950s (Sipple 1962). The effects of these outbreaks are most observable when the cumulative effects of diseases, other phytophagous insects and droughts lead to high tree mortality (Hogg *et al.* 2002). However, as *M. disstria* populations reach their peak numbers, a decline eventually occurs. This is due to greatly increased parasitism and disease epidemics that run rampant throughout the population (Donaldson and Lindroth 2008). These factors bring down the *M. disstria* population to far more manageable numbers. A cost assessment of controlling for *M. disstria* has shown that in Saskatchewan and New Brunswick alone, there would be a social benefit ranging from $7.9 to $22.0 million in reducing their population numbers (Niquidet *et al.* 2016). In addition, when the primary hosts are in short supply during these outbreak years, *M. disstria* will target other hosts, which can lead to growth and widespread die back if coupled with other biotic factors.

1.4 Thesis objectives

Forest tent caterpillars show intraspecific variation among geographic regions, with various life history differences associated to hosts, geographic structuring and outbreak dynamics. These life history differences may be due to plasticity or adaptive divergence between populations. To start disentangling these processes, I needed to characterize the underlying genetic population structure of *M. disstria* populations. Specifically, I asked which factors influence the population genetic structure in *M. disstria* in Canada. Evidence of varying regional outbreak dynamics and life history differences hints at underlying genetic differences between populations. I sampled egg masses and larvae on different host plants in Ontario, Quebec, Alberta and Saskatchewan, and characterized the genomic variation among these populations.

Overall, my thesis identified the relative contributions of several extrinsic factors shaping the genetic diversity and population structure of *M. disstria* based on large-scale continental divisions between central and eastern Canada, variation between the primary hosts, in addition to structuring based on regional forest zones.

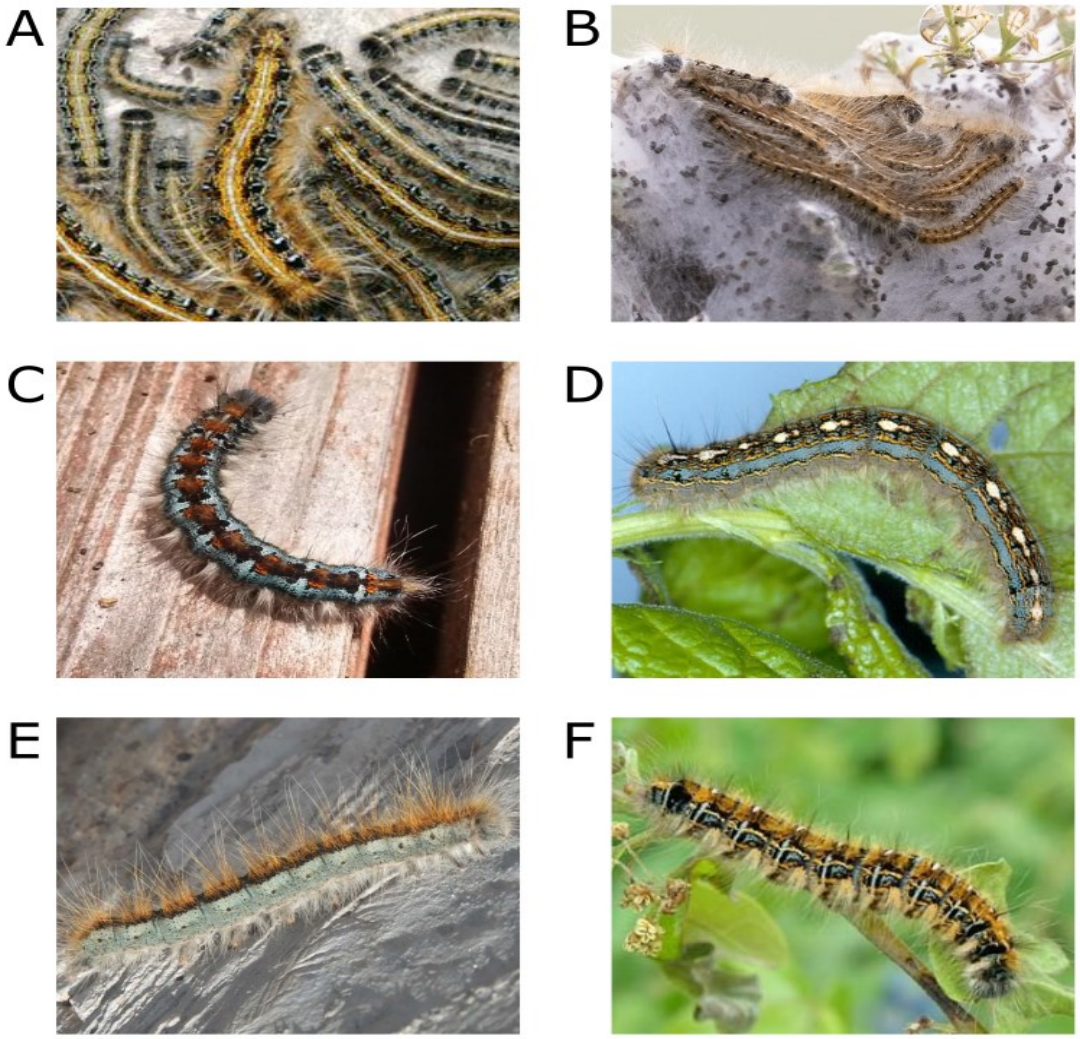


Figure 1.1 North American *Malacosoma* species. A: *M. americanum* (image courtesy of Nolie Schneider), B: *M. californica* (image courtesy of Sylvia Wright), C: *M. constricta* (image courtesy of Sylvia Wright), D: *M. disstria* (image is part of the Canadian National Collection; permission for use granted by curator Chris Schmidt), E: *M. incurva* (image courtesy of Robert Webster), F: *M. tigris* (image courtesy of Cynthia Van Den Broeke). All images obtained from the North American Moth Photographers Group at the Mississippi Entomological Museum.

1.5 Literature cited

Arimura GI, Huber DPW, Bohlmann J. 2004. Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa x deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (-)-germacrene D synthase, PtdTPS1. *The Plant Journal.* 37, 603-616. https://doi.org/10.1111/j.1365-313X.2003.01987.x

Bankhead-Dronnet S, Perdereau E, Kutnik M, Dupont S, Bagneres AG. 2015. Spatial structuring of the population genetics of a European subterranean termite species. *Ecology and Evolution.* 5, 3090-3102. https://doi.org/10.1002/ece3.1566

Braby MF, Pierce NE, Vila R. 2007. Phylogeny and historical biogeography of the subtribe Aporiina (Lepidoptera: Pieridae): implications for the origin of Australian butterflies. *Biological Journal of the Linnean Society.* 90, 413-440. https://doi.org/10.1111/j.1095-8312.2007.00732.x

Brandt JP, Cerezke HF, Mallett KI, Volney WJA, Weber JD. 2003. Factors affecting trembling aspen (*Populus tremuloides* Michx) health in the boreal forest of Alberta, Saskatchewan, and Manitoba, Canada. *Forest Ecology and Management*. 178, 287-300. https://doi.org/10.1016/S0378-1127(02)00479-6

Carlyle-Moses DE, Price AG. 2006. Growing-season stemflow production within a deciduous forest of southern Ontario. *Hydrological Processes*. 20, 3651-3663. https://doi.org/10.1002/hyp.6380

Carroll SP, Marler M, Winchell R, Dingle H. 2003. Evolution of cryptic flight morph and life history differences during host race radiation in the soapberry bug, *Jadera haematoloma* Herrich-Schaeffer (Hemiptera: Rhopalidae). *Annals of the Entomological Society of*

America. 96, 135-143. https://doi.org/10.1603/0013-8746(2003)096[0135:EOCFMA]2.0.CO;2

Chu ND, Kaluziak ST, Trussell GC, Vollmer SV. 2014. Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus*. *Molecular Ecology*. 23, 1863-1873. https://doi.org/10.1111/mec.12681

Cooke BJ, Lorenzetti F. 2006. The dynamics of forest tent caterpillar outbreaks in Quebec, Canada. *Forest Ecology and Management*. 226, 110-121. http://doi.org/10.1016/j.foreco.2006.01.034

Cooke BJ, Roland J. 2018. Early 20th century climate-driven shift in the dynamics of forest tent caterpillar outbreaks. *American Journal of Climate Change*. 7, 253-270. http://doi.org/10.4236/ajcc.2018.72017

Costa JT, Ross KG. 1994. Hierarchical genetic structure and gene flow in macrogeographic populations of the eastern tent caterpillar (*Malacosoma americanum*). *Evolution*. 48, 1158-1167. http://doi.org/10.1111/j.1558-5646.1994.tb05302.x

de Quieroz K. 2007. Species concepts and species delimitation. *Systematic Biology*. 56, 879-886. https://doi.org/10.1080/10635150701701083

Despland E, Hamzeh S. 2004. Ontogenetic changes in social behaviour in the forest tent caterpillar, *Malacosoma disstria*. *Behavioral Ecology and Sociobiology*. 56, 177-184. https://doi.org/10.1007/s00265-004-0767-8

Despland E, Huu AL. 2006. Pros and cons of group living in the forest tent caterpillar: separating the roles of silk and of grouping. *Entomologia Experimentalis et Applicata*. 122, 181-189. https://doi.org/10.1111/j.1570-7458.2006.00512.x

Despland E, Noseworthy M. 2006. How well do specialist feeders regulate nutrient intake? Evidence from a gregarious tree-feeding caterpillar. *Journal of Experimental Biology*. 209, 1301-1309. https://doi.org/10.1242/jeb.02130

Dillon ME, Lozier JD. 2019. Adaptation to the abiotic environment in insects: the influence of variability on ecophysiology and evolutionary genomics. *Current Opinion in Insect Science*. 36, 131-139. https://doi.org/10.1016/j.cois.2019.09.003

Donaldson JR, Lindroth RL. 2008. Effects of variable phytochemistry and budbreak phenology on defoliation of aspen during a forest tent caterpillar outbreak. *Agricultural and Forest Entomology*. 10, 399-410. https://doi.org/10.1111/j.1461-9563.2008.00392.x

Downey MH, Nice CC. 2013. A role for both ecology and geography as mechanisms of genetic differentiation in specialized butterflies. *Evolutionary Ecology*. 27, 565-578. https://doi.org/10.1007/s10682-012-9626-7

Endara MJ, Weinhold A, Cox JE, Wiggins NL, Coley PD, Kursar TA. 2015. Divergent evolution in antiherbivore defences within species complexes at a single Amazonian site. *Journal of Ecology*. 103, 1107-1118. https://doi.org/10.1111/1365-2745.12431

Evenden ML, Whitehouse CM, Jones BC. 2015. Resource allocation to flight in an outbreaking forest defoliator *Malacosoma disstria. Physiological Ecology.* 44, 835-845. https://doi.org/10.1093/ee/nvv055

Ferrari J, Godfray HCJ, Faulconbridge AS, Prior K, Via S. 2006. Population differentiation and genetic variation in host choice among pea aphids from eight host plant genera. *Evolution*: 60, 1574-1584. https://doi.org/10.1554/06-024.1

Fitzgerald TD. 1995. *The Tent Caterpillars*. Ithica (NY): Cornell University Press. 303 p.

Franclemont JG. 1973. *The moths of America north of Mexico: including Greenland. Fascicle 20.1. Mimallonoidea Mimallonidae and Bombycoidea Apatelodidae, Bombycidae,* Lasiocampidae. London (UK): EW Classey Limited and RBD Publications Inc. 86 p.

Franklin MT, Myers JH, Cory JS. 2014. Genetic similarity of island populations of tent caterpillars during successive outbreaks. *PLOS One*. 9, 1-9. https://doi.org/10.1371/journal.pone.0096679

Franklin MT, Ritland CE, Myers JH, Cory JS. 2012. Multiple mating and family structure of the western tent caterpillar, *Malacosoma californicum pluviale*: impact on disease resistance. *PLOS One*. 7, 1-8. https://doi.org/10.1371/journal.pone.0037472

Fricke C, Arnqvist G. 2007. Rapid adaptation to a novel host in a seed beetle (*Callosobruchus maculatus*): the role of sexual selection. *Evolution.* 61, 440-454. https://doi.org/10.1111/j.1558-5646.2007.00038.x

Fuentealba A, Pureswaran D, Beauce E, Despland E. 2017. How does synchrony with host plant affect the performance of an outbreaking insect defoliator? *Oecologia*. 184, 847-857. https://doi.org/10.1007/s00442-017-3914-4

Futuyma DJ, Saks ME. 1981. The effect of variation in host plant on the growth of an oligophagous insect, *Malacosoma americanum* and its polyphagous relative *Malacosoma disstria*. *Entomologia Experimentalis et Applicata.* 30, 163-168. https://doi.org/10.1111/j.1570-7458.1981.tb03091.x

Goldblum D, Rigg LS. 2005. Tree growth response to climate change at the deciduous-boreal forest ecotone, Ontario, Canada. *Canadian Journal of Forest Research*. 35, 2709-2718. https://doi.org/10.1139/X05-185

Gratton P, Konopinski MK, Sbordoni V. 2008. Pleistocene evolutionary history of the clouded apollo (*Parnassius mnemosyne*): genetic signatures of climate cycles and a "time-dependent" mitochondrial substitution rate. *Molecular Ecology*. 17, 4248-4262. https://doi.org/10.1111/j.1365-294X.2008.03901.x

Gray DR, Ostaff DP. 2012. Egg hatch of forest tent caterpillar (Lepidoptera: Lasiocampidae) on two preferred host species. *The Canadian Entomologist*. 144, 790-797. https://doi.org/10.4039/tce.2012.73

Harvey GT. 1983. A geographical cline in egg weights in *Choristoneura fumiferana* (Lepidoptera: Tortricidae) and its significance in population dynamics. *The Canadian Entomologist*. 115, 1103-1108. https://doi.org/10.4039/Ent1151103-9

Hogg EH, Brandt JP, Kochtubajda B. 2002. Growth and dieback of aspen forests in northwestern Alberta, Canada, in relation to climate and insects. *Canadian Journal of Forest Research*. 32, 823-832. https://doi.org/10.1139/X01-152

Jamieson MA, Schwartzberg EG, Raffa KF, Reich PB, Lindroth RL. 2015. Experimental climate warming alters aspen and birch phytochemistry and performance traits for an outbreak insect herbivore. *Global Change Biology*. 21, 2698-2710. https://doi.org/10.1111/gcb.12842

Kebli H, Brais S, Kernaghan G, Drouin P. 2012. Impact of harvesting intensity on wood-inhabiting fungi in boreal aspen forests of Eastern Canada. *Forest Ecology and Management*. 279, 45-54. http://doi.org/10.1016/j.foreco.2012.05.028

Lait LA, Hebert PDN. 2018. Phylogeographic structure in three North American tent caterpillar species (Lepidoptera: Lasiocampidae): *Malacosoma americana*, *M. californica*, and *M. disstria*. *PeerJ*. 6, 2-23. https://doi.org/10.7717/peerj.4479

Lapointe FJ, Rissler LT. 2005. Congruence, consensus, and the comparative phylogeography of codistributed species in California. *The American Naturalist*. 166, 290-299. https://doi.org/10.1086/431283

Loxdale H, Lushai G, Harvey JA. 2010. The evolutionary improbability of 'generalism' in nature, with special reference to insects. *Biological Journal of the Linnean Society.* 103, 1-18. https://doi.org/10.1111/j.1095-8312.2011.01627.x

McClure M, Despland E. 2010. Collective foraging patterns of field colonies of *Malacosoma disstria* caterpillars. *The Canadian Entomologist.* 142, 473-480. https://doi.org/10.4039/n10-001

Miller WE. 2006. Forest tent caterpillar: mating, oviposition, and adult congregation at town lights during a northern Minnesota outbreak. *Journal of the Lepidopterists' Society*. 60, 156-160

Meseguer AS, Coeur D'acier A, Genson G, Jousselin E. 2015. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus. *Journal of Biogeography*. 42, 1482-1492. https://doi.org/10.1111/jbi.12531

Nakadai R. 2017. Species diversity of herbivorous insects: a brief review to bridge the gap between theories focusing on the generation and maintenance of diversity. *Ecological Research*. 32, 811-819. https://doi.org/10.1007/s11284-017-1500-1

Nicol RW, Arnason JT, Helson B, Abou-Zaid MM. 1997. Effect of host and non-host trees on the growth and development of the forest tent caterpillar, *Malacosoma disstria* (Lepidoptera: Lasiocampidae). *The Canadian Entomologist*. 129, 991-999. https://doi.org/10.4039/Ent129991-6

Niese JN, Strong TF. 1992. Economic and tree diversity trade-offs in managed northern hardwoods. *Canadian Journal of Forest Research*. 22, 1807-1813. https://doi.org/10.1139/x92-236

Niquidet K, Tang J, Peter B. 2016. Economic analysis of forest insect pests in Canada. *The Canadian Entomologist*. 148, 357-366. https://doi.org/10.4039/tce.2015.27

Nosil P, Crespi BJ. 2006. Experimental evidence that predation promotes divergence in adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America*. 103, 9090-9095. https://doi.org/10.1073/pnas.0601575103

Nosil P. 2012. *Ecological Speciation*. Oxford (UK): Oxford University Press. 280 p.

Nyman T, Farrell BD, Zinovjev AG, Vikberg V. 2006. Larval habits, host-plant associations and speciation in nematine sawflies (Hymenoptera: Tenthredinindae). *Evolution*. 60, 1622-1637. https://doi.org/10.1554/05-674.1

Ohshima I. 2007. Host race formation in the leaf-mining moth *Acrocercops transecta* (Lepidoptera: Gracillariidae). *Biological Journal of the Linnean Society*. 93, 135-145. https://doi.org/10.1111/j.1095-8312.2007.00949.x

Parry D, Goyer RA, Lenhard GJ. 2001. Macrogeographic clines in fecundity, reproductive allocation, and offspring size of the forest tent caterpillar *Malacosoma disstria*. *Ecological Entomology*. 26, 381-291. https://doi.org/10.1046/j.1365-2311.2001.00319.x

Parry D, Goyer RA. 2004. Variation in the suitability of host tree species for geographically discrete populations of forest tent caterpillar. *Environmental Entomology*. 33, 1479-1487. https://doi.org/10.1603/0046-225X-33.5.1477

Parry D, Spence JR, Volney WJA. 1998. Budbreak phenology and natural enemies mediate survival of first-instar forest tent caterpillar (Lepidoptera: Lasiocampidae). *Environmental Entomology*. 27, 1368-1374. https://doi.org/10.1093/ee/27.6.1368

Peccoud J, Ollivier A, Plantegenest M, Simon J. 2009. A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proceedings of the National Academy of Sciences of the United States*. 106, 7495-7500. https://doi.org/10.1073/pnas.0811117106

Progar RA, Rinella MJ, Fekedulegn D, Butler L. 2010. Nuclear polyhedrosis virus as a biological control agent for *Malacosoma americanum* (Lepidoptera: Lasiocampidae). *Journal of Applied Entomology*. 134, 641-646. https://doi.org/10.1111/j.14390418.2010.01508.x

Rundle HD, Nosil P. 2005. Ecological speciation. *Ecology Letters*. 8, 336-352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science*. 323, 737-742. https://doi.org/10.1126/science.1160006

Schmidt BC, Roland J, Wakarchuk D. 2003. Evaluation of synthetic pheromones for monitoring forest tent caterpillar (Lepidoptera: Lasiocampidae) populations. *Environmental Entomology*. 32, 214-219. https://doi.org/10.1603/0046-225X-32.1.214

Shiojiri K, Takabayashi J. 2003. Effects of specialist parasitoids on oviposition preference of phytophagous insects: encounter-dilution effects in a tri-trophic interaction. *Ecological Entomology*. 28, 573-578. https://doi.org/10.1046/j.1365-2311.2003.00539.x

Sipple WL. 1962. Outbreaks of the forest tent caterpillar, *Malacosoma disstria* Hbn., a periodic defoliator of broad-leaved trees in Ontario. *The Canadian Entomologist*. 94, 408-416. https://doi.org/10.4039/Ent94408-4

Stark EJ, Harper JD. 1982. Pupal mortality in forest tent caterpillar (Lepidoptera: Lasiocampidae): causes and impact on populations in southwestern Alabama. *Environmental Entomology*. 11, 1071-1077. https://doi.org/10.1093/ee/11.5.1071

Trudeau M, Mauffette Y, Rochefort S, Han E, Bauce E. 2010. Impact of host tree on forest tent caterpillar performance and offspring overwintering mortality. *Environmental Entomology*. 39, 498-504. https://doi.org/10.1603/en09139

Williams DJ, Langor DW. 2011. Distribution, species composition, and incidence of egg parasitoids of the forest tent caterpillar (Lepidoptera: Lasiocampidae), during a widespread outbreak in the Canadian prairies. *The Canadian Entomologist*. 143, 272-278. https://doi.org/10.4039/n11-009

Chapter 2

Population structuring of a widespread forest pest, the forest tent caterpillar

Authors:

Kyle L Snape[1,2], Amanda D Roe[2], and Felix AH Sperling[1]

[1] Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2R3

[2] Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste. Marie, ON P6A 2E6

2.1 Summary

Phytophagous insects are highly diverse, which is often attributed to coevolution between insects and host plants. However, other factors also contribute to this diversity, such as regional adaptations, historic biogeography, or tri-trophic interactions. We used the forest tent caterpillar (*Malacosoma disstria* Hübner*),* a widespread forest defoliator in North America, to identify the different extrinsic factors shaping population structure in this forest pest. We collected *M. disstria* from four of its main larval hosts (trembling aspen, sugar maple, red oak and white birch) in eastern Canada, as well as from aspen in Alberta and Saskatchewan. We genotyped 130 specimens using genome-wide SNPs and mtDNA. We found no genetic evidence of larval host races but did detect regional genomic variation and large-scale continental differentiation between eastern and central populations. This indicates that *M. disstria* populations are structured based on geography and ecozone variation, rather than larval hosts, suggesting that host-driven differentiation is not the primary factor underlying *M. disstria* genetic diversity and shaping its population structure.

2.2 Introduction

Phytophagous insects represent 50% of all eukaryotic species (Hardy and Otto 2014) and many evolutionary forces drive this diversification. Many factors (Fuller *et al.* 2005; Stireman III and Singer 2018; Yu *et al.* 2018), have all likely played a role in shaping the intraspecific variation in insect lineages (Nosil and Crespi 2006). However, the relative contributions of these evolutionary forces are often overlooked, leading to a simplified understanding of insect diversification. Thus, to gain a more complete understanding of the processes shaping insect diversity, it is necessary to tease apart the roles different evolutionary forces play in shaping intraspecific variation and population structuring within insect species.

Host use, host race formation, and the coevolution between insects and their host plants has been considered a key force driving insect diversification and ecological speciation (Dres and Mallet 2002; Blair *et al.* 2005; Stireman III *et al.* 2005; Janz *et al.* 2006; Scheffer and Hawthrone 2007; Hardy and Otto 2014). Host race formation occurs as a subset of an insect population diverges due to preferential host selection and adaptation (Pappers *et al.* 2002). As individuals become associated with the new host, gene flow may become restricted due to changes in mate choice (Ferrari *et al.* 2006), phenology (Medina *et al.* 2012), or oviposition preference (Downey and Nice 2013). Feedback loops would further reduce gene flow leading to genetic and functional adaptations to the new hosts (Matsubayashi *et al.* 2011). Over time, host specialization and subsequent host race formation occurs, leading to sympatric coexistence of the two host lineages (Dres and Mallet 2002). Host race formation, however, is only one of several evolutionary processes leading to intraspecific variation.

Just as host race formation can result in diversification, other biotic interactions can influence population structuring and gene flow. For example, tri-trophic interactions (Johnson

2008) or predator-prey relationships (Nosil and Crespi 2006). In addition, abiotic conditions, such as temperature or photoperiod, exert significant influence on development and survival (Hartley *et al.* 2010) and can lead to regionally adapted and genetically structured populations as well (Rochefort *et al.* 2011; Sniegula *et al.* 2014). If, for example, these abiotic traits follow an environmental cline, a genetic cline may develop, characterised by weak global structuring with strong differentiation among regional populations (Wellenreuther *et al.* 2011). Smaller populations are maladapted to local conditions and promote greater differentiation (Gosden *et al.* 2011). On large geographical scales, genetic variation may be due to gene flow increasing fitness at the limits of a species' range (Sexton *et al.* 2011). Therefore, it is critical to examine multiple evolutionary processes collectively, rather than in isolation, to identify the factors leading to intraspecific variation within an insect herbivore.

The forest tent caterpillar (Lepidoptera: Lasiocampidae: *Malacosoma disstria* Hübner) serves as an excellent model with which to explore intraspecific diversification. The species is a major forest defoliator with a geographic distribution spanning multiple forest ecozones across North America (Hartmann and Messier 2011). This pest shows intraspecific variation in both life history traits and genetic variation; *M. disstria* has been documented on at least 15 different hosts (Fitzgerald 1995) and shows regional preferences for these host plants (Charbonneau *et al.* 2012). Spring emergence is regionally synchronized with budbreak in local hosts (Gray and Ostaff 2012), although it feeds predominantly on maple and aspen. Previous work has shown that host plants significantly impact *M. disstria* fitness (Nicol *et al.* 1997; Parry *et al.* 2001; Trudeau *et al.* 2010; Parry and Goyer 2004). In addition, genetic surveys of mitochondrial diversity have shown complex phylogeographic structuring (Lait and Hebert 2018) but could not explain this complexity. Collectively, these studies suggest that several evolutionary forces are contributing

to intraspecific variation and structuring of *M. disstria* populations, although their relative contributions are currently unknown.

We sampled *M. disstria* populations across Canada to quantify the relative contribution of large-scale geography, larval host specialization and regional adaptation in shaping population structure. Specifically, we sampled *M. disstria* from central and eastern Canada, to assess whether large-scale biogeography influences population structuring and explains the complex genetic diversity previously documented within the species. We also sampled from four host plants in eastern Canada to assess if larval host specialization could help explain the complex genetic diversity observed within populations. Finally, we explored whether *M. disstria* populations showed regional specializations by sampling populations from three forest ecozones in eastern Canada. We collected mitochondrial sequence data from the cytochrome c oxidase subunit I (COI) DNA barcode region to link our results with previous studies. We also used double-digest restriction DNA sequencing (ddRAD) to generate genome-wide single nucleotide polymorphisms (SNPs) to quantify overall genomic diversity within *M. disstria* populations. These data allowed us to identify the dominant evolutionary forces shaping the genomic structure of *M. disstria* populations and provided insight into the processes shaping this pest's diversification.

2.3 Methods

2.3.1 Sample collection and processing

M. disstria females lay a single egg band and emerging larvae form a cohesive family group. Given this family structure, we tried to sample multiple egg bands per site (n=10-20) and a single individual per egg band. Egg bands were collected from neighboring trees, in addition to

multiple bands from the same tree, and recently emerged larvae were collected from the same colony. We sampled Ontario and Quebec *M. disstria* populations from three eastern ecozones: boreal forests, temperate mixed forests, and temperate deciduous forest ecozones. The boreal forest ecozone is defined by the presence of the boreal shield and is characterized by a mixture of conifers and aspen dominated stands, average annual temperatures of 0.8°C, and average precipitation of 890mm (Kebli *et al.* 2012). The mixed forest zone, an approximately 50 km transition zone, is where conifer forests mix with the northern hardwood forests along valleys and hills, with annual rainfall of 727.4mm, and temperatures varying from -14.8°C in winters to 14.8°C in the summer (Goldblum and Rigg 2013). The eastern temperate deciduous forest is defined by the diversity of trees that make up the mixedwood plains. It is characterized by deciduous dominated forests, where a humid climate has average annual temperatures of 8°C, and annual precipitation of 785mm (Carlyle-Moses and Price 2006).

We relied heavily on collaborators to send egg bands and recently emerged larvae (Table 2.1). Collecting methods and shipping varied between collaborators, with some egg bands arriving individually while other egg bands from a single locality arrived mixed together at the Great Lakes Forestry Centre (GLFC) in Sault Ste Marie, Ontario. This complicated the separation of families, as some larvae hatched in transit and intermixed with larvae from neighboring egg bands. Egg bands were separated upon arrival, and larvae were fed fresh, locally collected foliage of the same ovipositional host tree species that the egg band was collected from. All egg bands were reared in the Insect Production and Quarantine Facility (IPQL) at the GLFC. Rearing conditions were constant at +27ºC, 55% R.H. and 16:8h L:D based on the IPQL rearing protocol. Host information was recorded for each specimen and egg band (Appendix 2.1, 2.2). Since females deposit a single egg mass onto a selected host (McClure *et al.* 2010) and the

gregarious larvae remain on their natal host until the fourth instar (Batzer *et al.* 1995), 22 to 45 days depending on environmental conditions (Witter *et al.* 1972), we were able to record the primary host for each collection. Specimens were collected predominantly across eastern Canada, but were also sampled from Lac la Biche, Alberta and Saskatoon, Saskatchewan (Figure 2.1; see Appendix 2.3 for simplified flowchart of laboratory methods). Egg bands and larvae were reared to their 3rd - 5th instar on their original host, while also subsampling and preserving 1st and 2nd instars if colony collapse seemed to be occurring, and then stored in 100% ethanol and frozen at -20°C. We dissected the head capsule and upper thorax from each larva and removed the digestive tract from the thorax to eliminate plant and microbial contamination. We also continued to rear a subset of larvae from each population from a single egg mass, to the adult stage for DNA extraction. We froze each at -20°C and dissected thorax tissue from each specimen for extraction. We extracted genomic DNA from each specimen (n=181; Table 2.1) using the Qiagen DNeasy Blood and Tissue Extraction kit (QIAGEN, Hilden, Germany;) with modified manufacturer's specifications. We removed up to 25 mg of tissue into a sterile 1.5 mL microcentrifuge tube, and added 180 μL Buffer ATL, then ground with a pestle. Then we added 20 μL proteinase K, vortexed and centrifuged briefly before incubating at 56°C overnight. The next day, 4 μL ribonuclease (RNase) was added and vortexed, incubated for 2 minutes at room temperature, then vortexed for 15 seconds. 200 μL Buffer AL was added and vortexed, then incubated at 70°C for 10 minutes. 200 μL EtOH (96-100%) was added, vortexed, transferred to a spin column, and centrifuged for 1 minute at 8,000 rpm. 500 μL Buffer AW1 was added, centrifuged for an additional minute, and with the flow through discarded. 500 μL Buffer AW2 was added and centrifuged at 14,000 rpm for 3 minutes before adding 25 μL Buffer EB. It was then incubated for 1 minute at room temperature and centrifuged for 1 minute. An additional 25

μL Buffer EB was added, incubated for 1 minute at room temperature, and centrifuged for 1 minute. All DNA extractions were then purified using ethanol precipitation and resuspended in Millipore water and stored at -20°C. The quality and quantity of each sample was calculated using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) and Qubit fluorometer (Thermo Fisher Scientific), respectively.

2.3.2 Mitochondrial phylogeography

We amplified the DNA barcoding fragment of cytochrome c oxidase I (COI) for 130 specimens to permit direct comparisons with previously published data (Lait and Hebert 2018). The mtDNA region of 658 bp was amplified using the LepF and LepR primers based on a modified protocol from Hebert *et al.* (2003). Each polymerase chain reaction (PCR) contained 9.76 μL of dH2O, 2 μL of 10x buffer, 2 μL of MgCl2, 0.4 μL of deoxynucleoside triphosphate (dNTPs), 0.4 μL of LepF primer, 0.4 μL of LepR primer, 0.04 μL of TopTaq DNA polymerase (QIAGEN, Hilden, Germany), and 5 μL of template DNA. The PCR thermal cycle consisted of one cycle at 94°C for 2 minutes; 35 cycles of 94°C for 2 minutes, 45°C for 30 seconds, 72°C for 2 minutes, and a final cycle of 72°C for 5 minutes. After completion, the thermocycler then held at 4°C. Afterwards, all samples were exonuclease-shrimp alkaline phosphatase (EXOSAP) purified with 0.02 μL of exonuclease (EXO), 0.2 μL of shrimp alkaline phosphatase (SAP) and 1.78 μL of dH2O in addition to 5 μL of PCR product in a thermocycler for one cycle at 37°C for 25 minutes, and 80°C for 15 minutes. MtDNA samples were submitted for Sanger sequencing to the Molecular Biology Service Unit (MBSU) at the University of Alberta, on an Applied Biosystems 3730. All sequences were aligned, consensus sequences generated, and quality checked by eye using Geneious Prime v.2019.2.3 (https://www.geneious.com/). We combined

our data (Appendix 2.4) with an additional 139 sequences previously published by Lait and Hebert (2018; Appendix 2.5), obtained from the Barcode of Life Data System (BOLD; http://www.boldsystems.org). The combined data set was aligned using MAFFT online v.7 (Katoh *et al.* 2017; https://mafft.cbrc.jp/alignment/server/) (default settings) and trimmed to 658 bp in Mesquite v.3.6 (Maddison and Maddison 2018; https://www.mesquiteproject.org/). We constructed Templeton, Crandall and Sing (TCS; also known as statistical parsimony) haplotype networks (Templeton *et al.* 1995; Clement *et al.* 2002) to assess phylogenetic relationships among mtDNA haplotypes. Haplotype networks were created in PopART version 1.7 (Bandelt *et al.* 1999, Leigh and Bryant 2015).

2.3.3 Double digest restriction site sequencing

We submitted 181 specimens to the Molecular Biology Service Unit (MBSU; University of Alberta) for double digest restriction-site associated DNA sequencing (ddRAD). The Peterson *et al.* (2012) protocol was used to prepare ddRAD libraries using *PstI-MspI* restriction enzymes on the Illumina NextSeq 500 platform in one run, without the size selection step.

2.3.4 SNP filtering

Raw reads were processed on SHARCNET (www.sharcnet.ca; Compute Canada). DdRAD data was demultiplexed using STACKS version 2.3e (Catchen *et al.* 2013), with the following parameters: m=3, M=2, n=1; a single, random SNP per 100 bp to remove linkage; and removed the 8 bp index sequence. CUTADAPT version 2.5 (Martin 2011) was used to trim 5 bp from the 5' end of each read to prevent erroneous SNP calling. We retained reads with phred scores >30 and those that passed Ilumina's internal quality control. The *de novo* was built using

loci with a minimum read depth of five and using custom perl wrappers. Using the STACKS pipeline, we used *ref_map* to assemble it, before calling SNPs. We then used the *populations* program in STACKS by assigning all individuals to a single population. We excluded individuals with > 80% missing data and loci in < 95% of the individuals. SNPs were further filtered using VCFtools version 0.1.14 (Danecek *et al.* 2011). We retained SNPs that were biallelic, had a minor allele frequency of 0.05 and with a maximum missing rate of 5% across all individuals.

2.3.5 Detecting siblings

Preliminary analysis of our data revealed unexpected groupings that lacked a host or large-scale geographic explanation (Appendix 2.6). On closer examination, specific localities had tight clusters of individuals (Appendix 2.6). Given the colonial behaviour of *M. disstria*, we screened our data for potential siblings that might have been unintentionally sampled. Using SNPRelate version 1.2.1 (https://github.com/zhengxwen/SNPRelate; Zheng *et al.* 2012) we assessed the kinship of individuals collected at the same locality. Related individuals were defined as pairs from the same locality with a kinship coefficient of ≥0.2. If we detected a potential sibling, then we retained only one specimen from each set of related individuals. For diploid organisms, 0.25 kinship coefficient is the expected value for full siblings (Manichaikul *et al.* 2010), but we adjusted the threshold to 0.20 for a more stringent cut off, as pairings below 0.25 were also identified as siblings.

2.3.6 SNP population analyses

We explored population structuring among the *M. disstria* populations at three different scales. First, we examined continental population structuring by analysing all *M. disstria* samples (Ontario, Quebec, Saskatchewan and Alberta) regardless of host. Second, we compared the population structure of *M. disstria* from four tree species (sugar maple, trembling aspen, red oak, white birch) in eastern Canada to assess whether these populations formed distinct genetic clusters based on their larval host. Finally, we assessed regional population structuring from *M. disstria* specimens collected off aspen trees in eastern Canada, the only host with samples collected throughout the eastern range. This was to determine whether *M. disstria* populations aligned with the three forest ecozones within the eastern study range: boreal forest, temperate mixed forest, and temperate deciduous forest (Baldwin *et al.* 2018).

We quantified population structuring using two multivariate analyses, principal component analysis (PCA) and discriminant analysis of principal components (DAPC; Jombart *et al.* 2010) and a Bayesian clustering approach, Structure (Pritchard *et al.* 2000). A PCA is a tool for exploratory data analysis that reduces the complexity of large data sets and allows clusters to be visualized as scatterplots depicting genetic relatedness between populations in successive dimensions (Jolliffe and Cadima 2016). A DAPC, on the other hand, uses a PCA as a prior step to transform the data and then enables identification of genetic divergences among *a priori* groups, maximizing the between group variation while minimizing the variation within groups (Jombart *et al.* 2010). Structure is a Bayesian clustering program that applies *Markov Chain Monte Carlo* (MCMC) estimations to determine the number of subpopulations that exist in a data set. The likelihood of relatedness among individuals is randomly assigned, and through many iterations of randomly assigning individuals to a group, the membership probabilities of

individuals in a population can be determined (Porras-Hurtado *et al.* 2013). Both PCA and DAPC analyses were conducted using v2.1.2 of the adegenet package (Jombart 2008) and visualized using ggplot2 3.2.1 (Wickham 2016). We used Structure v.2.3.4 (Pritchard *et al.* 2000) for our Bayesian analyses. Prior to the Structure analyses, we used PGDSpider v.2.1.1.5 to convert files from VCF to GENEPOP and STR formats (Lischer and Excoffier 2012) (http://www.cmpg.unibe.ch/software/PGDSpider/). Using the *LocPrior* option, we tested values of K (the number of subpopulations within the total population) from 1 to 5, with 10 replicated per K, each run with a burn-in period of 250,000 and 1,000,000 MCMC generations. Runs were visualized in CLUMPAK (Kopelman *et al.* 2015; http://clumpak.tau.ac.il/index.html). We identified the optimal value of K using the Pritchard (Pritchard *et al.* 2002) and Evanno methods (Evanno *et al.* 2005) and then using CLUMPAK. The Pritchard method is the original Structure program method and assumes there are an unknown number of subpopulations (K). Individuals are randomly assigned to groups, with the potential number of subpopulations supplied as input (Janes *et al.* 2017). The program runs many iterations for each successive K value, and plots an average of all estimated likelihood of K values. The K value selected is where the data plots plateau. The Evanno method was developed later, to complement the estimated likelihood of K output by the Pritchard method. The Evanno method uses a second-order rate of change of the likelihood of K, with the maximum value output being the true number of populations in a dataset (Janes *et al.* 2017). It detects the uppermost value of K in a population. Therefore, with the Pritchard method providing a more accurate estimate, the Evanno method is usually used to identify K most easily. Structure admixture plots were visualized in R Studio v.1.2.5001, using ggplot2 (Wickham 2016; https://ggplot2.tidyverse.org).

2.4 Results

2.4.1 Mitochondrial phylogeography

We sequenced 658 bp of the COI gene, corresponding to the DNA barcode region, for 130 *M. disstria* specimens (Appendix 2.4). We combined these with an additional 139 sequences from Lait and Hebert (2018; Appendix 2.5). We did not observe any mtDNA structuring by host plant in samples with host data (Figure 2.2). However, we found complex geographic structuring within the combined data set. MtDNA haplotypes formed distinct genetic clusters, but these were not restricted to a single geographic region; several clusters show haplotype groupings from distant provinces (AB/SK + ON/QC; ON/QC + Atlantic + E.US + S.US; BC + AB/ SK; ON/QC + Atlantic; Figure 2.2).

2.4.2 SNP population analyses

2.4.2.1 Continental population structure

Our continental data set contained 130 individuals from across Canada (Table 2.2; 202,579 raw loci, average read depth =23.6, and minimum read depth =10.54). We resolved three distinct clusters in our multivariate analyses (Figure 2.3), and we inferred an optimal K of K=3 (Pritchard) or K=2 (Evanno) in our Structure analyses (Figure 2.3C; Appendix 2.7) (Pritchard *et al.* 2000; Evanno, *et al.* 2005). Central Canada (n=18 specimens; Alberta, Saskatchewan) formed a distinct grouping, although one specimen from Saskatchewan was admixed from distant eastern Canadian populations, unlike the other central specimens (Figure 2.3A, B). Specimens collected from northern Ontario grouped separately from the other eastern Canadian specimens.

2.4.2.2 Larval host races

Our larval host race data set contained 112 individuals from eastern Canada (Table 2.2; 6 red oak, 27 sugar maple, 75 trembling aspen and 4 white birch). We found no evidence of host races among these populations using multivariate analyses (Figure 2.4A, B), rather these analyses indicate that the specimens form a single population. We did, however, find support for two populations in our Structure analysis. Using Structure (Figure 2.4C), we determined that K = 2 was optimal (Appendix 2.8) based on of Pritchard *et al.* (2000) and Evanno *et al.* (2005).

2.4.2.3 Regional population structure

To assess regional population structure, we examined genetic variation among aspen feeding *M. disstria* in eastern Canada (Table 2.2). By focusing on this single host, we were able to explore structuring throughout Ontario and Quebec and remove potential host effects. We observed regional spatial structuring among aspen feeding *M. disstria* in all three analyses (Figure 2.5, Appendix 2.9). *M. disstria* separated into two distinct clusters, which corresponded with the boreal forest ecozone and the temperate ecozones (mixed and deciduous). Specimens were consistently more genetically similar along east-west axes running from Quebec to western Ontario than they were along shorter axes running from southern to northern Ontario. This corresponded to forest ecozones, with eastern boreal specimens forming a separate cluster from specimens collected in the temperate mixed and temperate deciduous ecozone.

2.5 Discussion

Many ecological factors impact intraspecific variation, with different evolutionary processes working at different spatial scales to structure genetic variation within species. We

found no evidence of genetic structure between populations on different larval hosts in eastern Canada. However, we did detect regional and continental population structuring, demonstrating that multiple evolutionary processes are influencing the genetic structure of *M. disstria* throughout Canada.

2.5.1 Continental population structure

Population structure in *M. disstria* was dominated by continental differentiation. Our ddRAD data demonstrated a strong east-west separation among *M. disstria* populations (Figure 2.3). One specimen from Saskatoon, Saskatchewan, was found to be admixed with others in the eastern population (Figure 2.3A, B). Our results indicated two major groups, eastern and central, with enough variation within eastern specimens to drive separation between southern Ontario and northern Ontario populations (Figure 2.3). The east-west divide in *M. disstria* is not only observed in its genetic variation. Outbreak dynamics appear to differ between eastern and western populations. Eastern populations experience outbreaks approximately every 10 years (Cooke and Lorenzetti 2006), while western populations are asynchronous (Cooke and Roland 2018). Other organisms have also demonstrated an east-west separation. Spruce budworm (*Choristoneura fumiferana* Clemens) have also expressed similar patterns, which has been speculated to be a result of ice-age refugia and populations spreading along expanding host ranges (Lumley *et al.* 2020). Another organism, the red-headed fall webworm (*Hyphantria cunea* Drury) also demonstrates similar patterns in population structuring, with western populations being distinct from eastern, with possible expansion emanating from a west-central location separated by changing river corridors (Vidal *et al.* 2019). Often these population divergences are attributed to expansion range from ancient refugia, that was ultimately limited by a geological

event facilitating a permanent barrier to gene flow. In addition to the separation detected in SNPs, COI data can also be a useful tool in detecting population structuring on a large geographic scale.

Mitochondrial phylogeography showed complex geographic structuring, in contrast to the clear division between eastern and western populations using the genome-wide SNP markers. Phylogeographic structure primarily divided along eastern, central and western populations (Figure 2.2), similar to previously published results (Lait and Hebert 2018). These observations, however, contradict our genomic results, which suggest pronounced geographic structuring. This discordance is possibly due to our genomic data having greater resolution with >9000 SNPs compared with <100 variable sites across mtDNA, and therefore detecting global genetic variation patterns across the genome. In addition, mtDNA is haploid and maternally inherited, having an effective population 1/4th the size of nuclear DNA (Toews and Brelsford 2012). This means mtDNA should exhibit differences before nuclear DNA. However, the differences noticed might be the result of a complex evolutionary history, with males from distant populations introducing nuclear genes, that resulted in higher nuclear DNA variation (Bensch *et al.* 2006). Males are primarily responsible for dispersal, they have been documented to fly up to approximately 3 km (Evenden *et al.* 2015), and with wind current assistance, travel hundreds of kilometers (Fitzgerald 1995). These dispersal events originating from multiple northern ice-age glacial refugia, could have contributed to the observed patterns, with different regional populations contributing to higher genetic diversity due to gene flow during receding glaciers (van Els *et al.* 2012). Females on the other hand have limited dispersal, only performing a brief pre-ovipositional flight, would have required a greater length of time for mtDNA gene flow to migrate from ice-age glacial refugia. Therefore, as males are the main dispersers, nuclear DNA

movement would have been primarily through these individuals. However, relying on a single marker alone could lead to misrepresentation of a species' population structuring (Kodandaramaiah *et al.* 2013). Through interspecific gene flow, foreign alleles are introduced into a species' gene pool, which impact mtDNA to a greater degree due to a lack of recombination (Funk and Omland 2003). On the other hand, genome-wide data are less sensitive to introgression than single markers, as single markers are less constrained by linkage (Funk and Omland 2003). In addition, Lepidoptera have previously been documented for mtDNA introgression (Sperling 1993; Kodandaramaiah *et al.* 2013). For this reason, we incorporated ddRAD data to survey variation across the genome, and either serve as additional support to corroborate the mtDNA observations or provide the basis for an alternative explanation.

The large-scale break between eastern and central populations could be attributed to past geological events. During the last glacial period, most of Canada was covered in large ice sheets. These glaciers were responsible for separating species into isolated refugia across North America, leading to genetically distinct populations (Shafer *et al.* 2010). As the glaciers retreated, the glacial lake Lake Agassiz formed across the wider present-day Manitoba and surrounding lands (Murton *et al.* 2010). Similar to possible spruce budworm recolonization events (Lumley *et al.* 2020), *M. disstria* populations may have also spread back into their previous historic ranges.

It is worth noting that our specimen sampling was limited. We only sampled a small part of the overall range of *M. disstria*. Without sufficient population size (localities and/or individuals), rare alleles may be missed in the population (Hale *et al.* 2012). Thus, sampling size should be sufficient for allele frequencies to represent the true population. Historically, 30 individuals per population have been used for individual gene markers or microsatellites to

obtain a clear sense of genetic variation at a site (Hale *et al.* 2012). However, when large numbers of SNPs are used (as in our study), six to eight individuals can be sufficient to quantify within-population variation (Nazareno *et al.* 2017). As few as two specimens can even be enough, providing a sufficient number of SNPs are used (Nazareno *et al.* 2017). We aimed for 10-20 individuals per population, but outbreak intensity, the variable nature of the specimen collection, and high disease and mortality during rearing limited the number of specimens available for genotyping. Therefore, large errors can occur when allele frequencies are small, leading to an inaccurate sample that misrepresents the population. In addition to large scale geography testing, we also examined regional host race formation, particularly among four hosts of eastern Canada.

2.5.2 Larval host races

M. disstria feed on a range of host plants. Functional trait differences between larval hosts have been detected by multiple authors, which led us to hypothesize that *M. disstria* was composed of distinct host races. However, we found no genomic evidence to support host race specialization in *M. disstria* within eastern Canada (Figure 2.4). Larvae on different hosts, particularly those from the same location (n=4), were similar and not distinct (Figure 2.4A, B). While our multivariate analyses clearly showed overlapping genomic variation, our Structure analysis indicated that two distinct populations were optimal using both the Pritchard and Evanno methods (Pritchard *et al.* 2000; Evanno, *et al.* 2005; Appendix 2.8). This two-population result may be due to the K=2 conundrum recently described by Janes *et al.* (2017), which describes the tendency of Structure analyses to report K=2 even when there is no structure in the data. Oversplitting of single populations is a common problem (Janes *et al.* 2017). In our case,

other approaches (i.e. multivariate analyses) were used to clarify population structuring (Janes *et al.* 2017). While this explains our results from the Evanno method, the Pritchard method also indicated K=2. For the Pritchard method, K should be also selected with knowledge of the system's biology (Gilbert 2016). In addition, uneven sampling across populations can affect the results of Structure, which can result in underrepresented populations being combined even though they lack genetic relatedness (Meirmans 2018).

Functional differences exist between *M. disstria* on different host species, despite a lack of population structure. It has been observed that larvae preferentially feed on aspen, rather than maple foliage (Panzuto *et al.* 2001). In addition, aspen fed larvae grew faster than those on sugar maple, and larvae on aspen had larger pupae than those on maple (Trudeau *et al.* 2010). Among the main hosts of *M. disstria*, trembling aspen contains twice the amount of sugars compared to sugar maple, while also having higher levels of carbohydrates and tremulacin then sugar maple foliage (Panzuto *et al.* 2001). These have been linked to increased insect performance, while sugar maples possess high levels of tannins, which are detrimental to protein digestion in many insects. Since we failed to detect host-based population structure, these differences may be attributed to phenotypic plasticity (Gorur *et al.* 2005). Other species have demonstrated similar patterns, where geography overrides host effects. Fall webworm populations are structured due to geographic distance rather than host (Vidal *et al.* 2019). In another species, the oriental fruit moth (*Grapholia molesta* Busck), geography was a stronger influence on structuring populations than the fruit orchard where larvae originated (Silva-Brandao *et al.* 2015). These trends could come in the form of limited availability of hosts, where synchrony between emergence and budbreak is promoted, due to the detrimental effects of older foliage causing a negative response in larvae (Gray and Ostaff 2012). It also remains unclear if host preference is a result of the host

on which they matured, and whether they exhibit diversified risk spreading (Gray and Ostaff 2012). Alternatively, since *M. disstria* is a generalist insect, female adults may run into an information processing constraint (Bernays and Wcislo 1994). During low density, females tend to be at high canopy levels, while during low density, females will mate in in the forest understory, which allows larvae to find other food sources during outbreaks (Miller 2006). As females will begin mating hours after emerging from their pupae, they will tend to mate and oviposit relatively close to where they emerged (Evenden *et al.* 2015). This occurs because a variety of chemical stimuli associated with a wider range of hosts can lead to alternative host selection (Egan and Funk 2006). As no host race formation was detected, we examined regional forest ecozone structuring as well in eastern Canada.

2.5.3 Regional population structure

Population structure and genetic diversity can be structured at different spatial scales. While we did not detect differences among hosts at the same location, we found evidence of regional structuring among *M. disstria* populations in eastern Canada (Figure 2.5) that aligned with three forest ecozones: temperate deciduous forest, boreal forest and temperate mixed forest (Baldwin *et al.* 2018). Boreal and deciduous forests are divided by a transitional zone between the maple dominated deciduous forests of southern Ontario and Quebec, and the northern coniferous forests. There, red maple, eastern white pine, trembling aspen, white and red ash trees are mixed, with no major barriers to gene flow present. However, *M. disstria* populations show evidence of a non-gradual genomic cline, with rapid north-south change. This latitudinal shift in genomic diversity occurs more rapidly than population changes from east to west. Without a major barrier to keep the two populations apart, gene flow should prevent the development or

maintenance of genomic population structure. However, this is not the case. Genetic structure does exist in *M. disstria* populations between northern and southern populations.

Regional variation can often be explained through local environmental adaptations. Whereas warmer climates promote smaller body size, smaller eggs, and greater numbers of eggs, northern climates tend to favour the opposite, promoting greater emphasis on survivability of the individual, rather then maximizing larger brood numbers. This pattern has been observed in *M. disstria* populations, with northern populations having larger, fewer eggs, and larger neonates than southern populations (Parry *et al.* 2001). In *M. disstria,* females have limited flight dispersal, typically for oviposition after mating (Evenden *et al.* 2015), and mtDNA, being maternally inherited, can only be passed on through mother to offspring. With low dispersal, mtDNA would only be able to move through limited female flight dispersal, or during outbreak dynamics, where female larvae migrate to other host trees. In other species, this north-south variation is also linked to cold survival, such as in spruce budworm where northern population egg sizes were nearly double in mass and had fewer eggs compared to southern populations (Harvey 1983). These differences between northern and southern populations could explain the variation between *M. disstria* populations located in the three ecozones. Host availability would change among the ecozones, as host type, distribution and density would change across ecozones. Within an ecozone, forest stands become dense, often consisting of the primary regional host, while in transition zones, host tree distribution becomes more interspersed. In addition, natural enemies across various ecozones can play a role in shaping population structuring, by driving cyclic population dynamics of some insect populations while regulating others (Klemola *et al.* 2002). Predator populations are affected by local habitats, and the local predator community may play a role in furthering regional structuring, as bird species and

parasitoids alike will respond differently among different forest types (Nixon and Roland 2012). Even with small geographic distances separating populations, sufficient local adaptations have formed in the regional populations to enable survival in the different habitats and preventing gene flow from neighbouring groups.

2.6 Conclusion

Intraspecific variation and population structuring are products of numerous interacting evolutionary forces. It was necessary to examine these evolutionary processes collectively to adequately explain the population structuring we observed in *M. disstria*. We had initially speculated that the functional differences observed in *M. disstria* were due to undetected host races. This was not the case. Instead it was geography, both regional and continental, that had a greater influence on the population structure of *M. disstria*. While we did not explain the underlying fitness differences previously observed among larvae on different host plants, our study highlights the importance of disentangling multiple sources of variation. Populations were genetically differentiated between ecozones, with even greater differences based on geography rather than host. These regional differences may underlie regional outbreak dynamics and can serve as a foundation for further exploration of *M. disstria* population structure. More intensive sampling efforts to fill in the transition zone between eastern and central samples are needed to clarify the population boundaries within this species. Identifying the factors associated with population structuring within a species can provide a roadmap for further examining species life histories and the mechanisms of intraspecific diversification. In addition, better management decisions could be supported, since central and eastern populations show numerous variations in life history and genetic traits, and control of outbreaks in eastern Canada should be examined

through a regional context, as the population structuring there indicates that not all *M. disstria* populations are the same.

2.7 Data accessibility

COI sequences are available on Genbank:

https://www.ncbi.nlm.nih.gov/nuccore/?term=MT791498:MT791627[accn]

Custom STACKS perl wrappers are at https://github.com/muirheadk/GBS_analysis_pipeline

DNA voucher specimens will be stored at Natural Resources Canada, Sault Ste Marie, Ontario.

Table 2.1 Summary of *M. disstria* collection data: province, locality, host plant, coordinates, egg numbers if available, date collected, and number of individuals sampled. Pre *n* is number of specimens initially submitted for sequencing. Post *n* is number of specimens that remain after siblings and read quality were filtered.

Prov.	Locality	Host	Lat	Long	Collected	Egg Band#	Pre *n*	Post *n*
AB	Lac la Biche	Aspen	54.421	-111.57	02-May-18	18	18	14
SK	Saskatoon	Aspen	n/a	n/a	May-18	n/a	6	4
QC	Lac Duparquet	Aspen	48.5	-79.2	Jun-18	n/a	6	1
QC	Montebello	Maple	45.7	-74.8	Jun-18	n/a	4	4
ON	Bancroft	Maple	44.887	-77.747	16-May-18	17	3	2
ON	Constance Lake	Aspen	49.814	-84.189	May-18	10	10	1
ON	Elliot Lake	Aspen	46.339	-82.542	09-May-18	19	9	5
ON	Foots Bay	Aspen	45.166	-79.76	May-18	n/a	7	6
ON	Hearst	Aspen	49.731	-83.913	09-May-18	23	7	5
ON	Kapuskasing	Aspen	49.644	-82.302	08-May-18	13	8	5
ON	Kenora	Aspen	49.759	-94.476	11-May-18	11	6	2
ON	Killarney	Birch	46.011	-81.401	May-18	n/a	4	4
ON	Lanark	Maple	44.82	-76.45	05-Jun-18	4	4	4
ON	Lanark	Aspen	44.82	-76.45	05-Jun-18	6	3	3
ON	Latchford	Aspen	47.333	-79.81	May-18	23	7	4

ON	Little Current	Aspen	45.879	-81.899	May-18	3	4	2
ON	Marten River	Aspen	46.668	-79.728	May-18	1	5	5
ON	Nairn	Aspen	46.302	-81.678	08-May-18	36	10	6
ON	Oak Shores	Oak	44.588	-78.428	29-May-18	n/a	2	2
ON	Ottawa	Maple	n/a	n/a	May-18	6	1	1
ON	Ottawa	Aspen	n/a	n/a	May-18	5	1	1
ON	Parry Sound	Maple	45.378	-80.044	09-May-18	15	3	2
ON	Rabbit Lake	Aspen	46.93	-79.726	11-May-18	24	9	7
ON	Sault Ste Marie	Oak	46.508	-84.302	07-Jun-18	7	4	4
ON	Sault Ste Marie	Maple	46.508	-84.302	07-Jun-18	11	7	7
ON	Sault Ste Marie	Aspen	46.508	-84.302	07-Jun-18	19	18	15
ON	St Joseph Island	Maple	46.192	-84.042	03-Jun-18	10	7	6
ON	St Joseph Island	Aspen	46.192	-84.042	03-Jun-18	12	7	7
ON	Wharncliffe	Maple	46.538	-83.538	May-18	1	1	1

Table 2.2 Summary of the three SNP data sets: Continental, Larval host and Regional. Missing data can result in information loss at higher percentages, affecting the degree of accuracy.

Dataset Name	Number of Samples	Number of Localities	General Location	Hosts	SNPs	Missing Data
Continental	130	24	AB, SK, ON, QC	Aspen, Birch, Maple, Oak	9,427	1.55%
Larval host	112	22	ON, QC	Aspen, Birch, Maple, Oak	9,285	1.51%
Regional	75	16	ON, QC	Aspen	8,865	1.45%

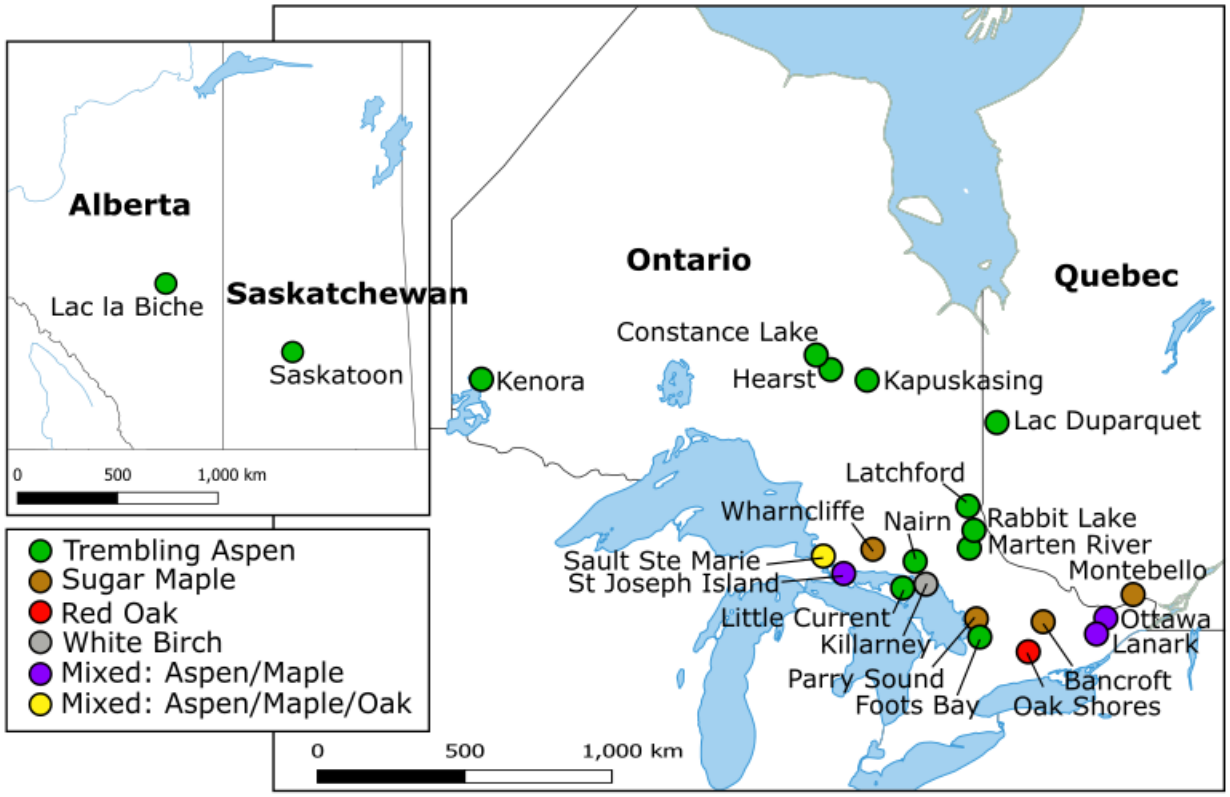


Figure 2.1 *M. disstria* collection sites across Canada. Colours indicate larval host. Specimens were collected by various collaborators and shipped to the Great Lakes Forestry Centre in Sault Ste Marie for rearing. Map prepared in QGIS version 3.12 and base map obtained from https://www.naturalearthdata.com/downloads/.

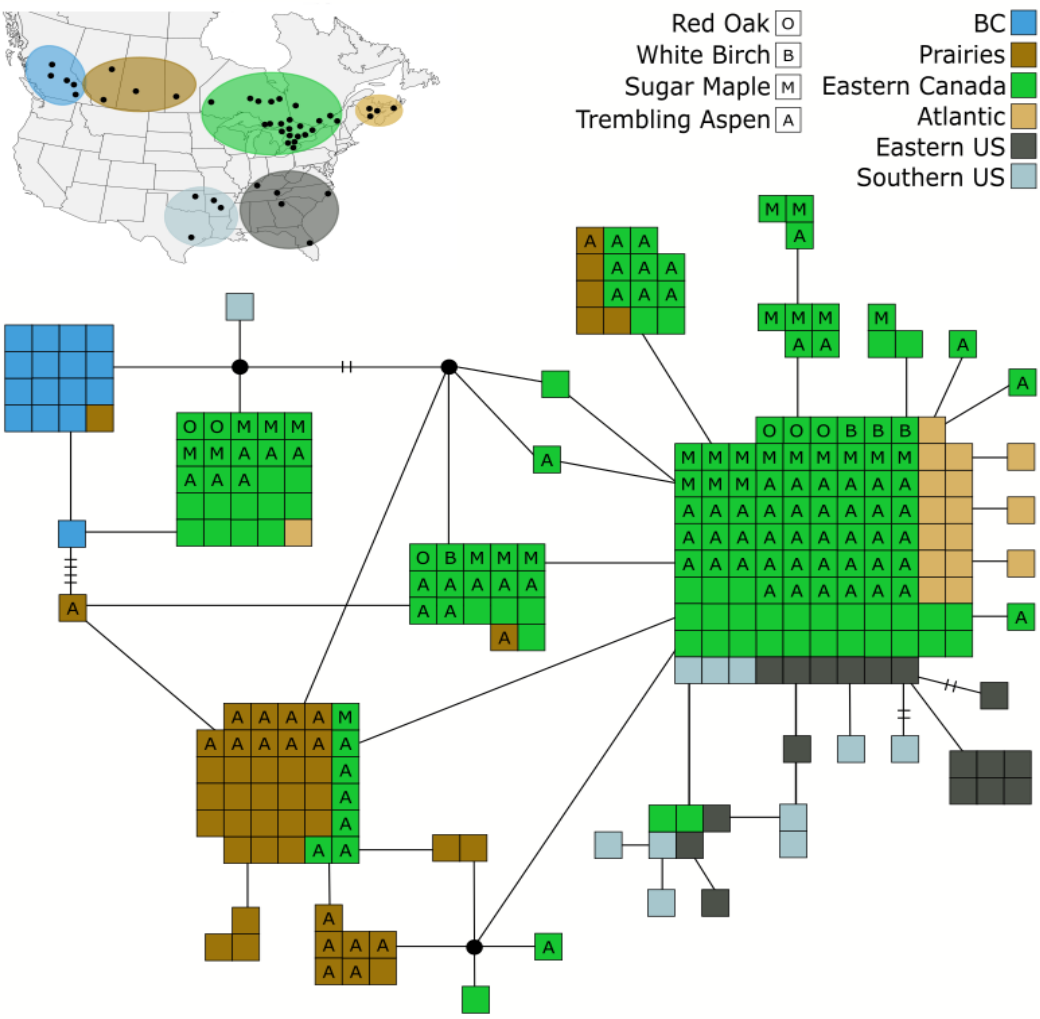


Figure 2.2 Templeton, Crandall and Sing (TCS) Haplotype network of 269 *M. disstria* mitochondrial DNA (mtDNA) cytochrome oxidase subunit 1 (COI) sequences from across Canada and the southeastern United States. A total of 130 *M. disstria* sequences were from specimens collected for this study, while 139 sequences were downloaded from BOLD (Lait and Hebert 2018). Each square represents a single individual. Letters in boxes indicate hosts, in addition to being specimens collected in this study. BOLD specimens lack host information. Black dots represent unsampled, potential haplotypes. Small black lines are mutational steps.

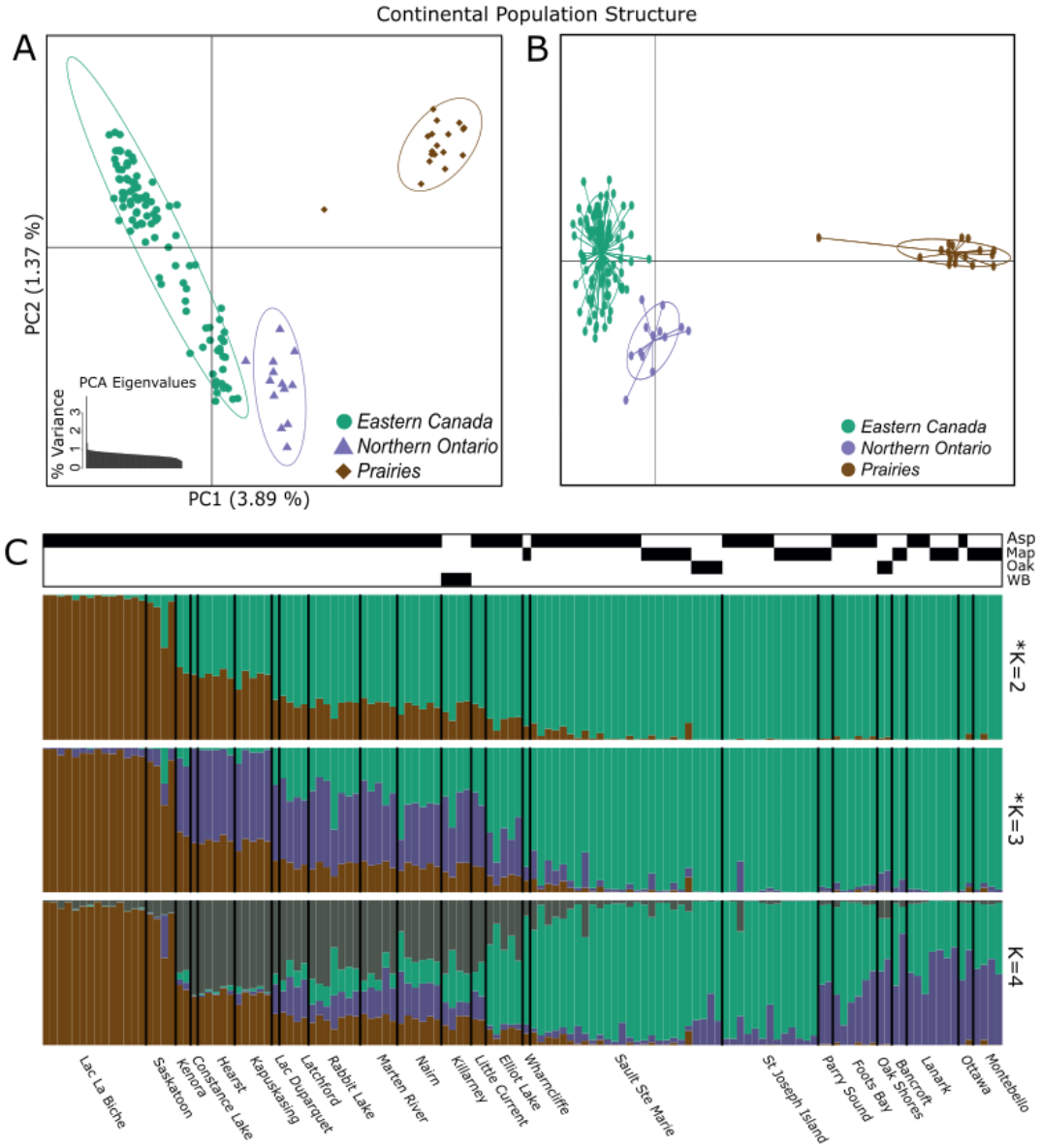


Figure 2.3 (A) Principal component analysis (PCA), (B) discriminant analysis of principal components (DAPC) and (C) Structure analysis of 130 *M. disstria* specimens from Alberta to Quebec. Black bars above the Structure plot (C) indicate locality specimen host. (C) Structure analysis assigns individuals, based on genomic similarity, into populations. Structure analysis indicated K=3 (Pritchard) and K=2 (Evanno) as optimal.

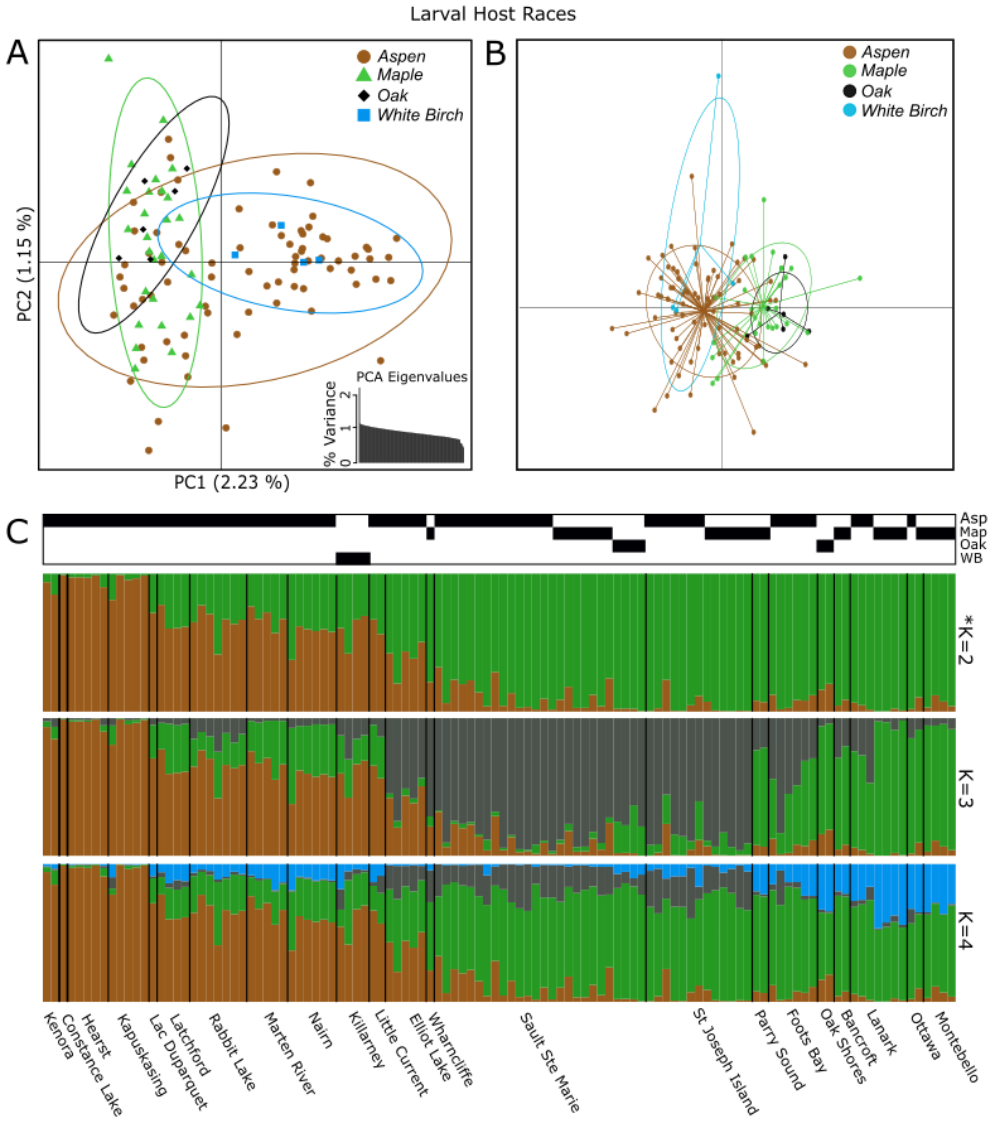


Figure 2.4 (A) Principal component analysis (PCA), (B) discriminant analysis of principal components (DAPC) and (C) Structure analysis of 112 *M. disstria* specimens from Ontario and Quebec. (A) Ellipses are 95% confidence intervals. Black bars above the Structure plot (C) indicate larval host of each locality specimen. Structure analysis assigns individuals into populations, based on genomic similarity. K=2 was optimal, both with the Pritchard and the Evanno methods.

Figure 2.5 (A) Principal component analysis (PCA), (B) discriminant analysis of principal components (DAPC) and (C) Structure analysis of 75 *M. disstria* specimens from Ontario and Quebec with aspen as larval hosts. Colours and groupings are based on Baldwin *et al.* 2018. (A) Ellipses are 95% confidence intervals. Black bars above the Structure plot (C) indicate locality specimen host. (C) Structure analysis assigns individuals, based on genomic similarity, into populations. K=2 was optimal, both with the Pritchard and the Evanno methods.

2.8 Literature Cited

Baldwin K, Allen L, Basquill S, Champan K, Downing D, Flyyn N, MacKenzie W, Major M, Meades W, Meidinger D, Morneau C, Saucier JP, Thorpe J, Uhlig P. 2018. Vegetation zones of Canada: a biogeoclimatic perspective. [Map] Scale 1:5,000,000. Natural Resources Canada, Canadian Forest Service. Sault Ste Marie, ON. http://www.cnvc-cnvc.ca

Bandelt H, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*. 16, 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Batzer HO, Martin MP, Mattson WJ, Miller WE. 1995. The forest tent caterpillar in aspen stands: distribution and density estimation of four life stages in four vegetation strata. *Forest Science.* 41, 99-121. https://doi.org/10.1093/forestscience/41.1.99

Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology*. 15, 161-171. https://doi.org/10.1111/j.1365-294X.2005.02766.x

Bernays EA, Wcislo WT. 1994. Sensory capabilities, information processing, and resources specialization. *The Quarterly Review of Biology.* 69, 187-204. https://doi.org/10.1093/aesa/65.1.25

Blair CP, Abrahamson WG, Jackman JA, Tyrrell L. 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. *Evolution*. 59, 304-316. https://doi.org/10.1111/j.0014-3820.2005.tb00991.x

Catchen J, Hohenlohe P, Bassham S, Amores A, Cresko W. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*. 22, 3124-3140. https://doi.org/10.1111/mec.12354

Charbonneau D, Lorenzetti F, Doyon F, Mauffette Y. 2012. The influence of stand and landscape characteristics on forest tent caterpillar (*Malacosoma disstria*) defoliation dynamics: the case of the 1999-2002 outbreak in northwestern Quebec. *Canadian Journal of Forest Research*. 42, 1827-1836. https://doi.org/10.1139/x2012-126

Clement M, Snell Q, Walke P, Posada D, Crandall K. 2002. TCS: estimating gene genealogies. *Proceedings of the 16th International Parallel and Distributed Processing Symposium*. 184-191. https://doi.org/10.1109/IPDPS.2002.1016585

Cooke BJ, Lorenzetti F. 2006. The dynamics of forest tent caterpillar outbreaks in Quebec, Canada. *Forest Ecology and Management*. 226, 110-121. http://doi.org/10.1016/j.foreco.2006.01.034

Cooke BJ, Roland J. 2018. Early 20th century climate-driven shift in the dynamics of forest tent caterpillar outbreaks. *American Journal of Climate Change*. 7, 253-270. http://doi.org/10.4236/ajcc.2018.72017

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker R, Lunter G, Marth G, Sherry ST, McVean G, Durbin R. 2011. 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*. https://vcftools.github.io/license.html

Downey MH, Nice CC. 2013. A role for both ecology and geography as mechanisms of genetic differentiation in specialized butterflies. *Evolutionary Ecology*. 27, 565-578. https://doi.org/10.1007/s10682-012-9626-7

Dres M, Mallet J. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 357, 471-492. https://doi.org/10.1098/rstb.2002.1059

Egan SP, Funk DJ. 2006. Individual advantages to ecological specialization: insights on cognitive constraints from three conspecific taxa. *Proceedings of the Royal Society B: Biological Sciences*. 273, 843-848. https://doi.org/10.1098/rspb.2005.3382

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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Evenden ML, Whitehouse CM, Jones BC. 2015. Resource allocation to flight in an outbreaking forest defoliator *Malacosoma disstria. Physiological Ecology.* 44, 835-845. https://doi.org/10.1093/ee/nvv055

Ferrari J, Godfray HCJ, Faulconbridge AS, Prior K, Via S. 2006. Population differentiation and genetic variation in host choice among pea aphids from eight host plant genera. *Evolution*. 60, 1574-1584. https://doi.org/10.1554/06-024.1

Fitzgerald TD. 1995. *The Tent Caterpillars*. Ithica (NY): Cornell University Press. 303 p.

Fuller S, Schwarz M, Tierney S. 2005. Phylogenetics of the allodapine bee genus *Braunsapis*: Historical biogeography and long-range dispersal over water. *Journal of Biogeography*. 32, 2135-2144. https://doi.org/10.1111/j.1365-2699.2005.01354.x

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*. 34, 397-423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421

Gilbert KJ. 2016. Identifying the number of population clusters with Structure: problems and solutions. *Molecular Ecology Resources*. 16, 601-603. https://doi.org /10.1111/1755-0998.12521

Gorur G, Lomanco C, Mackenzie A. 2005. Phenotypic plasticity in host-plant specialisation in *Aphis fabae*. *Ecological Entomology*. 30, 657-664. https://doi.org/10.1111/j.0307-6946.2005.00742.x

Gosden TP, Stoks R, Svensson EI. 2011. Range limits, large-scale biogeographic variation, and localized evolutionary dynamics in a polymorphic damselfly. *Biological Journal of the Linnean Society*. 102, 775-785. https://doi.org/10.1111/j.1095-8312.2011.01619.x

Gray DR, Ostaff DP. 2012. Egg hatch of forest tent caterpillar (Lepidoptera: Lasiocampidae) on two preferred host species. *The Canadian Entomologist*. 144, 790-797. https://doi.org/10.4039/tce.2012.73

Hale ML, Burg TM, Steeves TE. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PLoS One*. 9, 1-10. https://doi.org/10.1371/journal.pone.0045170

Hardy NB, Otto SP. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society B: Biological Sciences*. 281, 1-10. https://doi.org/10.1098/rspb.2013.2960

Hartley S, Krushelnycky PD, Lester PJ. 2010. Integrating physiology, population dynamics and climate to make multi-scale predictions for the spread of an invasive insect: the Argentine ant at Haleakala National Park, Hawaii. *Ecography*. 33, 83-94. https://doi.org/10.1111/j.1600-0587.2009.06037.x

Hartmann H, Messier C. 2011. Interannual variation in competitive interactions from natural and anthropogenic disturbances in a temperate forest tree species: implications for ecological interpretation. *Forest Ecology and Management*. 261, 1936-1944. https://doi.org/10.1016/j.foreco.2011.02.018

Harvey GT. 1983. A geographical cline in egg weights in *Choristoneura fumiferana* Lepidoptera: Tortricidae) and its significance in population dynamics. *The Canadian Entomologist*. 115, 1103-1108. https://doi.org/10.4039/Ent1151103-9

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*. 270, 313-321. https://doi.org/10.1098/rspb.2002.2218

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL. 2017. The K=2 conundrum. *Molecular Ecology*. 26, 3594-3602. https://doi.org/10.1111/mec.14187

Janz N, Nylin S, Wahlberg N. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. BMC *Evolutionary Biology*. 6, 1-10. https://doi.org/10.1186/1471-2148-6-4

Johnson MTJ. 2008. Bottom-up effects of plant genotype on aphids, ants, and predators. *Ecology*. 89, 145-154. https://doi.org/10.1890/07-0395.1

Jolliffe IT, Cadima J. 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 374, 1-16. http://doi.org/10.1098/rsta.2015.0202

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24, 1403-1405. http://doi.org/10.1093/bioinformatics/btn129

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 11, 1-15. https://doi.org/10.1186/1471-2156-11-94

Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*. 20, 1160-1166. https://doi.org/10.1093/bib/bbx108

Kodandaramaiah U, Simonsen TJ, Bromilow S, Wahlberg N, Sperling F. 2013. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution*. 3, 5167-5176. https://doi.org/10.1002/ece3.886

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*. 15, 1179-1191. https://doi.org/10.1111/1755-0998.12387

Lait LA, Hebert PDN. 2018. Phylogeographic structure in three North American tent caterpillar species (Lepidoptera: Lasiocampidae): *Malacosoma americana*, *M. californicum*, and *M. disstria*. *PeerJ*. 6, 2-23. https://doi.org/10.7717/peerj.4479

Leigh JW, Bryant D. 2015. PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*. 6, 1110-1116. https://doi.org/10.1111/2041-210X.12410

Lischer HEL, Excoffier L. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*. 28, 298-299. https://doi.org/10.1093/bioinformatics/btr642

Lumley LM, Pouliot E, Laroche J, Boyle B, Brunet BMT, Levesque RC, Sperling FAH, Cusson M. 2020. Continent-wide population genomic structure and phylogeography of North America's most destructive conifer defoliator, the spruce budworm (*Choristoneura fumiferana*). *Ecology and Evolution*. 10, 914-927. https://doi.org/10.1002/ece3.5950

Maddison WP, Maddison DR. 2019. Mesquite: a modular system for evolutionary analysis. Version 3.61 http://www.mesquiteproject.org

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 26, 2867-2873. https://doi.org/10.1093/bioinformatics/btq559

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 17, 10-12. https://doi.org/10.14806/ej.17.1.200.

Matsubayashi KW, Kahono S, Katakura H. 2011. Divergent host plant specialization as the critical driving force in speciation between populations of a phytophagous ladybird beetle. Jo*urnal of Evolutionary Biology*. 24, 1421-1432. https://doi.org/10.1111/j.1420-9101.2011.02274.x

McClure M, Cannell E, Despland E. 2010. Thermal ecology and behavior of the nomadic social forager *Malacosoma disstria. Physiological Entomology*. 36, 120-127. https://doi.org/10.1111/j.1365-3032.2010.00770.x

Medina RF, Reyna SM, Bernal JS. 2012. Population genetic structure of a specialist leafhopper on *Zea*: likely anthropogenic and ecological determinants of gene flow. *Entomologia Experimentalis et Applicata*. 142, 223-235. https://doi.org/10.1111/j.1570-7458.2012.01220.x

Meirmans PG. Subsampling reveals that unbalanced sampling affects Structure results in a multi-species dataset. *Heredity*. 122, 276-287. https://doi.org/ 10.1038/s41437-018-0124-8

Miller WE. 2006. Forest tent caterpillar: mating, oviposition, and adult congregation at town lights during a northern Minnesota outbreak. *Journal of the Lepidopterists' Society*. 60, 156-160

Murton JB, Bateman MD, Dallimore SR, Teller JT, Yang Z. 2010. Identification of Younger Dryas outburst flood path from Lake Agassiz to the Arctic Ocean. *Nature.* 464, 740-743. https://doi.org/10.1038/nature08954

Nazareno AG, Bemmels JB, Dick CW, Lohmann LG. 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*. 17, 1136-1147. https://doi.org/10.1111/1755-0998.12654

Nicol RW, Arnason JT, Helson B, Abou-Zaid MM. 1997. Effect of host and non-host trees on the growth and development of the forest tent caterpillar, *Malacosoma disstria* (Lepidoptera: Lasiocampidae). *The Canadian Entomologist*. 129, 991-999. https://doi.org/10.4039/Ent129991-6

Nixon AE, Roland J. 2012. Generalist predation on forest tent caterpillar varies with forest stand composition: an experimental study across multiple life stages. *Ecological Entomology*. 37, 13-23. https://doi.org/ 10.1111/j.1365-2311.2011.01330.x

Nosil P, Crespi BJ. 2006. Experimental evidence that predation promotes divergence in adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America*. 103, 9090-9095. https://doi.org/10.1073/pnas.0601575103

Panzuto M, Lorenzetti F, Mauffette Y, Albert PJ. 2001. Perception of aspen and sun/shade sugar maple leaf soluble extracts by larvae of *Malacosoma disstria*. *Journal of Chemical Ecology*. 27, 1963-1978. https://doi.org/10.1023/A:1012282618332

Pappers SM, Van der Velde G, Ouborg NJ. 2002. Host preference and larval performance suggest host race formation in *Galerucella nymphaeae*. *Oecologia*. 130, 433-440. https://doi.org/10.1007/s00442-001-0822-3

Parry D, Goyer RA, Lenhard GJ. 2001. Macrogeographic clines in fecundity, reproductive allocation, and offspring size of the forest tent caterpillar *Malacosoma disstria*. *Ecological Entomology*. 26, 381-291. https://doi.org/10.1046/j.1365-2311.2001.00319.x

Parry D, Goyer RA. 2004. Variation in the suitability of host tree species for geographically discrete populations of forest tent caterpillar. *Environmental Entomology*. 33, 1479-1487. https://doi.org/10.1603/0046-225X-33.5.1477

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoesktra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*. 7, 1-11. https://doi.org/10.1371/journal.pone.0037135

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MV. 2013. An overview of Structure: applications, parameter settings, and supporting software. *Frontiers in Genetics*. 4, 1-13. https://doi.org/ 10.3389/fgene.2013.00098

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

Rochefort S, Berthiaume R., Hebert C, Charest M, Bauce E. 2011. Effect of temperature and host tree on cold hardiness of hemlock looper eggs along a latitudinal gradient. *Journal of Insect Physiology*. 57, 751-759. https://doi.org/10.1016/j.jinsphys.2011.02.013

Scheffer SJ, Hawthorne DJ. 2007. Molecular evidence of host-associated genetic divergence in the holly leafminer *Phytomyza glabricola* (Diptera: Agromyzidae): apparent discordance among marker systems. *Molecular Ecology*. 16, 2627-2637. https://doi.org/10.1111/j.1365-294X.2007.03303.x

Sexton JP, Strauss SY, Rice KJ. 2011. Gene flow increases fitness at the warm edge of a species' edge. *Proceedings of the National Academy of Sciences of the United States of America.* 108, 11704-11709. https://doi.org/10.1073/pnas.1100404108

Shafer ABA, Cullingham CI, Cote SD, Coltman DW. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*. 19, 4589-4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Silva-Brandao KL, Silva OABN, Brandao MM, Omoto C, Sperling FAH. 2015. Genotyping by sequencing approach indicates geographic distance as the main factor affecting genetic structure and gene flow in Brazilian populations of *Grapholita molesta*. *Evolutionary Applications*. 8, 476-485. https://doi.org/10.1111/eva.12257

Sniegula S, Drobniak SM, GołaB MJ, Johansson F. 2014. Photoperiod and variation in life history traits in core and peripheral populations in the damselfly *Lestes sponsa*. *Ecological Entomology*. 39, 137–148. https://doi.org/10.1111/een.12076

Sperling FAH. 1993. Mitochondrial DNA variation and Haldane's rule in the *Papilio glaucus* and *P. troilus* species groups. *Heredity*. 71, 227-233. https://doi.org/10.1038/hdy.1993.130

Stireman III JO, Nason JD, Heard SB. 2005. Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a

goldenrod-insect community. *Evolution.* 60, 2573-2587. https://doi.org/10.1111/j.0014-3820.2005.tb00970.x

Stireman III JO, Singer MS. 2018. Tri-trophic niches of insect herbivores in an era of rapid environmental change. *Current Opinion in Insect Science*. 29, 117-125. https://doi.org/10.1016/j.cois.2018.07.008

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence Data. III. Cladogram estimation. *Genetics.* 132, 619-633.

Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*. 21, 3907-3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Trudeau M, Mauffette Y, Rochefort S, Han E, Bauce E. 2010. Impact of host tree on forest tent caterpillar performance and offspring overwintering mortality. *Environmental Entomology*. 39, 498-504. https://doi.org/10.1603/en09139

van Els P, Cicero C, Klicka J. 2012. High latitudes and high genetic diversity: phylogeography of a widespread boreal bird, the gray jay (*Perisoreus canadensis*). *Molecular Phylogenetics and Evolution*. 63, 456-465. https://doi.org/10.1016/j.ympev.2012.01.019

Vidal MC, Quinn TW, Stireman JO, Tinghitella RM, Murphy SM. 2019. Geography is more important than host plant use for the population genetic structure of a generalist insect herbivore. *Molecular Ecology*. 28, 4317-4334. https://doi.org/10.1111/mec.15218

Wellenreuther M, Sanchez-Guillen RA, Cordero-Rivera A, Svensson EI, Hansson B. 2011. Environmental and climatic determinants of molecular diversity and genetic population

structure in a coenagrionid damselfly. *PLoS One*. 6, 1-16. https://doi.org/10.1371/journal.pone.0020440

Wickham H. 2016. *ggplot2: elegant graphics for data analysis*. New York (NY): Springer-Verlag. 253 p.

Witter JA, Kulman HM, Hodson AC. 1972. Life tables for the forest tent caterpillar. *Annals of the Entomological Society of America*. 65, 25-31. https://doi.org/10.1093/aesa/65.1.25

Yu H, Liang D, Tian E, Zheng L, Kjellberg F. 2018. Plant geographic phenotypic variation drives diversification in its associated community of a phytophagous insect and its parasitoids. *BMC Evolutionary Biology*. 18, 1-11. https://doi.org/10.1186/s12862-018-1239-5

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 28, 3326-3328. https://doi.org/10.1093/bioinformatics/bts606

Chapter 3

General Conclusions

3.1 Thesis summary

While intraspecific variation is often attributed to host race formation (Dres and Mallet 2002; Blair *et al.* 2005), many other factors (Nosil and Crespi 2006; Braby *et al.* 2007; Dillon and Lozier 2019) contribute to within-species diversity. Thus, to develop a more thorough understanding of the interacting forces shaping intraspecific diversification, the contributions of these various factors need to be teased apart to allow a better understanding of evolutionary processes. In this thesis, I determined the factors affecting intraspecific diversification by examining the population structuring of *Malacosoma disstria* across part of its Canadian range. Previous studies had demonstrated that outbreak dynamics vary between eastern and western populations (Cooke and Lorenzetti 2006), while larval host plants impact larval fitness and lead to life history differences (Nicol *et al.* 1997; Parry *et al.* 2001; Trudeau *et al.* 2010). I hypothesized that these functional differences are due to cryptic population structure within *M. disstria*. I used mitochondrial DNA (mtDNA) to compare my work with previous studies on *M. disstria* population structuring, and used double digest restriction-site associated DNA (ddRAD) sequencing to survey genome-wide SNP markers in larvae sampled from multiple hosts, across several forest ecozones and from eastern and central Canada (Chapter 2).

Within-species structuring can exhibit variation at large scales. In Chapter 2, I observed substantial population structuring between eastern and central *M. disstria* populations across Canada (Figure 2.3). Previous work by Lait and Hebert (2018) found limited population structuring in mtDNA but observed that Canadian *M. disstria* populations were divided among three main clusters; a British Columbia group, a central Alberta/Saskatchewan/Manitoba group,

and an eastern group comprising Ontario/Quebec/New Brunswick/Nova Scotia. My mtDNA samples provided additional support for these clusters. However, several haplotypes from eastern Canada group within the central (prairie) cluster. There are also several central haplotypes clustered with eastern Canada. To explain the east-west divide that exists west of the Ontario-Manitoba border, historical geological events should be considered. During the Pleistocene, 20,000 years ago, ice age glaciers forced species into distinct refugia across continental North America (Shafer *et al.* 2010). Populations were isolated between refugia and became genetically distinct over time. As glaciers receded, eastern and western populations were kept isolated by the formation of Lake Agassiz across present day Manitoba, western Ontario and the north-central United States (Murton *et al.* 2010). As trees recolonized glacier-free land, *M. disstria* populations may have followed the expansion of their hosts, similar to hypothesized spruce budworm expansion events (Lumley *et al.* 2020). The eastern and western divisions that I observed from mtDNA and SNP data could be attributed in part to these events (Chapter 2; Figure 2.2, 2.3).

Just as population structuring can be a result of large-scale geographic events, plant hosts also influence population structuring. As host differences have been observed previously in the literature (Nicol *et al.* 1997; Parry and Goyer 2004; Trudeau *et al.* 2010), host races can be expected. However, I found that this was not the case, since population structuring lacked a strong host affiliation (Chapter 2; Figure 2.4). It has been previously suggested that temperate zone generalists are comprised of a species complex consisting of host-specialists (Bickford *et al.* 2006). However, my results in Chapter 2 indicate that not all generalists are cryptic host specialists, as demonstrated by the greater influence of geography in shaping population structuring (Vidal *et al.* 2019). This has also been observed in other species, such as the fall

webworm, where geographic distance accounted for greater genetic variation than host plant differences (Vidal *et al.* 2019). Oriental fruit moth population structure shows a stronger association with geography than with different larval hosts, which are various kinds of fruit trees (Silva-Brandao *et al.* 2015).

Bernays and Wcislo (1994) proposed the information processing constraint, where a generalist will be less efficient when searching for resources than a specialist is. An overabundance of stimuli caused by chemical signals from many hosts leads to decreased fitness as information processing capabilities are overloaded and confuse the individual. As a result, some adults encountering different hosts will simply use the various host plants intermixed within an environment, such as the warty leaf beetle (*Neochlamisus bebbianae* Brown) laying eggs on alternative hosts (Egan and Funk 2006). In the case of female *M. disstria*, which experience only one egg laying event, related larvae generally develop on a single natal plant. In this case, adult females make the host-use decision for future larvae but, due to host availability, could select sub-optimal hosts. Although no host races were detected, *M. disstria* populations did show regional variation congruent with forest ecozones.

Regional adaptations can also lead to population structuring among insects. In Chapter 2, I explored regional structuring among eastern *M. disstria* populations, based on ecozone designations of Baldwin *et al.* (2018). Specifically, aspen *M. disstria* populations were sampled in three eastern forest ecozones: temperate deciduous, temperate mixed and boreal forests. By examining only the larvae sampled from aspen sites, I was able to identify regional population structuring without host affects obscuring the results. Boreal *M. disstria* populations were distinct from those collected in the temperate forests (deciduous and mixed) (Chapter 2; Figure

2.5). This suggests that vegetation, temperature and habitat changes along a latitudinal gradient may have a stronger influence on *M. disstria* populations than larval host species (Chapter 2).

Local environmental conditions, such as understory vegetation, soil, moisture levels and temperatures can all impact regional populations, as they influence forest habitat that can affect local diversity (Werner and Raffa 2000). Over successive generations, local adaptations can emerge in a subset of a population. For example, due to the colder climate of northern forests, local species populations will tend to express morphological differences. Egg size varies, with fewer yet larger eggs laid as a result of more energy being spent on individual offspring survival in colder climates, as opposed to southern populations where adults will invest more in egg quantity (Parry *et al.* 2001). These patterns have been observed in other species, such as spruce budworm, where variation between northern and southern populations is attributed to enhancing larval cold survival (Harvey 1983).

As a consequence of neighboring ecozone populations that differ from each other, proximity is not always an indicator of genetic similarity. Genetic differences may still occur in continuous habitat with limited or no barriers to species movement (Ehrich and Stenseth 2001). At small spatial scales, *Carabus nemoralis* (Muller) and *Carabus punctatoauratus* (Germar) both have population structuring in local forested areas separated by only 1 to 2 km (Brouat *et al.* 2003). It was hypothesized that *C. nemoralis* population structuring could be the result of founder events, caused by extinction and recolonization by few individuals. *C. punctatoauratus* population structuring however may be limited by microhabitats, where humidity, roadways or canopy cover could all act as barriers to dispersal (Brouat *et al.* 2003). Similar to forest ecozones separating *M. disstria* populations, small scale barriers such as local microhabitats and re-colonization events could be influencing population structuring. In addition, the aforementioned

cold survival mechanisms present in boreal forest insect populations, which are adapted to tolerate more extreme winter conditions than occur in milder temperate forests, can result in a natural separation of populations.

For many studies, a single marker is used to sample a species and determine relationships (Dupuis *et al.* 2012). In Chapter 2, I compared genomic SNPs to the mtDNA COI gene sequence to examine large-scale continental population structuring. My results indicated substantial discordance between SNP data and mtDNA, although east-west structuring was observed in both cases. The discordance in *M. disstria* is likely the result of a complex evolutionary history. Introgression, which is hybridization between distant populations, may have contributed to the variation observed in mtDNA compared to the SNP data (Bensch *et al.* 2006; Kodandaramaiah *et al.* 2013). That is why I used ddRAD data in addition to mtDNA, to provide a genome-wide marker for assessing population structure and relatedness and to avoid false positives from a single marker.

My results provide better understanding of intraspecific diversification, especially within *M. disstria*, and demonstrate the need to disentangle multiple factors influencing speciation. These results thereby also highlight the need to use combinations of molecular techniques, as the narrative implied by limited results may be misleading and not necessarily indicate true population structure.

3.2 Future research

Forest tent caterpillars are fascinating insects. They show complex social and behavioural characteristics, and their massive, synchronous outbreaks generate both interest and disgust among the general public. Despite being comparatively well known, little was known about the

evolutionary processes shaping population structure across their range. I showed that their populations are structured by regional and historic processes but found little evidence of host races in eastern Canada (Chapter 2).

Over the course of this thesis I have identified a number of research directions that could prove fruitful. While I identified continental geographic samples, my sampling was restricted primarily to eastern Canada (Ontario and Quebec) and I only had limited sampling from the rest of the distribution of *M. disstria*. However, my genome-wide SNP data provided greater clarity of population structuring than previously published work using a single mitochondrial region (Lait and Hebert 2018). Therefore, to fully quantify continental variation in *M. disstria*, it would be vital to survey populations throughout its entire range. Uncovering the population boundaries in Manitoba would provide additional clarity, while sampling from the Atlantic provinces and British Columbia would further provide insight into the population structuring of *M. disstria* and allow a better comparison with the large-scale survey by Lait and Hebert (2018).

As my thesis focused on intraspecific diversification using genetic and genomic markers, a morphological study should be incorporated into any future continent-wide study of *M. disstria*. Previous work on morphology has examined wing area, with ongoing decreases in wing area throughout the duration of the flight season, which is believed to be the result of declining and changing foliage quality during the larval growing stages (Jones and Evenden 2008). Late emerging larvae can be exposed to increased plant chemical defenses, while possessing lower nutritional value. This lower quality food impacts larval development and size, resulting in smaller adults. *M. disstria* have also been found to have polymorphic melanism, with high levels of intraspecific variation in male colour (Ethier and Despland 2012). While wing size and colouration provide many benefits for mating male moths, including thermoregulatory

advantages and predator avoidance (Etheir and Despland 2012), the connection with hosts and geographical distribution remains unknown.

I initially thought that morphological traits may be linked to genetic variation, and therefore associated with host or geography. I originally set out to examine adult morphology, but high levels of disease in the *M. disstria* populations led to high larval mortality, so many individuals did not survive to adulthood. However, throughout rearing I imaged larvae in their 4th and 5th instars when possible. The data could serve as an additional avenue to explore variation among *M. disstria* populations and any future population study should incorporate a simultaneous morphological component.

In addition, genetic differences in microsporidia strains have been found between northern and southern populations of *M. disstria* (unpublished, Dr. George Kyei-Poku, Great Lakes Forestry Centre), and pursuing the susceptibility of different populations to disease is a worthwhile avenue for research. As different regional populations feed on varying hosts, the nutritional value of leaf foliage may play a role in disease survivability and resistance. Another avenue of interest would be *M. disstria* pheromones. With the increasing economic importance of *M. disstria*, pheromone traps have been developed to monitor populations (Schmidt *et al.* 2003). However, just as there are varying life history traits, variation in regional pheromone use is possible. In *Dendroctonus rufipennis* (Kirby), it was observed that there was geographic variation in occurrence and bioactivity of several pheromones, where traps were only effective with regionally specific chemicals (Borden *et al.* 1996). Such behavior patterns are plausible in a generalist such as *M. disstria,* which has demonstrated many other variable life history traits associated with regional geography.

In summary, disentangling the factors influencing intraspecific diversification provides further insight into evolutionary processes. Understanding the roles that hosts and geography play in shaping population structure is useful to many facets of biology and systematics. *Malacosoma disstria,* being the infamous forest pest that it is, is relatively easy to study and acquire specimens for, which helps toward resolving some of the mysteries surrounding speciation. However, *M. disstria* phylogeography still leaves many avenues of research to document and explore. This thesis is but a small step in understanding the full picture of *M. disstria* and enhancing our understanding of factors influencing intraspecific variation.

3.3 Literature cited

Baldwin K, Allen L, Basquill S, Champan K, Downing D, Flyyn N, MacKenzie W, Major M, Meades W, Meidinger D, Morneau C, Saucier JP, Thorpe J, Uhlig P. 2018. Vegetation zones of Canada: a biogeoclimatic perspective. [Map] Scale 1:5,000,000. Natural Resources Canada, Canadian Forest Service. Sault Ste Marie, ON. http://www.cnvc-cnvc.ca

Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology*. 15, 161-171. https://doi.org/10.1111/j.1365-294X.2005.02766.x

Bernays EA, Wcislo WT. 1994. Sensory capabilities, information processing, and resources specialization. *The Quarterly Review of Biology.* 69, 187-204. https://doi.org/10.1093/aesa/65.1.25

Blair CP, Abrahamson WG, Jackman JA, Tyrrell L. 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. *Evolution*. 59, 304-316. https://doi.org/10.1111/j.0014-3820.2005.tb00991.x

Borden JH, Gries G, Chong LJ, Werner RA, Holsten EH, Wiser H, Dixon EA, Cerezke HF. 1996. Regionally-specific bioactivity of two new pheromones for *Denroctonus rufipennis* (Kirby) (Col., Scolytidae). *Journal of Applied Entomology*. 120, 321-326. https://doi.org/10.1111/j.1439-0418.1996.tb01613.x

Braby MF, Pierce NE, Vila R. 2007. Phylogeny and historical biogeography of the subtribe Aporiina (Lepidoptera: Pieridae): implications for the origin of Australian butterflies. *Biological Journal of the Linnean Society.* 90, 413-440. https://doi.org/10.1111/j.1095 8312.2007.00732.x

Brouat C, Sennedot F, Audiot P, Leblois R, JY Rasplus. 2003. Fine-scale genetic structure of two carabid species with contrasted levels of habitat specialization. *Molecular Ecology*. 12, 1731-1745. https://doi.org/10.1046/j.1365-294X.2003.01861.x

Carlyle-Moses DE, Price AG. 2006. Growing-season stemflow production within a deciduous forest of southern Ontario. *Hydrological Processes*. 20, 3651-3663. https://doi.org/10.1002/hyp.6380

Cooke BJ, Lorenzetti F. 2006. The dynamics of forest tent caterpillar outbreaks in Quebec, Canada. *Forest Ecology and Management*. 226, 110-121. http://doi.org/10.1016/j.foreco.2006.01.034

Dillon ME, Lozier JD. 2019. Adaptation to the abiotic environment in insects: the influence of variability on ecophysiology and evolutionary genomics. *Current Opinion in Insect Science*. 36, 131-139. https://doi.org/10.1016/j.cois.2019.09.003

Dres M, Mallet J. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 357, 471-492. https://doi.org/10.1098/rstb.2002.1059

Dupuis JR, Roe AD, Sperling FAH. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology.* 21, 4422-4436. https://doi.org/10.1111/j.1365-294X.2012.05642.x

Egan SP, Funk DJ. 2006. Individual advantages to ecological specialization: insights on cognitive constraints from three conspecific taxa. *Proceedings of the Royal Society B: Biological Sciences*. 273, 843-848. https://doi.org/10.1098/rspb.2005.3382

Ehrich D, Stenseth NC. 2001. Genetic structure of Siberian lemmings (*Lemmus sibiricus*) in a continuous habitat: Large patches rather than isolation by distance. *Heredity*. 86, 716-730. https://doi.org/10.1046/j.1365-2540.2001.00883.x

Ethier J, Despland E. 2012. Effects of polymorphic melanism and larval diet on life history traits of *Malacosoma disstria* moths. *Journal of Insect Physiology.* 58, 67-74. https://doi.org/10.1016/j.jinsphys.2011.09.018

Harvey GT. 1983. A geographical cline in egg weights in *Choristoneura fumiferana* Lepidoptera: Tortricidae) and its significance in population dynamics. *The Canadian Entomologist*. 115, 1103-1108. https://doi.org/10.4039/Ent1151103-9

Jones BC, Evenden ML. 2008. Ecological applications of pheromone trapping of *Malacosoma disstria* and *Choristoneura conflictana*. *The Canadian Entomologist*. 140, 573-581. https://doi.org/10.4039/N08-013

Klemola T, Tanhuanpaa M, Korpimaki E, Ruohomaki K. 2002. Specialist and generalist natural enemies as an explanation for geographical gradients in population cycles of northern herbivores. *OIKOS.* 99, 83-94. https://doi.org/10.1034/j.1600-0706.2002.990109.x

Kodandaramaiah U, Simonsen TJ, Bromilow S, Wahlberg N, Sperling F. 2013. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution*. 3, 5167-5176. https://doi.org/10.1002/ece3.886

Lait LA, Hebert PDN. 2018. Phylogeographic structure in three North American tent caterpillar species (Lepidoptera: Lasiocampidae): *Malacosoma americana*, *M. californicum*, and *M. disstria*. *PeerJ*. 6, 2-23. https://doi.org/10.7717/peerj.4479

Lumley LM, Pouliot E, Laroche J, Boyle B, Brunet BMT, Levesque RC, Sperling FAH, Cusson M. 2020. Continent-wide population genomic structure and phylogeography of North America's most destructive conifer defoliator, the spruce budworm (*Choristoneura fumiferana*). *Ecology and Evolution*. 10, 914-927. https://doi.org/10.1002/ece3.5950

Murton JB, Bateman MD, Dallimore SR, Teller JT, Yang Z. 2010. Identification of Younger Dryas outburst flood path from Lake Agassiz to the Arctic Ocean. *Nature.* 464, 740-743. https://doi.org/10.1038/nature08954

Nicol RW, Arnason JT, Helson B, Abou-Zaid MM. 1997. Effect of host and non-host trees on the growth and development of the forest tent caterpillar, *Malacosoma disstria* (Lepidoptera: Lasiocampidae). *The Canadian Entomologist*. 129, 991-999. https://doi.org/10.4039/Ent129991-6

Parry D, Goyer RA, Lenhard GJ. 2001. Macrogeographic clines in fecundity, reproductive allocation, and offspring size of the forest tent caterpillar *Malacosoma disstria*. *Ecological Entomology*. 26, 381-291. https://doi.org/10.1046/j.1365-2311.2001.00319.x

Parry D, Goyer RA. 2004. Variation in the suitability of host tree species for geographically discrete populations of forest tent caterpillar. *Environmental Entomology*. 33, 1479-1487. https://doi.org/10.1603/0046-225X-33.5.1477

Schmidt BC, Roland J, Wakarchuk D. 2003. Evaluation of synthetic pheromones for monitoring forest tent caterpillar (Lepidoptera: Lasiocampidae) populations. *Environmental Entomology*. 32, 214-219. https://doi.org/10.1603/0046-225X-32.1.214

Shafer ABA, Cullingham CI, Cote SD, Coltman DW. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*. 19, 4589-4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Silva-Brandao KL, Silva OABN, Brandao MM, Omoto C, Sperling FAH. 2015. Genotyping by sequencing approach indicates geographic distance as the main factor affecting genetic structure and gene flow in Brazilian populations of *Grapholita molesta*. *Evolutionary Applications*. 8, 476-485. https://doi.org/10.1111/eva.12257

Trudeau M, Mauffette Y, Rochefort S, Han E, Bauce E. 2010. Impact of host tree on forest tent caterpillar performance and offspring overwintering mortality. *Environmental Entomology*. 39, 498-504. https://doi.org/10.1603/en09139

Vidal MC, Quinn TW, Stireman JO, Tinghitella RM, Murphy SM. 2019. Geography is more important than host plant use for the population genetic structure of a generalist insect herbivore. *Molecular Ecology*. 28, 4317-4334. https://doi.org/10.1111/mec.15218

Werner SM, Raffa KF. 2000. Effects of forest management practices on the diversity of ground-occurring beetles in mixed northern hardwood forests of the Great Lakes region. *Forest Ecology and Management*. 139, 135-155. https://doi.org/10.1016/S0378-1127(99)00341-2

Biography

I was born on January 1st, 1987 at the Ajax Pickering Hospital, in Ajax, Ontario. I was the first of two children (my younger sister, Terra-Lynn, was born in 1988) to my parents, Shelley and Harry Snape. Growing up in Oshawa, Ontario, I spent great amounts of time playing video games, reading dinosaur books and exploring the outdoors with friends as I grew up, cultivating a varied love for paleontology, entomology, technology and adventure.

It was during this time my first experiences with entomology began, when I first captured and observed a praying mantis on the driveway. Later, I would collect grasshoppers and bees on the school playground, while at the cottage I was often playing with dragonflies. Otherwise I was often carrying a new dinosaur book to school, reading it over and over. I had Jurassic park memorized. It was around this time, I found my love for technology, stemming from an old DOS computer my father let me play on, and the NES console my mother gave me.

After high school, I studied technology and game design at the University of Ontario Institute of Technology in 2006. However, it was short lived, as a desk-bound career was not of interest to me long term, and I switched to Algonquin College in 2008 to study Outdoor Adventure - something that still shocks my family to this day for such a radical change. It was through this program that I picked up an odd assortment of skills and interests that continue to this day, from rafting and kayaking, to winter camping and survival skills.

After graduation in December 2009, I was recommended for a position working at Bell Technical Solutions, from my sister's future husband, Sam Snyders. There I became a union representative, a health and safety representative, a field trainer for new fiber technicians and ended up as Local 1996-O recording secretary. An interesting way to put together some of the odd assortment of skills I had developed from my times at the University of Ontario and

Algonquin College. I was with the company for several years and began to plan a career change, when chance occurred, and I was able to visit China and Japan. After which I accepted an offer of admission to Algoma University, in Sault Ste Marie, Ontario to study biology in 2013.

This lead to four interesting years of my undergraduate program, where again, due to varied skills I had picked up over the years, I become audio/video editor for the Algoma University Radio, Science Representative for the Algoma Student Union, Resident Advisor in the campus dormitories, and participated on research projects led by Dr. Schamp, Dr. Antunes and Dr. Foote. This led to my undergraduate thesis with Dr. Schamp and Dr. Antunes on plant root mycorrhizal fungi. It was also during this time, I happened to take a class with my future graduate supervisor, Dr. Amanda Roe. After graduation in summer 2017, Dr. Roe remembered my time as one of her students and asked if I would work for her at her research lab at Natural Resources Canada in Sault Ste Marie. With her support and connections at the University of Alberta, she introduced me to Dr. Felix Sperling, who would later become my other supervisor and mentor.

All of this led to me attending the University of Alberta in January 2018, to pursue a Master of Science, in Systematics and Evolution, studying *Malacosoma disstria* ecology. Although knowing so little at first, Dr. Roe and Dr. Sperling never gave up on me, and always encouraged me. I still feel as if I know so little, but the skills they instilled in me made me a better researcher, and I cannot thank either of them enough for all of the time and effort they have put into me, in getting me to where I am today.

Bibliography

Arimura GI, Huber DPW, Bohlmann J. 2004. Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa x deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (-)-germacrene D synthase, PtdTPS1. *The Plant Journal.* 37, 603-616. https://doi.org/10.1111/j.1365-313X.2003.01987.x

Baldwin K, Allen L, Basquill S, Champan K, Downing D, Flyyn N, MacKenzie W, Major M, Meades W, Meidinger D, Morneau C, Saucier JP, Thorpe J, Uhlig P. 2018. Vegetation zones of Canada: a biogeoclimatic perspective. [Map] Scale 1:5,000,000. Natural Resources Canada, Canadian Forest Service. Sault Ste Marie, ON. http://www.cnvc-cnvc.ca

Bandelt H, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*. 16, 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Bankhead-Dronnet S, Perdereau E, Kutnik M, Dupont S, Bagneres AG. 2015. Spatial structuring of the population genetics of a European subterranean termite species. *Ecology and Evolution.* 5, 3090-3102. https://doi.org/10.1002/ece3.1566

Batzer HO, Martin MP, Mattson WJ, Miller WE. 1995. The forest tent caterpillar in aspen stands: distribution and density estimation of four life stages in four vegetation strata. *Forest Science.* 41, 99-121. https://doi.org/10.1093/forestscience/41.1.99

Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology*. 15, 161-171. https://doi.org/10.1111/j.1365-294X.2005.02766.x

Bernays EA, Wcislo WT. 1994. Sensory capabilities, information processing, and resources specialization. *The Quarterly Review of Biology.* 69, 187-204. https://doi.org/10.1093/aesa/65.1.25

Blair CP, Abrahamson WG, Jackman JA, Tyrrell L. 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. *Evolution*. 59, 304-316. https://doi.org/10.1111/j.0014-3820.2005.tb00991.x

Borden JH, Gries G, Chong LJ, Werner RA, Holsten EH, Wiser H, Dixon EA, Cerezke HF. 1996. Regionally-specific bioactivity of two new pheromones for *Denroctonus rufipennis* (Kirby) (Col., Scolytidae). *Journal of Applied Entomology*. 120, 321-326. https://doi.org/10.1111/j.1439-0418.1996.tb01613.x

Braby MF, Pierce NE, Vila R. 2007. Phylogeny and historical biogeography of the subtribe Aporiina (Lepidoptera: Pieridae): implications for the origin of Australian butterflies. *Biological Journal of the Linnean Society.* 90, 413-440. https://doi.org/10.1111/j.1095-8312.2007.00732.x

Brandt JP, Cerezke HF, Mallett KI, Volney WJA, Weber JD. 2003. Factors affecting trembling aspen (*Populus tremuloides* Michx) health in the boreal forest of Alberta, Saskatchewan, and Manitoba, Canada. *Forest Ecology and Management*. 178, 287-300. https://doi.org/10.1016/S0378-1127(02)00479-6

Brouat C, Sennedot F, Audiot P, Leblois R, JY Rasplus. 2003. Fine-scale genetic structure of two carabid species with contrasted levels of habitat specialization. *Molecular Ecology*. 12, 1731-1745. https://doi.org/10.1046/j.1365-294X.2003.01861.x

Carroll SP, Marler M, Winchell R, Dingle H. 2003. Evolution of cryptic flight morph and life history differences during host race radiation in the soapberry bug, *Jadera haematoloma*

Herrich-Schaeffer (Hemiptera: Rhopalidae). *Annals of the Entomological Society of America*. 96, 135-143. https://doi.org/10.1603/0013-8746(2003)096[0135:EOCFMA]2.0.CO;2

Catchen J, Hohenlohe P, Bassham S, Amores A, Cresko W. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*. 22, 3124-3140. https://doi.org/10.1111/mec.12354

Charbonneau D, Lorenzetti F, Doyon F, Mauffette Y. 2012. The influence of stand and landscape characteristics on forest tent caterpillar (*Malacosoma disstria*) defoliation dynamics: the case of the 1999-2002 outbreak in northwestern Quebec. *Canadian Journal of Forest Research*. 42, 1827-1836. https://doi.org/10.1139/x2012-126

Chu ND, Kaluziak ST, Trussell GC, Vollmer SV. 2014. Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus*. *Molecular Ecology*. 23, 1863-1873. https://doi.org/10.1111/mec.12681

Clement M, Snell Q, Walke P, Posada D, Crandall K. 2002. TCS: estimating gene genealogies. *Proceedings of the 16th International Parallel and Distributed Processing Symposium*. 184-191. https://doi.org/10.1109/IPDPS.2002.1016585

Cooke BJ, Lorenzetti F. 2006. The dynamics of forest tent caterpillar outbreaks in Quebec, Canada. *Forest Ecology and Management*. 226, 110-121. http://doi.org/10.1016/j.foreco.2006.01.034

Cooke BJ, Roland J. 2018. Early 20th century climate-driven shift in the dynamics of forest tent caterpillar outbreaks. *American Journal of Climate Change*. 7, 253-270. http://doi.org/10.4236/ajcc.2018.72017

Costa JT, Ross KG. 1994. Hierarchical genetic structure and gene flow in macrogeographic populations of the eastern tent caterpillar (*Malacosoma americanum*). *Evolution*. 48, 1158-1167. http://doi.org/10.1111/j.1558-5646.1994.tb05302.x

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker R, Lunter G, Marth G, Sherry ST, McVean G, Durbin R. 2011. 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*. https://vcftools.github.io/license.html

de Quieroz K. 2007. Species concepts and species delimitation. *Systematic Biology*. 56, 879-886. https://doi.org/10.1080/10635150701701083

Despland E, Hamzeh S. 2004. Ontogenetic changes in social behaviour in the forest tent caterpillar, *Malacosoma disstria*. *Behavioral Ecology and Sociobiology*. 56, 177–184. https://doi.org/10.1007/s00265-004-0767-8

Despland E, Huu AL. 2006. Pros and cons of group living in the forest tent caterpillar: separating the roles of silk and of grouping. *Entomologia Experimentalis et Applicata*. 122, 181-189. https://doi.org/10.1111/j.1570-7458.2006.00512.x

Despland E, Noseworthy M. 2006. How well do specialist feeders regulate nutrient intake? Evidence from a gregarious tree-feeding caterpillar. *Journal of Experimental Biology*. 209, 1301-1309. https://doi.org/10.1242/jeb.02130

Dillon ME, Lozier JD. 2019. Adaptation to the abiotic environment in insects: the influence of variability on ecophysiology and evolutionary genomics. *Current Opinion in Insect Science*. 36, 131-139. https://doi.org/10.1016/j.cois.2019.09.003

Donaldson JR, Lindroth RL. 2008. Effects of variable phytochemistry and budbreak phenology on defoliation of aspen during a forest tent caterpillar outbreak. *Agricultural and Forest Entomology*. 10, 399-410. https://doi.org/10.1111/j.1461-9563.2008.00392.x

Downey MH, Nice CC. 2013. A role for both ecology and geography as mechanisms of genetic differentiation in specialized butterflies. *Evolutionary Ecology*. 27, 565-578. https://doi.org/10.1007/s10682-012-9626-7

Dres M, Mallet J. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 357, 471-492. https://doi.org/10.1098/rstb.2002.1059

Dupuis JR, Roe AD, Sperling FAH. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology.* 21, 4422-4436. https://doi.org/10.1111/j.1365-294X.2012.05642.x

Egan SP, Funk DJ. 2006. Individual advantages to ecological specialization: insights on cognitive constraints from three conspecific taxa. *Proceedings of the Royal Society B: Biological Sciences*. 273, 843-848. https://doi.org/10.1098/rspb.2005.3382

Ehrich D, Stenseth NC. 2001. Genetic structure of Siberian lemmings (*Lemmus sibiricus*) in a continuous habitat: Large patches rather than isolation by distance. *Heredity*. 86, 716-730. https://doi.org/10.1046/j.1365-2540.2001.00883.x

Endara MJ, Weinhold A, Cox JE, Wiggins NL, Coley PD, Kursar TA. 2015. Divergent evolution in antiherbivore defences within species complexes at a single Amazonian site. *Journal of Ecology*. 103, 1107-1118. https://doi.org/10.1111/1365-2745.12431

Ethier J, Despland E. 2012. Effects of polymorphic melanism and larval diet on life history traits of *Malacosoma disstria* moths. *Journal of Insect Physiology.* 58, 67-74. https://doi.org/10.1016/j.jinsphys.2011.09.018

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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Evenden ML, Whitehouse CM, Jones BC. 2015. Resource allocation to flight in an outbreaking forest defoliator Malacosoma disstria. Physiological Ecology. 44, 835-845. https://doi.org/10.1093/ee/nvv055

Ferrari J, Godfray HCJ, Faulconbridge AS, Prior K, Via S. 2006. Population differentiation and genetic variation in host choice among pea aphids from eight host plant genera. *Evolution*: 60, 1574-1584. https://doi.org/10.1554/06-024.1

Fitzgerald TD. 1995. *The Tent Caterpillars*. Ithica (NY): Cornell University Press. 303 p.

Franclemont JG. 1973. *The moths of America north of Mexico: including Greenland. Fascicle 20.1. Mimallonoidea Mimallonidae and Bombycoidea Apatelodidae, Bombycidae,* Lasiocampidae. London (UK): EW Classey Limited and RBD Publications Inc. 86 p.

Franklin MT, Myers JH, Cory JS. 2014. Genetic similarity of island populations of tent caterpillars during successive outbreaks. *PLOS One*. 9, 1-9. https://doi.org/10.1371/journal.pone.0096679

Franklin MT, Ritland CE, Myers JH, Cory JS. 2012. Multiple mating and family structure of the western tent caterpillar, *Malacosoma californicum pluviale*: impact on disease resistance. *PLOS One*. 7, 1-8. https://doi.org/10.1371/journal.pone.0037472

Fricke C, Arnqvist G. 2007. Rapid adaptation to a novel host in a seed beetle (*Callosobruchus maculatus*): the role of sexual selection. *Evolution.* 61, 440-454. https://doi.org/10.1111/j.1558-5646.2007.00038.x

Fuller S, Schwarz M, Tierney S. 2005. Phylogenetics of the allodapine bee genus *Braunsapis*: Historical biogeography and long-range dispersal over water. *Journal of Biogeography*. 32, 2135-2144. https://doi.org/10.1111/j.1365-2699.2005.01354.x

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*. 34, 397-423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421

Fuentealba A, Pureswaran D, Beauce E, Despland E. 2017. How does synchrony with host plant affect the performance of an outbreaking insect defoliator? *Oecologia*. 184, 847-857. https://doi.org/10.1007/s00442-017-3914-4

Futuyma DJ, Saks ME. 1981. The effect of variation in host plant on the growth of an oligophagous insect, *Malacosoma americanum* and its polyphagous relative *Malacosoma disstria*. *Entomologia Experimentalis et Applicata.* 30, 163-168. https://doi.org/10.1111/j.1570-7458.1981.tb03091.x

Gilbert KJ. 2016. Identifying the number of population clusters with Structure: problems and solutions. *Molecular Ecology Resources*. 16, 601-603. https://doi.org /10.1111/1755-0998.12521

Goldblum D, Rigg LS. 2005. Tree growth response to climate change at the deciduous-boreal forest ecotone, Ontario, Canada. *Canadian Journal of Forest Research*. 35, 2709-2718. https://doi.org/10.1139/X05-185

Gorur G, Lomanco C, Mackenzie A. 2005. Phenotypic plasticity in host-plant specialisation in *Aphis fabae*. *Ecological Entomology*. 30, 657-664. https://doi.org/10.1111/j.0307-6946.2005.00742.x

Gosden TP, Stoks R, Svensson EI. 2011. Range limits, large-scale biogeographic variation, and localized evolutionary dynamics in a polymorphic damselfly. *Biological Journal of the Linnean Society*. 102, 775-785. https://doi.org/10.1111/j.1095-8312.2011.01619.x

Gratton P, Konopinski MK, Sbordoni V. 2008. Pleistocene evolutionary history of the clouded apollo (*Parnassius mnemosyne*): genetic signatures of climate cycles and a "time-dependent" mitochondrial substitution rate. *Molecular Ecology*. 17, 4248-4262. https://doi.org/10.1111/j.1365-294X.2008.03901.x

Gray DR, Ostaff DP. 2012. Egg hatch of forest tent caterpillar (Lepidoptera: Lasiocampidae) on two preferred host species. *The Canadian Entomologist*. 144, 790-797. https://doi.org/10.4039/tce.2012.73

Hale ML, Burg TM, Steeves TE. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PLoS One*. 9, 1-10. https://doi.org/10.1371/journal.pone.0045170

Hardy NB, Otto SP. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society B: Biological Sciences*. 281, 1-10. https://doi.org/10.1098/rspb.2013.2960

Hartley S, Krushelnycky PD, Lester PJ. 2010. Integrating physiology, population dynamics and climate to make multi-scale predictions for the spread of an invasive insect: the Argentine

ant at Haleakala National Park, Hawaii. *Ecography*. 33, 83-94. https://doi.org/10.1111/j.1600-0587.2009.06037.x

Hartmann H, Messier C. 2011. Interannual variation in competitive interactions from natural and anthropogenic disturbances in a temperate forest tree species: implications for ecological interpretation. *Forest Ecology and Management*. 261, 1936-1944. https://doi.org/10.1016/j.foreco.2011.02.018

Harvey GT. 1983. A geographical cline in egg weights in *Choristoneura fumiferana* (Lepidoptera: Tortricidae) and its significance in population dynamics. *The Canadian Entomologist*. 115, 1103-1108. https://doi.org/10.4039/Ent1151103-9

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*. 270, 313-321. https://doi.org/10.1098/rspb.2002.2218

Hogg EH, Brandt JP, Kochtubajda B. 2002. Growth and dieback of aspen forests in northwestern Alberta, Canada, in relation to climate and insects. *Canadian Journal of Forest Research*. 32, 823-832. https://doi.org/10.1139/X01-152

Jamieson MA, Schwartzberg EG, Raffa KF, Reich PB, Lindroth RL. 2015. Experimental climate warming alters aspen and birch phytochemistry and performance traits for an outbreak insect herbivore. *Global Change Biology*. 21, 2698-2710. https://doi.org/10.1111/gcb.12842

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL. 2017. The K=2 conundrum. *Molecular Ecology*. 26, 3594-3602. https://doi.org/10.1111/mec.14187

Janz N, Nylin S, Wahlberg N. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. BMC *Evolutionary Biology*. 6, 1-10. https://doi.org/10.1186/1471-2148-6-4

Johnson MTJ. 2008. Bottom-up effects of plant genotype on aphids, ants, and predators. *Ecology*. 89, 145-154. https://doi.org/10.1890/07-0395.1

Jolliffe IT, Cadima J. 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 374, 1-16. http://doi.org/10.1098/rsta.2015.0202

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 11, 1-15. https://doi.org/10.1186/1471-2156-11-94

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24, 1403-1405. https://doi.org/10.1093/bioinformatics/btn129

Jones BC, Evenden ML. 2008. Ecological applications of pheromone trapping of *Malacosoma disstria* and *Choristoneura conflictana*. *The Canadian Entomologist*. 140, 573-581. https://doi.org/10.4039/N08-013

Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*. 20, 1160-1166. https://doi.org/10.1093/bib/bbx108

Kebli H, Brais S, Kernaghan G, Drouin P. 2012. Impact of harvesting intensity on wood-inhabiting fungi in boreal aspen forests of Eastern Canada. *Forest Ecology and Management*. 279, 45-54. http://doi.org/10.1016/j.foreco.2012.05.028

Klemola T, Tanhuanpaa M, Korpimaki E, Ruohomaki K. 2002. Specialist and generalist natural enemies as an explanation for geographical gradients in population cycles of northern herbivores. *OIKOS.* 99, 83-94. https://doi.org/10.1034/j.1600-0706.2002.990109.x

Kodandaramaiah U, Simonsen TJ, Bromilow S, Wahlberg N, Sperling F. 2013. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution*. 3, 5167-5176. https://doi.org/10.1002/ece3.886

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*. 15, 1179-1191. https://doi.org/10.1111/1755-0998.12387

Lait LA, Hebert PDN. 2018. Phylogeographic structure in three North American tent caterpillar species (Lepidoptera: Lasiocampidae): *Malacosoma americana*, *M. californica*, and *M. disstria*. *PeerJ*. 6, 2-23. https://doi.org/10.7717/peerj.4479

Lapointe FJ, Rissler LT. 2005. Congruence, consensus, and the comparative phylogeography of codistributed species in California. *The American Naturalist*. 166, 290-299. https://doi.org/10.1086/431283

Leigh JW, Bryant D. 2015. PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lischer HEL, Excoffier L. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*. 28, 298-299. https://doi.org/10.1093/bioinformatics/btr642

Loxdale H, Lushai G, Harvey JA. 2010. The evolutionary improbability of 'generalism' in nature, with special reference to insects. *Biological Journal of the Linnean Society.* 103, 1-18. https://doi.org/10.1111/j.1095-8312.2011.01627.x

Lumley LM, Pouliot E, Laroche J, Boyle B, Brunet BMT, Levesque RC, Sperling FAH, Cusson M. 2020. Continent-wide population genomic structure and phylogeography of North America's most destructive conifer defoliator, the spruce budworm (*Choristoneura fumiferana*). *Ecology and Evolution*. 10, 914-927. https://doi.org/10.1002/ece3.5950

Maddison WP, Maddison DR. 2019. Mesquite: a modular system for evolutionary analysis. Version 3.61 http://www.mesquiteproject.org

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 26, 2867-2873. https://doi.org/10.1093/bioinformatics/btq559

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 17, 10-12. https://doi.org/10.14806/ej.17.1.200.

Matsubayashi KW, Kahono S, Katakura H. 2011. Divergent host plant specialization as the critical driving force in speciation between populations of a phytophagous ladybird beetle. Jo*urnal of Evolutionary Biology*. 24, 1421–1432. https://doi.org/10.1111/j.1420-9101.2011.02274.x

McClure M, Cannell E, Despland E. 2010. Thermal ecology and behavior of the nomadic social forager *Malacosoma disstria. Physiological Entomology*. 36, 120-127. https://doi.org/10.1111/j.1365-3032.2010.00770.x

McClure M, Despland E. 2010. Collective foraging patterns of field colonies of *Malacosoma disstria* caterpillars. *The Canadian Entomologist.* 142, 473-480. https://doi.org/10.4039/n10-001

Medina RF, Reyna SM, Bernal JS. 2012. Population genetic structure of a specialist leafhopper on *Zea*: likely anthropogenic and ecological determinants of gene flow. *Entomologia Experimentalis et Applicata*. 142, 223-235. https://doi.org/10.1111/j.1570-7458.2012.01220.x

Meirmans PG. Subsampling reveals that unbalanced sampling affects Structure results in a multi-species dataset. *Heredity*. 122, 276-287. https://doi.org/ 10.1038/s41437-018-0124-8

Meseguer AS, Coeur D'acier A, Genson G, Jousselin E. 2015. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus. *Journal of Biogeography*. 42, 1482-1492. https://doi.org/10.1111/jbi.12531

Miller WE. 2006. Forest tent caterpillar: mating, oviposition, and adult congregation at town lights during a northern Minnesota outbreak. *Journal of the Lepidopterists' Society*. 60, 156-160

Murton JB, Bateman MD, Dallimore SR, Teller JT, Yang Z. 2010. Identification of Younger Dryas outburst flood path from Lake Agassiz to the Arctic Ocean. *Nature.* 464, 740-743. https://doi.org/10.1038/nature08954

Nakadai R. 2017. Species diversity of herbivorous insects: a brief review to bridge the gap between theories focusing on the generation and maintenance of diversity. *Ecological Research*. 32, 811-819. https://doi.org/10.1007/s11284-017-1500-1

Nazareno AG, Bemmels JB, Dick CW, Lohmann LG. 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*. 17, 1136-1147. https://doi.org/10.1111/1755-0998.12654

Nicol RW, Arnason JT, Helson B, Abou-Zaid MM. 1997. Effect of host and non-host trees on the growth and development of the forest tent caterpillar, *Malacosoma disstria* (Lepidoptera: Lasiocampidae). *The Canadian Entomologist*. 129, 991-999. https://doi.org/10.4039/Ent129991-6

Niese JN, Strong TF. 1992. Economic and tree diversity trade-offs in managed northern hardwoods. *Canadian Journal of Forest Research*. 22, 1807-1813. https://doi.org/10.1139/x92-236

Niquidet K, Tang J, Peter B. 2016. Economic analysis of forest insect pests in Canada. *The Canadian Entomologist*. 148, 357-366. https://doi.org/10.4039/tce.2015.27

Nixon AE, Roland J. 2012. Generalist predation on forest tent caterpillar varies with forest stand composition: an experimental study across multiple life stages. *Ecological Entomology*. 37, 13-23. https://doi.org/ 10.1111/j.1365-2311.2011.01330.x

Nosil P, Crespi BJ. 2006. Experimental evidence that predation promotes divergence in adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America*. 103, 9090-9095. https://doi.org/10.1073/pnas.0601575103

Nosil P. 2012. *Ecological Speciation*. Oxford (UK): Oxford University Press. 280 p.

Nyman T, Farrell BD, Zinovjev AG, Vikberg V. 2006. Larval habits, host-plant associations and speciation in nematine sawflies (Hymenoptera: Tenthredinindae). *Evolution*. 60, 1622-1637. https://doi.org/10.1554/05-674.1

Ohshima I. 2007. Host race formation in the leaf-mining moth *Acrocercops transecta* (Lepidoptera: Gracillariidae). *Biological Journal of the Linnean Society*. 93, 135-145. https://doi.org/10.1111/j.1095-8312.2007.00949.x

Panzuto M, Lorenzetti F, Mauffette Y, Albert PJ. 2001. Perception of aspen and sun/shade sugar maple leaf soluble extracts by larvae of *Malacosoma disstria*. *Journal of Chemical Ecology*. 27, 1963-1978. https://doi.org/10.1023/A:1012282618332

Pappers SM, Van der Velde G, Ouborg NJ. 2002. Host preference and larval performance suggest host race formation in *Galerucella nymphaeae*. *Oecologia*. 130, 433-440. https://doi.org/10.1007/s00442-001-0822-3

Parry D, Goyer RA, Lenhard GJ. 2001. Macrogeographic clines in fecundity, reproductive allocation, and offspring size of the forest tent caterpillar *Malacosoma disstria*. *Ecological Entomology*. 26, 381-291. https://doi.org/10.1046/j.1365-2311.2001.00319.x

Parry D, Goyer RA. 2004. Variation in the suitability of host tree species for geographically discrete populations of forest tent caterpillar. *Environmental Entomology*. 33, 1479-1487. https://doi.org/10.1603/0046-225X-33.5.1477

Parry D, Spence JR, Volney WJA. 1998. Budbreak phenology and natural enemies mediate survival of first-instar forest tent caterpillar (Lepidoptera: Lasiocampidae). *Environmental Entomology*. 27, 1368-1374. https://doi.org/10.1093/ee/27.6.1368

Peccoud J, Ollivier A, Plantegenest M, Simon J. 2009. A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proceedings of the National Academy of Sciences of the United States*. 106, 7495-7500. https://doi.org/10.1073/pnas.0811117106

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoesktra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*. 7, 1-11. https://doi.org/10.1371/journal.pone.0037135

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MV. 2013. An overview of Structure: applications, parameter settings, and supporting software. *Frontiers in Genetics*. 4, 1-13. https://doi.org/ 10.3389/fgene.2013.00098

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

Progar RA, Rinella MJ, Fekedulegn D, Butler L. 2010. Nuclear polyhedrosis virus as a biological control agent for *Malacosoma americanum* (Lepidoptera: Lasiocampidae). *Journal of Applied Entomology*. 134, 641-646. https://doi.org/10.1111/j.14390418.2010.01508.x

Rochefort S, Berthiaume R., Hebert C, Charest M, Bauce E. 2011. Effect of temperature and host tree on cold hardiness of hemlock looper eggs along a latitudinal gradient. *Journal of Insect Physiology*. 57, 751-759. https://doi.org/10.1016/j.jinsphys.2011.02.013

Rundle HD, Nosil P. 2005. Ecological speciation. *Ecology Letters*. 8, 336-352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Scheffer SJ, Hawthorne DJ. 2007. Molecular evidence of host-associated genetic divergence in the holly leafminer *Phytomyza glabricola* (Diptera: Agromyzidae): apparent discordance among marker systems. *Molecular Ecology*. 16, 2627-2637. https://doi.org/10.1111/j.1365-294X.2007.03303.x

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science*. 323, 737-742. https://doi.org/10.1126/science.1160006

Schmidt BC, Roland J, Wakarchuk D. 2003. Evaluation of synthetic pheromones for monitoring forest tent caterpillar (Lepidoptera: Lasiocampidae) populations. *Environmental Entomology*. 32, 214-219. https://doi.org/10.1603/0046-225X-32.1.214

Sexton JP, Strauss SY, Rice KJ. 2011. Gene flow increases fitness at the warm edge of a species' edge. *Proceedings of the National Academy of Sciences of the United States of America.* 108, 11704-11709. https://doi.org/10.1073/pnas.1100404108

Shafer ABA, Cullingham CI, Cote SD, Coltman DW. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*. 19, 4589-4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Shiojiri K, Takabayashi J. 2003. Effects of specialist parasitoids on oviposition preference of phytophagous insects: encounter-dilution effects in a tri-trophic interaction. *Ecological Entomology*. 28, 573-578. https://doi.org/10.1046/j.1365-2311.2003.00539.x

Silva-Brandao KL, Silva OABN, Brandao MM, Omoto C, Sperling FAH. 2015. Genotyping by sequencing approach indicates geographic distance as the main factor affecting genetic structure and gene flow in Brazilian populations of *Grapholita molesta*. *Evolutionary Applications*. 8, 476-485. https://doi.org/10.1111/eva.12257

Sipple WL. 1962. Outbreaks of the forest tent caterpillar, *Malacosoma disstria* Hbn., a periodic defoliator of broad-leaved trees in Ontario. *The Canadian Entomologist*. 94, 408-416. https://doi.org/10.4039/Ent94408-4

Sniegula S, Drobniak SM, GołaB MJ, Johansson F. 2014. Photoperiod and variation in life history traits in core and peripheral populations in the damselfly *Lestes sponsa*. *Ecological Entomology*. 39, 137-148. https://doi.org/10.1111/een.12076

Sperling FAH. 1993. Mitochondrial DNA variation and Haldane's rule in the *Papilio glaucus* and *P. troilus* species groups. *Heredity*. 71, 227-233. https://doi.org/10.1038/hdy.1993.130

Stark EJ, Harper JD. 1982. Pupal mortality in forest tent caterpillar (Lepidoptera: Lasiocampidae): causes and impact on populations in southwestern Alabama. *Environmental Entomology*. 11, 1071-1077. https://doi.org/10.1093/ee/11.5.1071

Stireman III JO, Nason JD, Heard SB. 2005. Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod-insect community. *Evolution.* 60, 2573-2587. https://doi.org/10.1111/j.0014-3820.2005.tb00970.x

Stireman III JO, Singer MS. 2018. Tri-trophic niches of insect herbivores in an era of rapid environmental change. *Current Opinion in Insect Science*. 29, 117-125. https://doi.org/10.1016/j.cois.2018.07.008

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence Data. III. Cladogram estimation. *Genetics.* 132, 619-633.

Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*. 21, 3907-3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Trudeau M, Mauffette Y, Rochefort S, Han E, Bauce E. 2010. Impact of host tree on forest tent caterpillar performance and offspring overwintering mortality. *Environmental Entomology*. 39, 498-504. https://doi.org/10.1603/en09139

van Els P, Cicero C, Klicka J. 2012. High latitudes and high genetic diversity: phylogeography of a widespread boreal bird, the gray jay (*Perisoreus canadensis*). *Molecular Phylogenetics and Evolution*. 63, 456-465. https://doi.org/10.1016/j.ympev.2012.01.019

Vidal MC, Quinn TW, Stireman JO, Tinghitella RM, Murphy SM. 2019. Geography is more important than host plant use for the population genetic structure of a generalist insect herbivore. *Molecular Ecology*. 28, 4317-4334. https://doi.org/10.1111/mec.15218

Wellenreuther M, Sanchez-Guillen RA, Cordero-Rivera A, Svensson EI, Hansson B. 2011. Environmental and climatic determinants of molecular diversity and genetic population structure in a coenagrionid damselfly. *PLoS One*. 6, 1-16. https://doi.org/10.1371/journal.pone.0020440

Werner SM, Raffa KF. 2000. Effects of forest management practices on the diversity of ground-occurring beetles in mixed northern hardwood forests of the Great Lakes region. *Forest Ecology and Management*. 139, 135-155. https://doi.org/10.1016/S0378-1127(99)00341-2

Wickham H. 2016. *ggplot2: elegant graphics for data analysis*. New York (NY): Springer-Verlag. 253 p.

Williams DJ, Langor DW. 2011. Distribution, species composition, and incidence of egg parasitoids of the forest tent caterpillar (Lepidoptera: Lasiocampidae), during a widespread outbreak in the Canadian prairies. *The Canadian Entomologist*. 143, 272-278. https://doi.org/10.4039/n11-009

Witter JA, Kulman HM, Hodson AC. 1972. Life tables for the forest tent caterpillar. *Annals of the Entomological Society of America*. 65, 25-31. https://doi.org/10.1093/aesa/65.1.25

Yu H, Liang D, Tian E, Zheng L, Kjellberg F. 2018. Plant geographic phenotypic variation drives diversification in its associated community of a phytophagous insect and its parasitoids. *BMC Evolutionary Biology*. 18, 1-11. https://doi.org/10.1186/s12862-018-1239-5

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 28, 3326-3328. https://doi.org/10.1093/bioinformatics/bts606

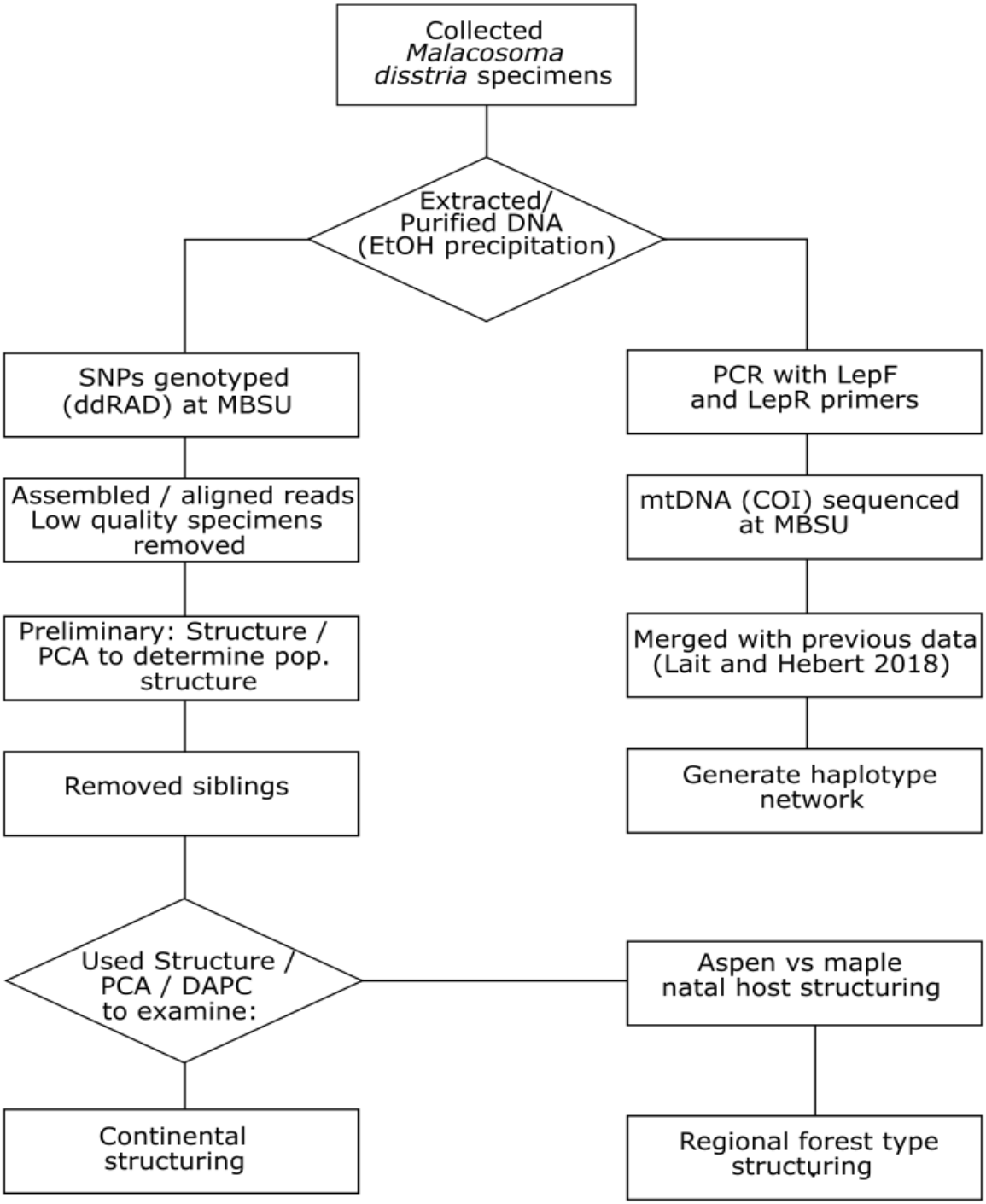
Appendix 2.1 *M. disstria* collection event data. Filtered *M. disstria* specimens only. Associated specimen ID are in Appendix 2.2. Associated Genbank ID are in Appendix 2.4.

Prov	Locality	Collection point	Lat	Long	Host	Collection date	Collector
AB	Lac la Biche	Lac La Biche	54.421	-111.571	Aspen	02-May-18	Tyler Nelson, Kyle Snape
SK	Saskatoon	Saskatoon	n/a	n/a	Aspen	May-18	Tyler Wist
ON	Kenora	5th St. local park	49.759	-94.476	Aspen	11-May-18	Kirstin Hicks
ON	Constance Lake	Calstock Bypass (6 km NW of Constance Lake)	49.814	-84.189	Aspen	May-18	Lia Fricano
ON	Hearst	O'Connor Road, 5 km S Hwy 11	49.731	-83.913	Aspen	09-May-18	Lia Fricano
ON	Kapuskasing	33 km north on Fred Flatt Road	49.644	-82.302	Aspen	08-May-18	Lia Fricano
QC	Lac Duparquet	Lac Duparquet	48.5	-79.2	Aspen	Jun-18	Joshua Jarry
ON	Latchford	Latchford	47.333	-79.81	Aspen	May-18	Chris McVeety
ON	Rabbit Lake	Rabbit Lake Rd	46.93	-79.726	Aspen	11-May-18	Chris McVeety
ON	Marten River	Bidwell Rd off Hwy 11	46.668	-79.728	Aspen	May-18	Chris McVeety
ON	Nairn	Sand Bay Road	46.302	-81.678	Aspen	08-May-18	Ariel Ilic
ON	Killarney	Killarney Prov Park	46.011	-81.401	Birch	May-18	Ariel Ilic
ON	Little Current	Batman Campgr.	45.879	-81.899	Aspen	May-18	Ariel Ilic
ON	Elliot Lake	Depot Lake	46.339	-82.542	Aspen	09-May-18	Mike Francis
ON	Wharncliffe	Jobammageeshig Lake, Hwy 129	46.538	-83.437	Maple	01-May-18	Mike Francis
ON	Sault Ste Marie (1)	Grand View Public School	46.51	-84.267	Aspen	07-Jun-18	Amanda Roe, Kyle Snape
ON	Sault Ste Marie (2)	Queen Elizabeth Public School	46.506	-84.306	Maple	07-Jun-18	Amanda Roe, Kyle Snape
ON	Sault Ste Marie (3)	John Rhodes Ctr.	46.508	-84.302	Oak	07-Jun-18	Amanda Roe, Kyle Snape

ON	St Joseph Island (1)	St Joseph Island	46.192	-84.042	Aspen	03-Jun-18	Kyle Snape, Reshma Jose
ON	St Joseph Island (2)	St Joseph Island	46.192	-84.042	Maple	03-Jun-18	Kyle Snape, Reshma Jose
ON	Parry Sound	Kinsmen Park	45.378	-80.044	Maple	09-May-18	Ariel Ilic
ON	Foots Bay	Chown Road	45.166	-79.76	Aspen	May-18	Ariel Ilic
ON	Oak Shores	221 Kennedy Dr	44.588	-78.428	Oak	29-May-18	Vanessa Chaimbrone
ON	Bancroft	2798 Old Hastings Rd	44.887	-77.747	Maple	16-May-18	Vanessa Chaimbrone
ON	Lanark (1)	Christie Lake	44.82	-76.45	Aspen	05-Jun-18	Christi Jaeger
ON	Lanark (2)	Christie Lake	44.82	-76.45	Maple	05-Jun-18	Christi Jaeger
ON	Ottawa (1)	Ottawa	n/a	n/a	Aspen	May-18	Unknown
ON	Ottawa (2)	Ottawa	n/a	n/a	Maple	May-18	Unknown
QC	Montebello	Kenauk Resort	45.7	-74.8	Maple	Jun-18	Anne-Sophie Caron

Appendix 2.2 *M. disstria* specimen ID list. Filtered *M. disstria* specimens only. Associated collection sites are in Appendix 2.1. Associated Genbank ID are in Appendix 2.4.

Collection Locality	Specimen ID
Lac la Biche	KS059, KS063, KS092, KS102, KS111, KS114, KS118, KS125, KS131, KS134, KS158, KS197, KS203, KS212
Saskatoon	KS140, KS189, KS230, KS239
Kenora	KS168, KS174
Constance Lake	KS047
Hearst	KS069, KS074, KS152, KS162, KS207
Kapuskasing	KS076, KS106, KS149, KS167, KS190
Lac Duparquet	KS229
Latchford	KS077, KS095, KS148, KS200
Rabbit Lake	KS009, KS013, KS018, KS055, KS057, KS139, KS143
Marten River	KS096, KS113, KS138, KS188, KS201
Nairn	KS003, KS006, KS016, KS037, KS072, KS159
Killarney	KS085, KS186, KS194, KS196
Little Current	KS101, KS218
Elliot Lake	KS004, KS017, KS058, KS060, KS166
Wharncliffe	KS214
Sault Ste Marie (1)	KS068, KS070, KS078, KS080, KS081, KS087, KS089, KS091, KS097, KS100, KS112, KS132, KS133, KS183, KS192
Sault Ste Marie (2)	KS082, KS124, KS126, KS141, KS198, KS216, KS217
Sault Ste Marie (3)	KS079, KS104, KS116, KS215
St Joseph Island (1)	KS107, KS115, KS117, KS129, KS136, KS170, KS213
St Joseph Island (2)	KS088, KS090, KS161, KS184, KS205, KS234
Parry Sound	KS086, KS157
Foots Bay	KS012, KS050, KS083, KS144, KS163, KS219
Oak Shores	KS015, KS120
Bancroft	KS228, KS241
Lanark (1)	KS019, KS025, KS045
Lanark (2)	KS001, KS022, KS027, KS034
Ottawa (1)	KS044
Ottawa (2)	KS041
Montebello	KS051, KS147, KS169, KS233



Appendix 2.3 A simplified flowchart of our laboratory and analytical methods.

Appendix 2.4 Genbank ID for *M. disstria* collection data. Filtered *M. disstria* specimens only. Associated collection event data is in Appendix 2.1 and collection ID are in Appendix 2.2.

Genbank ID	Specimen ID	Prov	Locality	Lat	Long
MT791498	KS001	ON	Lanark	44.82	-76.45
MT791499	KS003	ON	Nairn	46.302	-81.678
MT791500	KS004	ON	Elliot	46.339	-82.542
MT791501	KS006	ON	Nairn	46.302	-81.678
MT791502	KS009	ON	Rabbit Lake	46.93	-79.726
MT791503	KS012	ON	Foots Bay	45.166	-79.76
MT791504	KS013	ON	Rabbit Lake	46.93	-79.726
MT791505	KS015	ON	Oak Shores	44.588	-78.428
MT791506	KS016	ON	Nairn	46.302	-81.678
MT791507	KS017	ON	Elliot	46.339	-82.542
MT791508	KS018	ON	Rabbit Lake	46.93	-79.726
MT791509	KS019	ON	Lanark	44.82	-76.45
MT791510	KS022	ON	Lanark	44.82	-76.45
MT791511	KS025	ON	Lanark	44.82	-76.45
MT791512	KS027	ON	Lanark	44.82	-76.45
MT791513	KS034	ON	Lanark	44.82	-76.45
MT791514	KS037	ON	Nairn	46.302	-81.678
MT791515	KS041	ON	Ottawa	n/a	n/a
MT791516	KS044	ON	Ottawa	n/a	n/a
MT791517	KS045	ON	Lanark	44.82	-76.45
MT791518	KS047	ON	Constance Lake	49.814	-84.189
MT791519	KS050	ON	Foots Bay	45.166	-79.76
MT791520	KS051	QC	Montebello	45.7	-74.8
MT791521	KS055	ON	Rabbit Lake	46.93	-79.726
MT791522	KS057	ON	Rabbit Lake	46.93	-79.726
MT791523	KS058	ON	Elliot	46.339	-82.542
MT791524	KS059	AB	Lac La Biche	54.421	-111.571
MT791525	KS060	ON	Elliot	46.339	-82.542
MT791526	KS063	AB	Lac La Biche	54.421	-111.571
MT791527	KS068	ON	Sault Ste Marie	46.51	-84.267
MT791528	KS069	ON	Hearst	49.731	-83.913
MT791529	KS070	ON	Sault Ste Marie	46.51	-84.267
MT791530	KS072	ON	Nairn	46.302	-81.678
MT791531	KS074	ON	Hearst	49.731	-83.913
MT791532	KS076	ON	Kapuskasing	49.644	-82.302

MT791533	KS077	ON	Latchford	47.333	-79.81
MT791534	KS078	ON	Sault Ste Marie	46.51	-84.267
MT791535	KS079	ON	Sault Ste Marie	46.508	-84.302
MT791536	KS080	ON	Sault Ste Marie	46.51	-84.267
MT791537	KS081	ON	Sault Ste Marie	46.51	-84.267
MT791538	KS082	ON	Sault Ste Marie	46.506	-84.306
MT791539	KS083	ON	Foots Bay	45.166	-79.76
MT791540	KS085	ON	Killarney	46.011	-81.401
MT791541	KS086	ON	Parry Sound	45.378	-80.044
MT791542	KS087	ON	Sault Ste Marie	46.51	-84.267
MT791543	KS088	ON	St Joseph Island	46.192	-84.042
MT791544	KS089	ON	Sault Ste Marie	46.51	-84.267
MT791545	KS090	ON	St Joseph Island	46.192	-84.042
MT791546	KS091	ON	Sault Ste Marie	46.51	-84.267
MT791547	KS092	AB	Lac La Biche	54.421	-111.571
MT791548	KS095	ON	Latchford	47.333	-79.81
MT791549	KS096	ON	Marten River	46.668	-79.728
MT791550	KS097	ON	Sault Ste Marie	46.51	-84.267
MT791551	KS100	ON	Sault Ste Marie	46.51	-84.267
MT791552	KS101	ON	Little Current	45.879	-81.899
MT791553	KS102	AB	Lac La Biche	54.421	-111.571
MT791554	KS104	ON	Sault Ste Marie	46.508	-84.302
MT791555	KS106	ON	Kapuskasing	49.644	-82.302
MT791556	KS107	ON	St Joseph Island	46.192	-84.042
MT791557	KS111	AB	Lac La Biche	54.421	-111.571
MT791558	KS112	ON	Sault Ste Marie	46.51	-84.267
MT791559	KS113	ON	Marten River	46.668	-79.728
MT791560	KS114	AB	Lac La Biche	54.421	-111.571
MT791561	KS115	ON	St Joseph Island	46.192	-84.042
MT791562	KS116	ON	Sault Ste Marie	46.508	-84.302
MT791563	KS117	ON	St Joseph Island	46.192	-84.042
MT791564	KS118	AB	Lac La Biche	54.421	-111.571
MT791565	KS120	ON	Oak Shores	44.588	-78.428
MT791566	KS124	ON	Sault Ste Marie	46.506	-84.306
MT791567	KS125	AB	Lac La Biche	54.421	-111.571
MT791568	KS126	ON	Sault Ste Marie	46.506	-84.306
MT791569	KS129	ON	St Joseph Island	46.192	-84.042
MT791570	KS131	AB	Lac La Biche	54.421	-111.571
MT791571	KS132	ON	Sault Ste Marie	46.51	-84.267

MT791572	KS133	ON	Sault Ste Marie	46.51	-84.267
MT791573	KS134	AB	Lac La Biche	54.421	-111.571
MT791574	KS136	ON	St Joseph Island	46.192	-84.042
MT791575	KS138	ON	Marten River	46.668	-79.728
MT791576	KS139	ON	Rabbit Lake	46.93	-79.726
MT791577	KS140	SK	Saskatoon	n/a	n/a
MT791578	KS141	ON	Sault Ste Marie	46.506	-84.306
MT791579	KS143	ON	Rabbit Lake	46.93	-79.726
MT791580	KS144	ON	Foots Bay	45.166	-79.76
MT791581	KS147	QC	Montebello	45.7	-74.8
MT791582	KS148	ON	Latchford	47.333	-79.81
MT791583	KS149	ON	Kapuskasing	49.644	-82.302
MT791584	KS152	ON	Hearst	49.731	-83.913
MT791585	KS157	ON	Parry Sound	45.378	-80.044
MT791586	KS158	AB	Lac La Biche	54.421	-111.571
MT791587	KS159	ON	Nairn	46.302	-81.678
MT791588	KS161	ON	St Joseph Island	46.192	-84.042
MT791589	KS162	ON	Hearst	49.731	-83.913
MT791590	KS163	ON	Foots Bay	45.166	-79.76
MT791591	KS166	ON	Elliot	46.339	-82.542
MT791592	KS167	ON	Kapuskasing	49.644	-82.302
MT791593	KS168	ON	Kenora	49.759	-94.476
MT791594	KS169	QC	Montebello	45.7	-74.8
MT791595	KS170	ON	St Joseph Island	46.192	-84.042
MT791596	KS174	ON	Kenora	49.759	-94.476
MT791597	KS183	ON	Sault Ste Marie	46.51	-84.267
MT791598	KS184	ON	St Joseph Island	46.192	-84.042
MT791599	KS186	ON	Killarney	46.011	-81.401
MT791600	KS188	ON	Marten River	46.668	-79.728
MT791601	KS189	SK	Saskatoon	n/a	n/a
MT791602	KS190	ON	Kapuskasing	49.644	-82.302
MT791603	KS192	ON	Sault Ste Marie	46.51	-84.267
MT791604	KS194	ON	Killarney	46.011	-81.401
MT791605	KS196	ON	Killarney	46.011	-81.401
MT791606	KS197	AB	Lac La Biche	54.421	-111.571
MT791607	KS198	ON	Sault Ste Marie	46.506	-84.306
MT791608	KS200	ON	Latchford	47.333	-79.81
MT791609	KS201	ON	Marten River	46.668	-79.728
MT791610	KS203	AB	Lac La Biche	54.421	-111.571

MT791611	KS205	ON	St Joseph Island	46.192	-84.042
MT791612	KS207	ON	Hearst	49.731	-83.913
MT791613	KS212	AB	Lac La Biche	54.421	-111.571
MT791614	KS213	ON	St Joseph Island	46.192	-84.042
MT791615	KS214	ON	Wharncliffe	46.538	-83.437
MT791616	KS215	ON	Sault Ste Marie	46.508	-84.302
MT791617	KS216	ON	Sault Ste Marie	46.506	-84.306
MT791618	KS217	ON	Sault Ste Marie	46.506	-84.306
MT791619	KS218	ON	Little Current	45.879	-81.899
MT791620	KS219	ON	Foots Bay	45.166	-79.76
MT791621	KS228	ON	Bancroft	44.887	-77.747
MT791622	KS229	QC	Duparquet	48.5	-79.2
MT791623	KS230	SK	Saskatoon	n/a	n/a
MT791624	KS233	QC	Montebello	45.7	-74.8
MT791625	KS234	ON	St Joseph Island	46.192	-84.042
MT791626	KS239	SK	Saskatoon	n/a	n/a
MT791627	KS241	ON	Bancroft	44.887	-77.747

Appendix 2.5 BOLD ID for *M. disstria* collection data from Lait and Hebert 2018.

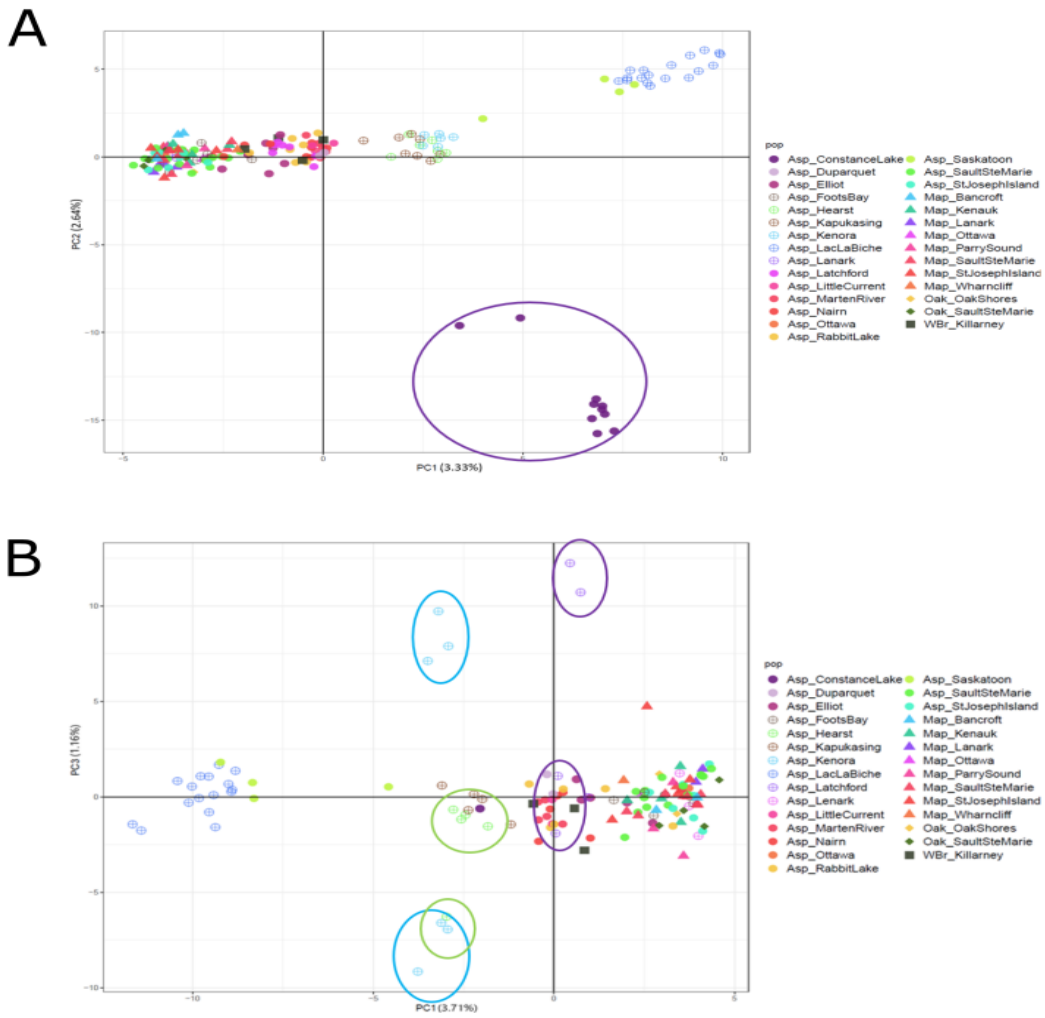
Species	BOLD ID	Prov/State	Lat	Long
M. disstria	BIOUG03567-G01	Alberta	49.11	-113.82
M. disstria	BIOUG03610-F10	Alberta	49.08	-113.88
M. disstria	08BBLEP-02575	Alberta	49.08	-113.88
M. disstria	08BBLEP-02576	Alberta	49.08	-113.88
M. disstria	08BBLEP-02578	Alberta	49.08	-113.88
M. disstria	08BBLEP-02588	Alberta	49.08	-113.88
M. disstria	08BBLEP-03307	Alberta	49.05	-113.91
M. disstria	08BBLEP-03335	Alberta	49.05	-113.91
M. disstria	08BBLEP-04315	Alberta	49.11	-113.84
M. disstria	08BBLEP-04366	Alberta	49.11	-113.84
M. disstria	08BBLEP-04893	Alberta	49.08	-113.88
M. disstria	08BBLEP-04924	Alberta	49.08	-113.88
M. disstria	08BBLEP-04925	Alberta	49.08	-113.88
M. disstria	08BBLEP-04926	Alberta	49.08	-113.88
M. disstria	08BBLEP-04931	Alberta	49.08	-113.88
M. disstria	08BBLEP-04934	Alberta	49.08	-113.88
M. disstria	08BBLEP-04935	Alberta	49.08	-113.88
M. disstria	08BBLEP-04939	Alberta	49.08	-113.88
M. disstria	08BBLEP-02277	Alberta	49.08	-113.88
M. disstria	08BBLEP-02316	Alberta	49.08	-113.88
M. disstria	10BBCLP-0424	British Columbia	50.99	-118.16
M. disstria	10BBCLP-0429	British Columbia	50.63	-116.06
M. disstria	10BBCLP-0430	British Columbia	50.99	-118.16
M. disstria	10BBCLP-0434	British Columbia	50.99	-118.16
M. disstria	HLC-20368	British Columbia	51.41	-117.48
M. disstria	HLC-21131	British Columbia	51.02	-118.21
M. disstria	HLC-21135	British Columbia	51.02	-118.21
M. disstria	HLC-21137	British Columbia	51.02	-118.21
M. disstria	HLC-21138	British Columbia	51.02	-118.21

M. disstria	HLC-21875	British Columbia	51.09	-117.92
M. disstria	CGWC-0495	British Columbia	51.95	-122.4
M. disstria	CGWC-0501	British Columbia	51.92	-122.29
M. disstria	CGWC-0503	British Columbia	52.58	-122.2
M. disstria	CGWC-0504	British Columbia	52.58	-122.2
M. disstria	CGWC-0505	British Columbia	52.58	-122.2
M. disstria	CGWC-0506	British Columbia	52.58	-122.2
M. disstria	BIOUG03567-F11	Manitoba	50.68	-99.9
M. disstria	moth58.02SA	New Brunswick	45.08	-67.07
M. disstria	moth87.02SA	New Brunswick	45.08	-67.07
M. disstria	04HBL00794	New Brunswick	45.08	-67.07
M. disstria	Moth 412.03SA	New Brunswick	45.08	-67.07
M. disstria	Moth 459.03SA	New Brunswick	45.08	-67.07
M. disstria	MNBTT-3236	New Brunswick	45.92	-66.63
M. disstria	MNBTT-3237	New Brunswick	45.92	-66.63
M. disstria	MNBTT-3238	New Brunswick	45.92	-66.63
M. disstria	MNBTT-3239	New Brunswick	45.92	-66.63
M. disstria	MNBTT-3240	New Brunswick	45.92	-66.63
M. disstria	MNBTT-908	New Brunswick	46	-66.18
M. disstria	MNBTT-909	New Brunswick	46	-66.18
M. disstria	09BBELE-0267	Nova Scotia	46.81	-60.77
M. disstria	09BBELE-0302	Nova Scotia	46.81	-60.77
M. disstria	09BBELE-0598	Nova Scotia	46.83	-60.61
M. disstria	09BBELE-0284	Nova Scotia	46.81	-60.77
M. disstria	BIOUG09673-F10	Ontario	44.85	-79.87
M. disstria	BIOUG09852-C06	Ontario	44.85	-79.87
M. disstria	BIOUG10646-F02	Ontario	44.85	-79.87
M. disstria	BIOUG10646-F03	Ontario	44.85	-79.87
M. disstria	BIOUG10646-F04	Ontario	44.85	-79.87
M. disstria	BIOUG10646-F05	Ontario	44.85	-79.87
M. disstria	BIOUG11791-F05	Ontario	44.85	-79.87

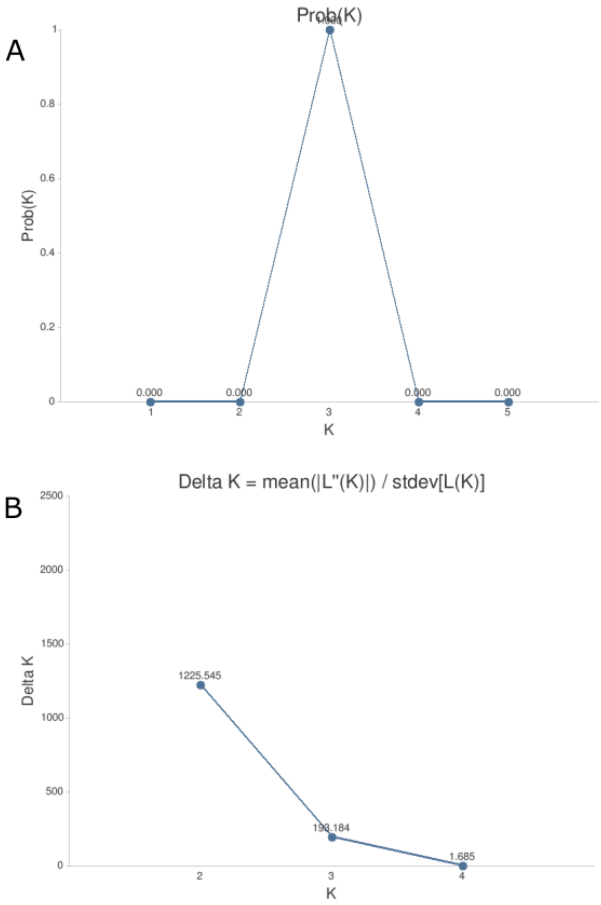
M. disstria	BIOUG11791-F06	Ontario	44.85	-79.87
M. disstria	moth994.01	Ontario	43.54	-80.13
M. disstria	BIOUG22324-G02	Ontario	43.37	-80.36
M. disstria	BIOUG22569-B02	Ontario	43.37	-80.36
M. disstria	BIOUG22569-B03	Ontario	43.37	-80.36
M. disstria	BIOUG22569-B04	Ontario	43.37	-80.36
M. disstria	BIOUG22569-B05	Ontario	43.37	-80.36
M. disstria	BIOUG21896-A08	Ontario	43.37	-80.35
M. disstria	04HBL005114	Ontario	43.54	-80.13
M. disstria	04HBL005145	Ontario	43.54	-80.13
M. disstria	04HBL005146	Ontario	43.54	-80.13
M. disstria	04HBL005206	Ontario	43.54	-80.13
M. disstria	2005-ONT-554	Ontario	44.53	-77
M. disstria	2005-ONT-560	Ontario	44.53	-77
M. disstria	2005-ONT-561	Ontario	44.53	-77
M. disstria	2005-ONT-569	Ontario	44.53	-77
M. disstria	2005-ONT-570	Ontario	44.53	-77
M. disstria	2005-ONT-571	Ontario	44.53	-77
M. disstria	2005-ONT-572	Ontario	44.53	-77
M. disstria	2005-ONT-578	Ontario	44.53	-77
M. disstria	2005-ONT-579	Ontario	44.53	-77
M. disstria	Moth4379.03	Ontario	43.54	-80.13
M. disstria	Moth4416.03	Ontario	43.54	-80.13
M. disstria	Moth4432.03	Ontario	43.54	-80.13
M. disstria	Moth4516.03	Ontario	43.54	-80.13
M. disstria	Moth4603.03	Ontario	43.54	-80.13
M. disstria	2006-ONT-0835	Ontario	44.53	-77
M. disstria	2006-ONT-0893	Ontario	44.53	-77
M. disstria	2006-ONT-0920	Ontario	44.53	-77
M. disstria	2006-ONT-1441	Ontario	44.53	-77
M. disstria	2006-ONT-1442	Ontario	44.53	-77

M. disstria	BIOUG06756-F07	Ontario	45.52	-78.42
M. disstria	BIOUG06756-D09	Ontario	45.46	-78.8
M. disstria	BL606	Ontario	n/a	n/a
M. disstria	BL696	Ontario	n/a	n/a
M. disstria	BL816	Ontario	n/a	n/a
M. disstria	LEP041295	Ontario	n/a	n/a
M. disstria	LEP041296	Ontario	n/a	n/a
M. disstria	BIOUG23415-H08	Ontario	43.37	-80.36
M. disstria	BIOUG23203-C08	Ontario	43.37	-80.35
M. disstria	DH001568	Quebec	45.69	-73.09
M. disstria	DH005298	Quebec	n/a	n/a
M. disstria	BIOUG03567-H08	Saskatchewan	53.85	-106.08
M. disstria	BIOUG03567-H09	Saskatchewan	53.85	-106.08
M. disstria	BIOUG03567-H10	Saskatchewan	53.85	-106.08
M. disstria	BIOUG03567-H11	Saskatchewan	53.85	-106.08
M. disstria	BIOUG04567-A02	Saskatchewan	53.85	-106.08
M. disstria	BIOUG04567-A06	Saskatchewan	53.85	-106.08
M. disstria	BIOUG04566-A07	Saskatchewan	53.85	-106.08
M. disstria	BIOUG04566-A08	Saskatchewan	53.85	-106.08
M. disstria	10BBLEP-00512	Arkansas	35.08	-92.55
M. disstria	10BBLEP-00513	Arkansas	35.37	-93.34
M. disstria	10BBLEP-00514	Arkansas	35.37	-93.34
M. disstria	10BBLEP-00515	Arkansas	35.37	-93.34
M. disstria	CNCNoctuoidea13836	Florida	n/a	n/a
M. disstria	06-JKA-0702	Georgia	n/a	n/a
M. disstria	06-SUSA-0166	Kentucky	37.01	-88.54
M. disstria	06-SUSA-0173	Kentucky	37.01	-88.54
M. disstria	06-SUSA-0175	Kentucky	37.01	-88.54
M. disstria	06-NCCC-1094	North Carolina	34.77	-76.76
M. disstria	06-NCCC-1095	North Carolina	34.77	-76.76
M. disstria	06-NCCC-1096	North Carolina	34.77	-76.76

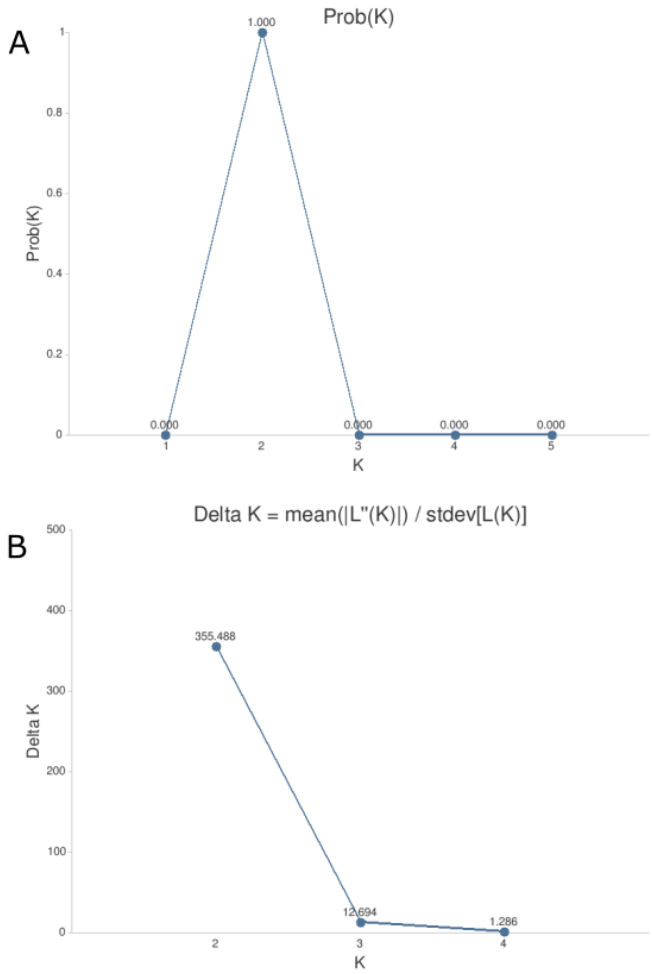
M. disstria	MDOK-2057	Oklahoma	36.74	-95.95
M. disstria	MDOK-1495	Oklahoma	36.74	-95.95
M. disstria	MDOK-2368	Oklahoma	36.74	-95.95
M. disstria	MDOK-2411	Oklahoma	36.74	-95.95
M. disstria	MDOK-2415	Oklahoma	36.74	-95.95
M. disstria	BIOUG02884-A02	Tennessee	35.69	-83.5
M. disstria	BIOUG02884-A03	Tennessee	35.69	-83.5
M. disstria	BIOUG02884-A04	Tennessee	35.69	-83.5
M. disstria	BIOUG02884-A05	Tennessee	35.69	-83.5
M. disstria	BIOUG02884-A07	Tennessee	35.69	-83.5
M. disstria	BIOUG02884-A11	Tennessee	35.69	-83.5
M. disstria	BIOUG03567-H02	Tennessee	35.69	-83.5
M. disstria	BIOUG03567-H04	Tennessee	35.69	-83.5
M. disstria	BIOUG03567-H05	Tennessee	35.69	-83.5
M. disstria	BIOUG03567-H07	Tennessee	35.69	-83.5
M. disstria	TAMUICEGR-0827	Texas	30.59	-96.25
M. disstria	TAMUICEGR-0828	Texas	30.59	-96.25



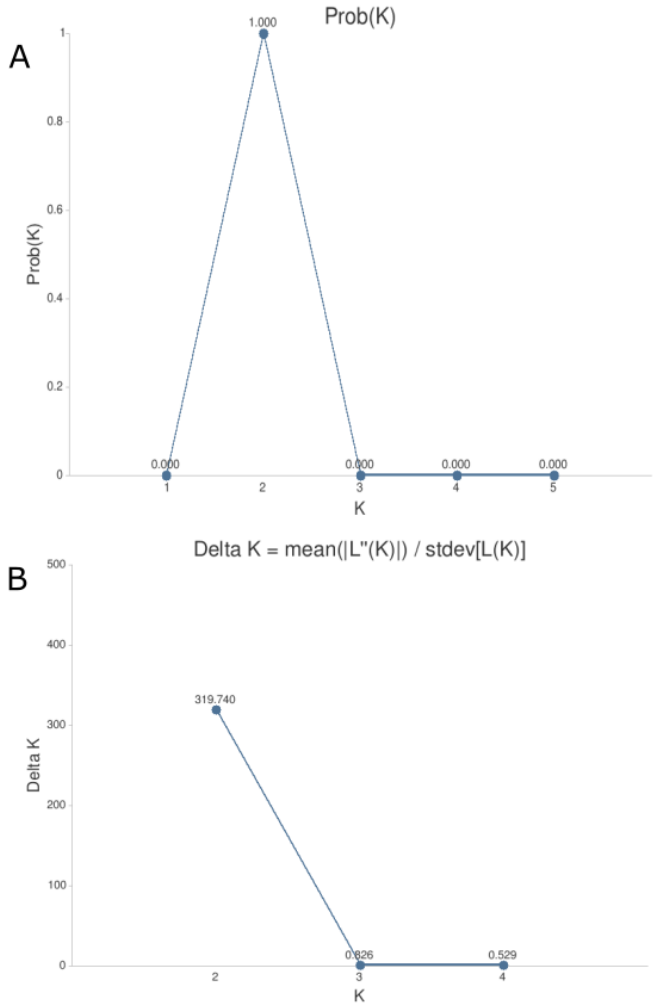
Appendix 2.6 (A) Preliminary principal component analysis (PC1 vs PC2) of SNP variation in *M. disstria* samples from all sites (Alberta to Quebec) and all hosts, siblings left included. Note that AB and SK samples are the most divergent. The grouping of Constance Lake specimens was an effect of the families of siblings included in the data (in ellipses). (B) Preliminary principal component analysis (PC1 vs PC3) of SNP variation in *M. disstria* samples from of all sites (Alberta to Quebec) and all hosts, siblings left included. Note that Kenora, Hearst and Latchford each have 2 clusters on PC3 (in ellipses).



Appendix 2.7 K support for continental population structuring. (A) Pritchard method and (B) Evanno method. The Pritchard method assumes there are an unknown number of clusters present, and individuals are assigned to population groups. The likelihood of K is determined through multiple iterations of the program, then plotting the average of the estimated natural log. The Evanno method, on the other hand, is useful for assessing K, the likely number of populations present. This method, however, describes the uppermost clustering level possible and can lead to the K=2 conundrum (Janes et al. 2017).



Appendix 2.8 K support for larval host races. (A) Pritchard method and (B) Evanno method. The Pritchard method assumes there are an unknown number of clusters present, and individuals are assigned to population groups. The likelihood of K is determined through multiple iterations of the program, then plotting the average of the estimated natural log. The Evanno method, on the other hand, is useful for assessing K, the likely number of populations present. This method, however, describes the uppermost clustering level possible and can lead to the K=2 conundrum (Janes et al. 2017).



Appendix 2.9 K support for regional population structuring. (A) Pritchard method and (B) Evanno method. The Pritchard method assumes there are an unknown number of clusters present, and individuals are assigned to population groups. The likelihood of K is determined through multiple iterations of the program, then plotting the average of the estimated natural log. The Evanno method, on the other hand, is useful for assessing K, the likely number of populations present. This method, however, describes the uppermost clustering level possible and can lead to the K=2 conundrum (Janes et al. 2017).