University of Windsor

Scholarship at UWindsor

Electronic Theses and Dissertations | Theses, Dissertations, and Major Papers

2011

Speciation with gene flow in the Daphnia pulex species complex

Anna Constantin
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation

Constantin, Anna, "Speciation with gene flow in the Daphnia pulex species complex" (2011). *Electronic Theses and Dissertations*. 447.
https://scholar.uwindsor.ca/etd/447

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

Speciation with gene flow in the *Daphnia pulex* species complex

by

Anna Constantin

A Thesis
Submitted to the Faculty of Graduate Studies
through the Great Lakes Institute for Environmental Research
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2010

© 2010 Anna Constantin

Speciation with gene flow in the *Daphnia pulex* species complex

by

Anna Constantin

APPROVED BY:

____
T.J. Crease
Department of Integrative Biology, University of Guelph

____
T.E. Pitcher
Department of Biological Sciences, University of Windsor

____
D.D. Heath
Great Lakes Institute for Environmental Research, University of Windsor

____
M.E. Cristescu, Advisor
Great Lakes Institute for Environmental Research, University of Windsor

____
G.D. Haffner, Chair of Defense
Great Lakes Institute for Environmental Research, University of Windsor

October 14, 2010

DECLARATION OF CO-AUTHORSHIP/ PREVIOUS PUBLICATION

I. Co-Authorship Declaration

I hereby declare this thesis incorporates the outcome of a joint research undertaken in collaboration with T.J. Crease and C.E. Cáceres under the supervision of professor M.E. Cristescu. This collaboration is covered in Chapter 2 of the thesis. The key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contribution of co-authors was primarily in an advisory capacity.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-authors to include the above materials in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publication

Chapter 2 includes one original paper by A. Constantin, T.J. Crease, C.E. Cáceres and M.E. Cristescu titled "Ecological speciation with gene flow in the *Daphnia pulex* species complex" which will be submitted for publication in a peer reviewed journal.

I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owners to include such materials in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Divergent selection between contrasting habitats has the potential to drive adaptive divergence and the evolution of reproductive isolation in the face of initially high gene flow. This work explores the genetic divergence in a young ecological species pair, *Daphnia pulex* and *Daphnia pulicaria*, during habitat transition events, by surveying 363 individuals from 9 lakes and 8 ponds in Southern Ontario and Michigan. I conducted a phylogenetic and population genetics study using the mitochondrial *NADH dehydrogenase* 5 (*ND*5) gene, the nuclear *Lactate dehydrogenase* A (*Ldh-A)* locus, and 21 microsatellite markers. A discordant phylogenetic signal between nuclear and mitochondrial markers suggests a prolonged history of hybridization and introgression between lake and pond species. Population genetic analysis, based on nuclear markers, reflects a low level of contemporary gene flow, clear genetic differentiation between pond and lake populations, and additional substructure within lakes, suggesting the existence of strong habitat isolating barriers between ponds and lakes.

DEDICATION

To my mom and dad

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Melania Cristescu, for taking me on as her graduate student and encouraging me throughout these two years. She has taught me many things about science, the academic world, and how to balance work with family life. I also thank the members of my graduate committee, Trevor Pitcher and Daniel Heath, for their guidance and suggestions, and especially Teresa Crease for all her ideas and continuous feedback. Carla Cáceres and Michael Allen helped with much of my sampling and their knowledge of these habitats sparked much interest and stimulating discussions. Thank you also to Allyson Hutchens from the Litchman lab for help with sampling. All the undergraduate students and volunteers in the Cristescu lab have helped tremendously with *Daphnia* culturing and maintenance. Special thanks go to Dan Bock, James Vaillant, Ryan Walter, Sen Xu, and Aibin Zhan without whose knowledge and assistance this study would not have been successful. I am grateful to the rest of the Cristescu, Heath, and MacIsaac lab members for conversations that clarified my thinking on this and other matters, and especially for their friendship, which meant a great deal to me. Thank you to Russell Hepburn at the Environmental Genomics facility of the Great Lakes Institute for Environmental Research for performing all the sequencing and genotyping for my project and to Mary Lou Scratch for taking care of all administrative tasks and being so helpful in so many ways.

I would like to thank my parents and my mother-in-law for all their help and encouragement and especially thank my one and only, Alexandru Constantin, for supporting and encouraging me to pursue this degree. Without his love and encouragement, I would not have finished the degree.

TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP.. iii

ABSTRACT... iv

DEDICATION...v

ACKNOWLEDGEMENTS vi

LIST OF TABLES ix

LIST OF FIGURE…………………………………………………………………..xi

CHAPTER

I. INTRODUCTION

Ecological speciation with gene flow………………………………1

References……………………….…………………………………...5

II. SPECIATION WITH GENE FLOW IN *DAPHNIA PULEX* AND *DAPHNIA PULICARIA*

Introduction

Speciation in *Daphnia*…………………………………………..…9

Daphnia pulex and *Daphnia pulicaria*……………………..…11

Divergence and speciation between *Daphnia pulex* and *Daphnia pulicaria*……………………………………….…13

Materials and Methods

Sample collection……………. ………………………………15

Sexuality tests………………………………………………...…16

Mitochondrial DNA amplification………………………...….17

Nuclear *Lactate dehydrogenase* survey………………………18

Microsatellite survey…………………………...………..….18

Phylogenetic analyses……………………………….………..19

Population genetics analyses…………………………………..21

Results

Phylogenetic analysis………………………………………24

Genotypes at the *Ldh-A* locus……………………………...…25

Population genetics analysis……………………………….....25

Linkage disequilibrium……………………………………….26

Population differentiation and genetic distance………...........27

Discussion
Divergence with gene flow………………………………..…29
Discordance between nuclear *Ldh* and mitochondrial data.…..32
What multi locus nuclear data reveals about lake and
pond *Daphnia*…………………………..…………..……..33
Difference in linkage disequilibrium between lakes and
ponds…………………………………………………..…34
Population differentiation between the different habitats.........35
Discordance between nuclear and mitochondrial
phylogenies……………………………….……………...37
Conclusions…………………………………………..….38
References…………………………………………………..…72

III. CONCLUSIONS AND FUTURE DIRECTIONS
Identifying genomic regions under divergent selection……...…84
References...87

APPENDIX
Supplementary material 90

VITA AUCTORIS………………………………………………….……..91

LIST OF TABLES

Table 2.1A Habitat location and sampling size for pond *Daphnia pulex* populations. Habitat locations, mitochondrial and nuclear profiling for the study of 271 *Daphnia* isolates with *ID,* location code, *N1*, number of individuals analyzed using mitochondrial marker ; *N2*, number of individuals analyzed using 21 microsatellite markers and mitochondrial marker; *Ldh*, nuclear lactate dehydrogenase A profiling, SS- homozygous slow, FF- homozygous fast, SF-heterozygous; mtDNA, mitochondrial profiling, ppx, panarctic *Daphnia pulex*, wpc, western *Daphnia pulicaria,*; Cld, indicates which *Daphnia pulex* clade the individuals belong to in phylogenetic analysis based on mtDNA (figure 2.2); Rep, reproduction mode, CP-cyclical parthenogenesis, OP-obligate parthenogenesis; Prov/St, province or state of habitat, MI-Michigan, ON-Ontario…………………….....**39**

Table 2.1B Habitat location and sampling size for lake *Daphnia pulicaria* populations with *ID,* location code, *N1*, number of individuals analyzed using mitochondrial marker only; *N2*, number of individuals analyzed using 21 microsatellite markers and mitochondrial marker; *Ldh*, nuclear lactate dehydrogenase A profiling, FF- homozygous fast; mtDNA, mitochondrial profiling ppx, panarctic *Daphnia pulex*, wpc, western *Daphnia pulicaria,* epc, eastern *Daphnia pulicaria,*; Cld, indicates which *Daphnia pulex* clade the individuals belong to in the phylogenetic tree based on mtDNA (figure 2.2); Rep, reproduction mode, CP-cyclical parthenogenesis, OP-obligate parthenogenesis; Prov/St, province or state of habitat, MI-Michigan, ON-Ontario, IL-Illinois……..…….**40**

Table 2.2 Genetic diversity indexes for 2 pond and 3 lake populations based on a 687 bp sequence of the mitochondrial *NADH dehydrogenase* 5 gene...**41**

Table 2.3 Genetic diversity at 21 microsatellite loci for 6 populations of *Daphnia pulex* and *Daphnia pulicaria*. *N*, sample size; *A*, number of alleles; A_r, allele richness; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient; P_{HW}, exact *P*-value for Hardy-Weinberg equilibrium test; r, frequency of null allele,“-“ indicates no null allele detected. Values in bold indicate a deviation from HWE after

sequential Bonferroni correction. Abbreviations for the different populations are given in tables 2.1A, 2.1B……………………………………………………………………**42**

Table 2.4 Pairwise F_{ST} estimates between five populations of *Daphnia pulex* and *Daphnia pulicaria* based on 21 microsatellite loci. All F_{ST} values are significantly different from 0 (P<0.05). Abbreviations for the different populations are given in tables 2.1A, 2.1B...**49**

Table 2.5 Migration rates (Nm, number of migrants per generation) between lake and pond populations. Results are averaged over 21 microsatellite loci. Source populations are listed by column, recipient populations listed by row. Abbreviations for the different populations are given in tables 2.1A, 2.1B……………………………………………**50**

Table 2.6 Linkage disequilibrium between all pairs of loci in each population of *Daphnia pulex* and *Daphnia pulicaria*. Abbreviations for the different populations are given in tables 2.1A, 2.1B, with Law, all individuals from Lawrence Lake; Law1, Lawrence lake individuals with *Daphnia pulicaria* mitochondrial DNA; Law2, Lawrence lake individuals with *Daphnia pulex* mitochondria DNA. Values are P-values and those in bold are significant after Bonferroni correction…………………………………………**51**

Table 2.7 Linkage disequilibrium between 21 microsatellite loci and mitochondrial profiles of *Daphnia pulex* and *Daphnia pulicaria*. Abbreviations for the different populations are given in table 2.1, with Law, all individuals from Lawrence Lake; Law2, Lawrence lake individuals with *Daphnia pulex* mitochondria DNA; mtDNA, indicates LD between a given microsatellite locus and an association with any mitochondrial profile (either panarctic *D. pulex* clade or western *D. pulicaria* clade); clade A/B, indicates any LD between a microsatellite locus and either clade A or clade B (figure 2.2). Values are P-values and those in bold are significant after Bonferroni correction...**61**

LIST OF FIGURES

Figure 2.1 Distribution of collection sites in Michigan and Ontario. Lake habitats are denoted by black circles and pond habitats are denoted by white circles. The smaller circles represent habitats from which a low number of individuals were collected (3-14 individuals from each site) and only screened with the mitochondrial *NADH dehydrogenase 5* gene and *Lactate dehydrogenase* markers. Larger circles indicate populations that have been sampled intensely (>35 individuals from each site) and were screened with microsatellite, *LDH*, and mitochondrial markers......................................**64**

Figure 2.2 A Neighbor-joining (NJ) phylogeny of the mitochondrially encoded *NADH dehydrogenase 5* gene (*ND5*) of *Daphnia pulex* and *Daphnia pulicaria* from Illinois, Michigan, and Ontario. The label code of each haplotype identifies the haplotype number, it's location (see table 2.1), followed by *lactate dehydrogenase A* profile (FF-homozygous fast, SS-homozygous slow, SF-heterozygous slow/fast) and habitat (lake denoted by black circles, ponds denoted by white circles), * indicates populations reproducing by obligate parthenogenesis, all other populations reproduce by cyclical parthenogenesis. Numbers before and after dashes beside nodes represent bootstrap support with 1000 replicates and posterior probabilities, respectively. The tree was rooted with European *D. pulex* (GenBank accession number DQ235231). Only values $\geq$ 50% are shown. The small panel on the left represents a NJ *ND5* phylogeny of the *D. pulex* complex (Colbourne *et al*. 1998), branches in bold correspond to the two major clades in the phylogenetic reconstruction of this study..**65**

Figure 2.3 Unrooted statistical parsimony network of mitochondrial ND5 haplotypes of 363 *Daphnia pulex* and *Daphnia pulicaria* isolates from Illinois, Michigan, and Ontario. The network was estimated under the 95% statistical limits of parsimony using the algorithm of Templeton *et al.*(1992). Numbered circles represent the label code of each haplotype. The area of the haplotype circles is scaled to represent the relative frequency of that haplotype. The outline of each circle identifies the specific habitat(s) from which each haplotype was collected. Lake populations are shaded in grey, pond populations are white with CP, cyclical parthenogenesis, and OP, obligate parthenogenesis. Small black

circles on the branches represent a single nucleotide difference between haplotypes and are hypothetical haplotypes. Only Haplotype 2 is found in both lakes and ponds. The box corresponds to clade B in figure 2.2……………………………………………………..**67**

Figure 2.4 Unrooted neighbour-joining phylogram based on genetic distances at 21 microsatellite loci in 101 *Daphnia pule*x and 165 *Daphnia pulicaria* from Michigan and Ontario. The two distinct, monophyletic groups correspond to pond and lake habitats. Different symbols identify specific habitats. Solid lines identify lineages that share *D. pulex* mitochondria while dashed lines identify *D. pulicaria* mitochondria…………..…**68**

Figure 2.5 Number of alleles at 21 microsatellite loci in 101 *Daphnia pulex* collected from Solomon and Disputed ponds and 215 *Daphnia pulicaria* isolates sampled from Warner, Lawrence, and Three Lakes II…………………………………………………**69**

Figure 2.6 Results of a Bayesian STRUCTURE analysis of variation at 21 microsatellite loci. A) Genotypes of *Daphnia pulex* (101) and *Daphnia pulicaria* (165) from 5 habitats with best support for K = 2. B) Genotypes of *Daphnia pulicaria* (165) from 3 habitats with best support for K = 2. Each individual's multilocus genotype is represented by a thin vertical line, which is partitioned into K shaded segments that represent the individual's probability of belonging to each of the genetic clusters. Black lines separate different populations, which are labeled below the figure. Five STRUCTURE runs at each K-value produced nearly identical individual membership coefficients. The figures show the highest probability runs………………………………………………………..**70**

Figure 2.7 Genetic structure of *Daphnia* represented by a principle component analysis of variation at 21 microsatellite loci. A) 101 genotypes of *Daphnia pulex* and 165 genotypes of *Daphnia pulicaria* from 2 ponds and 3 lakes. B) 165 genotypes of *Daphnia pulicaria* from lakes, with Lawrence lake divided into Law1, *Daphnia pulicaria* mitochondrial profile and Law2, *Daphnia pulex* mitochondrial profile. Different symbols represent the different populations...**71**

CHAPTER I

INTRODUCTION

Ecological speciation with gene flow

Ecological processes are central to the formation of new species when barriers to gene flow evolve between populations as a result of ecologically-based divergent selection (Schluter and Conte 2009; Rundle and Nosil 2005) such as in a habitat transition event. Habitat isolation is usually based on the inability of a species to use another species' environment, and rests on genetically based differences in fitness associated with habitat use (Coyne and Orr 2004). The process of populations becoming differentially adapted to occupy distinct habitats or utilize different resources while reproductive isolation develops incidentally is called by-product speciation (Rice 1987; Rice and Hostert 1993; Rundle and Whitlock 2001). Several laboratory experiments have simulated by-product speciation using *Drosophila* (Kilias *et al*. 1980; Dodd 1989; Rice and Salt 1990) and the yeast *Saccharomyces cerevisiae* (Dettman *et al*. 2007). However, examples of speciation by habitat isolation from nature are rare (Schluter 2002), mainly because it is difficult to assess if habitat separation is the main mechanism that reduces gene flow during the incipient stage of speciation (Coyne and Orr 2004). However, a few cases of ecological speciation have been explored in which ecological factors were the main driving force of speciation. For example, the freshwater amphipod, *Hyalella azteca* exhibits substantial adaptive, genetically based phenotypic variation among populations that occupy distinct habitat types and likely experienced recent ecological speciation (McPeek and Wellborn 1998). Lakes with fish contain a small-bodied form of the amphipod, and fishless ponds and marshes contain a large-bodied form (Wellborn 1994). Different morphs of the three-

spined stickleback fish, *Gasterosteus aculeatus*, evolved independently across multiple lake-stream habitat transitions that usually coincide with limnetic-benthic ecotones (Berner *et al*. 2009). These ecological populations have diverged with gene flow within a few thousand generations and make a case for ecological speciation in a parapatric context.

Speciation in the face of gene flow is generally thought to be difficult, because gene flow constrains population differentiation and prevents the evolution of strong reproductive isolation (Mayr 1963; Coyne and Orr 2004). However, recent studies show that strong natural selection may promote local adaptation and ecological speciation, even in the face of extensive gene flow. Niemiller *et al.* (2008) present phylogenetic evidence from nuclear and mitochondrial genealogies suggesting that the Tennessee cave salamander (*Gyrinophilus palleucus*) originated from its sister species, the surface-dwelling spring salamander (*Gyrinophilus porphyriticus)* via divergence with gene flow. In the sympatric host races of the larch budmoth, *Zeirphera diniana*, evidence from RFLP markers show that strong divergent selection acts on a few linkage groups, while the selectively neutral part of the genome is subjected to homogenizing gene flow between races (Emelianov *et al*. 2004). The two forms are considered host races rather than full species because of the potential for hybridization, but sympatric differentiation is maintained by selection. In another study of the African malaria mosquito, *Anopheles gambiae*, which is divided into two sympatric, partially isolated subtaxa, the M and S form, a genome scan revealed that differentiation between the two forms is only present in three small regions of the genome (Turner *et al*. 2005). These regions of differentiation likely contain genes

responsible for the ecological and behavioral isolation between the M and S form of the mosquito.

In ecological speciation with gene flow, divergence can occur in some genes even if there is significant exchange of other regions (Hey 2006; Via 2009). A simple model by Hey (2006) proposes that hybrids carry a full set of genes from each population, but backcross hybrids do not, and so it is possible for some genes to pass between populations if backcross hybrids vary in their fitness depending on which genes they carry. An extension of this idea is the "transporter" hypothesis (Schluter and Conte 2009) which proposes that in the early stages of divergence, standing variation of one population is maintained by recurrent gene flow from another population. A slightly different model is proposed by Via (2009), where she describes the genome of sister species in early speciation as having a mosaic nature, where ecologically important genomic regions resist gene exchange, while gene flow continues over most of the genome. Evidence for gene flow may be revealed by discordance between different gene genealogies often caused by hybridization (Wang *et al*. 1997; Dopman *et al*. 2005; Bull *et al*. 2006; Putman *et al*. 2007; Chen *et al.* 2009), and suggests a history of divergence with gene flow.

Despite the above mentioned studies, demonstrating divergence with gene flow remains somewhat difficult because weak genetic differentiation between taxa could be due to recent divergence, gene flow, or a combination of both (Nosil 2008). Our understanding of the genetics of ecological speciation is very limited (Rundle and Nosil 2005; Schluter 2009) and future work on the ecological and genetic factors reducing gene flow can help increase our understanding of the conditions that facilitate divergence in the face of gene

flow (Nosil 2008). To accomplish this, new model systems are needed to better understand the evolutionary forces driving speciation with gene flow, such as during habitat transition.

Studying the concept of speciation is best achieved through comparative studies of evolutionary young lineages where the process of strengthening reproductive isolation is still active (Bernatchez 2004; Via 2009). *Daphnia* (Crustacea: Branchipoda) has been used as a model organism in many diverse areas of biology (Peters and de Bernandi 1987) and its wide geographic distribution across many aquatic environments, easy cultivation under controlled conditions, as well as the availability of many genomic resources, makes it also an ideal study system for studies of speciation. In this study, I am using two ecological sister species (*Daphnia pulex* and *Daphnia pulicaria*) which are part of the *Daphnia pulex* species complex to study speciation with gene flow during habitat transition between lakes and ponds. By conducting both a phylogenetic and population genetics study and using a variety of different genomic markers, I evaluate contemporary and historical patterns of gene flow between and among the two ecological sister species. I also examine the population structure and the colonization history of these species. Ultimately, this study introduces *Daphnia* as a model system for the study of speciation with gene flow during habitat transition and reveals interesting findings about speciation in freshwater organisms.

REFERENCES

Bernatchez L. 2004. Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes). In *Evolution illuminated: salmon and their relatives,* A.P. Hendry and S.C. Stearns, editors. Oxford University Press, Toronto, pp. 175-207.

Berner D., A. Grandchamp, and A.P. Hendry. 2009. Variable progress toward ecological speciation in parapatry: Stickleback across eight lake-stream transitions. Evolution 63: 1740-1753.

Bull V., M. Beltrán, C.D. Jiggins, W.O. McMillan, E. Bermingham, and J. Mallet. 2006. Polyphyly and gene flow between non-sibling Heliconius species. BMC Biology 4:11.

Chen W., K. Bi, and J. Fu. 2009. Frequent mitochondrial gene introgression among high elevation Tibetan megophryid frogs revealed by conflicting gene genealogies. Molecular Ecology 18: 2856-2876.

Coyne J.A. and H.A. Orr. 2004. Speciation. Sinauer Associates, MA, U.S.A.

Dettman J.R., C. Sirjusingh, L.M. Kohn, and J.B. Anderson. 2007. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. Nature 447: 585-588.

Dodd D.M. 1989. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. Evolution 43: 1380-1311.

Dopman E.B., L. Pérez, S.M. Bogdanowicz, and R.G. Harrison. 2005. Consequences of reproductive barriers from genealogical discordance in the European corn borer. Proceeds of the National Academy of Science 102: 14706-14711.

Emelianov I., F. Marec, and J. Mallet. 2004. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proceeds of the Royal Society of London 271: 97-105.

Hey J. 2006. Recent advances in assessing gene flow between diverging populations and species. Current Opinion in Genetics & Development 16: 592-596.

Kilias G., S.N. Alahiotis, and M. Pelecanos. 1980. A multifactorial genetic investigation of speciation theory using *Drosophila melanogaster*. Evolution 34: 730-737.

Lynch M. 1985. Speciation in the Cladocera. Vert. Internat. Verein Limnol. 22: 3116-3123

Mayr E. 1963. Animal species and Evolution. Harvard University Press, Cambridge, Massachusetts.

Mcpeek M.A. and G.A. Wellborn. 1998. Genetic variation and reproductive isolation among phenotypically divergent amphipod populations. Limnology and Oceanography 43:1162-1169.

Niemiller M.L., B.M. Fitzpatrick, and B.T. Miller. 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. Molecular Ecology 17: 2258-2275.

Nosil P. 2008. Speciation with gene flow could be common. Molecular Ecology 17: 2103-2106.

Omilian A.R. and M. Lynch. 2009. Patterns of interspecific DNA variation in the *Daphnia* nuclear genome. Genetics 182: 325-336.

Peters R.H. and R. de Bernandi. 1990. *Memorie dell'Istituto Italiano di Idrobiologia*, Vol. 45: *Daphnia*. Verbania Palaza, Palanza, Italy.

Putman A.S., J.M. Scriber, and P. Andolfatto. 2007. Discordant divergence times among Z-chromosome regions between two ecologically distinct swallowtail butterfly species. Evolution 61: 912-927.

Rice W.R. 1987. Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. Evolutionary Ecology 1: 301-314.

Rice W.R. and E.E. Hostert. 1993. Laboratory experiments on speciation: what have we learned in 40 years? Evolution 47: 1637-1653.

Rice W.R. and G. Salt. 1990. The evolution of reproductive isolation as a correlated character under sympatric conditions: experimental evidence. Evolution 44: 1140-1152.

Rundle H.D. and P. Nosil. 2005. Ecological speciation. Ecology Letters 8: 336–352.

Rundle H.D. and M.C. Whitlock. 2001. A genetic interpretation of ecologically dependent isolation. Evolution 55: 198-201.

Schluter D. 2002. The ecology of adaptive radiation. Oxford University Press, Oxford, U.S.A.

Schluter D. 2009. Evidence for ecological speciation and its alternative. Science 323: 737-741.

Schluter D. and G. Conte. 2009. Genetics and ecological speciation. Proceeds of the National Academy of Science 106: 9955-9962.

Turner T.L., M.W. Hahn, and S.V. Nuzhdin. 2005. Genomic islands of speciation in Anapheles gambiae. Public Library of Science Biology 3:e285

Via S. 2009. Natural selection in action during speciation. Proceeds of the National Academy of Science 106: 9939-9946.

Wang R.L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics 147: 1091-1106.

Wellborn G.A. 1994. Size-based predation and prey life histories: A comprehensive study of freshwater amphipod populations. Ecology 75: 2104-2117.

CHAPTER II

SPECIATION WITH GENE FLOW IN *DAPHNIA PULEX* AND *DAPHNIA PULICARIA**

INTRODUCTION

Speciation in *Daphnia*

The relative contribution of geography and ecology to the diversification of freshwater organisms is little understood. While allopatric isolation is considered the main mechanism of speciation in zooplankton species such as *Daphnia* (Adamowicz *et al*. 2009), colonization of new aquatic habitats has been also proposed to initiate many speciation events in cladocerans (Lynch 1985). The genus *Daphnia* (Cladocera) is a group of widespread freshwater crustaceans which includes about 200 species (Colbourne *et al*. 1997), of which 34 species inhabit North America (Hebert 1995). It is believed that this genus originated over 200 million years ago, during the Mesozoic (Colbourne and Hebert 1996), and fossil records from Australia confirm that the genus has been in existence for at least 70 million years and closely related genera have existed for at least 120 million years (Fryer 1991). The genus includes 3 subgenera (*Daphnia*, *Hyalodaphnia*, and *Ctenodahnia*) comprised of about 15 species complexes (Colbourne and Hebert 1996) that possess strong dispersal abilities due to their diapausing eggs being encased in a modification of the female's carapace known as an ephippium, typical of Anomopoda zooplankton.

**This chapter is the outcome of joint research as stated in the declaration of co-authorship page.*

Attempts to understand taxonomic relationships within the genus *Daphnia* have been limited by the constrained morphological diversity and dramatic phenotypic plasticity of this group (Hebert 1978; Dodson 1989; Lampert 1994; Ghadouani and Pinel-Alloul 2002), the occurrence of interspecific hybrids (Taylor and Hebert 1992; Hebert and Finston 1996; Spaak 1997; Weider *et al*. 1999), the total suppression of sexual reproduction in some groups (Crease *et al*. 1989; Crease and Lynch 1991; Hebert *et al*. 1993), and the occurrence of polyploidy (Dufresne and Hebert 1994; Adamowicz et sl. 2002; Mergeay *et al*. 2008; Vergilino *et al*. 2009). All these factors make the establishment of species boundaries difficult. Allozyme analyses have traditionally been used to distinguish between species in the *Daphnia pulex* complex (Hebert 1987) and more recently, sequence analysis provided more insight into the evolutionary history of this group (e.g. Colbourne and Hebert 1996; Adamowicz *et al*. 2009). Lynch (1985) was the first to propose an explicit mechanism of speciation for cladocerans where he argues for a combined role of founder effect and adaptive divergence in *Daphnia* speciation. According to his model, speciation via the founder effect is much more likely to occur if it is accompanied by a shift in environment since this can facilitate the development of reproductive isolation through different selective pressures in different habitats. De Meester *et al*. (2002) extended this idea and argued that once a population is locally adapted, a strong colonization "priority effect" reduces much of the gene flow between differently adapted aquatic habitats. This priority effect is achieved by founder events, rapid population growth and local adaptation upon colonization, resource monopolization, and the buildup of large resting egg banks, which together resists the

persistence of newly invading genotypes and results in high genetic subdivision and speciation.

Daphnia pulex and *Daphnia pulicaria*

It has been suggested that habitat transitions followed by local adaptation played an important role in the evolution of the *Daphnia pulex* species complex (Lynch *et al*. 1999; Pfrender *et al*. 2000), which includes several ecological species inhabiting a variety of different freshwater habitats (Adamowicz *et al*. 2009). Some of the species in this complex include: *Daphnia middendorffiana* which is an arctic lake species (Hobaek and Weider 1999), *Daphnia tenebrosa* which inhabits both ponds and lakes in the Arctic (Edmondson 1955), *Daphnia melanica* that is found in sand dune ponds (Hebert 1995), *Daphnia pulex* which is a temperate pond species, and *Daphnia pulicaria* which is one of the most widely distributed North American lake species (Hebert 1995). The two sister species, *Daphnia pulex* and *Daphnia pulicaria*, are estimated to have diverged ~82,000 years ago but still experience significant levels of gene flow (Omilian and Lynch 2009). Hybrids of the two species can be successfully produced in laboratory settings (Heier and Dudycha 2009), and can be found in nature in disturbed, deforested ponds, and generally reproduce by obligate parthenogenesis (Hebert and Crease 1983). The opportunity for gene flow between lake and pond populations is high because *Daphnia* can easily disperse across wide distances (Cohen and Shurin 2003) when its long term dormant eggs, that are enclosed with an ephippial case, are transported by wind, rain (Cáceres and Soluk 2002), or animal vectors (Allen 2009). However, despite *Daphnia*'s ability to

disperse between water bodies, genetic data indicates low levels of achieved gene flow between lake and pond populations (Pfrender *et al*. 2000).

It has been proposed that barriers to gene flow between lake and pond *Daphnia* are likely ecologically based (Lynch 1985; Heier and Dudycha 2009) and divergent selection between *D. pulex* and *D. pulicaria* populations should be substantial. Several studies have found that lakes and ponds have different physical and biotic conditions (Wellborn *et al*. 1996), and that *Daphnia* in these habitats differs significantly in its life history traits (Dudych and Tessier 1999; Dudycha 2003; Dudycha 2004). *Daphnia pulex* is present in shallow, fishless, temporary ponds for a short period of time in the spring, while *D. pulicaria* populations can persist in stratified lakes year-round (Cáceres and Tessier 2004). In lakes, *Daphnia* populations feed on phytoplankton and are usually exposed to predation by fish, while in temporal ponds, they also feed on detritus, experience mainly invertebrate predation and experience, anoxia, and complete freezing (Colbourne *et al*. 1997). In the presence of fish, *D. pulicaria* inhabits the cold hypolimnetic region to avoid fish predation and competition from other *Daphnia* species (Wright and Shapiro 1990), while in the absence of fish, *D. pulicaria* largely feeds in the epilimnetic waters (Werner *et al*. 1977) and has been observed to be up to 3 times more abundant (Leibold 1991). Additionally, sediment egg banks contain a larger volume of resting eggs in lakes than in ponds (Cáceres and Tessier 2004), and this is likely caused by a lower hatching rate in the lakes due to differences in environmental cues between lake and pond habitats (Cáceres and Tessier 2003). Pond *Daphnia* grow faster and have shorter life spans (Dudych and Tessier 1999; Dudycha 2003; Dudycha 2004), experience greater changes in density, and have greater early reproductive output than lake *Daphnia* (Dudycha

2004). All these differences between pond *D. pulex* and lake *D. pulicaria* indicate that the two species have diverged ecologically and make for a good study system of ecological speciation with gene flow involving habitat transition between lakes and ponds.

Divergence and speciation between *Daphnia pulex* and *Daphnia pulicaria*

Models for gene exchange between the ecologically distinct *Daphnia pulex* and *Daphnia pulicaria* have previously been proposed. Pfrender and colleagues (2000) suggest that in Oregon, permanent lake populations periodically colonize temporary ponds following floods and must then quickly adapt to an ephemeral habitat (Pfrender *et al*. 2000). Despite this long-term gene flow, *D. pulex* and *D. pulicaria* in Oregon form monophyletic clades (based on 10 allozyme and 6 microsatellite loci, Morgan *et al*. 2001). However, an allozyme screen has been used in the past as a diagnostic marker to distinguish between pond *D. pulex* and lake *D. pulicaria* (Hebert *et al*. 1989; Hebert *et al*. 1993), where pond individuals are usually homozygous for the "slow" (S) allele and lake individuals are homozygous for the "fast" (F) allele. Additionally, a recent study of variation at six nuclear protein-coding loci indicates that *Daphnia pulex* and *Daphnia pulicaria* form distinct genetic clusters and are also monophyletic with respect to their closest relative, *D. arenata* (Omilian and Lynch 2009). The study also reports high levels of gene flow between *D. pulex* and *D. pulicaria*. However, based on mitochondrial data, as much as 19% sequence divergence separates the different lineages found within this complex (Colbourne et al. 1998). North American *D. pulex* and *D. pulicaria* belong to the same major clade within this complex, with North American *D. pulicaria* consisting of several species that include, polar, western, and eastern *D. pulicaria* lineages.

Since *D. pulex* and *D. pulicaria* are not completely reproductively isolated but clearly occupy different habitats, there is a need to study the extent of genetic mixing, both between and among habitats, to better understand the colonization and evolutionary history of *Daphnia*. For this study, I most often use Van Valen's (1976) ecological species concept that defines a species as a lineage which occupies an adaptive zone minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range. Using the mitochondrial *ND5* gene, the nuclear *Ldh-A* gene, and 21 microsatellite markers, I explore the evolutionary consequences of habitat transition events in the *Daphnia pulex* complex. I present a phylogenetic and population genetics study using 363 *Daphnia* isolates collected from natural ponds and lakes in Southern Ontario and Michigan. Specifically, I evaluate the extent of gene flow between lake and pond *D. pulicaria* and *D. pulex* populations in Southern Michigan and Ontario, examine the population structure and explore the history of the lake species.

MATERIALS AND METHODS

Sample collection

Lake samples were collected by towing a plankton net vertically through the deepest part of each lake, while pond samples were taken with a dip net from shore. After collection, single individual female *Daphnia* were placed in separate 250 ml beakers and allowed to reproduce parthenogenetically to establish clonal lines, hereafter referred to as isolates. The isolates were maintained in filtered river water at 15 - 18°C with a 12-h light, 12-h dark photoperiod and fed every 3-4 days with a combination of the microalgae species *Nannochloropsis* and *Tetraselmis* (Reed Mariculture) diluted in ddH_2O. After several weeks, 6-10 clonal individuals were collected from each beaker and immediately stored at -20°C.

D. pulex and *D. pulicaria* were collected from a total of 15 habitats (9 lakes and 8 ponds) across Michigan, Illinois, and Ontario (figure 2.1). The mitochondrial ND5 gene was used to conduct a phylogenetic study on a large number of isolates (363) with low sampling (3-14 individuals) per habitat (tables 2.1A, 2.1B). In contrast, the population genetic survey was based on a focal geographic area (southwestern Ontario and Michigan) and on an intensive sampling of 3 lakes (165 isolates) and 2 ponds (101 isolates) with 35-86 isolates per habitat. The 165 lake *D. pulicaria* isolates were collected in July 2008 and May 2009 from three permanent lakes; Lawrence and Warner Lakes in Barry County and Three Lakes II in Kalamazoo County all located in southwestern Michigan, USA (figure 2.1, table 2.1B). All three lakes are hard water lakes with small surface area (<30 ha), relatively deep (>10m), thermally stratified (Leibold and Tessier

1991), and have similar zooplankton communities dominated by *D. pulicaria* and *D. galeata mendotae* (Haney and Hall 1975; Leibold and Tessier 1991). Warner Lake also contains *D. retrocurva* (Leibold and Tessier 1991). The 101 *D. pulex* isolates were collected from two temporary ponds, Disputed and Solomon, located in Southern Ontario and Michigan, respectively, in the spring and early summer of 2007 2008, and 2009 (table 2.1A).

Sexuality tests

During optimal conditions, cyclically parthenogenetic (CP) females produce diploid eggs by apomixis which develop into genetically identical daughters. Certain environmental cues, such as warm temperatures and crowding can induce the production of males and haploid diapausing eggs, which need to be fertilized (Hebert and Crease 1983). Some populations reproduce by obligate parthenogenesis (OP), in which case the diapausing eggs are also produced by apomixis and do not require fertilization. Unlike the apomictic eggs, which develop directly into juveniles in the female's brood pouch, the diapausing eggs are expelled into an ephippium, where they can remain dormant for days or decades (Heier and Dudycha 2009). *D. pulicaria* shows large between-population variation in the magnitude of investment in dormancy or sex (Cáceres and Tessier 2004), while *D. pulex* is more consistent and produces dormant eggs every year before its temporary habitat dries up.

Since *D. pulex* is known to consist of cyclically parthenogenetic (CP) populations, obligately parthenogenetic (OP) populations, as well as populations with mixed reproductive strategies (Hebert and Crease 1983), extensive sexuality tests were

conducted on all pond isolates to determine their reproductive strategy. Single females were isolated and their mode of reproduction was determined using the method of Innes *et al.* (1986). Since lake populations were previously reported to be reproducing solely by CP (Tessier and Leibold 1997), sexuality tests were performed on a subset of 10 isolates from each lake to confirm that their mode of reproduction was indeed CP.

Mitochondrial DNA amplification

DNA was extracted from isolate cultures using the CTAB protocol described by Doyle and Doyle (1987) and the final yield of DNA was resuspended in 100 µl of H_2O. An 897 bp fragment of the NADH dehydrogenase 5 (ND5) gene was amplified using the forward primer: 5’GGGGTGTATCTATTAATTCG 3' and reverse primer: 5’ATAAAACTCCAATCAACCTTG 3' (Colbourne *et al.* 1998). PCR was carried out in a 25 µl volume consisting of 1.5 µl DNA template, 1X PCR buffer with 0.25 mM of $MgCl_2$, 2.5 units of Taq polymerase, 0.1 µM of dNTP, and 0.08 mM of each of the forward and reverse primers. The thermal cycle program included an initial denaturation step of 3 min at 95°C followed by 5 cycles of 35 s denaturation at 94°C, 35 s annealing at 54°C, 40 s extension at 72°C followed by 30 cycles of 35 s at 94°C, 35 s at 50°C, and 40 s at 72°C, with a final extension at 72°C for 4 min. PCR products were verified on a 1% agarose gel and sequenced with the forward primer using BigDye terminator sequencing chemistry. The reactions were resolved on an ABI 3130XL genetic analyzer (Applied Biosystems). Sequences were inspected and aligned using CODONCODE ALIGNER 2.0 (CodonCode Corporation, Dedham, MA) and manually corrected.

Nuclear *Lactate dehydrogenase* survey

Previous surveys of allozyme variation (Hebert *et al*. 1989; Hebert *et al*. 1993) have shown that lake populations are generally fixed for an electrophoretically "fast" (F) allele at the *Lactate dehydrogenase* A locus (Cristescu *et al.* 2008). Pond populations are either fixed for a "slow" (S) allele or are SF heterozygotes. SF heterozygotes have been reported to reproduce by OP (Innes *et al*. 1986) and been considered F1-generation hybrids of *D. pulex* and *D. pulicaria* (Hebert *et al*. 1993; Hebert and Finston 2001). Allele specific primers (Crease *et al*. 2010) were used to determine the *Ldh-A* genotype of each isolate (table AA). Primers that amplify the F allele are LdhAF-F; 5'GAGCGATTTAACGTTGCGCCT' and LdhAF-R: 5'GGACGACTTGTGTGTGAATTTC. Primers that amplify the S allele are LdhAS-F; 5'GAGCGATTTAACGTTGCGCCC3' and LdhAS-R: 5'GGACGACTTGTGTGTGAATTTG3'. Each isolate was tested with both sets of primers. PCR reactions and cycling conditions were the same as those used for ND5 amplification. Alleles were resolved on a 1.5% agarose gel. To confirm the results, fifteen individuals were additionally analyzed using the traditional method of allozyme electrophoresis (Hebert and Beaton 1989).

Microsatellite survey

Twenty one unlinked and previously mapped microsatellite markers were chosen from different linkage groups of the *D. pulex* linkage map (Cristescu *et al*. 2006) and were used to genotype 266 isolates from three lake and two pond populations. The forward, sequence-specific primers were 5'-extended with the M13(-21) oligonucleotide,

according to the method described by Schuelke (2000). The PCR was performed in 12 µl reactions with 0.98 µl DNA template, 1X PCR buffer with 25 nmol of $MgCl_2$, 0.5 units of Taq polymerase, 2.5 nmol of dNTP, 1 pmol of forward primer, 2 pmol of reverse primer, and 2 pmol of a universal fluorescently-labeled M13(-21) primer (NED, PET, FAM, VIC). A touchdown PCR was used to reduce nonspecific amplification. Thermal cycle programs include an initial denaturation step of 3 min at 95°C followed by 10 cycles of 35 s denaturation at 94°C, 35 s at 60°C with the annealing temperature decreased by 1°C every cycle during each of the 9 following cycles, 45 s extension at 72°C followed by 30 cycles of 35 s at 94°C, 35 s at 53°C, and 45 s at 72°C, with a final extension at 72°C for 10 min. Reactions were denatured for 5 min at 90°C, quickly cooled on ice and resolved on an ABI 3130 XL automated sequencer with GeneScan™ - 500 LIZ™ internal size standard. Genotypes were scored using GENEMAPPER v4.0 (Applied Biosystems) and verified manually by eye.

Phylogenetic analyses

Unique mitochondrial ND5 haplotypes were identified using DnaSP v.5.0 (Librado and Rozas 2009). Genetic diversity for mtDNA was characterized by the standard indices of haplotype diversity and nucleotide diversity using DNASP v 5.00.07 (Rozas *et al*. 2003). Phylogenetic analyses were performed using neighbor-joining (NJ) and Bayesian inference (BI) methods. Based on the phylogeny of the *D. pulex* complex constructed by Adamowicz *et al*. (2009), European *Daphnia pulex* (GenBank accession number DQ235231) was chosen as an outgroup. MODELTEST 3.7 (Posada and Crandall 1998) was used to select the best-fit model of sequence substitution (HKY+G). Neighbor-

joining phylogenetic analysis was conducted in MEGA 4.0 (Tamura *et al*. 2007) based on nucleotide distances corrected using the Tamura–Nei model (Tamura & Nei 1993) with a gamma rate distribution (0.3241). Confidence level for the topology of the tree was estimated using bootstrap analyses with 1000 replicates. Bayesian phylogenetic analyses were performed in MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003). All searches used random starting trees and employed four independent runs. Trees were sampled every 100 generations for 6 million generations and the first 25% of all the trees were discarded as burn-in. The 50% majority rule consensus tree was generated from the remaining trees and the posterior probability of each node was calculated as the percentage of trees recovering any particular node.

Since I was interested at looking at the close relationship of mitochondrial haplotypes between pond and lake individuals, a network was generated using all ND5 haplotypes from the panarctic *Daphnia pulex* clade and MI lake *Daphnia pulicaria* clade and excluding the western *Daphnia pulicaria* clade using TCS 1.0 (Clement *et al*. 2000). The program estimates genealogical relationships among sequences at the population level using the 95% statistical parsimony algorithm (Templeton *et al*. 1992).

A NJ phylogeny was constructed based on microsatellites by calculating the commonly used Nei's standard genetic distance *Dm* (1972) between genotypes of all pairs of individuals in POPULATIONS 1.2.30 (Langella 1999).

Population genetics analyses

Allelic richness (A_r) at microsatellite loci was measured in each of the three lake and two pond populations as the number of alleles independent of sample size using FSTAT v. 2.9.3.2 (Goudet 2001). Weir and Cockerham's (1984) inbreeding coefficient (F_{IS}) was calculated for each population and linkage disequilibrium (LD) was measured between all pairs of loci in each population using GENEPOP online v.4.0.10 (Raymond and Rousset 1995). Alleles with a frequency of less than 10% were removed from linkage disequilibrium analysis because the rare alleles often give false positives.
A probability test with Markov chain (1000 dememorization steps, 100 batches, and 1000 iterations per batch) was conducted to determine the likelihood of two pairs of loci being in linkage disequilibrium. Significance levels were determined after Bonferonni correction of P-values ($P<0.00048$).

Observed heterozygosity (H_O) and unbiased estimates of expected heterozygosity from Hardy-Weinberg assumptions (H_E) as well as P-values for tests of Hardy-Weinberg equilibrium (HWE) were calculated using ARLEQUIN version 3.1 (Excoffier *et al.* 2005). Tests for deviations from HWE used Markov chain (1000 dememorization steps, 100 batches, 1000 iterations per batch) and sequential Bonferonni correction was applied to determine significant P-values. The presence of null alleles was tested with the software MICRO-CHECKER version 2.2.0 (van Oosterhout *et al.* 2004).

Repeated multilocus genotypes were detected using GENALEX v. 6 (Peakall and Smouse 2006). Repeated genotypes were removed from the dataset for all subsequent analyses because clonal amplification of genotypes can influence data interpretation

(Sunnucks *et al*. 1997). Microsatellite data analysis was performed with one representative of each genotype (i.e. clonal copies removed).

Pairwise estimates of the fixation index, F_{ST} from Weir and Cockerham (1984) were calculated as a measure of genetic differentiation among populations and tested for a significant departure from zero using permutation procedures in ARLEQUIN ver. 3.0 (Excoffier *et al*. 2005). To examine fine-scale genetic patterns between lakes and ponds and among the three lake populations, I used GENALEX v.6 (Peakall and Smouse 2006) to construct a Principal Coordinates Analysis (PCA) to explore multivariate patterns of molecular diversity relative to populations.

To further determine if there was genetic structure between ponds and lakes and among the lakes, a Bayesian inference of population structure was conducted using STRUCTURE v. 2.3.1 (Pritchard *et al*. 2000; Falush *et al*. 2003). This program uses multilocus genotypic data to define a set of populations with distinct allele frequencies, hereafter referred to as clusters, and assign individuals probabilistically to these defined clusters without prior knowledge of sampling location. Two separate analyses were conducted, the first included 2 ponds and 3 lakes and the second included only the 3 lakes. For the lake and pond analysis I assessed likelihoods for models with the number of clusters (K) ranging from K = 1 to K = 5 (total number of populations) and for lakes K ranged from K = 1 to K = 3. For each value of K, I carried out 5 independent Markov Chain Monte Carlo (MCMC) runs with 100,000 generations discarded as burn-in followed by an additional 1,000,000 generations and results were consistent across runs. The optimal number of clusters was estimated by comparing the log-likelihood of the

data given the number of clusters [ln $P(X|K)$] (Pritchard *et al.* 2000) and by examining the standardized second order rate change of ln $P(X|K)$ (ΔK) (Evanno *et al.* 2005). Individual multilocus genotypes were then assigned to a cluster according to the HWE criteria (Pritchard *et al.* 2000).

To illustrate the historical dispersal patterns between sites, a Bayesian method was used to calculate emigration and immigration rates with MIGRATE version 3.0.3 (Beerli 2008). The number of migrants (*Nm*) per generation was calculated as $\theta_i M_i$, where θ_i equals $xN_e^{(i)}\mu$ and M_i equals m_i/μ. Among the parameters, x is the inheritance parameter; $N_e^{(i)}$ is the effective population size; μ is the mutation rate per locus per generation; and m_i is the immigration rate. For my analysis, x was set as 4. This value is commonly used for nuclear gene data, and other parameters were estimated from the data by the program. A Brownian motion mutation model was used. I used 10 short chains (10,000 iterations) and 3 long chains (1,000,000 iterations) with 50,000 iterations discarded as an initial 'burn-in' for the Bayesian search strategy. MIGRATE assumes that all interbreeding populations have been sampled, despite this limitation, I have chosen to use this software since it allows to estimate both emigration and immigration rates between all the populations.

RESULTS

Phylogenetic Analyses

The 687 bp long mitochondrial ND5 sequence-alignment of the 363 isolates contained 539 conserved sites and 148 variable sites of which 88 were parsimony-uninformative. There were more haplotypes found among the two ponds (25) than among the three lakes (17; table 2.2) and haplotype diversity was slightly higher for ponds (0.89) than for lakes (0.82). However, nucleotide diversity was lower for ponds (0.006) than for lakes (0.025). The 50 unique haplotypes identified formed two well supported clades that correspond to the panarctic *D. pulex* (ppx) and western *D. pulicaria* (wpc) lineages identified by Colbourne *et al*. (1998). All isolates collected from ponds grouped within the ppx clade while lake isolates were found either in the ppx clade or in the wpc clade (figure 2.2). All but one of the isolates from Three Lakes II had a ppx mitochondrial profile (table 2.1B). Four isolates from Warner Lake had ppx mtDNA with the rest of the isolates having wpc mtDNA. In Lawrence Lake, 47% of the isolates were found to have ppx mtDNA and 53% had wpc mtDNA.

The network of the ppx clade displayed a star-shaped pattern. Two separate groups were detected within the network (figure 2.3) that corresponded with the clades observed in the NJ and BI analysis (figure 2.2). The most common haplotype, haplotype 2 (ppx), was the only one found in both lakes and ponds. Many of the haplotypes differed from haplotype 2 by only 1-4 nucleotide differences, while a distinct group of lake haplotypes (corresponding with clade B in figure 2.2) differed from haplotype 2 by at least 8

nucleotides. Only lake isolates were found among this more distinct group, except some lake-pond hybrid (*LDH* heterozygotes) sampled from Windsor ponds.

A neighbor-joining (NJ) phylogram was constructed based on 21 microsatellite loci (figure 2.4) and shows a clear separation with no overlap between lake and pond habitats. All of the pond isolates group into one clade, while all the lake isolates group into another. Lake isolates with the two different mitochondrial profiles group together.

Genotypes at the *Ldh-A* locus

Despite the occurrence of both ppx and wpc mtDNA in lake populations, all lake isolates were homozygous for the F allele at the *Ldh-A* locus and are referred to as *Daphnia pulicaria*. Moreover, all pond isolates that were determined to reproduce by CP were homozygous for the S allele and are referred to as *Daphnia pulex*. Out of the 136 total pond individuals screened in the large phylogenetic survey, 19 were determined to be OP and were either homozygous (SS) or heterozygous (SF) at the *Ldh-A* locus (table 2.1A, figure 2.2).

Population genetics analyses

The total number of alleles at each microsatellite locus ranged from 1 to 10 (table 2.3, figure 2.5). The allelic richness for the ponds ranged from 2.000 to 8.863, and for the lakes from 1.000 to 4.974. The observed heterozygosity for each microsatellite locus in each population ranged from 0 to 0.925. In total, 50 private alleles were found among the two pond populations and 6 private alleles in the three lakes, all with a frequency below 25% except for one allele at locus d153 in Solomon Pond, which had a frequency of 68%.

At locus *d174* on linkage group I, all lake populations were fixed for the same allele. Additionally, genetic diversity was very low in the lake populations at loci d015 and d111, on linkage groups II and VI respectively. Fifty-four out of 105 tests of HWE across all populations and loci were nominally significant ($P<0.05$) and 25 were significant after sequential Bonferroni adjustment (table 2.3). Each lake had ~3-4 loci (14-19% of loci) out of HWE. Disputed and Solomon ponds had 3 and 9 loci (14%, 43%) out of HWE respectively. Almost all loci that were out of HWE in the two ponds showed heterozygote deficiency (8 out of 9 in Solomon Pond and 3 out of 3 in Disputed Pond). There was no clear pattern of heterozygote deficiency or excess in the lakes. Low levels (10-24%) of null alleles were detected among the lakes and ponds (table 2.3). Isolates that had the same genotype at all 21 microsatellite markers were identified and clones were removed for all subsequent microsatellite analysis.

Linkage disequilibrium

The test for linkage disequilibrium (LD) between pairs of microsatellite loci indicated that lake populations have higher numbers of loci in LD than pond populations. After Bonferroni correction for multiple tests ($P<0.00048$), Lawrence Lake had the highest number of pairs of loci in LD with 34 out of 210 pairs, Warner Lake had 18 out of 210 pairs in linkage disequilibrium, and Three Lakes II had the lowest number with 10 out of 210 pairs of loci (table 2.6). The two pond populations had much lower levels of linkage disequilibrium; Solomon Pond had 3 pairs out of 210 pairs and Disputed Pond had 0 out of 210 pairs of loci in disequilibrium. A separate analysis was conducted for Lawrence Lake based on its mitochondrial profile; individuals with ppx mtDNA (Law2) had higher

levels of LD (31 pairs) than individuals with wpc mtDNA (Law1, 13 pairs). Locus pairs d186-d006 and d186-d148 were in LD in all the lake habitats. In addition 6 loci were found to be in LD with either of the three mitochondrial clades (A and B, or C) in Lawrence Lake; d027, d029, d186, d006, d148, d016 (table 2.7). Ten loci were identified (d070, d027, d087, d127, d029, d186, d042, d006, d148, d016) to be in LD with either mitochondrial haplotypes within clade A or B.

Population differentiation and genetic distance

Microsatellite markers revealed marked genetic differentiation among the lake and pond habitats. For example, pairwise F_{ST} values between the ponds and lakes ranged from 0.438 to 0.481 (table 2.4). F_{ST} values among lakes were between 0.076 and 0.147 between ponds was 0.080. A separate analysis comparing the two mitochondrial groups in Lawrence Lake (Law1 and Law2) revealed an F_{ST} value of 0.109 between the two groups within the same lake (APPENDIX table S.1). All F_{ST} values were significantly different from 0 ($P<0.05$). Principal component analysis (PCA) indicated the existence of two clusters corresponding to ponds and lakes (figure 2.7A). The PCA analysis of only the lake isolates did not show any pattern of population subdivision.

The STRUCTURE analysis based on microsatellites indicated the highest posterior probability for two clusters, corresponding to the pond and lake groups (figure 2.6A). The lakes-only analysis indicated that there are two distinct clusters within the lakes (figure 2.6B). The method recommended by Evanno *et al*. (2005) confirmed two genetic clusters for both the global data set and also for the lakes subset. Warner Lake was mostly part of one cluster, Three Lakes II was mostly part of another cluster, and Lawrence Lake was a

mixture of the two clusters. In Lawrence lake (where there is a mixture of individuals with both ppx and ppc mtDNA), there was no correlation between the clustering pattern observed in STRUCTURE and the mitochondrial type of each individual.

Analysis of emigration and immigration rates (number of migrants/generation=*Nm*) revealed that the highest level of gene flow can be observed among ponds with *Nm* of 1.5 migrants/generation, while the lowest level of gene flow was observed between lakes and ponds with *Nm* ranging from 0.48 to 1.34 (table 2.5). The number of migrants among the lakes was variable and ranged from 0.47 to 0.80.

DISCUSSION

Divergence with gene flow

Mitochondrial gene introgression may be detected without any evidence for nuclear gene mixing when hybrids from one habitat successfully introgress into a different habitat type and are subject to strong selection pressures at the nuclear genome level. It is striking that two divergent mitochondrial lineages occur in 5 out of 9 lakes surveyed in this study (figure 2.2). *Daphnia pulex* mitochondrial DNA lineage (ppx mtDNA) was previously found in lakes from Michigan (Crease *et al*. 1989), the arctic (Dufresne and Hebert 1997) and western Canada (Crease *et al*. 1997). However, no previous study detected the presence of both ppx mtDNA and wpc mtDNA within the same habitat. In this study, all lakes examined from Michigan (Three Lakes II, Warner, Lawrence, Bassett, and Mill) had both mitochondrial types. It is likely that the large sample size enabled me to detect both mitochondrial types since each lake had one common type and one rare type, except for Lawrence Lake which had both mitochondrial types in equal proportions. Of course, I cannot rule out the possibility that the occurrence of the different mitochondrial types within the same lake is not a common occurrence in lake *Daphnia pulicaria* and that the lake system presented here is unique.

Mitochondrial introgression between two young species can reveal historical patterns of gene flow and can shed light on possible habitat transition events. The mitochondrial phylogenetic reconstruction revealed three monophyletic clades with high statistical support (figure 2.2): a Western *D. pulicaria* clade (clade C), a more diverse panarctic *D. pulex* clade (clade A), and a third previously unrecognized *D. pulicaria* clade (clade B)

which is very closely related to panarctic *D. pulex*. The Western *D. pulicaria* clade consists of only lake isolates. Clade A includes all the pond isolates and some of the lake isolates, including the most common haplotype (haplotype 2, figure 2.5) that was the only one shared between lakes and ponds. Clade B mostly consists of lake isolates and pond-lake hybrids (SF *LDH* heterozygotes) from ponds in Windsor. The ppx mtDNA haplotypes of lake *D. pulicaria* do not form a monophyletic group relative to the *D. pulex* cluster, indicating that at least three independent habitat transition events from ponds to lakes have occurred. The first corresponds to basal clade C and likely represents the initial establishment of lake *Daphnia* from ponds. The second event corresponds to clade B, which is more recent than the well established Western *D. pulicaria* clade (clade C). Since clade B is highly statistically supported and contains only lake individuals and F1 hybrids, it seems reasonable to suggest that this is a separate *D. pulicaria* clade that may be genetically distinct from other *D. pulicaria* groups. I call this clade, MI lake *Daphnia pulicaria*. Clade A corresponds to the most recent transition event from ponds to lakes, where the presence of shared haplotypes between lake and pond suggests that transition or hybridization events are ongoing. An overall pattern of historical gene flow from ponds to lakes is supported by the mitochondrial phylogeny.

This proposed scenario of multiple, unidirectional habitat transitions is further supported by the results of the network analyses. The ND5 haplotype network exhibits a star-shaped pattern with two major groups recovered (figure 2.3). These two groups correspond to the two clades (A and B) identified in the NJ phylogenetic analyses (figure 2.2). Group A displays a typical star shape with haplotype 2 appearing to be the ancestral haplotype and including a total of 35 haplotypes. A single ancestral haplotype often gives rise to

multiple descendant haplotypes yielding a haplotype tree with true multifurcations (Posada and Crandall 2001), as is seen with the network in this study. The large number of derived haplotypes in this group suggests population expansion within this species. The second group enclosed by a rectangle in figure 2.3 includes only one F1 hybrid haplotype and 7 lake haplotypes that are divided into two groups. The network indicates that this group is separated by at least 8 mutation steps and suggests that the transition from lakes to ponds of lineages with type B mitochondria happened earlier than did transitions from lakes to ponds of lineages with type A.

Differences between a species' mitochondrial gene genealogy and its nuclear gene genealogy can provide initial support for divergence with gene flow. The higher mutation rate of microsatellite markers than mitochondrial markers indicate that microsatellite data reflect a more contemporary pattern, whereas mitochondrial data show a more historical perspective. The nuclear microsatellite phylogram based on allele frequencies (figure 2.4) shows two distinct clades corresponding to lake and pond habitats, indicating that populations of lakes and ponds are currently diverging. The pond isolates show more diversity than the lake isolates and this same pattern can also be observed from the mitochondrial data (figure 2.3). Since the lake and pond clades in the microsatellite phylogram (figure 2.4) do not overlap, it is reasonable to conclude that isolates with either mitochondrial haplotype found in the lakes are interbreeding and that these mitochondrial patterns are remnants of past colonization events. Despite introgression in the mitochondria, nuclear data clearly supports divergence between lakes and ponds and the formation of the two incipient lineages supports a case of divergence with gene flow.

Discordance between nuclear *Ldh* and mitochondrial data

Past work on allozymes in pond and lake *Daphnia* indicate that certain nuclear loci show a consistent pattern corresponding to each type of habitat (Hebert *et al*. 1989; Hebert *et al*. 1993) and this study is strongly concordant with previous findings. The *Ldh-A* genotype of all the lake isolates, regardless of their mitochondrial lineage, has a typical lake profile (FF). The *Ldh-A* genotype of the isolates was also consistent with the microsatellite data since it provided evidence of discordance between the mitochondrial data and a nuclear coding region. Based on nucleotide variation analysis at the *Ldh-A* locus of *Daphnia*, Crease *et al.* (2010) found that this locus is under strong purifying selection in the lakes and the occurrence of a selective sweep in lake populations was associated with the appearance of the fast (F) allele at *Ldh-A*. Lactate dehydrogenase (LDH) catalyses the interconversion of pyruvate and lactate, is involved in the terminal step of anaerobic glycolysis, and the conversion of lactate to glucose in gluconeogenesis (Powers *et al*. 1991). In the fish *Fundulus heteroclitus*, LDH enzyme activity was found to change with temperature (Crawford and Powers 1989) and differences in *Ldh* gene expression exist between populations adapted to different thermal habitats (Schulte *et al*. 2000). *Ldh* may directly affect many biological functions such as: differences in oxygen consumption, metabolic flux, developmental rate, hatching time, swimming performance, survival at elevated temperatures (Powers and Schulte 1998), and tolerance to hydrostatic pressure (Nishiguchi *et al*. 2010). This study again points out the importance of the *LDH* locus to *Daphnia*'s survival in a pond versus a lake habitat.

What multi locus nuclear data reveals about lake and pond *Daphnia*

Relatively lower numbers of alleles at each locus (figure 2.5) as well as lower allelic richness in lakes (table 2.3) compared to the ponds may be due to genetic drift or a recent colonization of these lakes. This is in agreement with a previous study showing lower nucleotide diversity levels in lake *Daphnia* than in pond *Daphnia* at six protein coding loci (Omilian and Lynch 2009). However, the presence of three different mitochondrial lineages in the lakes may be contributing to the higher levels of mitochondrial nucleotide diversity observed in these lakes compared to the ponds. Locus *d174* may be located in a potential region involved in lake adaptation as this locus was found to be fixed for the same allele in all the lakes examined. Locus *d174* is found in the exon region of a zinc-finger protein (wFleaBase), which is a class of proteins that are involved in DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding (Laity *et al*. 2001). Further work on this genomic region is needed to determine whether it is in fact under strong natural selection in the lakes and whether the fixation at this locus was caused by drift or a selective sweep which is consistent with the locus being under strong natural selection.

Most of the microsatellite loci were in HWE (81-86%) in the lakes and ponds, except Solomon pond (57% loci in HWE). Genotype frequencies closer to HWE are typical of a CP, randomly mating pond population (Morgan *et al*. 2001) and deviations from HWE due to homozygous excess is indicative of inbreeding, which is the pattern seen in Solomon pond. The cyclical parthenogenetic life history of *Daphnia* makes it possible for populations to experience prolonged periods of clonal selection (Morgan *et al*. 2001), and

since Solomon pond was the only habitat sampled later in the season, clonal selection likely accounts for the HWE deviations in this pond. The occurrence of null alleles in this study as suggested by MICRO-CHECKER is quite low and is not detected at many loci that are out of HWE, therefore I suggest that null alleles have a minimal impact on the results of this study.

Difference in linkage disequilibrium between lakes and ponds

I found higher levels of linkage disequilibrium (LD) in the lake populations than in the ponds (table 2.6A) and there are three possible reasons to explain this observed pattern: (1) sampling strategy, (2) clonal selection, and (3) natural selection. (1) Sampling size or strategy may cause the high LD observed in the lakes because lake habitats are much larger than ponds and contain many more *Daphnia* individuals, and it is thus more difficult to get an accurate representation of the lake population, even though there is a lower effective population size observed in lakes compared to ponds (Omilian and Lynch 2009). (2) The cyclical parthenogenetic life history of *Daphnia* makes it possible for populations to experience prolonged periods of clonal selection (Morgan *et al*. 2001), especially in the lakes since *D. pulicaria* is present in the water column for extended periods of time before engaging in sexual reproduction. Since lake individuals engage in sex less often than pond populations (Cáceres and Tessier 2004) LD may decay more slowly in lakes than in ponds. (3) High LD in the lake populations may indicate that certain combinations of alleles are particularly favored by natural selection in one environment, but not in the other, such as is often seen in ecological species (Schluter 2009). Furthermore, LD is found between certain microsatellite markers and specific

mitochondrial type (ppx or wpc; table 2.7), and between some markers and the two different mtDNA clades (clades A and B). This indicates that the three clades (A, B, and C) may be on different evolutionary trajectories in the lakes.

Population differentiation between the different habitats

Based on the F_{ST} estimates (table 2.4) this study suggests low levels of gene flow between lakes and ponds, and based on migration rates (table 2.5) a slightly higher level of gene flow from ponds to lakes than the other way around is evident. However, the gene flow estimates based on frequency data of 21 microsatellite markers are lower than the estimate of Omilian and Lynch (2009) based on 6 nuclear coding regions. Gene flow among the ponds is much higher (1.5 migrants/generation) than among the lakes (0.5- 0.8 migrants/generation), indicating that, despite their close geographic proximity, habitat segregation among the lakes is common. The F_{ST} values between the two groups in Lawrence Lake (based on mitochondrial type) was low (0.109), but significantly different from 0 (table S.1 in appendix), indicating that there may still be some distinction between the two mitochondrial groups within this lake.

Based on the results of STRUCTURE analysis (figure 2.6), the nuclear neighbor-joining (NJ) phylogram (figure 2.4), as well as the principal component analysis (figure 2.7A), it can be clearly seen that lakes and ponds form two distinct groups. I also detected two genetic clusters within the lakes from a separate STRUCTURE analysis (figure 2.7B). However, this clustering pattern within the lakes was not as pronounced as the distinction between lakes and ponds. Although this genetic clustering observed in the lakes is not likely caused by local adaptations to food resources (Allen *et al*. 2010), other differences

may be responsible for this pattern. For example, the lack of complete homogenization of nuclear genomes in some lake individuals, compared to more established or older lake groups may be causing this partitioning within the lakes.

Difference in predation pressures and conditions in the lakes may account for the observed genetic pattern. The invertebrate *Chaoborus* preys mainly on *Daphnia* in both lakes and ponds, while the Bluegill Sunfish (*Lepomis macrochirus*) is the most abundant planktivorous fish species in the surveyed lakes (Werner *et al.* 1977; Osenberg *et al.* 1988; Werner and Hall 1988) and is not found in ponds. Three lakes II has the highest *Chaoborus* density compared to the other lakes (Leibold and Tessier 1991) and has dystrophic conditions and high dissolved organic content (Haney and Hall 1975; Leibold and Tessier 1991), much like the pond environment. *Daphnia* experiences the strongest predation pressure from fish (*Lepomis macrochirus*) in Warner Lake and lowest in Lawrence Lake (Osenberg *et al.* 1988; Leibold and Tessier 1991).

In addition to differences between the lakes in this study, each lake contains two habitats: the shallow, warm epilimnion, where fish predation is high (Hall and Werner 1977; Werner and Hall 1988) and the deeper colder, anoxic hypolimnion, where fish are usually absent. Since clonal habitat and depth specialization is common in *D. pulicaria*, it may be useful to conduct future work exploring the relationship between the genetic clustering observed in this study and habitat partitioning within the same lake habitat. Although the sampling protocol in this study was consistent across all the lakes, it does not allow such an analysis.

Discordance between multiple nuclear markers and mitochondrial phylogenies

Describing and interpreting historical and contemporary patterns of divergence between species is one of the principal goals of evolutionary biology. However, for recently diverged populations or species with incomplete reproductive isolation, gene genealogies from different markers may be discordant, which often supports a history of divergence with gene flow. The analysis of nuclear microsatellite markers indicates a high F_{ST} value between lakes and ponds (table 2.4) with low levels of gene flow (table 2.5), and a clustering pattern separating lakes and ponds (figures 2.5, 2.6, 2.7). At the same time, the mitochondrial phylogenetic analysis indicates that some lake haplotypes group with pond haplotypes (figure 2.2). This pattern suggests that introgression of pond *D. pulex* is occurring in the lakes and the low levels of mitochondrial sequence divergence between ppx mtDNA haplotypes in lakes and ponds indicates that these events are very recent.

Pfrender and colleagues (2000) proposed that much of the subdivision within ponds in Oregon is due to some populations containing "lake-like" nuclear alleles. They have suggested that permanent lake lineages periodically colonize temporary ponds following floods and quickly adapt to an ephemeral habitat (Pfrender *et al.* 2000), likely through hybridization and introgression with the resident pond lineage. My data shows that the most common mitochondrial haplotype is shared between ponds and lakes (figure 2.3) and this suggests a pattern of pond individuals invading lakes and introgressing into the lake population, since at the nuclear level all the lakes group together (figures 2.5, 2.6, 2.7). It is quite possible that certain individuals in ponds already contain "lake-like" alleles and can more easily migrate and introgress into a permanent lake habitat. The

results of this study show that gene flow is occurring in both directions (from ponds to lakes and from lakes to ponds), but introgression of pond migrants into lakes happens more easily than the introgression of lake migrants into ponds.

Conclusions

This study explores the evolutionary consequences of habitat transition events in lake and pond *Daphnia*. The mitochondrial phylogenetic survey revealed the occurrence of three different mitochondrial lineages within lake *D. pulicaria*, which likely correspond to three separate habitat transition events into the lakes. This finding based on the mitochondrial ND5 marker is in contrast with the phylogenetic signal revealed by the nuclear markers that consistently group *Daphnia* based on habitat. The strong discordant phylogenetic signal between nuclear and mitochondrial markers suggests that hybridization and introgression of pond *D. pulex* genes into the *D. pulicaria* genome has been occurring in the lakes and that some of these events are relatively recent. Additionally, the detection of two genetic units within the lakes needs further investigation to determine the cause of this genetic subdivision within the lakes. Despite historical evidence for hybridization and gene flow revealed by phylogenetic analysis between lake and pond populations, population genetic data indicates low levels of contemporaneous gene flow suggesting the existence of strong habitat isolating barriers between ponds and lakes. The results of this study point to a divergence with gene flow scenario for the speciation of pond *D. pulex* and lake *D. pulicaria*.

Table 2.1A Habitat location and sampling size for pond *Daphnia pulex* populations. Habitat locations, mitochondrial and nuclear profiling for the study of 271 *Daphnia* isolates with *ID,* location code, *N1*, number of individuals analyzed using mitochondrial marker ; *N2*, number of individuals analyzed using 21 microsatellite markers and mitochondrial marker; *Ldh*, nuclear lactate dehydrogenase A profiling, SS- homozygous slow, FF- homozygous fast, SF-heterozygous; mtDNA, mitochondrial profiling, ppx, panarctic *Daphnia pulex*, wpc, western *Daphnia pulicaria,*; Cld, indicates which *Daphnia pulex* clade the individuals belong to in phylogenetic analysis based on mtDNA (figure 2.2); Rep, reproduction mode, CP-cyclical parthenogenesis, OP-obligate parthenogenesis; Prov/St, province or state of habitat, MI-Michigan, ON-Ontario.

Ponds	**ID**	**Prv/St**	**Lat**	**Long**	**N1**	**N2**	**Ldh**	**mtDNA**	**Cld**	**Rep**
Disputed	Disp	ON	42.175	-83.035	52	50	SS	ppx	A	CP
Solomon	Sol	MI	42.719	-85.388	53	51	SS	ppx	A	CP
Canard 1	Can1	ON	42.12	-82.98	13	-	SS/SF	ppx	A	CP/OP
Canard 2	Can2	ON	42.16	-83.02	3	-	SF	ppx	A	OP
Canard 3	Can3	ON	42.12	-82.92	6	-	SS/SF	ppx	A	CP/OP
Gesto	Ges	ON	42.13	-82.88	2	-	SF	ppx	A	CP/OP
West Gull	WG	MI	42.41	-85.44	4	-	SS	ppx	A	OP
Grimey	Grm	MI	42.31	-85.36	3	-	SS	ppx	A	OP
Total					136	101				

Table 2.1B Habitat location and sampling size for lake *Daphnia pulicaria* populations with *ID,* location code, *N1*, number of individuals analyzed using mitochondrial marker only; *N2*, number of individuals analyzed using 21 microsatellite markers and mitochondrial marker; *Ldh*, nuclear lactate dehydrogenase A profiling, FF- homozygous fast; mtDNA, mitochondrial profiling ppx, panarctic *Daphnia pulex*, wpc, western *Daphnia pulicaria,* epc, eastern *Daphnia pulicaria,*; Cld, indicates which *Daphnia pulex* clade the individuals belong to in the phylogenetic tree based on mtDNA (figure 2.2); Rep, reproduction mode, CP-cyclical parthenogenesis, OP-obligate parthenogenesis; Prov/St, province or state of habitat, MI-Michigan, ON-Ontario, IL-Illinois.

Lakes	**ID**	**Prv/St**	**Lat**	**Long**	**N1**	**N2**	**Ldh**	**mtDNA**	**Cld**	**Rep**
Lawrence	Law	MI	42.26	-85.21	86	86	FF	40 ppx 46 wpc	A/B	CP
Three Lakes II	3L2	MI	42.21	-85.26	38	35	FF	37 ppx, 1wpc	A/B	CP
Warner	Warn	MI	42.28	-85.31	63	44	FF	4 ppx, 59wpc	A/B	CP
Bassett	Bas	MI	42.40	-85.29	12	-	FF	9 ppx, 3 wpc	-	CP
Mill	Mill	MI	42.27	-85.15	14	-	FF	13 ppx, 1 wpc	-	CP
Long	Lng	IL	40.14	-87.44	4	-	FF	ppx	A	CP
Sportsman	Spm	IL	40.14	-87.44	3	-	FF	ppx	A/B	CP
Clear	Clr	IL	40.14	-87.44	3	-	FF	ppx	A	CP
Big Gull	BG	ON	44.88	-78.75	3	-	FF	2 wpc, 1epc	-	CP
Total					226	165		113 ppx 112 wpc 1epc		

Table 2.2 Genetic diversity indexes for 2 pond and 3 lake populations based on a 687 bp sequence of the mitochondrial *NADH dehydrogenase* 5 gene.

	Number of isolates	**Number of haplotypes**	**Haplotype diversity**	**Nucleotide diversity**
Ponds	**108**	**25**	**0.887**	**0.006**
Solomon	54	17	0.793	0.006
Disputed	54	8	0.751	0.002
Lakes	**192**	**17**	**0.820**	**0.025**
Warner	63	7	0.313	0.005
Three Lakes II	37	6	0.751	0.009
Lawrence	92	8	0.765	0.025

Table 2.3 Genetic diversity at 21 microsatellite loci for 6 populations of *Daphnia pulex* and *Daphnia pulicaria*. *N*, sample size; *A*, number of alleles; A_r, allele richness; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient; P_{HW}, exact *P*-value for Hardy-Weinberg equilibrium test; r, frequency of null allele. Values in bold indicate a deviation from HWE after sequential Bonferroni correction. Abbreviations for the different populations are given in tables 2.1A, 2.1B.

Locus	**Index**	**Disp**	**Sol**	**Warn**	**Law**	**3L2**
d070	*N*	49	48	45	87	35
	A/A_r	5/4.931	6/5.296	3/2.604	4/3.985	3/2.994
	H_O	0.6735	0.5625	0.8444	0.9081	0.4571
	H_E	0.7259	0.6340	0.5136	0.6475	0.4932
	F_{IS}	0.0729	0.1138	-0.6771	-0.4058	0.0465
	P_{HW}	0.0879	**0.0080**	**0.0000**	**0.0000**	0.0845
	r	-	**-**	**-**	**-**	-
d027	*N*	51	50	45	86	35
	A/A_r	8/6.627	7/6.550	3/2.941	3/3.000	3/3.000
	H_O	0.6275	0.7400	0.4222	0.3372	0.4000
	H_E	0.7164	0.6600	0.5271	0.5295	0.5694
	F_{IS}	0.1252	-0.1226	0.1997	0.3645	0.2688
	P_{HW}	0.2938	0.5870	0.3280	**0.0004**	0.0062
	r	-	-	-	0.158	-
d117	*N*	50	49	45	86	35

Locus	**Index**	**Disp**	**Sol**	**Warn**	**Law**	**3L2**
d117	A/A_r	7/6.193	7/6.238	3/2.604	2/2.000	3/3.000
	H_O	0.3200	0.5102	0.4889	0.4070	0.4286
	H_E	0.3683	0.5742	0.5061	0.3527	0.4302
	F_{IS}	0.1323	0.1124	0.0066	0.1549	0.0113
	P_{HW}	0.1221	**0.0009**	0.6329	0.2182	**0.0021**
	r	-	**-**	-	-	**-**
d078	*N*	51	50	45	84	35
	A/A_r	10/7.646	9/8.436	1/1.000	2/1.991	1/1.000
	H_O	0.8040	0.6	-	0.0238	-
	H_E	0.7321	0.6946	-	0.0237	-
	F_{IS}	-0.9920	0.1373	-	-0.0061	-
	P_{HW}	0.0278	**0.0001**	-	1.0000	-
	r	-	**-**	-	-	-
d087	*N*	48	46	43	82	35
	A/A_r	7/5.812	8/7.502	3/2.261	2/2.000	2/2.000
	H_O	0.6042	0.4783	0.0465	0.4390	0.2000
	H_E	0.7347	0.7131	0.0462	0.3582	0.2273
	F_{IS}	0.1792	0.3318	-0.0056	-0.227	0.1250
	P_{HW}	0.0364	**0.0002**	1	0.0576	0.4450
	r	-	0.157	-	-	-
d088	*N*	50	46	45	85	35
	A/A_r	5/4.989	7/6.591	4/3.588	4/3.894	3/3.000
	H_O	0.6800	0.5000	0.2444	0.6	0.4857
	H_E	0.7313	0.5538	0.2257	0.4862	0.4207

Locus	Index	Disp	Sol	Warn	Law	3L2
d088	F_{IS}	0.0708	0.0980	-0.0782	-0.2358	-0.1677
	P_{HW}	0.1994	0.2999	1.0000	**0.0000**	1.0000
	r	-	-	-	**-**	-
d166	*N*	49	41	43	85	35
	A/A_r	5/4.998	6/5.915	3/2.866	4/3.883	2/1.806
	H_O	0.4898	0.5854	0.5814	0.5294	0.0286
	H_E	0.72691	0.7552	0.5152	0.5165	0.0286
	F_{IS}	0.3285	0.2271	-0.1495	-0.0252	-
	P_{HW}	0.0022	**0.0006**	0.0074	0.0111	1.0000
	r	0.156	0.111	-	-	-
d050	*N*	50	46	45	85	35
	A/A_r	4/3.160	3/2.630	1/1.000	2/2.000	2/1.999
	H_O	0.1400	0.4348	-	0.0000	0.0571
	H_E	0.2331	0.4589	-	0.0685	0.1093
	F_{IS}	0.4019	0.0531	-	1.0000	0.4815
	P_{HW}	0.0363	0.4493	-	**0.0000**	0.0877
	r	0.141	-	-	0.186	-
d127	*N*	50	41	45	80	34
	A/A_r	4/3.925	3/3.000	2/1.992	4/4.000	3/2.973
	H_O	0.5600	0.4878	0.0222	0.3500	0.6765
	H_E	0.5727	0.4264	0.1061	0.4380	0.5342
	F_{IS}	0.0224	-0.1461	0.7930	0.2020	-0.2911
	P_{HW}	0.0189	0.5564	0.0030	**0.0000**	0.0090
	r	-	-	0.182	**-**	-

Locus	Index	Disp	Sol	Warn	Law	3L2
d105	N	49	41	40	81	34
	A/A_r	6/5.020	3/3.000	4/3.897	3/3.000	3/3.000
	H_O	0.4082	0.4878	0.3000	0.2840	0.2941
	H_E	0.5306	0.5658	0.5117	0.6050	0.5953
	F_{IS}	0.2326	0.1393	0.4295	0.5322	0.4777
	P_{HW}	0.0024	0.0235	0.0014	**0.0000**	**0.0000**
	r	0.112	-	0.205	0.232	0.207
d029	N	50	48	44	86	34
	A/A_r	5/4.971	6/5.604	3/2.999	3/3.000	3/3.000
	H_O	0.5400	0.5833	0.7727	0.7209	0.7353
	H_E	0.6956	0.6985	0.5687	0.5695	0.5272
	F_{IS}	0.2254	0.1663	-0.3985	-0.2679	-0.4133
	P_{HW}	0.0180	0.0169	**0.0000**	0.0125	0.0010
	r	0.101	-	**-**	-	-
d174	N	50	47	45	87	35
	A/A_r	7/6.377	9/7.945	1/1.000	1/1.000	1/1.000
	H_O	0.5400	0.7021	-	-	-
	H_E	0.7640	0.8003	-	-	-
	F_{IS}	0.2953	0.1238	-	-	-
	P_{HW}	0.0016	0.2065	-	-	-
	r	0.141	-	-	-	-
d015	N	48	29	39	87	34
	A/A_r	11/8.863	8/8.000	2/1.690	1/1.000	2/1.829
	H_O	0.5417	0.4483	0.0256	-	0.0294

Locus	Index	Disp	Sol	Warn	Law	3L2
d015	H_E	0.7730	0.7066	0.0256	-	0.0294
	F_{IS}	0.3015	0.3697	-	-	-
	P_{HW}	**0.0000**	0.0016	1.0000	-	1.0000
	r	0.138	0.182	-	-	-
d186	N	46	38	44	86	35
	A/A_r	8/7.356	7/6.943	4/4	4/4	5/4.805
	H_O	0.5217	0.3947	0.7045	0.7209	0.5429
	H_E	0.7076	0.8193	0.7325	0.6832	0.5019
	F_{IS}	0.2648	0.5216	0.0092	-0.0556	-0.0752
	P_{HW}	**0.0003**	**0.0000**	**0.0005**	0.0191	0.9338
	r	0.103	0.247	**-**	-	-
d111	N	44	34	-	86	35
	A/A_r	8/7.870	4/4.000	1/1.000	2/1.884	1/1.000
	H_O	0.8182	0.4118	-	0.0116	-
	H_E	0.8130	0.6054	-	0.0116	-
	F_{IS}	-0.0065	0.3231	-	0.0000	-
	P_{HW}	0.8666	**0.0004**	-	1.0000	-
	r	-	0.147	-	-	-
d153	N	45	36	39	87	35
	A/A_r	6/5.998	2/2.000	3/2.652	2/1.874	1/1.000
	H_O	0.7778	0.5278	0.0769	0.0115	-
	H_E	0.8040	0.4409	0.0756	0.0115	-
	F_{IS}	0.0330	-0.2004	-0.0174	0.0000	-
	P_{HW}	0.4323	0.2813	1.0000	1.0000	-

Locus	Index	Disp	Sol	Warn	Law	3L2
d153	*r*	-	-	-	-	-
d042	*N*	32	32	37	85	35
	A/A_r	6/5.884	4/3.906	2/2.000	4/3.989	3/3.000
	H_O	0.1563	0.2500	0.5405	0.5765	0.3429
	H_E	0.3120	0.4896	0.4976	0.5093	0.5453
	F_{IS}	0.5032	0.4934	-0.1559	-0.1328	0.3564
	P_{HW}	0.0059	0.0011	0.7334	0.0024	0.0235
	r	0.179	0.219	-	-	-
d006	*N*	49	48	44	86	35
	A/A_r	8/7.525	8/7.032	3/2.998	5/4.974	3/2.806
	H_O	0.7551	0.5625	0.7500	0.5930	0.2571
	H_E	0.7690	0.7886	0.5465	0.5947	0.2306
	F_{IS}	0.0182	0.2889	-0.4112	0.0028	-0.1131
	P_{HW}	0.2495	**0.0000**	0.0226	0.0519	1.0000
	r	-	0.139	-	-	-
d148	*N*	48	46	42	87	34
	A/A_r	5/4.391	4/3.726	3/3.000	3/3.000	4/4.000
	H_O	0.1667	0.1087	0.5714	0.4828	0.7941
	H_E	0.3805	0.1842	0.6004	0.4757	0.7239
	F_{IS}	0.5646	0.4125	0.0349	-0.0149	-0.0749
	P_{HW}	**0.0000**	**0.0001**	0.0006	0.5694	**0.0000**
	r	0.233	0.131	-	-	**-**
d182	*N*	44	34	38	76	33
	A/A_r	5/4.644	5/4.853	3/3.000	4/4.000	3/2.853

Locus	Index	Disp	Sol	Warn	Law	3L2
d182	H_O	0.2955	0.5000	0.579	0.500	0.4849
	H_E	0.4582	0.6602	0.5632	0.4468	0.3800
	F_{IS}	0.3578	0.2455	-0.0828	-0.1201	-0.2956
	P_{HW}	0.0086	0.0058	0.0765	0.9397	0.3726
	r	0.160	-	-	-	-
d016	N	51	50	44	87	35
	A/A_r	4/3.567	7/5.738	3/2.999	3/3.000	3/2.994
	H_O	0.3137	0.7000	0.4318	0.3333	0.3143
	H_E	0.3896	0.6343	0.3790	0.4257	0.2787
	F_{IS}	0.1964	-0.1047	-0.1263	0.2181	-0.1257
	P_{HW}	0.0452	0.1350	0.8635	**0.0000**	1.0000
	r	-	-	-	**-**	-
Total number of loci out of HWE in each population		3	9	3	4	3
heterozygote deficiency		3	8	1	2	2

Table 2.4 Pairwise F_{ST} estimates between five populations of *Daphnia pulex* and *Daphnia pulicaria* based on 21 microsatellite loci. All F_{ST} values are significantly different from 0 ($P < 0.05$). Abbreviations for the different populations are given in tables 2.1A, 2.1B.

	Disp	**Sol**	**Warn**	**Law**
Sol	0.0790			
Warn	0.4750	0.4434		
Law	0.4810	0.4505	0.0764	
3L2	0.4669	0.4384	0.1466	0.0868

Table 2.5 Migration rates (Nm, number of migrants per generation) between lake and pond populations. Results are averaged over 21 microsatellite loci. Source populations are listed by column, recipient populations listed by row. Abbreviations for the different populations are given in tables 2.1A, 2.1B.

	Disp	**Sol**	**Warn**	**Law**	**3L2**
Disp	-	1.5206	1.1256	1.3294	1.1151
Sol	1.5333	-	1.1073	1.3376	1.0606
Warn	0.5278	0.5855	-	0.6715	0.7656
Law	0.5622	0.5966	0.8022	-	0.6426
3L2	0.5510	0.4829	0.4670	0.5539	-

Table 2.6 Linkage disequilibrium between all pairs of loci in each population of *Daphnia pulex* and *Daphnia pulicaria*. Abbreviations for the different populations are given in tables 2.1A, 2.1B, with Law, all individuals from Lawrence Lake; Law1, Lawrence lake individuals with *Daphnia pulicaria* mitochondrial DNA; Law2, Lawrence lake individuals with *Daphnia pulex* mitochondria DNA. Values are P-values and those in bold are significant after Bonferroni correction.

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d070	*d027*	0.3229	0.5593	0.0245	0.8807	0.2030	0.0233	0.6085
d070	*d117*	0.2108	0.0300	0.0143	0.0321	**0.0002**	0.3632	**0.0000**
d027	*d117*	0.1565	0.0473	**0.0002**	0.0298	0.1988	0.0000	0.1893
d070	*d078*	0.0411	0.6930	-	-	0.4492	1.0000	0.2788
d027	*d078*	0.5372	0.8615	-	-	0.4042	1.0000	0.1520
d117	*d078*	0.4079	0.8828	-	-	1.0000	0.5124	1.0000
d070	*d087*	0.2689	0.9627	0.0179	1.0000	**0.0004**	0.0089	0.0006
d027	*d087*	0.0269	0.9546	**0.0001**	0.2941	0.1061	0.0719	0.0442
d117	*d087*	0.0280	0.6893	0.0406	0.0103	0.6663	0.7873	0.1826
d078	*d087*	0.1104	0.8330	-	-	1.0000	1.0000	1.0000
d070	*d088*	0.2517	0.6070	0.0026	0.6859	0.2752	0.5186	0.5171
d027	*d088*	0.7142	0.5491	0.2117	0.0198	0.0040	0.0157	0.0004
d117	*d088*	0.8801	0.0146	0.6689	0.1766	0.0793	0.2341	0.2779
d078	*d088*	0.4496	0.4309	-	-	0.0835	0.1173	0.2352
d087	*d088*	0.0208	0.5617	0.048	1.0000	0.3014	0.5418	0.7930
d070	*d166*	0.0460	0.5438	0.5568	0.1113	**0.0000**	0.0152	**0.0000**

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d027	*d166*	0.5832	0.5820	1.0000	0.0032	0.0728	0.3111	0.1586
d117	*d166*	0.0821	0.3670	0.4714	0.0019	0.2326	0.6697	0.0319
d078	*d166*	**0.0000**	0.9180	-	-	0.7220	1.0000	1.0000
d087	*d166*	0.0036	0.8306	1.0000	0.6378	0.3100	0.0269	0.5444
d088	*d166*	0.0235	0.6939	0.0270	0.2714	0.3417	0.8761	0.0336
d070	*d050*	0.9059	0.3776	0.0114	-	0.3755	-	0.0639
d027	*d050*	0.9720	0.1275	0.4003	-	0.1282	-	0.3775
d117	*d050*	0.4386	0.1849	0.3625	-	0.2926	-	0.3390
d078	*d050*	0.1372	0.8918	-	-	1.0000	-	1.0000
d087	*d050*	0.0198	0.8653	0.2883	-	0.2700	-	0.0742
d088	*d050*	0.4377	0.5990	0.1793	-	0.4529	-	0.6120
d166	*d050*	0.0056	0.5164	1.0000	-	0.4679	-	0.4011
d070	*d127*	0.0367	0.5859	0.0122	0.2200	**0.0001**	0.0140	0.0009
d027	*d127*	0.6451	0.9584	0.0046	0.0424	**0.0000**	0.0006	0.0105
d117	*d127*	0.4410	0.7649	0.2507	0.1021	0.4101	0.1279	**0.0001**
d078	*d127*	0.0803	0.2779	-	-	1.0000	1.0000	1.0000
d087	*d127*	0.5803	0.1159	0.5398	0.0476	0.0329	0.0796	0.0610
d088	*d127*	0.3851	0.3616	0.3001	0.5544	0.4307	0.1512	0.5104
d166	*d127*	0.0051	0.4993	1.0000	0.7980	0.3006	0.6878	0.0128
d050	*d127*	0.2058	0.3092	0.0335	-	0.2046	-	0.4963
d070	*d105*	0.2639	0.2303	0.2371	0.6405	**0.0002**	0.0838	**0.0001**
d027	*d105*	0.6451	0.8585	**0.0000**	**0.0001**	**0.0000**	**0.0003**	0.0046
d117	*d105*	0.0232	0.1953	0.0062	0.3577	**0.0000**	0.0340	0.0090
d078	*d105*	0.1004	0.7109	-	-	0.8359	1.0000	1.0000

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d087	*d105*	0.0173	0.1392	0.0126	0.0210	0.0018	**0.0000**	0.2793
d088	*d105*	0.0891	0.9456	0.0073	**0.0002**	0.0354	0.1995	0.0784
d166	*d105*	0.1412	0.5308	1.0000	0.2421	**0.0000**	0.0038	**0.0002**
d050	*d105*	0.3062	0.7628	0.1204	-	0.0019	-	0.0146
d127	*d105*	0.1527	0.7016	0.0029	0.0527	0.0688	0.2648	0.0744
d070	*d029*	0.4172	0.5129	0.0227	0.3511	**0.0000**	**0.0000**	0.0006
d027	*d029*	0.6347	0.7234	0.0391	0.4485	0.0687	0.0225	0.0157
d117	*d029*	0.7364	0.4037	0.0251	0.0007	0.0105	0.0080	0.0945
d078	*d029*	**0.0000**	0.9579	-	-	0.2938	0.1381	1.0000
d087	*d029*	0.4728	0.0524	0.0735	0.0208	0.0137	0.0117	0.3971
d088	*d029*	0.2845	0.8694	0.6441	0.0340	**0.0003**	0.0304	0.0109
d166	*d029*	0.0541	0.1825	1.0000	0.9096	0.0079	0.1079	**0.0001**
d050	*d029*	0.3718	0.3077	0.0811	-	0.8126	-	1.0000
d127	*d029*	0.6961	0.2176	0.0768	0.0326	0.0251	0.0164	0.0167
d105	*d029*	0.2008	0.7404	0.0473	0.0887	0.3711	0.1590	0.0109
d070	*d174*	0.9178	0.0386	-	-	-	-	-
d027	*d174*	0.0308	0.8284	-	-	-	-	-
d117	*d174*	0.2003	0.6787	-	-	-	-	-
d078	*d174*	0.3648	0.7642	-	-	-	-	-
d087	*d174*	0.5073	0.5767	-	-	-	-	-
d088	*d174*	0.2837	0.8133	-	-	-	-	-
d166	*d174*	0.2761	0.5419	-	-	-	-	-
d050	*d174*	0.9769	0.2962	-	-	-	-	-
d127	*d174*	0.5408	0.5273	-	-	-	-	-

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d105	*d174*	0.5305	0.0362	-	-	-	-	-
d029	*d174*	0.2776	0.5581	-	-	-	-	-
d070	*d015*	0.2597	0.2912	0.5368	1.0000	-	-	-
d027	*d015*	0.6012	0.7128	0.3666	1.0000	-	-	-
d117	*d015*	0.8180	0.3031	1.0000	0.5221	-	-	-
d078	*d015*	0.8447	0.7610	-	-	-	-	-
d087	*d015*	0.3415	0.0613	1.0000	1.0000	-	-	-
d088	*d015*	0.1766	0.0456	0.2268	1.0000	-	-	-
d166	*d015*	0.7088	0.6017	1.0000	-	-	-	-
d050	*d015*	0.1922	0.2892	1.0000	-	-	-	-
d127	*d015*	0.0530	0.7471	1.0000	1.0000	-	-	-
d105	*d015*	0.8454	0.9000	0.5944	-	-	-	-
d029	*d015*	0.5335	0.0065	1.0000	1.0000	-	-	-
d174	*d015*	0.8019	0.7903	-	-	-	-	-
d070	*d186*	0.1842	0.9020	0.0216	0.0752	**0.0000**	**0.0000**	**0.0000**
d027	*d186*	0.8492	0.0661	0.0110	**0.0000**	0.0007	**0.0003**	0.1497
d117	*d186*	0.1164	0.6088	0.0223	**0.0000**	**0.0000**	**0**.0000	0.0028
d078	*d186*	0.1628	0.3822	-	-	1.0000	0.4670	0.2383
d087	*d186*	0.4085	0.2734	0.0006	0.0976	0.0068	0.4185	**0.0000**
d088	*d186*	0.8617	0.7961	0.0032	**0.0000**	0.1478	0.0297	0.8967
d166	*d186*	0.1635	0.5965	0.5461	**0.0000**	0.0052	0.4396	**0.0000**
d050	*d186*	0.0456	0.5383	0.1661	-	0.0359	-	0.1436
d127	*d186*	0.3035	0.4304	0.0015	0.0508	**0.0003**	0.0034	0.0069
d105	*d186*	0.7632	0.0970	0.2565	**0.0000**	**0.0000**	0.0491	**0.0000**

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d029	*d186*	0.2996	0.2360	0.0037	**0.0004**	**0.0000**	**0.0000**	0.0020
d174	*d186*	0.6384	0.3397	-	-	-	-	-
d015	*d186*	0.0157	0.8302	1.0000	0.1979	-	-	-
d070	*d111*	0.1264	0.8664	-	-	1.0000	1.0000	-
d027	*d111*	0.31514	0.9332	-	-	1.0000	1.0000	-
d117	*d111*	0.9872	0.5094	-	-	1.0000	0.5113	-
d078	*d111*	0.5769	0.3496	-	-	1.0000	1.0000	-
d087	*d111*	0.5433	0.2117	-	-	1.0000	1.0000	-
d088	*d111*	0.3055	0.1577	-	-	1.0000	0.5403	-
d166	*d111*	0.0144	0.7448	-	-	0.2083	0.2132	-
d050	*d111*	0.8156	0.2891	-	-	1.0000	-	-
d127	*d111*	0.4062	0.3817	-	-	1.0000	1.0000	-
d105	*d111*	0.3123	0.4123	-	-	-	-	-
d029	*d111*	0.4113	0.9564	-	-	0.2714	0.6345	-
d174	*d111*	0.3305	0.1897	-	-	-	-	-
d015	*d111*	0.3524	0.3293	-	-	-	-	-
d186	*d111*	0.5375	0.2074	-	-	1.0000	1.0000	-
d070	*d153*	0.8245	0.5564	-	0.4631	1.0000	1.0000	-
d027	*d153*	0.6554	0.6941	-	0.0408	1.0000	1.0000	-
d117	*d153*	0.4994	0.5325	-	0.7437	1.0000	0.5153	-
d078	*d153*	0.1445	0.3983	-	-	1.0000	1.0000	-
d087	*d153*	0.8846	0.4263	-	1.0000	1.0000	1.0000	-
d088	*d153*	0.0856	0.5474	-	1.0000	1.0000	0.5326	-
d166	*d153*	0.3243	0.6337	-	0.4361	0.2082	0.2159	-

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d050	*d153*	0.0619	0.2360	-	-	1.0000	-	-
d127	*d153*	0.1046	0.8428	-	1.0000	1.0000	1.0000	-
d105	*d153*	0.0598	0.1289	-	1.0000	-	-	-
d029	*d153*	0.1009	0.9890	-	0.2880	0.2721	0.6294	-
d174	*d153*	0.9791	0.8957	-	-	-	-	-
d015	*d153*	0.1512	0.9095	-	1.0000	-	-	-
d186	*d153*	0.5645	0.6139	-	0.0258	1.0000	1.0000	-
d111	*d153*	0.2717	**0.0000**	-	-	0.0111	0.0205	-
d070	*d042*	0.4281	0.8003	0.1740	0.9432	0.0028	0.2612	**0.0000**
d027	*d042*	0.5612	0.6547	0.0045	0.6522	0.0240	0.0008	0.1505
d117	*d042*	0.6636	0.9802	0.0009	0.0031	**0.0000**	**0.0000**	**0.0000**
d078	*d042*	0.4106	0.2724	-	-	0.6474	1.0000	1.0000
d087	*d042*	0.5893	0.0149	0.0169	0.1101	**0.0000**	1.0000	**0.0002**
d088	*d042*	0.2873	0.6088	0.1568	0.0123	0.3619	0.1758	0.1343
d166	*d042*	0.2866	0.9776	0.6124	0.8964	0.0273	0.2862	**0.0000**
d050	*d042*	0.3071	1.0000	0.4248	-	0.2205	-	0.1240
d127	*d042*	0.7419	0.3091	0.1115	0.3571	0.3423	0.0465	0.0005
d105	*d042*	0.1647	0.8220	0.0087	0.0466	0.0089	0.0853	**0.0000**
d029	*d042*	0.5841	0.3264	0.0399	0.1398	0.0040	0.0079	0.0283
d174	*d042*	0.5521	0.9993	-	-	-	-	-
d015	*d042*	0.3346	0.6411	0.3136	1.0000	-	-	-
d186	*d042*	0.6303	0.9528	0.0014	0.0255	**0.0000**	0.0060	**0.0000**
d111	*d042*	0.6504	0.3151	-	-	0.4616	0.3684	-
d153	*d042*	0.6174	0.1458	-	1.0000	0.4578	0.3713	-

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d070	*d006*	0.3025	0.8888	0.5581	1.0000	**0.0000**	0.0119	**0.0000**
d027	*d006*	**0.0002**	0.4137	0.0635	0.0028	0.0007	0.0177	0.0860
d117	*d006*	0.0396	0.8015	**0.0000**	0.0101	**0.0000**	**0.0000**	0.0012
d078	*d006*	0.1400	0.6920	-	-	0.1417	0.1733	0.5588
d087	*d006*	0.0087	0.1247	0.4939	0.0472	0.0117	0.4664	0.0020
d088	*d006*	0.2287	0.0107	0.2612	0.0453	**0.0004**	0.0442	0.0072
d166	*d006*	0.0922	0.8248	0.2506	0.0041	0.0072	0.2407	**0.0002**
d050	*d006*	0.8876	0.9059	1.0000	-	0.0402	-	0.1350
d127	*d006*	0.1710	0.3855	0.0033	0.0037	0.0328	0.1105	0.0091
d105	*d006*	0.0711	0.5077	0.0430	0.0022	0.0538	0.2513	0.0403
d029	*d006*	0.6334	0.7319	0.2264	**0.0003**	**0.0000**	0.0029	**0.0000**
d174	*d006*	0.6294	0.6163	-	-	-	-	-
d015	*d006*	0.1756	0.1151	1.0000	0.3594	-	-	-
d186	*d006*	0.6055	0.4057	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
d111	*d006*	0.0150	0.6942	-	-	0.2105	0.5200	-
d153	*d006*	0.6958	0.7531	-	0.2274	0.2192	0.5298	-
d042	*d006*	0.5237	0.0585	0.0334	0.5250	0.0024	0.2334	0.0008
d070	*d148*	0.4071	0.6425	**0.0001**	0.2957	**0.0000**	0.0477	**0.0000**
d027	*d148*	0.6459	0.8643	0.0200	0.0908	**0.0000**	0.1935	0.0021
d117	*d148*	0.2905	0.4904	0.0012	0.0302	0.9701	0.7968	0.0290
d078	*d148*	0.7111	0.8593	-	-	0.0674	0.4045	0.1278
d087	*d148*	0.4818	0.1293	**0.0000**	1.0000	**0.0000**	0.0108	**0.0000**
d088	*d148*	0.0572	0.8676	**0.0004**	0.0066	0.0529	0.6022	0.0373
d166	*d148*	0.2417	0.5081	1.0000	0.2846	**0.0001**	**0.0000**	0.1034

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d050	*d148*	0.0453	0.8378	0.0039	-	0.5816	-	0.0466
d127	*d148*	0.8829	0.0968	0.0011	1.0000	0.3235	0.4475	0.3366
d105	*d148*	0.1211	0.0594	0.0088	0.0595	0.0007	0.0025	0.1667
d029	*d148*	0.9187	0.4368	**0.0003**	0.0038	0.0006	0.0207	0.0147
d174	*d148*	0.2053	0.4296	-	-	-	-	-
d015	*d148*	0.6260	0.4428	0.2075	-	-	-	-
d186	*d148*	0.3483	0.0047	**0.0000**	**0.0000**	**0.0002**	0.3957	**0.0000**
d111	*d148*	0.5440	0.0022	-	-	1.0000	1.0000	-
d153	*d148*	0.0958	0.2399	-	1.0000	1.0000	1.0000	-
d042	*d148*	0.3380	0.1879	**0.0000**	0.0242	0.0013	0.1928	0.0005
d006	*d148*	0.4791	0.7868	0.0039	0.0011	**0.0000**	0.0341	**0.0000**
d070	*d182*	0.2607	0.6053	0.6641	0.0743	0.3042	0.2284	0.0058
d027	*d182*	0.3414	0.3855	0.2565	0.4561	0.3897	0.1062	0.1196
d117	*d182*	0.5163	0.5444	0.1368	0.0018	0.0271	0.0589	0.6740
d078	*d182*	0.4985	0.5963	-	-	0.5200	0.6052	0.3604
d087	*d182*	**0.0002**	0.6272	0.0137	0.4452	0.4739	0.3706	0.7298
d088	*d182*	0.3946	0.2184	0.0048	0.1231	0.9692	0.6622	0.5083
d166	*d182*	0.1366	0.7696	1.0000	0.8717	0.5958	0.2216	0.1543
d050	*d182*	0.2165	0.5605	0.2733	-	0.6915	-	1.0000
d127	*d182*	0.5559	0.4540	0.0584	0.4957	0.7420	0.8040	0.0341
d105	*d182*	0.3355	0.4129	0.0029	0.6343	0.0767	0.0014	0.3938
d029	*d182*	0.6340	0.7668	0.2606	0.0056	**0.0000**	**0.0004**	0.0266
d174	*d182*	0.9436	0.6218	-	-	-	-	-
d015	*d182*	0.5761	0.9197	0.4894	-	-	-	-

Locus1	Locus2	Sol	Disp	3L2	Warn	Law	Law1	Law2
d186	*d182*	0.2857	0.6430	0.0052	0.0026	0.0263	0.6570	0.025
d111	*d182*	0.8568	0.3196	-	-	-	-	-
d153	*d182*	0.0933	0.7906	-	0.5776	-	-	-
d042	*d182*	0.9108	0.2133	0.3996	0.2700	0.3789	0.1074	0.6918
d006	*d182*	0.3636	0.1176	0.0557	0.0046	**0.0000**	0.0010	**0.0002**
d148	*d182*	0.8357	0.0697	0.0135	**0.0000**	0.0904	0.0278	0.1542
d070	*d016*	0.0063	0.5472	0.1128	0.4966	**0.0000**	0.1819	**0.0000**
d027	*d016*	0.1507	0.2863	0.0838	0.0009	0.0933	0.3862	0.0883
d117	*d016*	0.2445	0.5058	0.1164	**0.0001**	0.0459	0.0541	0.0062
d078	*d016*	0.0699	0.6671	-	-	0.2497	1.0000	0.1556
d087	*d016*	0.0080	0.5075	0.3428	0.1348	0.0596	0.7882	0.0012
d088	*d016*	0.0328	0.6606	0.0134	**0.0000**	0.0613	**0.0001**	0.8740
d166	*d016*	0.0428	0.2395	0.3024	0.9943	0.0502	0.9683	**0.0005**
d050	*d016*	0.3655	0.1748	0.0086	-	0.0176	-	0.0188
d127	*d016*	0.3363	0.5462	0.0131	0.3428	0.0026	0.0015	0.1010
d105	*d016*	0.0648	0.8894	0.0011	0.0051	0.0272	0.9061	0.0266
d029	*d016*	0.8784	0.7738	0.0037	**0.0000**	0.0898	0.4834	0.0271
d174	*d016*	0.6099	0.2867	-	-	-	-	-
d015	*d016*	0.6288	0.6633	1.0000	1.0000	-	-	-
d186	*d016*	0.4745	0.8034	0.0011	**0.0000**	**0.0000**	0.0673	**0.0000**
d111	*d016*	0.0521	0.5582	-	-	0.2100	0.3083	-
d153	*d016*	0.7462	0.5308	-	0.4299	0.2091	0.2935	-
d042	*d016*	0.0137	0.3475	0.0747	**0.0000**	0.2597	1.0000	**0.0001**
d006	*d016*	0.2593	0.8048	0.0116	0.00121	**0.0000**	0.0245	**0.0000**

Locus1	**Locus2**	**Sol**	**Disp**	**3L2**	**Warn**	**Law**	**Law1**	**Law2**
d148	*d016*	0.2152	0.2878	0.0229	**0.0004**	0.1748	0.7001	**0.0000**
d182	*d016*	0.0732	0.9056	0.0019	0.0009	0.0474	0.2184	0.0010
Total loci in LD		3	0	10	18	34	31	13

Table 2.7 Linkage disequilibrium between 21 microsatellite loci and mitochondrial profiles of *Daphnia pulex* and *Daphnia pulicaria*. Abbreviations for the different populations are given in table 2.1, with Law, all individuals from Lawrence Lake; Law2, Lawrence lake individuals with *Daphnia pulex* mitochondria DNA; clade AB/C, indicates LD between a given microsatellite locus and an association with any mitochondrial profile (either panarctic *D. pulex* clade and MI lake *D. pulicaria* clade, clade A and B, or western *D. pulicaria* clade C); clade A/B, indicates any LD between a microsatellite locus and either clade A or clade B (figure 2.2). Values are P-values and those in bold are significant after Bonferroni correction.

Locus	Mitochondrial profile	3L2	Warn	Law	Law2
d070	clade AB/C	0.5531	1.0000	0.30622	-
d027	clade AB/C	0.6370	0.8252	**0.0000**	-
d117	clade AB/C	1.0000	1.0000	0.0603	-
d078	clade AB/C	-	-	1.0000	-
d087	clade AB/C	1.0000	1.0000	0.0005	-
d088	clade AB/C	0.2533	1.0000	0.0025	-
d166	clade AB/C	1.0000	1.0000	0.1976	-
d050	clade AB/C	1.0000	-	0.0921	-
d127	clade AB/C	1.0000	1.0000	0.0012	-
d105	clade AB/C	0.5971	0.8204	0.0098	-
d029	clade AB/C	1.0000	1.0000	**0.0000**	-

Locus	Mitochondrial profile	3L2	Warn	Law	Law2
d174	clade AB/C	-	-	-	-
d015	clade AB/C	1.0000	1.0000	-	-
d186	clade AB/C	1.0000	0.9317	**0.0002**	-
d111	clade AB/C	-	-	1.0000	-
d153	clade AB/C	-	1.0000	1.0000	-
d042	clade AB/C	0.0985	0.7217	0.0006	-
d006	clade AB/C	1.0000	1.0000	**0.0000**	-
d148	clade AB/C	0.2346	0.2257	**0.0000**	-
d182	clade AB/C	0.4998	0.6611	0.0091	-
d016	clade AB/C	1.0000	0.7183	**0.0001**	-
d070	clade A/B	0.0713	0.5346	**0.0000**	**0.0000**
d027	clade A/B	0.3420	0.4835	**0.0000**	0.5240
d117	clade A/B	0.1845	0.0140	0.0152	0.0265
d078	clade A/B	-	-	0.3964	0.3572
d087	clade A/B	**0.0006**	0.3131	**0.0003**	0.1612
d088	clade A/B	0.1429	0.0766	0.0280	0.3824
d166	clade A/B	1.0000	0.6658	0.0025	**0.0002**
d050	clade A/B	0.7456	-	0.0941	1.0000
d127	clade A/B	0.4774	0.4375	**0.0000**	0.0032
d105	clade A/B	0.1468	0.6444	0.0006	0.0251
d029	clade A/B	0.0397	**0.0003**	**0.0000**	0.0054

Locus	**Mitochondrial profile**	**3L2**	**Warn**	**Law**	**Law2**
d174	clade A/B	-	-	-	-
d015	clade A/B	1.0000	1.0000	-	-
d186	clade A/B	0.1382	0.7318	**0.0000**	**0.0000**
d111	clade A/B	-	-	1.0000	-
d153	clade A/B	-	1.0000	1.0000	-
d042	clade A/B	0.0028	0.7871	**0.0001**	0.0007
d006	clade A/B	0.0994	0.4332	**0.0000**	0.1201
d148	clade A/B	0.0093	0.3071	**0.0000**	0.0823
d182	clade A/B	0.1351	0.6785	0.0085	0.0249
d016	clade A/B	0.0071	0.0844	**0.0000**	0.0230
Nd5	clade A/B	0.0289	0.0020	**0.0000**	-
Total loci in LD with clade AB/C		-	-	6	-
Total loci in LD with clade A/B		1	1	10	3

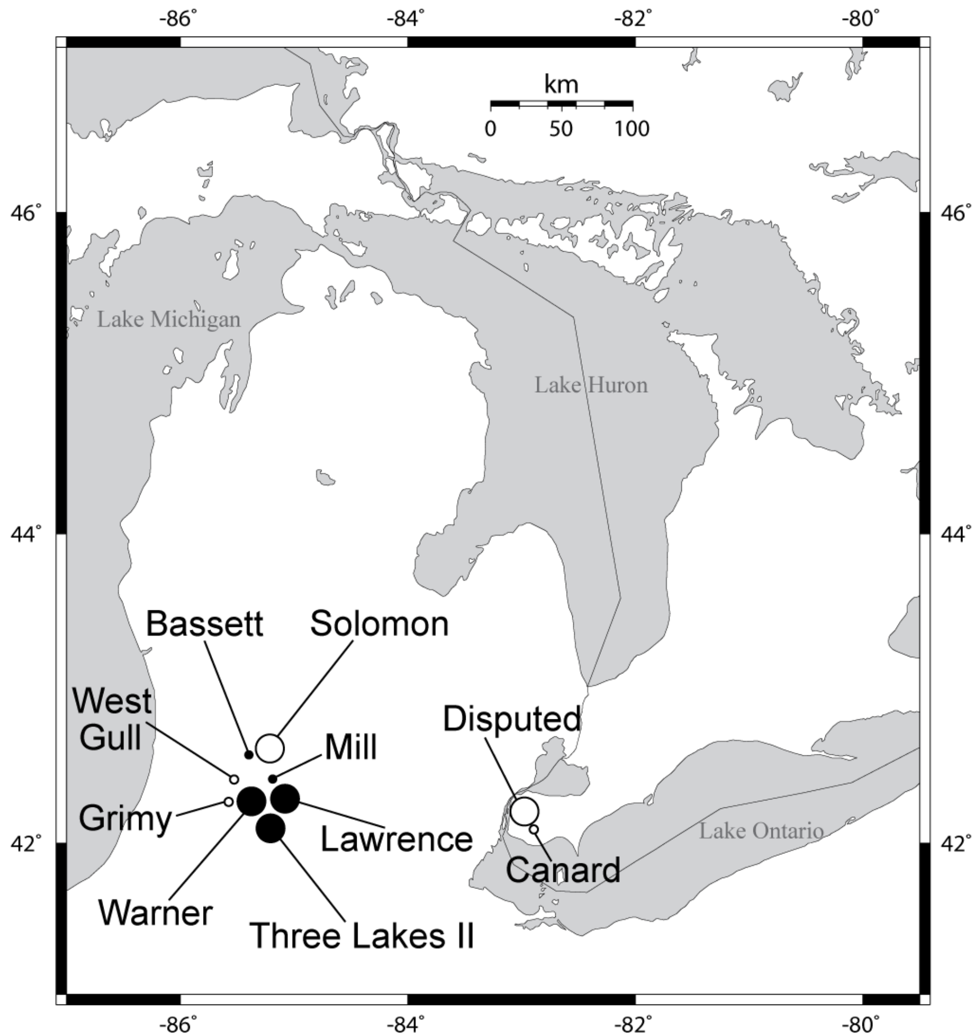

Figure 2.1 Distribution of collection sites in Michigan and Ontario. Lake habitats are denoted by black circles and pond habitats are denoted by white circles. The smaller circles represent habitats from which a low number of individuals were collected (3-14 individuals from each site) and only screened with the mitochondrial *NADH dehydrogenase 5* gene and *Lactate dehydrogenase* markers. Larger circles indicate populations that have been sampled intensely (>35 individuals from each site) and were screened with microsatellite, *LDH*, and mitochondrial markers.

32 Sol SS
41 Sol SS
35 Sol SS
44 Sol SS
45 Sol SS
6 3L2 FF
34 Sol SS
43 Sol SS
37 Sol SS
40 Sol SS
31 Sol SS
46 Sol SS
33 Sol SS
42 Sol SS
16 Disp SS
18 Disp SS
30 Sol SS
36 Sol SS
38 Sol SS
39 Sol SS
19 Disp SS
52 WG SS*
27 Law FF
25 Law FF
28 Law FF
24 Grm SS*
2 3L2/ Bas/ Mill/ Warn/ Disp SS/FF
15 Clr/ Lng/ Spm FF
5 3L2 FF
11 Can SF*
10 Can/ Ges SF*
23 Grm*SS
17 Disp SS
12 Disp/ Can/ WG* SS
20 Disp SS
14 Disp SS
3 Warn/ 3L2 FF
7 3L2 FF
47 Spm FF
13 Can*/Ges* SF
4 3L2 FF
29 Law FF
8 Law/ Bas FF
51 Warn FF
21 BG FF Eastern Daphnia pulicaria
22 BG FF
26 Law FF
48 Warn FF
9 Law/ Warn/ Bas/ Mill/ 3L2 FF
49 Warn FF
50 Warn FF
European Daphnia pulex
75/ 99
71/ 93
81/94
52/84
50/74
91/99
58/95
70/77
96/76
90/99
72/93
97/100
99/100
0.02
A Panarctic Daphnia pulex
B MI Lake Daphnia pulicaria
C Western Daphnia pulicaria
panarctic D. pulex
polar D. pulicaria
western D. pulicaria
eastern D. pulicaria
D. melanica
D. middendorfiana
European D. pulicaria
D. tenebrosa
European D. pulex

Figure 2.2 A Neighbor-joining (NJ) phylogeny of the mitochondrially encoded *NADH dehydrogenase 5* gene (*ND5*) of *Daphnia pulex* and *Daphnia pulicaria* from Illinois, Michigan, and Ontario. The label code of each haplotype identifies the haplotype number, its location (see table 2.1), followed by *lactate dehydrogenase A* profile (FF-homozygous fast, SS-homozygous slow, SF-heterozygous slow/fast) and habitat (lake denoted by black circles, ponds denoted by white circles), * indicates populations reproducing by obligate parthenogenesis, all other populations reproduce by cyclical parthenogenesis. Numbers before and after dashes, beside nodes, represent bootstrap support with 1000 replicates and posterior probabilities, respectively. The tree was rooted with European *D. pulex* (GenBank accession number DQ235231). Only values ≥ 50% are shown. The small panel on the left represents a NJ *ND5* phylogeny of the *D. pulex* complex (Colbourne *et al*. 1998), branches in bold correspond to the two major clades in the phylogenetic reconstruction of this study.

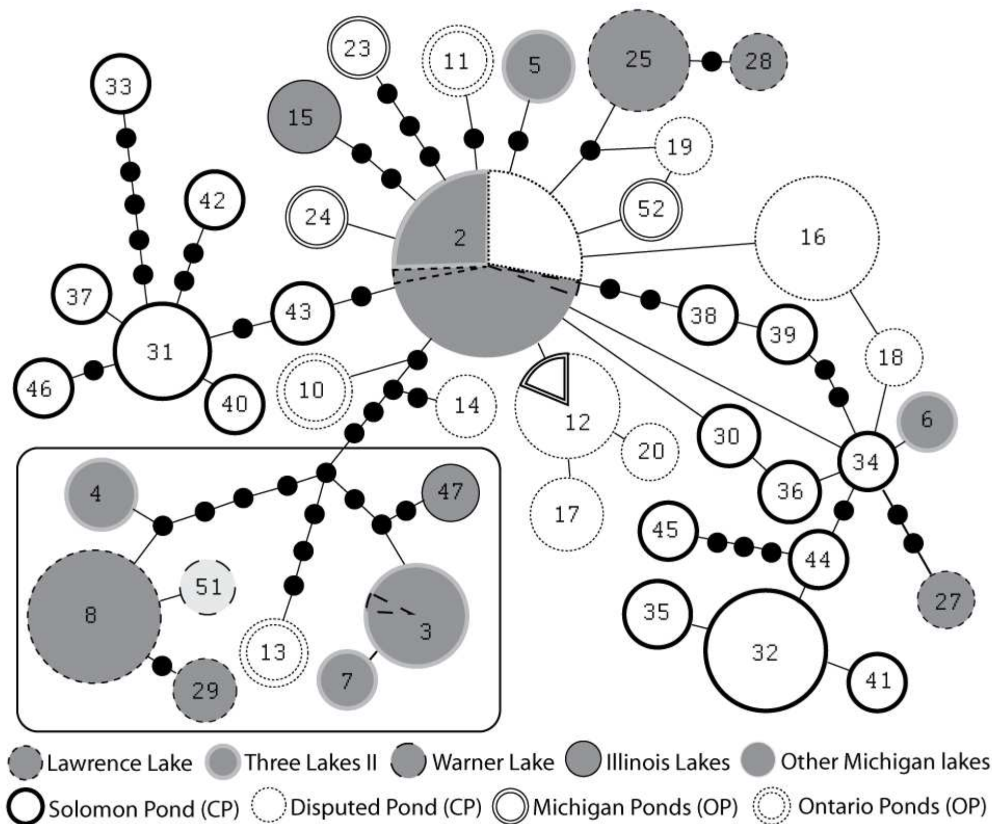

Figure 2.3. Unrooted statistical parsimony network of mitochondrial ND5 haplotypes of 363 *Daphnia pulex* and *Daphnia pulicaria* isolates from Illinois, Michigan, and Ontario. The network was estimated under the 95% statistical limits of parsimony using the algorithm of Templeton *et al.* (1992). Numbered circles represent the label code of each haplotype. The area of the haplotype circles is scaled to represent the relative frequency of that haplotype. The outline of each circle identifies the specific habitat(s) from which each haplotype was collected. Lake populations are shaded in grey, pond populations are white with CP, cyclical parthenogenesis, and OP, obligate parthenogenesis. Small black circles on the branches represent a single nucleotide difference between haplotypes and are hypothetical haplotypes. Only Haplotype 2 is found in both lakes and ponds. The box corresponds to clade B in figure 2.2.

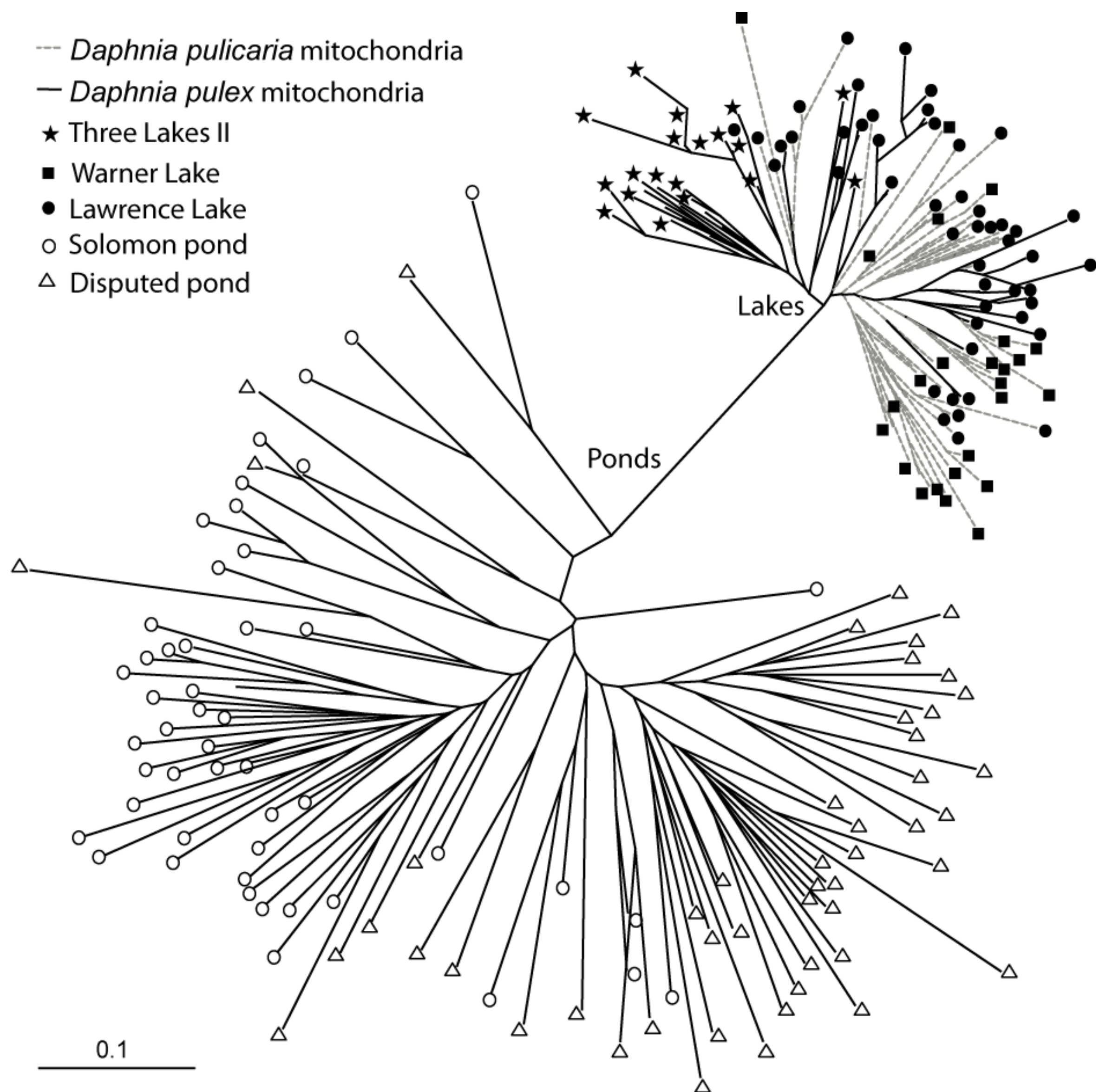

Figure 2.4 Unrooted neighbor-joining phylogram based on genetic distances at 21 microsatellite loci in 101 *Daphnia pulex* and 165 *Daphnia pulicaria* from Michigan and Ontario. The two distinct, monophyletic groups correspond to pond and lake habitats. Different symbols identify specific habitats. Solid lines identify lineages that share *D. pulex* mitochondria while dashed lines identify *D. pulicaria* mitochondria.

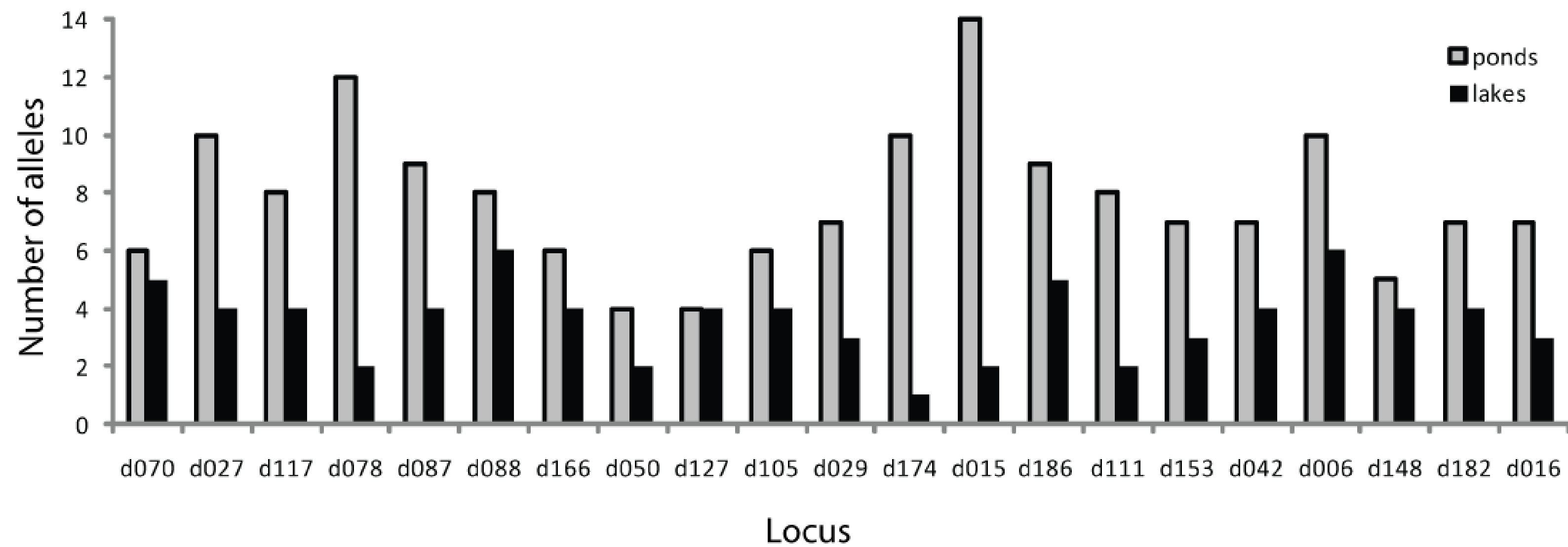

Figure 2.5 Number of alleles at 21 microsatellite loci in 101 *Daphnia pulex* collected from Solomon and Disputed ponds and 215 *Daphnia pulicaria* isolates sampled from Warner, Lawrence, and Three Lakes II.

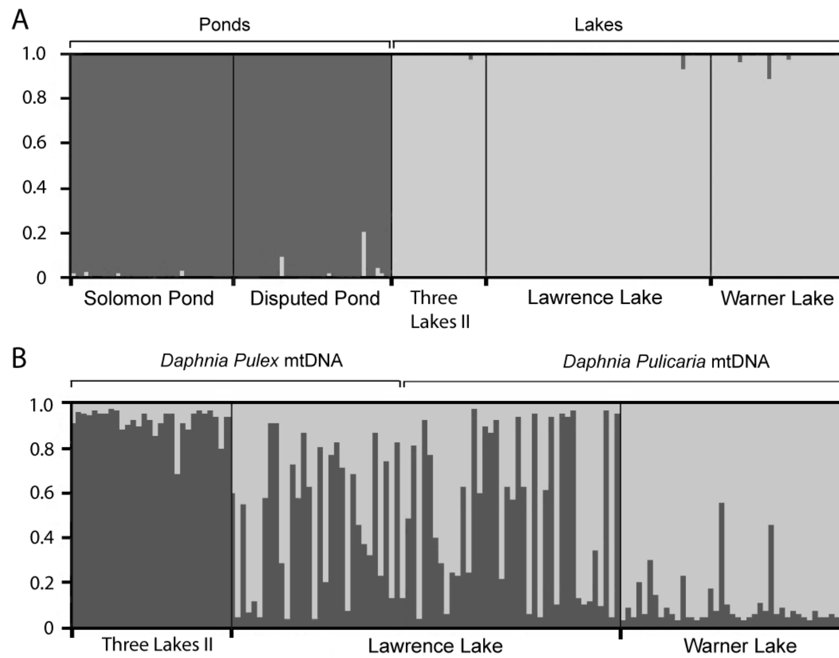

Figure 2.6. Results of a Bayesian STRUCTURE analysis of variation at 21 microsatellite loci. A) Genotypes of *Daphnia pulex* (101) and *Daphnia pulicaria* (165) from 5 habitats with best support for K = 2. B) Genotypes of *Daphnia pulicaria* (165) from 3 habitats with best support for K = 2. Each individual's multilocus genotype is represented by a thin vertical line, which is partitioned into K shaded segments that represent the individual's probability of belonging to each of the genetic clusters. Black lines separate different populations, which are labeled below the figure. Five STRUCTURE runs at each K-value produced nearly identical individual membership coefficients. The figures show the highest probability runs.

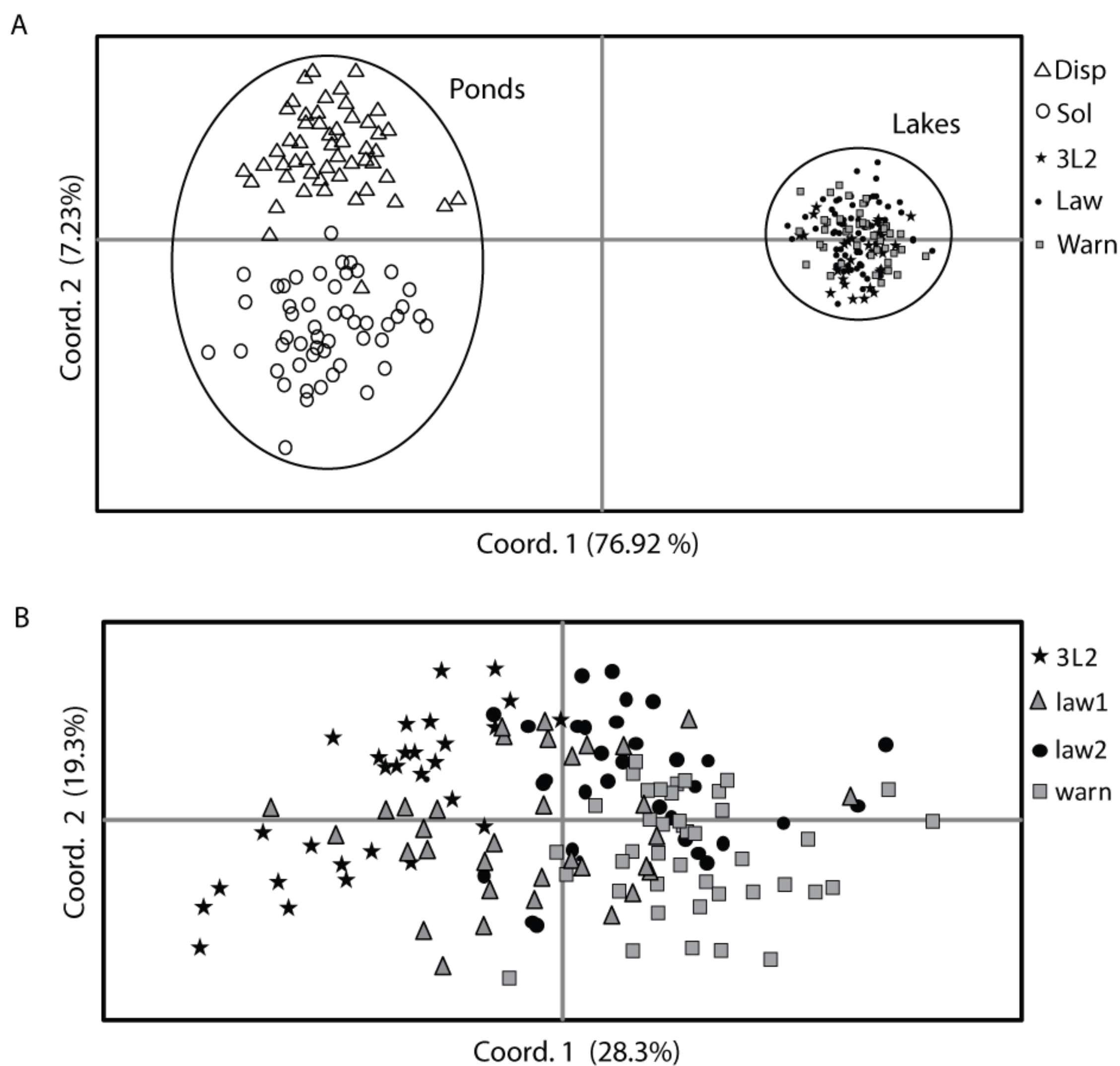

Figure 2.7. Genetic structure of *Daphnia* represented by a principle component analysis of variation at 21 microsatellite loci. A) 101 genotypes of *Daphnia pulex* and 165 genotypes of *Daphnia pulicaria* from 2 ponds and 3 lakes. B) 165 genotypes of *Daphnia pulicaria* from lakes, with Lawrence lake divided into Law1, *Daphnia pulicaria* mitochondrial profile and Law2, *Daphnia pulex* mitochondrial profile. Different symbols represent the different populations.

REFERENCES

Adamowicz S.J., R. Gregory, M.C. Marinone, and P.D.N. Hebert. 2002. New insights into the distribution of polyploidy *Daphnia*: the Holarctic revisited and Argentina explored. Molecular Ecology 11: 1209-1217.

Adamowicz S.J., A. Petrusek, J.K. Colbourne, P.D.N. Hebert, and J.D.S. Witt. 2009. The scale of divergence: a phylogenetic appraisal of intercontinental allopatric speciation in a passively dispersed freshwater zooplankton genus. Molecular Phylogenetic and Evolution 50:423-36

Allen M.R. 2009. Ecological and Evolutionary effects of dispersal on freshwater zooplankton. Ph.D. University of Illinois at Urbana-Champaign: USA

Allen M.R., R.A. Thum, and C.E. Cáceres. 2010. Does local adaptation to resources explain genetic differentiation among *Daphnia* populations? Molecular Ecology 19: 3076-3087.

Beerli P. 2008. MIGRATE version 3.0 [http://popgen.sc.fsu.edu/Migrate-n.html]. Computer program distributed by the author.

Cáceres C.E. and D.A. Soluk. 2002. Blowing in the wind: a field test of overland dispersal and colonization by aquatic invertebrates. Oecologia 131: 402-408.

Cáceres C.E. and A.J. Tessier. 2003. How long to rest: The ecology of optimal dormancy and environmental constraint. Ecology 84: 1189-1198.

Cáceres C.E. and A.J. Tessier. 2004. To sink or swim: Variable diapauses strategies among *Daphnia* species. Limnol. Oceanogr. 49: 1333-1340.

Clement M., D. Posada, K.A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657–1659.

Cohen G.M. and J.B. Shurin. 2003. Scale-dependence and mechanisms of dispersal in freshwater zooplankton. OIKOS 103: 603-617.

Colbourne J.K., T.J. Crease, L.J. Weider, P.D.N. Hebert, F. Dufresne, and A. Hobaek. 1998. Phylogenetics and evolution of the circumarctic species complex (Cladocera: *Daphnia pulex*). Biological Journal of the Linnean Society. 65: 347-365.

Colbourne J.K. and P.D.N. Hebert. 1996. The systematic of North American *Daphnia* (Crustacea: Anomopoda): a molecular phylogenetic approach. Philosophical transactions of the Royal Society of London B 351: 349-360.

Colbourne J.K., P.D.N. Hebert, and D.J. Taylor. 1997. Evolutionary origins of phenotypic diversity in *Daphnia*. In *Molecular Evolution and Adaptive Radiation*, T.J. Givnish and K.J. Sytsma, editors. Cambridge University Press, New York, pp. 163-188.

Crease T.J., R. Floyd, M.E. Cristescu, and D. Innes. 2010. *Lactate dehydrogenase* A and B variation in the *Daphnia pulex* species complex; evidence for a selective sweep in lake populations. In review for Molecular Ecology.

Crease T.J. and M. Lynch. 1991. Ribosomal DNA variation in *Daphnia pulex*. Molecular Biology and Evolution 8:620-640.

Crease T.J., S.K. Lee, S.L Yu, K. Spitze, N. Lehman, and M. Lynch. 1997. Allozyme and mtDNA variation in populations of the *Daphnia pulex* complex from both sides of the Rocky Mountains. Heredity 79:242-251.

Crease T.J., D.J, Stanton, P.D.N. Hebert. 1989. Polyphyletic origins of asexuality in *Daphnia pulex*. II Mitochondrial-DNA variation. Evolution. 43: 1016-1026.

Cristescu M.E., J.K. Colbourne, J. Radivojac, and M. Lynch. 2006. A microsatellite-based linkage map of the water flea *Daphnia pulex*: On the prospect of crustacean genomics. Genomics 88: 415-430.

Cristescu M.E., D.J. Innes, J.H. Stillman, and T.J. Crease. 2008. D- and L-*lactate dehydrogenases* during invertebrate evolution. BMC Evol. Biol. 8: 268.

De Meester L., A. Gómez, B. Okamura, and K. Schwenk. 2002. The monopolization hypothesis and the dispersal–gene flow paradox in aquatic organisms. Acta Oecologica 23: 121–135

Dodson S.I. 1989. Predator-induced reaction norms. Bioscience 39: 447-452.

Doyle J.J. and Doyle J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin19: 11-15.

Dudycha J.L. 2003. A multi-environment comparison of senescence between sister species of *Daphnia*. Oecologia 135: 555-563.

Dudycha J.L. 2004. Mortality dynamics of *Daphnia* in contrasting habitats and their role in ecological divergence. Freshwater biology 49: 505-514.

Dudycha J.L. and A.J. Tessier. 1999. Natural genetic variation of life span, reproduction, and juvenile growth in *Daphnia*. Evolution 53:1744-1756.

Dufresne F. and P.D.N. Hebert. 1994. Hybridization and origins of polyploidy. Proceeds of the Royal Society of London Series B 258: 141-146.

Dufresne F. and P.D.N. Hebert. 1997. Pleistocene glaciations and polyphyletic origins of polyploidy in an arctic cladoceran. Proceedings of the Royal Society of London Series B 264: 201-26.

Edmondson W.T. 1955. The seasonal life history of *Daphnia* in an arctic lake. Ecology 36: 439-455.

Evanno G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-20.

Excoffier, L. G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

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

Fryer G. 1991. A daphniid ephippium (Branchiopoda: Anomopoda) of Cretaceous age. Zoological Journal of the Linnean Society 102: 163-167.

Ghadouani A. and B. Pinel-Alloul. 2002. Phenotypic plasticity in *Daphnia pulicaria* as an adaptation to high biomass of colonial and filamentous cyanobacteria: experimental evidence. Journal of Plankton Research 24: 1047-1056.

Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www2.unil.ch/popgen/softwares/fstat.htm

Haney J.F. and D.J. Hall. 1975. Diel vertical migration and filter-feeding activities of *Daphnia*. Archiv fuer Hydrobiologie 75: 413-441.

Hebert P.D.N 1978. The population biology of *Daphnia* (Crustacea, Daphnidae). Biological Reviews 53: 387-426.

Hebert P.D.N. 1987. Genotypic characteristics of the Cladocera. Hydrobiologia 145: 183-193.

Hebert P.D.N. 1995. The *Daphnia* of North America: an illustrated fauna. CD-ROM. Distributed by the author. Department of Zoology, University of Guelph.

Hebert P.D.N. 1997. Cladocera as model systems in biology. Journal of Experimental Marine Biolology and Ecology 209: 310-311.

Hebert P.D.N. and M.J. Beaton. 1989. Methodologies for Allozyme Analysis using Cellulose Acetate Electrophoresis. Helena Laboratories, Beaumont, Texas.

Hebert P.D.N., M.J. Beaton, S.S Schwartz, and D.J. Stanton. 1989. Polyphyletic origins of asexuality in *Daphnia pulex*. I. Breeding-system variation and levels of clonal diversity. Evolution 43: 1004-1015.

Hebert P.D.N. and Crease T. 1983. Clonal diversity in populations of *Daphnia pulex* reproducing by obligate parthenogenesis. Heredity 51: 353-369.

Hebert P.D.N. and T.L. Finston. 1996. A taxonomic reevaluation of North American *Daphnia* (Crustacea: Cladocera). II New species in the *Daphnia pulex* group from the south-central United States and Mexico. Canadian Journal of Zoology 74: 632-653.

Hebert P.D.N. and T.L. Finston. 2001. Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group from the United States and Mexico. Heredity 87: 153-161.

Hebert P.D.N., S.S. Schwartz, R.D. Ward, and T.L. Finston. 1993. Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group. I. Breeding systems of Canadian populations. Heredity 70: 148–161.

Heier C.R. and J.L. Dudycha. 2009. Ecological speciation in a cyclic parthenogen: Sexual capability of experimental hybrids between *Daphnia pulex* and *Daphnia pulicaria*. Limnology 54: 492-502.

Hobaek A. and L.W. Weider. 1999. A circumpolar study of Arctic biodiversity: phylogeographic patterns in the *Daphnia pulex* complex. Ambio 28**:** 245–250.

Ronquist F. and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Innes D.J., S.S Schwartz, and P.D.N. Hebert. 1986. Genotypic diversity and variation in mode of reproduction among populations in the *Daphnia pulex* group. Heredity 57: 345-355

Laity J.H., B.M. Lee, and P.E. Wright. 2001. Zinc finger proteins: new insights into structural and functional diversity. Current Opinions in Structural Biology 11: 39-46.

Lampert W. 1994. Phenotypic plasticity of the filter screens in *Daphnia*: adaptation to a low-food environment. Limnology and Oceanography 39: 997-1006.

Langella O. 1999. POPULATIONS 1.2.30 (http://bioinformatics.org/~tryphon/populations/).

Leibold M.A. 1991. Trophic interactions and habitat segregation between competing *Daphnia* species. Oecologia 86: 510-520.

Leibold MA. and A.J. Tessier. 1991. Contrasting patterns of body size for *Daphnia* species that segregate by habitat. Oecologia 86: 342-348.

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

Lynch M. 1985. Speciation in the Cladocera. Vert. Internat. Verein Limnol. 22: 3116-3123.

Lynch M., M. Pfrender, K. Spitze, N. Lehman, J. Hicks, D. Allen, L. Latta, M. Ottene, F. Bogue, and J. Colbourne. 1999. The quantitative and molecular genetic architecture of a subdivided species. Evolution 53: 100-110.

Mergeay J., X. Aguilera, S. Declerck, A. Petrusek, T. Huyses, and L. De Meester. 2008. The genetic legacy of polyploidy Bolivian *Daphnia*: the tropical Andres as a

source for the North and South American D. pulicaria complex. Molecular Ecology 17: 1789-1800.

Morgan K.K., J. Hicks, K. Spitze, L. Latta, M.E. Pfrender, C.S. Weaver, M. Ottone, and M. Lynch. 2001. Patters of genetic architecture for life-history traits and molecular markers in a subdivided species. Evolution 55: 1753-1761.

Nei M. 1972. Genetic distance between populations. American Naturalist 106: 283–291.

Nishiguchi Y., N. Ito, and M.Okada. 2010. Structure and Function of *Lactate dehydrogenase* from Hagfish. Mar. Drugs 8: 594–607

Omilian A.R. and M. Lynch. 2009. Patterns of interspecific DNA variation in the *Daphnia* nuclear genome. Genetics 182: 325-336.

Van Oosterhout C., W.F. Hutchinson, D.P.M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Osenberg C.W., E.E. Werner, G.G. Mittelbach, and D.J. Hall. 1988. Groth patterns in Bluegill (Lrpomid mscrochirud) and Pumpkinseed (L. gibbosus) Sunfish: Environmental variation and the importance of ontogenetic niche shifts. Canadian Journal of Fisheries and Aquatic Sciences 45: 17-26.

Peakall R. and P.E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295.

Pfrender M.E., K. Spitze, and N. Lehman. 2000. Multi-locus genetic evidence for rapid ecologically based speciation in *Daphnia*. Molecular Ecology 9: 1717-1735.

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

Posada D. and K.A. Crandall. 2001. Interspecific gene genealogies: trees grafting into networks. Trends in ecology and Evolution 16: 37-45.

Powers D.A., T. Lauerman, D. Crawford, and L. DiMichele. 1991. Genetic mechanisms for adapting to a changing environment. Annual Review of Genetics 25: 629-659.

Powers D.A. and P.M. Schulte. 1989. Molecular basis of evolutionary adaptation in two latitudinally extreme populations of *Fundulus heteroclitus*. Proceeds of the National Academy of Science 86:9365–9369.

Powers D.A. and P.M. Schulte. 1998. Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: A multidisciplinary approach to address the million-year saga of a small fish. The Journal of Experimental Zoology 282: 71–94.

Pritchard J.K., M. Stephens, and P. J. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Raymond M. and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Heredity, 86:248-249
also available online http://genepop.curtin.edu.au/

Rozas J, J.C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18: 233-234.

Schulte P.M., H.C. Glémet, A.A. Fiebig, and D.A. Powers. 2000. Adaptive variation in *lactate dehydrogenase*-B gene expression: Role of a stress-responsive regulatory element. Proceeds of the National Academy of Sciences 97: 6597-6602.

Schluter D. 2009. Evidence for ecological speciation and its alternative. Science 323: 737-741.

Spaak P. 1997. Hybridization in the *Daphnia galeata* complex: Are hybrids locally produced? Hydrobiologia 360: 127-133.

Sunnucks P., P.J. De Barro, G. Lushai, N. Maclean, and D. Hales. 1997. Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages and host specialization. Molecular Ecology 6: 1059-1073.

Tamura K. and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10: 512–526.

Tamura K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Taylor D.J. and P.D.N. Hebert.1992. *Daphnia galeata mendotae* as a cryptic species complex with interspecific hybrids. Limnology and Oceanography 37: 658-665.

Templeton A.R., K.A. Crandall, and C.F. Sing. 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data.3. Cladogram estimation. Genetics 132: 619-633.

Tessier A.J. and M.A. Leibold. 1997. Habitat use and ecological specialization within lake *Daphnia* populations. Oecologia 109: 561-570.

Van Valen L. 1976. Ecological species, multispecies, and oaks. Taxon 25: 233-239.

Vergilino R., C. Belzile, and F. Dufresne. 2009. Genome size evolution and polyploidy in the *Daphnia pulex* complex (Cladocera: Daphniidea). Biological Journal of the Linnean Society 97: 68-79.

Weider L.J., A. Hobaek, P.D. N. Hebert, and T.J. Crease. 1999. Holarctic phylogeography of an asexual species complex- II. Allozmic variation and clonal structure in Arctic *Daphnia*. Molecular Ecology 8: 1-13.

Weir B.S. and C.C. Cockerham. 1984. Estimating F-statisics for the analysis of population structure. Evolution 38:1358-1370.

Wellborn G.A., D.K. Skelly, and E.E. Werner. 1996. Mechanisms creating community structure across a freshwater habitat gradient. Annual Review of Ecology, Evolution, and Systematics 27: 337-363.

Werner E.E. and D.J. Hall. 1988. Ontogenetic habitat shifts in Bluegill: The foraging rate-predation risk trade-off. Ecology 69: 1352-1366.

Werner E.E., D.J. Hall, D.R. Laughlin, D.J. Wagner, L.A. Wilsmann, and F.C Funk. 1977. Habitat partitioning in a freshwater fish community. Journal of the Fisheries Research Board of Canada 34 :360-370.

wFleaBase. *Daphnia* genes 2010. http://wfleabase.org/

Wright D. and J. Shapiro. 1990. Refuge availability: a key to understanding the summer disappearance of *Daphnia*. Freshwater Biology 24: 43-62.

CHAPTER III

CONCLUSIONS AND FUTURE DIRECTIONS

Identifying genomic regions under divergent selection

Identification of genes and genomic regions under divergent selection in natural populations has become one of the major goals in evolutionary genetics (Makinen *et al*. 2008), particularly in speciation. This study reveals that despite a prolonged history of hybridization and introgression between the ecological species, *Daphnia pulex* and *Daphnia pulicaria,* they are genetically diverged at the nuclear genome. In the face of gene flow, divergent selection may act on a few genomic regions, while other neutral regions homogenize, such as in the Z and E strain of the European corn borer moth (*Ostrinia nubilalis*) where only 1 out of the 5 genes examined, *Tpi*, was found to be divergent between the two moth strains (Dopman *et al*. 2005). A similar pattern can be observed between *Daphnia pulex* and *Daphnia pulicaria* at the *Ldh* locus, which is fixed for the fast allele in lake habitats (Hebert *et al*. 1989; Hebert *et al*. 1993; this study). A second potential genomic region identified in this study (locus *d174* on linkage group I) seems to show similar pattern of one allele being fixed in all lake populations examined. Genomic regions for which one or both species are exclusive groups may mark the footprint of recent selective sweeps, as is suggested by Crease *et al*. (2010) from their *Ldh* sequence analysis in lake and pond *Daphnia*. These selective sweep regions may be closely linked to “speciation genes” or genes involved in reproductive isolation of

ecologically diverging populations.

Despite *Daphnia pulex*'s genome being publically available (wFleaBase), our knowledge of potential regions under divergent selection is limited to the *Ldh* locus and perhaps locus *d174* identified in this study. A multi locus screen is the best approach for detecting selective sweeps when no prior information is available on possible candidate regions. This approach has been successfully used for detecting selective sweeps in several model species such as marine and freshwater three-spined stickleback, *Gasterosteus aculeatus* (Makinen *et al*. 2008); the common sunflower, *Helianthus annuus*, adapted to drought and salt tolerance (Kane and Rieseberg 2007); populations of the house mouse *Mus musculus* from different parts of Europe; and two *Drosophila* species that experienced an out of Africa habitat expansion; *Drosophila simulans* (Schöfl and Schlötterer 2004) and *Drosophila melangaster* (Kauer *et al*. 2003). These studies used a sufficiently high density of neutral markers to identify regions that have recently experienced a selective sweep.

The concept of hitchhiking can be used to pinpoint to specific regions of the genome that are under different selective pressures in different wild populations. Hitchhiking refers to the increase in frequency of neutral variations in a region of the genome that is closely linked to a locus under selection (Smith and Haigh 1974; Harr *et al*. 2002). Genomic regions under divergent selection are expected to show reduction of variation below neutral expectations and can indicate the presence of a "selective sweep" (Schlottere and Wiehe 1999), which may occur despite high levels of gene flow between diverging populations (Barton and Bengtsson 1986; Emelianov *et al*. 2004).

Daphnia is becoming an attractive new model system because various genomic tools have become available for this organism (Cristescu *et al*. 2006; wFleaBase). Microsatellites have been the marker of choice for various population genetics studies in the past (Chambers and Macavoy 2000; Ellegren 2000; Selkoe and Toonen 2006) due to their high mutation rate relative to the rest of the genome. They are particularly useful for inferring recent evolutionary events. For example, analysis of microsatellite variability offers a way to identify selective sweep regions and to ask whether they occur more often than expected by chance (Ihle *et al*. 2006). It is now possible to conduct a genome wide scan for signatures of selective sweeps in the *Daphnia* genome. This type of study can pinpoint regions of the genome experiencing divergent selection in lake verses pond habitats. The question of which genes or genomic regions facilitate the genetic adaptation of organisms to a new environment is central to ecological genetics and could contribute to a better understanding of how new species emerge.

REFERENCES

Barton N. and B.O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. Heredity 56: 357-376.

Chambers G.K. and E.S. MacAvoy. 2000. Microsatellites: consensus and controversy, Comp. Biochem. Physiol., B 126: 455–476.

Crease T.J., R. Floyd, M.E. Cristescu, and D. Innes. 2010. *Lactate dehydrogenase* A and B variation in the *Daphnia pulex* species complex; evidence for a selective sweep in lake populations. In review for Molecular Ecology.

Cristescu M.E.A, J.K. Colbourne, J. Radivojac, and M. Lynch. 2006. A microsatellite-based linkage map of the waterflea *Daphnia pulex*: On the prospect of crustacean genomics. Genomics 88: 415-430.

Dopman E.B., L. Pérez, S.M. Bogdanowicz, and R.G. Harrison. 2005. Consequences of reproductive barriers for genealogical discordance in the European corn borer. Proceeds of the National Academy of Science 102: 14706-14711.

Ellegren H. 2000. Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genetics 24*:*400–402.

Emelianov I., F. Marec, and J. Mallet. 2004. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proceeds of the Royal Society of London B 271: 97-105.

Harr B., M. Kauer, and C. Schlötterer. 2002. Hitchhiking mapping- a population-based fine mapping strategy for adaptive mutations in *Drosophila melangaster*. Proceeds of the National Academy of Science 99: 12949-12954.

Hebert P.D.N., M.J. Beaton, S.S Schwartz, and D.J. Stanton. 1989. Polyphyletic origins of asexuality in *Daphnia pulex*. I. Breeding-system variation and levels of clonal diversity. Evolution 43: 1004-1015.

Hebert P.D.N., S.S. Schwartz, R.D. Ward, and T.L. Finston. 1993. Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group. I. Breeding systems of Canadian populations. Heredity 70: 148–161.

Ihle S., I. Ravaoarimanana, M. Thomas, and D. Tautz. 2006. An analysis of signature of selective sweeps in natural populations of the house mouse. Molecular Biology and Evolution 23:790-797.

Kane N.C. and L.H. Rieseberg. 2007. Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annus*. Genetics 175: 1823-1834.

Kauer M., D. Dieringer, and C. Schlötterer. 2003. A microsatellite variability screen for positive selection associated with the “out of Africa” habitat expansion of *Drosophila melangaster*. Genetics 165: 1137-1148.

Makinen H.S., T. Shikano, J.M. Cano, and J. Merila. 2008. Hitchhiking mapping reveals a candidate genomic region for natural selection in Three-Spined Stickleback

chromosome VIII. Genetics 178: 453-465.

Schlötterer C. and T. Wiehe. 1999. Microsatellites, a neutral marker to infer selective sweeps, pp. 238-248 in Microsatellites- Evolution and Applications, edited by D.B. Goldstein and C. Schlötterer.

Schöfl G. and C. Schlötterer. 2004. Patterns of microsatellite variability among X chromosomes and autosomes indicate a high frequency of beneficial mutations in non-African *D. simulans*. Molecular Biology and Evolution, 21: 1384–1390.

Selkoe K.A. and R.J. Toonen. 2006. Microsatellites for Ecologists: A practical guide to using and evaluating microsatellite markers. Ecology Letters 9: 615-629.

Smith J.M. and J. Haigh. 1974. The hitch-hiking effect of a favorable gene. Genetics and Molecular Research 23:23-35.

wFleaBase. *Daphnia* genes 2010. http://wfleabase.org/

APPENDIX

Supplementary material

Table S.1. Pairwise F_{ST} estimates for all microsatellite loci generated using ARLEQUIN. Lawrence Lake is divided into two groups based on different mitochondrial profiles of individuals in this lake. Law1 individuals contain *Daphnia pulicaria* mtDNA, law2 individuals contain *Dapnia pulex* mtDNA, Law=(law1+ law2). All F_{ST} values are significant (P<0.05). Abbreviations for the different populations are given in table 1.

	Disp	**Sol**	**Warn**	**Law1**	**Law2**	**Law**
Sol	0.0797					
Warn	0.4751	0.4434				
Law1	0.4827	0.4570	0.0864			
Law2	0.4649	0.4310	0.1271	0.1094		
Law	0.4810	0.4505	0.0764			
3L2	0.4669	0.4384	0.1466	0.1415	0.0856	0.0868

VITA AUCTORIS

NAME:	Anna Constantin (Gindin)
PLACE OF BIRTH:	St. Petersburg, Russia
YEAR OF BIRTH:	1985
HOME TOWN:	Toronto, Ontario
EDUCATION:	Claude Watson Arts Program Earl Haig Secondary School, Toronto, Ontario Visual Arts Major 2000-2003
	York University, Toronto, Ontario BSc. Honors Biology 2003-2008
	Great Lakes Institute for Environmental Research University of Windsor, Windsor, Ontario 2008-2010