

Electronic Thesis and Dissertation Repository

9-5-2013 12:00 AM

Genetic and morphological variation of butterflies in relict habitats

Lindsay A. Crawford
The University of Western Ontario

Supervisor
Dr. Nusha Keyghobadi
The University of Western Ontario

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Lindsay A. Crawford 2013

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Ecology and Evolutionary Biology Commons](#)

Recommended Citation

Crawford, Lindsay A., "Genetic and morphological variation of butterflies in relict habitats" (2013).
Electronic Thesis and Dissertation Repository. 1600.
<https://ir.lib.uwo.ca/etd/1600>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Genetic and morphological variation of butterflies in relict habitats

(Thesis format: Integrated Article)

by

Lindsay Ann Crawford

Graduate Program in Biology with Environment & Sustainability

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Lindsay A. Crawford 2013

Abstract

Habitat fragmentation and loss are leading threats to global biodiversity and can alter patterns of dispersal, population dynamics, and genetics with implications for long-term species persistence. Most habitat fragmentation research has focused on recently fragmented species that historically occupied interconnected habitat patches. We know comparatively little about how naturally fragmented species may respond to habitat loss. For these species, local habitat patch quality may influence the dynamics and genetics of populations more than the structure of the surrounding landscape (e.g., degree of isolation of suitable habitat). I examined aspects of the ecology and evolution of populations inhabiting fragmented landscapes, using two butterfly species representing relict populations that are 1) recently fragmented by anthropogenic activities (Mormon metalmark, *Apodemia mormo*), and 2) naturally fragmented (bog copper, *Lycaena epixanthe*). I assessed patterns of genetic (amplified fragment length polymorphism, AFLP) and flight-related morphological variation, and their relationship to measures of surrounding landscape structure and local habitat quality.

Population genetic analysis of the anthropogenically fragmented Mormon metalmark revealed a high degree of spatial genetic structure, indicating limited gene flow, despite a small geographic scale (<20 km). Management of this endangered population should focus on increasing connectivity among the most isolated sub-populations and through urban areas. For the naturally fragmented bog copper, genetic diversity was explained by variables related to patch quality rather than landscape structure. Movement ability in the bog copper (inferred by flight morphology) appeared to depend on both local habitat conditions and the surrounding landscape. Also, using an AFLP-based genome scan approach, I identified signatures of selection in the bog copper associated with fine-scale landscape heterogeneity.

My work on the bog copper highlights the importance of considering the effects of local habitat conditions, in addition to habitat isolation, for conservation of fragmented populations.

Finally, I also reviewed the current literature (470 articles) to evaluate the quality of AFLP data used in ecological and evolutionary research. I discovered a pervasive lack of consistency and transparency in both the methods used to assess data reproducibility, and in the details of methodology presented. This work has identified an important publishing gap in molecular ecology research.

Keywords

Amplified fragment length polymorphism, Adaptive genetic variation, *Apodemia mormo*, Conservation genetics, Flight morphology, Fragmented habitats, Genetic diversity, Genotyping error rate, Habitat fragmentation, Habitat quality, Host plant, Isolation, Landscape heterogeneity, Lepidoptera, *Lycaena epixanthe*, Mobility, Outlier loci, Population genetics, Relict, Specialist.

Co-Authorship Statement

A version of Chapter 2 was published in *Molecular Ecology* with Daria Koscinski and Nusha Keyghobadi as co-authors. Dr. Koscinski and Dr. Keyghobadi contributed to study design, helped with data interpretation and provided editorial comments on the manuscript.

A version of Chapter 3 was published in *Conservation Genetics* with Sylvie Desjardins and Nusha Keyghobadi as co-authors. Dr. Desjardins provided logistical support in the field, and both co-authors contributed to study design, provided access to equipment, helped with data interpretation and provided editorial comments on the manuscript.

Chapters 4 - 6 are co-authored with Nusha Keyghobadi who contributed to study design, provided equipment, helped interpret data and provided editorial comments on the manuscripts.

Acknowledgments

First and foremost I would like to thank my supervisor Dr. Nusha Keyghobadi for her invaluable support and guidance throughout my degree. You have been an influential mentor to me both academically and personally, and I consider your leadership something to aspire to. To all the members of my lab, past and present: Liam Frappe, Heidi Keller, Chung-Yung (Christina) Kim, Dr. Daria Koscinski, Sheri Maxwell, Katie Millette, John O'Leary, Dr. Gordana Rasic, Noah Shapiro, Jenna Siu, Eliot Winkler and Kristina Zilic, thank you for your friendship and support. You have made my graduate experience truly special and something that I will always look back on with fond memories. In particular I would like to thank Sheri for teaching me the ins and outs of working with AFLPs and Katie for helping me to learn how to navigate the daunting world of R and model selection analyses. As well, I am deeply grateful to Daria whose collaboration, mentorship and friendship has been central to all my successes throughout my degree.

I would also like to thank the members of my advisory committee both past and present: Dr. Beth MacDougall-Shackleton, Dr. Brock Fenton, Dr. Cheryl Pearce and Dr. Adam Yates. Your valuable advice and expertise has greatly contributed to the development of my project.

I have been very lucky to have had the chance to work with many wonderful people throughout my degree and would like to thank all of the volunteers and research assistants who have helped with my lab and field work: Wesley Chick, Amy Dang, Jenna Donald, Orville Dyer, Jared Hobbs, Curtis Irvine, Heidi Keller, Chung-Yung (Christina) Kim, Katie Millette, George Pang, Kevin Schreiber, Dennis St-John, Lucy Reiss, Sue Seddon, Katelyn Weaver and Katie White. As well I am also extremely appreciative to the following people

who offered technical and logistical support, service and access to equipment as well as provided valuable advice towards my project: Dr. Brian Branfireun, Jason Dromboskie, Aren Fischer, Jennifer Hillyer, Angela Marinas, Dr. Jeremy McNeil, Justin Peters, Alan Noon, Dr. Brent Sinclair, Rick Stronks and the staff at Harkness Laboratory for Fisheries Research.

Financial support for myself and this research project was generously provided by The Centre for Environment and Sustainability, Canada Research Chairs Program, Government of Ontario, National Sciences and Engineering Research Council of Canada, The Xerces Society for Invertebrate Conservation, The University of Western University Department of Biology, Malcolm Ferguson Award in Life Science, and a partnership of the World Wildlife Fund and Environment Canada.

Finally, thank you to my family especially my parents Rob and Debbie, my sister Katie, my brother Scott and my parents-in-law Ian and Dianna for their constant love and encouragement. As well, I am forever indebted to my parents who fostered my curiosity about the natural world and who have always encouraged and provided support for me to pursue an education and career that challenges and excites me. To my husband Ross, throughout the trials and tribulations of my project, you have provided me with unwavering support and motivation to face my challenges. You have helped me to find pride in my accomplishments and confidence in times of self-doubt. This thesis would not have been possible without you and I thank you for your love and support.

Table of Contents

Abstract.....	ii
Co-Authorship Statement.....	iv
Acknowledgments.....	v
List of Tables	xii
List of Figures	xvi
List of Appendices	xviii
List of Symbols and Abbreviations.....	xix
List of Software Packages.....	xxi
Chapter 1	1
1 Introduction and literature review	1
1.1 Habitat fragmentation	1
1.2 Relict species and populations	2
1.3 Habitat fragmentation versus fragmented habitats	3
1.4 Role of habitat quality in fragmented landscapes	5
1.5 Evaluating the evolutionary potential of fragmented populations	6
1.6 The role of movement in fragmented landscapes	7
1.7 Butterflies: important flagship species and a model system for the study of habitat fragmentation	9
1.8 Population genetic studies of butterflies using AFLPs and non-lethal tissue sampling	10
1.9 The quality of AFLP data in ecological and evolutionary research.....	13
1.10Dissertation structure	14
1.11Literature Cited	18
Chapter 2.....	35

2	A call for more transparent reporting of error rate: The quality of AFLP data in ecological and evolutionary research	35
2.1	Introduction.....	35
2.2	Methods.....	38
2.2.1	Literature search.....	38
2.2.2	Data collection	40
2.3	Results.....	42
2.3.1	Overall trends.....	42
2.3.2	Comparisons among ‘reproducible’ studies.....	48
2.4	Discussion.....	49
2.4.1	Genotyping error reporting in AFLP studies	49
2.4.2	Conclusions and recommendations.....	55
2.5	Literature Cited	59
3	Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (<i>Apodemia mormo</i>) revealed using AFLPs	63
3.1	Introduction.....	63
3.2	Methods.....	67
3.2.1	Study species.....	67
3.2.2	Sample collection.....	69
3.2.3	DNA extraction and AFLP analysis	69
3.2.4	Data analysis	73
3.3	Results.....	76
3.3.1	AFLP analysis and phenotype scoring.....	76
3.3.2	Genetic diversity	77
3.3.3	Population genetic structure.....	77
3.4	Discussion.....	82
3.4.1	AFLP analysis and phenotype scoring.....	82

3.4.2	Population genetic structure of the Mormon metalmark	83
3.4.3	Genetic variation.....	87
3.4.4	Future management and recommendations	87
3.5	Literature cited.....	89
Chapter 4.....		98
4	Local patch characteristics determine patterns of genetic diversity in a glacial relict, peatland specialist butterfly.....	98
4.1	Introduction.....	98
4.2	Methods.....	101
4.2.1	Study species.....	101
4.2.2	Study area.....	102
4.2.3	Sample collection.....	104
4.2.4	DNA extraction and AFLP marker selection.....	107
4.2.5	AFLP Analysis.....	108
4.2.6	Landscape evaluation.....	110
4.2.7	Data analysis	113
4.3	Results.....	116
4.3.1	AFLP analysis and phenotype scoring.....	116
4.3.2	Population structure	119
4.3.3	Genetic diversity	120
4.3.4	Effects of landscape	120
4.4	Discussion.....	131
4.4.1	AFLP analysis and phenotype scoring.....	131
4.4.2	Population genetic structure.....	132
4.4.3	Genetic variation.....	133
4.4.4	Landscape effects on genetic diversity	134

4.4.5	Management and Conservation Implications.....	139
4.5	Literature Cited	141
Chapter 5	154
5	Flight morphology corresponds to both broad- and fine-scale landscape structure in a highly specialized glacial relict butterfly (<i>Lycaena epixanthe</i>)	154
5.1	Introduction.....	154
5.2	Methods.....	159
5.2.1	Study species.....	159
5.2.2	Study area.....	160
5.2.3	Specimen collection and preparation	161
5.2.4	Measurement of morphological characters	162
5.2.5	Landscape evaluation.....	165
5.2.6	Statistical analyses	166
5.3	Results.....	169
5.4	Discussion.....	175
5.5	Literature Cited	181
Chapter 6	190
6	Molecular signatures of selection associated with fine-scale landscape heterogeneity in a relict butterfly, <i>Lycaena epixanthe</i>	190
6.1	Introduction.....	190
6.2	Methods.....	195
6.2.1	Study sites and data collection.....	195
6.2.2	Detecting potentially adaptive loci	200
6.2.3	Identifying landscape variables associated with putative outlier loci.....	201
6.3	Results.....	203
6.3.1	AFLP analysis and phenotype scoring.....	203
6.3.2	Outlier locus detection	203

6.3.3	Associations between landscape variables and putative outlier loci	204
6.4	Discussion	209
6.4.1	Detecting potentially adaptive loci using an F_{ST} -based outlier approach	210
6.4.2	Identifying landscape variables affecting potentially adaptive loci.....	211
6.4.3	Conclusion and future perspectives	213
6.5	Literature Cited	215
Chapter 7	225
7	General Discussion.....	225
7.1	AFLPs: an alternative marker system for assessments of neutral and adaptive genetic variation in butterflies	226
7.2	Genetics of relict populations in fragmented landscapes.....	228
7.2.1	Relict populations in anthropogenically fragmented habitats.....	228
7.2.2	Relict populations in naturally fragmented habitats –the importance of local habitat patch characteristics	231
7.3	Mobility in naturally fragmented relict populations	234
7.4	Conclusion	235
7.5	Literature Cited	237
Appendix A:	Laboratory protocol used to optimize AFLPs for <i>Lycaena epixanthe</i>	243
Appendix B:	Landscape data reported in Chapters 3-5	250
Appendix C:	Chapter 3 Supplementary Material.....	251
Appendix D:	Chapter 4 Supplementary Material	256
Appendix E:	Permission to reproduce published material.....	263
Curriculum Vitae	263

List of Tables

Table 2.1 Replicate sample size, as a percentage of the total number of samples in the dataset, for studies which assessed reproducibility of AFLP datasets.....	50
Table 2.2 The stage at which replicate samples were generated for studies which assessed reproducibility.....	51
Table 3.1 Summary of representative studies using AFLPs to measure genetic diversity in natural populations of Lepidoptera.	68
Table 3.2 Summary of the AFLP phenotype scoring results for all selective primer combinations.....	79
Table 3.3 Within sub-population and global genetic diversity for the British Columbia population of <i>Apodemia mormo</i>	79
Table 4.1 Collection record of the 13 studied populations of <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Ontario, Canada.....	106
Table 4.2 Summary of the AFLP phenotype scoring results for <i>Lycaena epixanthe</i> for all selective primer combinations following a < 4% mismatch error rate criterion.....	118
Table 4.3 Pairwise F_{ST} values for the surveyed populations and sub-populations (N = 15) of <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Ontario, Canada.	122
Table 4.4 Mean population genetic diversity for <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Canada..	126
Table 4.5 Effect of landscape and patch predictors on estimates of genetic diversity in <i>Lycaena epixanthe</i> for a) all 13 surveyed peatland sites, and b) for all sites except Mizzy Lake (ML).....	127
Table 5.1 Collection records for adult male and female butterflies of <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Ontario, Canada. UTM coordinates (17N) represent the centroid of butterfly capture.....	163

Table 5.2 Absolute mean morphology measurements (\pm SE) of adult male and female <i>Lycaena epixanthe</i> collected from eight peatlands in Algonquin Provincial Park, Ontario, Canada.....	171
Table 5.3 Summary of model selection results for morphological traits in male and female <i>Lycaena epixanthe</i>	172
Table 6.1 Collection record of the 15 studied populations of <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Ontario, Canada.....	198
Table 6.2 Details of the five selective primer combinations used and their contribution to the final AFLP dataset.	205
Table 6.3 AFLP loci identified as putatively under divergent selection for <i>Lycaena epixanthe</i>	206
Table 6.4 Summary of model selection results for the five candidate outlier loci identified to be under divergent selection by DFDIST and BayeScan.....	207
Table 6.5 The relative importance of landscape variables acting as potential drivers of genetic variation in the five candidate outlier loci for <i>Lycaena epixanthe</i>	208
Table A.1 Oligonucleotides used for the AFLP analysis of <i>Lycaena epixanthe</i>	245
Table A.2 Protocol for annelaing a) <i>EcoRI</i> and b) <i>MseI</i> adaptors.....	208
Table A.3 Restriction-ligation (R-L) protocol outlining for a) Enzyme Master Mix, b) R-L Master Mix, and c) R-L Reaction	208
Table A.4 PCR protocol for Pre-Selective AFLP amplification.....	208
Table A.5 PCR protocol for Selective AFLP amplification	208
Table B.1 Summary of the landscape variables collected for the 15 Algonquin Provincial Park populations of <i>Lycaena epixanthe</i>	208

Table C.1 Summary of the AFLP phenotype scoring results for all selective primer combinations	208
Table C.2 Within sub-population and global genetic diversity for the British Columbia population of <i>Apodemia mormo</i>	208
Table C.3 Genetic structure summary for the British Columbian population of <i>Apodemia mormo</i>	208
Table C.4 Robustness of the genetic barriers identified by BARRIER based on 100 permuted genetic distance matrices (F_{ST}).	208
Table C.5 Summary of the principal coordiante analysis (PCoA) of all studied sub-populations based on a genetic distance matrix (F_{ST}).....	208
Table C.6 Summary of the mantel tests results examining the relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (ln transformed) among all studied sub-populations	208
Table D.1 Capture record for individuals of <i>Lycaena epixanthe</i> surveyed in 13 peatlands in Algonquin Provincial Park.....	208
Table D.2 Pairwise F_{ST} values for 13 surveyed populations of <i>Lycaena epixanthe</i> in Algonquin Provincial Park, Ontario, Canada	208
Table D.3 Summary of the optimal AFLP phenotype scoring parameters and associated mismatch error rates and number of retained loci for all selective primer combinations.....	208
Table D.4 Summary of the overall genetic structure and intrapopulation genetic diversity results for <i>Lycaena epixanthe</i> in Algonquin Provincial Park	208
Table D.5 Summary of analysis of molecular variance (AMOVA) results for the Algonquin Provincial Park populations of <i>Lycaena epixanthe</i>	208
Table D.6 Summary of the principal coordinate analysis (PCoA) for all studied populations of <i>Lycaena epixanthe</i>	208

Table D.7 Summary of the mantel test results examining the relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (log transformed) among all studied populations of *Lycaena epixanthe* 208

List of Figures

Figure 2.1 AFLP studies surveyed, categorized by reporting of dataset reproducibility.	44
Figure 2.2 Within each taxonomic group, the proportion of surveyed AFLP studies where reproducibility was assessed and genotyping error rate was reported (<i>GE</i>), reproducibility was assessed but genotyping error rate was not reported (<i>NGE</i>) and where reproducibility was not reported at all (<i>NR</i>)..	45
Figure 2.3 Average number of polymorphic AFLP loci retained for final analysis in studies, grouped by reporting of dataset reproducibility.....	46
Figure 3.1 Locations of the 14 studied sub-populations of <i>Apodemia mormo</i> in the Similkameen River Valley, British Columbia, Canada.	70
Figure 3.2 Principal coordinate analysis (PCoA) of all 14 sub-populations of <i>Apodemia mormo</i>	80
Figure 3.3 Relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (ln transformed) among all studied sub-populations of <i>Apodemia mormo</i>	81
Figure 4.1 Map of the study system and sampling locations for <i>Lycaena epixanthe</i> outlining the location of Algonquin Provincial Park in Ontario, Canada (a), the study region within Algonquin Provincial Park (b) and the locations of the 13 studied populations (b, c).	105
Figure 4.2 Principal coordinate analysis of all surveyed <i>Lycaena epixanthe</i> populations in Algonquin Provincial Park based on a covariance matrix of pairwise F_{ST} values.....	123
Figure 4.3 Relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (\log_{10} transformed) among a) all 15 studied populations and sub-populations, b) western populations (BUG, DL, MLa, MLb, WH and WR) and c) eastern populations (BAB, EOS, MINa, MINb, OPL, SB and ZEN).	125

Figure 4.4 Relationships between peatland size (ha) and a) mean cranberry density, b) cranberry abundance, and c) the coefficient of variation in cranberry density, for peatlands in Algonquin Provincial Park, Canada..... 129

Figure 4.5 Relationship between catch per unit effort (number of individuals / hour) of *Lycaena epixanthe* and a) mean cranberry density and b) peatland size. 130

Figure 5.1 Map of the study system and sampling locations of *Lycaena epixanthe*: (a) location of Algonquin Provincial Park in Ontario, Canada; (b, c) location of the eight studied populations..... 164

Figure 6.1 Sampling locations of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada..... 199

List of Appendices

Appendix A: Laboratory protocol used to optimize AFLPs for <i>Lycaena epixanthe</i>	243
Appendix B: Landscape data used in Chapters 4-6	250
Appendix C: Supplementary material for Chapter 3	251
Appendix E: Supplementary material for Chapter 4.....	256
Appendix F: Permission to reproduce published material	263

List of Symbols and Abbreviations

ΔAIC_c	Difference of Akaike information criterion value between model x and the model with the lowest AIC value
β_j	Estimated parameter coefficient of parameter j
$\hat{\beta}_j$	Weighted average parameter estimate of parameter j
w_i	Akaike model weight
$w_+(i)$	relative variable importance
AIC	Akaike information criterion
AIC _c	Second order Akaike information criterion
AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
AFLP	Amplified fragment length polymorphism
Area	Total bog area (ha)
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
bp	Base pair
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
FDR	False discovery rate
For	Proportion of forest habitat
F _{ST}	Fixation index among sites (the degree of between- relative to within-site genetic variation)
GE	Genotyping error rate reported
GIS	Geographical information system
GPS	Geographical positioning system
H _e	Expected heterozygosity
IBD	Isolation-by-distance
ISI	Institute for Scientific Information
ISSR	Inter simple sequence repeat
logLik	Log likelihood
MgCl ₂	Magnesium chloride
NA	Information not available
NaCl	Sodium chloride
NGE	Reproducibility assessed but genotyping error rate not reported
NGS	Next-generation sequencing
n _f	Number of fixed loci
n _p	Number of private loci
NR	No assessment of reproducibility reported
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
pH	Power of hydrogen
PPL	Proportion of polymorphic loci
Prox	Mean proximity of potentially suitable wetland habitat
RAPD	Random amplified polymorphic DNA

rfu	Relative florescence units
RJ-MCMC	Reversible jump Markov Chain Monte Carlo
R-L	Restriction-ligation
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat (microsatellite)
TE	Tris-EDTA buffer
Tris	Tris(hydroxymethyl)aminomethane
Wat	Proportion of open water habitat
BAB	Bab Lake
BUG	Buggy
COS	Costello Creek
DL	Dizzy Lake
DT	D. Thompson Lake
EOS	Eos Lake
KB	Kearney
MIN	Minor Lake
ML	Mizzy Lake
OPL	Opeongo Lake
SB	Spruce Bog
SUN	Sunday Creek
WH	Wolf Howl Pond
WR	West Rose Lake
ZEN	Zenobia Lake

List of Software Packages

AFLPDAT	http://www.nhm.uio/no/english/research/ncb/aflpdat/
AFLPScore v.1.3b	http://www.shef.ac.uk/molecol/software/aflpscore
AFLP-SURV v.1.0	http://www.ulb.ac.be/sciences/lagev/aflp-surv.html
AICcmodavg v.1.30	http://CRAN.R-project.org/package=AICcmodavg
ArcGIS v.10.0	ESRI (Redlands, CA)
Arlequin v.3.1	http://cmpg.unibe.ch/software/arlequin3/
BARRIER v2.2	http://www.mnhn.fr/mnhn/ecoanthropologie/software/barrier.html
BayeScan v.2.1	http://cmpg.unibe.ch/software/bayescan/
EndNote x4	Thompson Reuters (New York)
FAFLPcalc	http://www.york.ac.uk/res/dasmahapatra/publications.html
FAMD v.1.25	http://www.famd.me.uk/famd.html
FRAGSTATS v.4.0	http://www.umass.edu/landeco/research/fragstats/fragstats.html
GenAIEx v.6.0	http://www.anu.edu.au/BoZo/GenAIEx/
GENEMAPPER v.4.0	Applied Biosystems (Forest City, CA)
IBM SPSS v.20	IBM Corp (Armonk, New York)
ImageJ v.1.5m	http://image.nih.gov/ij/
JMP v.8	SAS Institute Inc. (Cary, NC)
lme4 v.0.999999-2	http://lme4.r-forge.r-project.org/
Mcheza	http://popgen/net/soft/mcheza/
MuMIn v.1.9.5	http://r-forge.r-project.org/projects/mumin/
R v.3.0.0	R Development Core Team, http://www.r-project.org/
SPSS v. 16.0	SPSS Inc. (Chicago)

Chapter 1

1 Introduction and literature review

1.1 Habitat fragmentation

Habitat fragmentation is the breaking up of a previously continuous expanse of habitat into a number of smaller sized patches, isolated from one another by an unfavourable landscape matrix (Noss et al. 2006; Habel and Zachos 2012). Habitat fragmentation is often accompanied by habitat loss, and is considered a leading threat to biodiversity worldwide (Caughley 1994; Debinski and Holt 2000; Sala et al. 2000; Foley et al. 2005), with negative impacts that have been documented across a wide range of taxonomic groups including plants (Hobbs and Yates 2003), invertebrates (Didham et al. 1996), amphibians (Stuart et al. 2004), birds and mammals (Andr en 1994). Predicting whether populations in highly fragmented landscapes will be able to survive and evolve in small, isolated habitat patches has thus become a major research theme for conservation biologists (Fazey et al. 2005). Changes in landscape structure and composition associated with habitat fragmentation result in the alteration of many biological processes, including gene flow and genetic drift, which influence levels of genetic diversity within and among local populations (Keyghobadi 2007). Alterations in patterns of population genetic variation may, in turn, affect individual fitness and evolutionary potential, with implications for both short- and long-term population viability (Frankham et al. 2002). Consequently, conservation biologists have increasingly become concerned with studying the genetic impacts of habitat fragmentation on populations (Pertoldi et al. 2007).

While habitats can be fragmented by anthropogenic activities such as agriculture and urbanization, they can also be naturally fragmented due to abiotic (e.g., geology, climate) and biotic (e.g., presence of hosts, food) factors (Hampe and Jump 2011; Habel and Zachos 2012). The process of natural habitat fragmentation may take place slowly over longer time scales (e.g., millennia; Lomolino et al. 2010), in contrast to anthropogenic habitat fragmentation which typically occurs within a relatively short time period (i.e., centuries, decades or even years; Lindenmayer and Fischer 2006).

1.2 Relict species and populations

Species or populations which occur as small and geographically isolated remnants of a formerly more widespread distribution are known as ‘relicts’ (Cassel-Lundhagen 2010; Lomolino et al. 2010) and their formation can be attributed to either historical or contemporary habitat fragmentation. For example, many arctic-alpine species which are currently restricted to disjunct areas on mountain tops and/or more Northern latitudes (e.g., mountain avens flowering plant, *Dryas octopetala*, Skrede et al. 2006; violet copper butterfly, *Lycaena helle*, Habel et al. 2011) are relicts species whose once widespread distributions have been gradually restricted by warming temperatures and the accompanying loss of suitable habitat following the last glacial period (Habel et al. 2010). On the other hand, species which were until relatively recently quite prevalent and widespread may have become spatially restricted to remnant relict populations as result of human-induced habitat loss and fragmentation (e.g., North American brown bear, *Ursus arctos*, Paetkau et al. 1998; regal fritillary butterfly, *Speyeria idalia*, Keyghobadi et al. 2012). In some cases, the small, isolated nature of naturally formed relict habitat

patches may be further exacerbated by contemporary fragmentation processes (e.g., inland hypersaline environments in Western Mediterranean, Gómez et al. 2005). Due to their limited distributions and increasing rarity, many relict species and populations are now of high conservation concern (Habel et al. 2010). As the overall rate and manner in which habitats become fragmented (i.e., anthropogenic vs. natural) can cause divergent evolutionary and genetic trajectories for populations (MacDougall-Shackleton et al. 2011; Habel and Zachos 2012), it is vital that management strategies for relict species and populations consider the history of the landscape and the processes by which population isolation has arisen.

1.3 Habitat fragmentation versus fragmented habitats

Species which occur in naturally interconnected habitats are predicted to be negatively affected by sudden human-induced habitat fragmentation (Habel and Zachos 2012). In theory, smaller and more isolated habitat patches will support small populations which experience reduced gene flow and increased genetic drift as a result of decreased dispersal and small local effective population sizes, respectively (Keyghobadi 2007). Over time, genetic diversity within populations will erode as a result of increased genetic drift and reduced gene flow, at the same time increasing genetic differentiation among populations (Templeton et al. 1990; Frankham et al. 2010). Additionally, intrapopulation genetic diversity may be further lost and genetic differentiation increased through bottlenecks and local extinctions which often accompany population fragmentation (Gilpin 1991; Andersen et al. 2004; Keller et al. 2004; Broquet et al. 2010). The disruptions to the drift-gene flow equilibrium which occur in recently fragmented

populations are also predicted to lead to an increase in the occurrence of inbreeding and inbreeding depression, as well as the random fixation of deleterious mutations and a loss of adaptive potential (Lande 1998; Keller and Waller 2002; Frankham et al. 2010). Changes in these genetic processes will contribute to an overall decline in population viability as well as the ability to respond to future environmental change, and thus increase the probability of extinction (Reed and Frankham 2003; Frankham 2005a).

In contrast, organisms which occupy naturally fragmented habitats are typically specialist species adapted to living within a habitat-matrix mosaic (Habel and Zachos 2012). Because of their ecological specialisation, populations are geographically isolated in discrete habitat patches and experience limited gene flow. As a result, local populations are typically small in size and exhibit lower genetic diversity, as well as increased interpopulation genetic differentiation (Habel and Schmitt 2012). Interestingly, however, these specialist species appear to have the capacity to persist over long periods of time despite the genetic isolation of populations (Habel and Schmitt 2012). It is theorized that deleterious alleles may have been purged from these populations over many generations (Frankham et al. 2001; Habel and Zachos 2012), and that their current genetic make-up consists of a small number of alleles which are highly adapted to local habitat conditions (Watt et al. 2003; Karl et al. 2008). Thus, low genetic diversity in these populations is actually associated with the ability to resist the negative consequences of genetic bottlenecks, inbreeding or fluctuations in local population size (Crnokrak and Barrett 2002; Reed 2010) and means that populations are not dependent on gene flow for genetic refreshment (Habel and Schmitt 2012). Indeed, several studies have now documented long-term persistence of genetically depauperate populations living in

geographic isolation (e.g., black mangrove tree, *Aegiceras corniculatum*; Ge and Sun 1999; San Nicolas Island fox, *Urocyon littoralis dickeyi*, Aguilar et al. 2004; Chillingham cattle, *Bos taurus*, Visscher et al. 2001; Red Apollo butterfly, *Parnassius apollo*, Habel et al. 2012). However, while low genetic variation may not pose an immediate concern for populations living under isolated conditions, the adaptive scope of populations to respond to future environmental change will surely be limited.

Clearly landscape history is important to consider when studying the ecology and genetics of species inhabiting discrete habitat patches (Habel and Zachos 2012). Although many naturally patchy habitat types are now becoming further fragmented due to contemporary anthropogenic activities, the specialists which occupy these habitats may not be as negatively affected by increased habitat isolation as other species because they have essentially adapted to persist under conditions of low gene flow and genetic diversity. For such specialists, local habitat patch characteristics (e.g., habitat size and quality) may be more likely to affect population persistence than habitat isolation. A recent theoretical simulation supports this hypothesis (Ye et al. 2013), however, empirical evidence of this relationship is limited.

1.4 Role of habitat quality in fragmented landscapes

In the context of habitat fragmentation and species conservation, most theoretical and empirical research to date has focused on understanding how habitat patch area and isolation influence dispersal, population size and the likelihood of extinction (i.e., ‘patch area-isolation paradigm’; Hanski and Gaggiotti 2004; Prugh et al. 2008) with

considerably less emphasis on the role of within-habitat patch quality (Ye et al. 2013). However, there is an increasing body of evidence which indicates that heterogeneity in quality among habitat patches can be an extremely important driver of local population dynamics and long-term persistence in fragmented landscapes (Thomas et al. 2001; Baguette et al. 2011; Ye et al. 2013) and thus should be included in ecological models (Fleishman et al. 2002; Schooley and Branch 2007). Several studies have now demonstrated the negative impacts of decreased habitat quality on the abundance and distribution of populations occupying fragmented landscapes (e.g., Dennis and Eales 1997; Thomas et al. 2001; Fleishman et al. 2002; Krauss et al. 2004). In contrast, there have been very few studies to date which have sought to evaluate the effects of spatial heterogeneity in habitat patch quality on measures of population genetic structure and diversity (but see Porlier et al. 2009; Pitra et al. 2011; Alda et al. 2013). In theory, higher quality habitat patches should support larger populations which are more genetically diverse; however, empirical evidence of this relationship is still largely lacking (de Vere et al. 2009; Pitra et al. 2011). Understanding of the genetic consequences of changes in habitat quality in addition to habitat patch size and isolation will be enable us to more accurately predict the dynamics and long-term persistence of populations in fragmented landscapes.

1.5 Evaluating the evolutionary potential of fragmented populations

Given the low levels of genetic diversity characteristic of fragmented populations, both anthropogenically and naturally induced, a major challenge in species conservation is

how to maintain the evolutionary potential of populations so that they have the ability to respond to future environmental stressors such as climate change, introduced diseases or parasites (Frankham 2005b; Allendorf et al. 2010). Most studies of the genetic effects of habitat fragmentation to date have focused on neutral genetic variation through which the relative influences of genetic drift and gene flow on fragmented populations can be inferred (Keyghobadi 2007). However neutral markers do not evolve directly in response to selection, and whether estimates of neutral genetic diversity accurately reflect adaptive potential is debated (Luikart et al. 2003; Holderegger et al. 2008; Frankham 2010). Therefore, in order to fully understand the effects of habitat fragmentation on patterns and levels of genetic diversity, adaptive genetic variation should also be examined (Holderegger et al. 2010). This will provide for more accurate assessments of local population viability and predictions on whether populations have the capacity to respond to changing environmental conditions (Allendorf et al. 2010). This is an area of research which has up until recently been relatively underexplored. However, new genomic methods and analytical tools provided through the emerging disciplines of population and landscape genomics (Luikart et al. 2003; Storz 2005; Joost et al. 2007; Manel et al. 2010) now make it possible to more readily study patterns of adaptive genetic variation in natural populations of non-model organisms (Holderegger et al. 2008).

1.6 The role of movement in fragmented landscapes

Particularly for small, isolated populations inhabiting fragmented landscapes, dispersal is a key life-history trait which influences local and regional population dynamics, population genetics and adaptive evolution (Hanski 1999; Ronce 2007; Nitepõld et al.

2009; Benton and Bowler 2012). Dispersal ability or tendency may itself become altered due to changes in landscape structure however, thereby affecting many aspects of the ecology and evolution of populations. Therefore, understanding how movement ability or propensity is influenced by habitat fragmentation may allow us to better predict how the dynamics and genetics of local populations will respond to landscape change (Stevens and Coulon 2012). Increasing isolation of habitat patches could potentially select for individuals with more mobile phenotypes who are able to traverse longer distances (e.g., Taylor and Merriam 1995). Conversely, because movement between habitat patches is physically costly and associated with many risks, increased isolation of patches may select for a decreased dispersal propensity (e.g., Dempster 1991; Schtickzelle et al. 2006). Such changes to the movement abilities of individuals among local habitat patches have the potential to either exacerbate or ameliorate the effects of habitat fragmentation on populations.

In addition, it has been recently suggested that spatial heterogeneity (e.g., in resource availability) within habitat patches can impose significant selective pressures on routine movements which can affect overall dispersal ability and morphology (Baguette and VanDyck 2007). ‘Routine’ or ‘station keeping’ movements are daily tasks which for the most part occur within a habitat patch, for example mate-location and foraging activities (Van Dyck and Baguette 2005). Local conditions within habitat patches may affect mobility, potentially in contrasting directions to the effects of surrounding landscape structure (Van Dyck and Baguette 2005; Turlure et al. 2010). While the effects of habitat fragmentation on movement and dispersal ability have been relatively well studied (e.g., Thomas et al. 1998; Hill et al. 1999a, b; Norberg and Leimar 2002; Merckx

et al. 2003; Vandewoestijne and Van Dyck 2011), we know comparatively very little regarding the influence of within-patch habitat heterogeneity (e.g., spatial variation in resource availability) on the evolution of movement ability (Turlure et al. 2010; Ye et al. 2013). Therefore, future studies which examine movement ability in relation to both landscape structure and local habitat characteristics will provide new insights into the factors shaping dispersal propensity of populations.

1.7 Butterflies: important flagship species and a model system for the study of habitat fragmentation

Butterflies (Order: Lepidoptera) are one of the most extensively studied and well-described groups of invertebrates (Gaston 1991). Due to their strict ecological requirements, butterflies are recognized as potentially valuable indicators of ecosystem health (McGeoch 2007) and climate change (Hellman 2002; Parmesan 2003). In recent years many species have experienced marked declines in abundance, and in some regions, such as the U.K., these losses have been well-documented and the causes well-understood. Consequently, butterflies are often the focus of conservation efforts as protection of their habitat is seen as a proxy for more general species conservation (e.g., Fox et al. 2011). In addition, in comparison to other threatened invertebrate taxa, butterflies are perceived as charismatic and aesthetically appealing animals, and have thus become an important flagship for raising awareness and developing new methodologies for invertebrate species conservation (Samways 1994; New 1997).

Butterflies have long been used as a model organism for the study of population biology (Gilpin and Hanski 1991; Hanski 1999; Ehrlich and Hanski 2004). Their utility in this field stems from the relative ease with which they can be observed and manipulated in both laboratory and field experiments and their well studied life-histories and general ecology (Boggs et al. 2003). As well, species within this taxonomic group are ecologically diverse and exhibit population structures ranging from tightly ‘closed’ to extremely ‘open’, possess dispersal abilities ranging from sedentary to migratory, and utilize resources which range in distribution from spatially restricted to widespread (Dover and Settele 2008 and references therein). Furthermore, these ecological traits can even vary considerably within species (Stevens et al. 2010). Much research has been conducted on the factors affecting butterfly dispersal, population genetic structure, population dynamics and spatial structure (e.g., Hill et al. 1996; Keyghobadi et al. 1999; Schtickzelle et al. 2002; Louy et al. 2007). Thus they are ideal candidate species with which to study the effects of habitat fragmentation on populations.

1.8 Population genetic studies of butterflies using AFLPs and non-lethal tissue sampling

Microsatellite molecular markers (short, tandemly repeating DNA sequences with a repeat motif of one to six nucleotides) are one of the most powerful and commonly applied marker systems in current population genetic and evolutionary research due to their high variability and presumed selective neutrality (Bruford and Wayne 1993; Zhang and Hewitt 2003). In most Lepidoptera species however, microsatellite sequences are associated with repetitive flanking regions which makes it extremely challenging to

isolate successfully a sufficient number of microsatellite markers (e.g., > 10 loci) for population genetic studies (Nève and MeglécZ 2000; Zhang 2004). Furthermore, scoring microsatellites and calculating population genetic estimates for many Lepidopteran species is often complicated by particularly high null-allele frequencies which must be accounted for (MeglécZ et al. 2004). Consequently many genetic studies of butterflies have employed other marker types including allozymes (e.g., Nève et al. 2008), mitochondrial DNA (mtDNA; e.g., Proshek et al. 2013), random amplified polymorphic DNAs (RAPDs; e.g., Vandewoestijne and Baguette 2004), and amplified fragment length polymorphisms (AFLPs; Takami et al. 2004; Collier et al. 2010; Leidner and Haddad 2010).

For studies of population genetic structure and diversity in non-model species, AFLPs are arguably the most suitable alternative marker system (Meudt and Clarke 2007). The AFLP technique (Vos et al. 1995) generates a large number of informative and reproducible multilocus markers (>100) which are widely distributed throughout the genome (Meudt and Clarke 2007). AFLPs are relatively quick and inexpensive to produce, and *a priori* sequence knowledge of the study organism is not required (Mueller and Wolfenbarger 1999; Bensch and Åkesson 2005). In comparison to other multilocus genomic techniques such as randomly amplified polymorphic DNA (RAPDs) and intersimple sequence repeats (ISSRs), AFLPs have been shown to be far more robust, informative and reproducible (Meudt and Clarke 2007 and references therein). Briefly, the AFLP protocol involves complete digestion of genomic DNA with restriction enzymes, followed by two cycles of selective polymerase chain reaction (PCR) amplifications and capillary electrophoresis of a subset of the fragments to produce a

multi-locus profile for every individual (Bensch and Åkesson 2005). Unlike PCR-based molecular systems where the target marker is directly amplified (e.g., microsatellites, mtDNA), the first step of the AFLP method involves a restriction and ligation, and thus requires a sufficient quantity (~ 100-1000 ng) and quality of template DNA (Meudt and Clarke 2007). AFLP markers are dominant, meaning that heterozygotes are not distinguishable from homozygotes (Meudt and Clarke 2007). Thus, a limitation of using AFLPs (or any dominant marker type) for population genetic analyses is that allele frequencies must be estimated (Krauss 2000). For AFLPs this is commonly accomplished through the use of a robust Bayesian method (Zhivotovsky 1999). As well, the di-allelic nature of AFLP loci means that individually, they are less informative than microsatellites. However, because such a large number of loci are typically generated, AFLPs have the statistical power to detect even small genetic differences and in studies examining taxonomic and population differentiation AFLPs have actually been shown to out-perform microsatellites (Perrie et al. 2003; Woodhead et al. 2005; Meudt and Clarke 2007). As well, an additional advantage to generating such a large number of markers which are scattered across the genome is that some loci may be linked to genes or genomic regions under selection (Bonin et al. 2007). Using genome scan approaches it is possible to identify those AFLP loci exhibiting signatures of selection, which can then be used for example to study relationships between adaptive genetic variation and environmental factors (Holderegger et al. 2010). Thus for non-model species when no prior genomic or phenotypic information is available, AFLPs are a particularly useful molecular tool for studying the role of selection in shaping patterns of genetic variation among populations (Meudt and Clarke 2007).

In genetic studies of butterflies, particularly for threatened species, non-lethal methods of tissue sampling (e.g., wing tissue sampling, leg removal) are increasingly being employed in an effort to minimize the effects of sampling on populations and justify the collection of larger sample sizes (Lushai et al. 2000; Keyghobadi et al. 2006, 2009; Vila et al. 2009; Hamm et al. 2010). Yet to my knowledge no AFLP studies to date have used DNA samples collected non-lethally. Recently, my colleagues and I demonstrated that small, non-lethally sampled, pieces of butterfly wing tissue yield DNA in concentrations adequate for the generation and analysis of AFLPs (Keyghobadi et al. 2009). We have also found non-lethal sampling to have no effect on individual survival, and flight and reproductive behaviours in two butterfly species thereby validating the use of these methods which have long been assumed to be non-detrimental (Koscinski et al. 2011; Crawford et al. 2013). For population genetic studies of butterflies therefore, not only do AFLPs represent a suitable alternative molecular marker to microsatellites, but through the use of non-lethally sampled tissue they are particularly useful for the study of endangered populations.

1.9 The quality of AFLP data in ecological and evolutionary research

An important caveat to consider when using AFLPs for population genetic studies is that differences in laboratory, peak-calling and locus-selection protocols can generate datasets varying widely in genotyping error rate (i.e., mismatch error rate of AFLP primer combinations), the number of loci used and, potentially, estimates of genetic diversity or differentiation. In my experience the majority of even recently published AFLP studies

do not provide clear details regarding the methodology followed, and do not quantify and report genotyping error. As a result, it can be challenging to make meaningful comparisons among studies. For example, in genetic studies of fragmented populations, researchers may be interested in evaluating how levels of genetic diversity and estimates of gene flow compare to other more connected populations. Without knowing the details of the methodologies used to generate the AFLP datasets and the overall quality of the loci used in genetic analyses (genotyping error rate) inter-study comparisons must be made with caution. Moreover, these problems also exist for other marker types such as microsatellites and next-generation sequencing techniques, particularly those which use restriction enzymes for fragment generation. Therefore, in order for genetic studies to be of value to conservation research it is important that all genotyping studies more transparently report the methodological details followed, and in particular the assessment of dataset reproducibility. Currently, I am not aware of any journals which enforce such standards for the publication of AFLP or any other genotype data.

1.10 Dissertation structure

My thesis consists of five data chapters which were designed as separate studies for independent publication. Chapters 2 and 3 have previously been published, and Chapters 4 – 6 will soon be submitted for publication. The data chapters represent a compilation of case studies focused on understanding aspects of the ecology and evolution of remnant butterfly populations inhabiting fragmented landscapes. I employ multiple methodological approaches including analyses of neutral and adaptive genetic variation (AFLP-based), and flight-related morphology, to examine how local habitat

characteristics and surrounding landscape structure contribute to the dynamics and genetics of fragmented populations. As well, my thesis also includes a review study which examines the use of AFLPs in ecological and evolutionary research, highlighting the challenges associated with employing this marker system for genetic studies.

In Chapter 2 (“A call for more transparent reporting of error rate: the quality of AFLP data in ecological and evolutionary research”; Crawford et al. 2012) I conduct a detailed and extensive literature review to quantify the extent to which AFLP studies provide adequate information on AFLP locus reproducibility. I review current molecular ecology literature (470 recently published AFLP articles) to determine the proportion of studies that report an error rate and follow established guidelines for assessing error. Based on my results, I highlight current gaps in the publication standards for AFLP data and provide recommendations for both researchers and publishers regarding ways to improve the quality and transparency of the data used in all genotyping-based studies.

In Chapter 3 (“Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs”; Crawford et al. 2011), I investigate the genetic structure and diversity of a relict Canadian population of the Mormon metalmark butterfly using, for the first time, AFLP markers generated from non-lethal samples of butterfly wing tissue. Within the last century, the geographic range of the Mormon metalmark in British Columbia, Canada has been considerably restricted due to human activities. Currently, only a single population occurs, distributed across a small number of fragmented habitat patches. The functional connectivity and genetic status of these remnant sub-populations has been unknown, making it difficult to predict future population trends and develop an effective management strategy. Thus, the

objective of this study was to assess patterns of neutral genetic variation within and among sub-populations in order to identify potential barriers to movement and genetic exchange which could be targeted in future recovery plans.

In Chapter 4 (“Local patch characteristics determine patterns of genetic diversity in a glacial relict, peatland specialist butterfly”) I used AFLP markers to describe patterns of population genetic structure and diversity among local populations of the bog copper (*Lycaena epixanthe*), a glacial relict butterfly endemic to naturally fragmented peatland habitats. Relict species which inhabit naturally fragmented habitats are adapted to living under conditions of limited gene flow and low genetic diversity. It has been recently proposed that for these habitat specialists, population viability may be more dependent on local habitat patch characteristics than on structural characteristics of the surrounding landscape. However, few empirical tests of this hypothesis exist. In this study, I first aimed to evaluate whether the bog copper exhibits the genetic characteristics of a habitat specialist, and secondly whether differences in intrapopulation genetic diversity could be best explained by local patch characteristics (habitat patch size and quality) rather than by the surrounding landscape structure (structural connectivity of potential habitat).

In Chapter 5 (“Flight morphology corresponds to both broad and fine-scale landscape structure in a highly specialized glacial relict butterfly (*Lycaena epixanthe*)”), I investigated variation in flight-related morphology among populations of the bog copper. Flight morphology characters (e.g., thorax mass and wing loading) are a reliable proxy of flight ability in many butterflies, and have been shown to respond rapidly to ecological and landscape change. For habitat specialists like the bog copper, flight and mobility may reflect ecological conditions within habitat patches rather than broad-scale landscape

structure. However, we currently know very little about how within-patch habitat heterogeneity influences movement ability and related aspects of morphological design. Therefore, the purpose of this study was to determine the relative influence of local habitat patch characteristics and surrounding landscape structure on inter-population variation in flight morphology in the bog copper.

In Chapter 6 (“Molecular signatures of selection associated with fine-scale landscape heterogeneity in a relict butterfly, *Lycaena epixanthe*”), I assessed whether local populations of the bog copper butterfly exhibited molecular evidence of local adaptation using an AFLP-based genome scan approach. For relict habitat specialists such as the bog copper, which are naturally geographically restricted to discrete habitat patches, fine-scale landscape heterogeneity may be an important driver of local adaptation among populations. However, few empirical studies have explored this question. My previous genetic and morphological analyses (Chapters 4 and 5) suggested that both habitat patch characteristics and aspects of surrounding landscape structure may be imposing selective pressures on populations of the bog copper. Thus, the objectives of this study were to 1) identify candidate AFLP loci potentially under divergent selection using an outlier approach and 2) test for associations between allele frequencies of these candidate loci and habitat and landscape variables identified as influential in my previous genetic and morphological studies.

Finally, in Chapter 7, I conclude my thesis with a general summary of the insights provided by my dissertation research.

1.11 Literature Cited

- Aguilar A, Roemer G, Deenham S, Binns M, Garcelon D and Wayne RK. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceeding of the National Academy of Sciences of the United States of America*, 101: 3490-3494.
- Alda F, Garcia J, Garcia JT and Suarez-Seoane S. 2013. Local genetic structure on breeding grounds of a long-distance migrant passerine: the bluethroat (*Luscinia svecica*) in Spain. *Journal of Heredity*, 104: 36-46.
- Allendorf FW, Hohenlohe PA and Luikart G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11: 679-709.
- Anderson LW, Fog K and Damgaard C. 2004. Habitat fragmentation causes bottlenecks and inbreeding the European tree frog (*Hyla arborea*). *Proceedings of the Royal Society of London B: Biological Sciences*, 271: 1293-1302.
- Andr n H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat – a review. *Oikos*, 71: 355-366.
- Baguette M, Clobert J and Schtickzelle N. 2011. Metapopulation dynamics of the bog fritillary butterfly: experimental changes in habitat quality induced negative density-dependent dispersal. *Ecography*, 34: 170-176.
- Baguette M and Van Dyck H. 2007. Landscape connectivity and animal behaviour: functional grain as a key determinant for dispersal. *Landscape Ecology*, 22: 1117-1129.

- Bensch S, Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.
- Benton TG and Bowler DE. 2012. Linking dispersal to spatial dynamics, in: Dispersal ecology and evolution, eds. Clobert J, Baguette M, Benton TG and Bullock JM. Oxford University Press, Oxford.
- Boggs CL, Watt WB and Ehrlich PR. 2003. Butterflies: Ecology and evolution taking flight. University of Chicago Press, London.
- Boitani L, Falcucci A, Maiorano L, Rondinini C. 2007. Ecological networks as conceptual frameworks or operational tools in conservation. *Conservation Biology*, 21: 1414-1422.
- Bonin A, Ehrlich D and Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16: 3737-3758.
- Broquet T, Angelone S, Jaquiere J, Joly P, Lena J-P, Lengagne T, Plenet S, Luquet E and Perrin N. 2010. Genetic bottlenecks driven by population disconnection. *Conservation Biology*, 24: 1596-1605.
- Bruford MW and Wayne RK. 1993. Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development*, 3: 939-943.
- Cassel-Lundhagen A. 2010. Peripheral relict populations of widespread species: evolutionary hotspots of just more of the same?, in: Relict species: phylogeography

and conservation biology, eds. Habel JC and Assmann T. Springer, Heidelberg, Germany.

Caughley G. 1994. Directions in Conservation Biology. *Journal of Animal Ecology*, 63: 215-244.

Collier N, Gardner M, Adams M, McMahon CR, Benkendorff K, Mackay DA. 2010. Contemporary habitat loss reduces genetic diversity in an ecologically specialized butterfly. *Journal of Biogeography*, 37: 1277-1287.

Crawford LA, Koscinski D, Watt KM, McNeil JN and Keyghobadi N. 2013. Mating and oviposition success of a butterfly are not affected by non-lethal tissue sampling. *Journal of Insect Conservation*. doi: 10.1007/s10841-013-9566-8.

Crokrak P and Barrett SCH. 2002. Purging the genetic load: a review of the experimental evidence. *Evolution*, 56: 2347-2358.

Debinski DM and Holt RD. 2000. A survey and overview of habitat fragmentation experiments. *Conservation Biology*, 14: 342-355.

Dempster JP. 1991. Fragmentation, isolation and mobility of insect populations, in: Conservation of insects and their habitats, eds. Collins NM and Thomas JA. Academic Press Inc., San Diego, CA.

de Vere N, Jongejans E, Plowman A and Williams E. 2009. Population size and habitat quality affect genetic diversity and fitness in the clonal herb *Cirsium dissectum*. *Oecologia*, 159: 59-68.

- Dennis RLH and Eales HT. 1997. Patch occupancy in *Coenonympha tullia* (Müller, 1976) (Lepidoptera: Satyrinae): habitat quality matters as much as patch size and isolation. *Journal of Insect Conservation*, 1: 167-176.
- Didham RK, Ghazoul J, Stork NE and Davis AJ. 1996. Insects in fragmented forests: a functional approach. *Trends in Ecology and Evolution*, 11: 255-260.
- Ehrlich PR, Hanski I. 2004. On the wings of checkerspots: a model system for population biology. Oxford University Press, Oxford.
- Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Reviews of Ecology, Evolution and Systematics*, 34: 487-515.
- Fazey I, Fischer J and Lindenmayer DB. 2005. What do conservation biologists publish? *Biological Conservation*, 124: 63-73.
- Fleishman E, Ray C, Sjögren-Gulve P, Boggs CL and Murphy DD. 2002. Assessing the roles of patch quality, area, and isolation in predicting metapopulation dynamics. *Conservation Biology*, 16: 706-716.
- Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N and Snyder PK. 2005. Global consequences of land use. *Science*, 309: 570-574.

- Fox R, Brereton TM, Asher J, Botham MS, Middlebrook I, Roy DB and Warren MS. 2011. The state of the UK's butterflies 2011. Butterfly Conservation and the Centre for Ecology and Hydrology, Wareham, UK.
- Frankham R. 2005a. Genetics and extinction. *Biological Conservation*, 126: 131-140.
- Frankham R. 2005b. Stress and adaptation in conservation genetics. *Journal of Evolutionary Biology*, 18: 750-755.
- Frankham R. 2010. Where are we in conservation genetics and where do we need to go? *Conservation Genetics*, 11: 661-663.
- Frankham R, Ballou JD and Briscoe DA. 2010. Introduction to Conservation Genetics, 2nd ed. Cambridge University Press, Cambridge.
- Frankham R, Gilligan DM, Morris D and Briscoe DA. 2001. Inbreeding and extinction: Effects of purging. *Conservation Genetics*, 2: 279-285.
- Gaston KJ. 1991. The magnitude of global insect species richness. *Conservation Biology*, 5:283-296
- Ge XJ and Sun M. 1999. Reproductive biology and genetic diversity of a cryptoviviparous mangrove *Aegiceras corniculatum* (Myrsinaceae) using allozyme and intersimple sequence repeat (ISSR) analysis. *Molecular Ecology*, 8: 2061-2069.
- Gilpin M. 1991. The genetic effective size of a metapopulation. *Biological Journal of the Linnean Society*, 42: 165-175.

- Gilpin M and Hanski I. 1991. Metapopulation dynamics: empirical and theoretical investigations. Academic Press Inc., London.
- Gómez R, Hurtado I, Suárez ML, Vidal-Abarca MR. 2005. Ramblas in south-east Spain: threatened and valuable ecosystems. *Aquatic Conservation*, 15, 387-402.
- Habel JC, Assmann T, Schmitt T and Avise JC. 2010. Relict species: from past to future, in: Relict species: phylogeography and conservation biology, eds. Habel JC and Assmann T. Springer, Heidelberg, Germany.
- Habel JC, Reuter M, Drees C, Pfaender J. 2012. Does isolation affect phenotypic variability and fluctuating asymmetry in the endangered Red Apollo? *Journal of Insect Conservation*, 16: 571-579.
- Habel JC, Rödder D, Schmitt T and Nève G. 2011. Global warming will affect the genetic diversity and uniqueness of *Lycaena helle* populations. *Global Change Biology*, 17: 194-205.
- Habel JC and Schmitt T. 2012. The burden of genetic diversity. *Biological Conservation*, 147: 270-274.
- Habel JC and Zachos FE. 2012. Habitat fragmentation versus fragmented habitats. *Biodiversity and Conservation*, 21: 2987-2990.
- Hampe A and Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics*, 42: 313-333.
- Hanski I. 1999. Metapopulation ecology. Oxford University Press, Oxford.

- Hanski I and Gaggiotti O. 2004. Ecology, genetics and evolution of metapopulations. Elsevier Academic, Amsterdam.
- Hamm CA, Aggarwal D and Landis DA. 2010. Evaluating the impact of non-lethal DNA sampling on two butterflies, *Vanessa cardui* and *Satyrodes eurydice*. *Journal of Insect Conservation*, 14: 11-18.
- Hellman JJ. 2002. Butterflies as model systems for understanding and predicting climate change, in: Wildlife responses to climate change, eds. Schneider SH and Root TL. Island Press, Washington, DC.
- Hill JK, Thomas CD and Blakeley DS. 1999a. Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121: 165-170.
- Hill JK, Thomas CD and Lewis OT. 1996. Effects of habitat patch size and isolation on dispersal by *Hesperia comma* butterflies: implications for metapopulation structure. *Journal of Animal Ecology*, 65: 725-735.
- Hill JK, Thomas CD and Lewis OT. 1999b. Flight morphology in fragmented populations of a rare British butterfly, *Hesperia comma*. *Biological Conservation*, 87: 277-283.
- Hobbs RJ and Yates CJ. 2003. Impacts of ecosystems fragmentation on plant populations: generalising the idiosyncratic. *Australian Journal of Botany*, 51: 471-488.
- Holderegger R, Böhler D, Gugerli F, Manel S. 2010. Landscape genetics of plants. *Trends in Plant Science*, 15: 675-683.

- Holderegger R, Herrmann D, Poncet B, Gugerli F, Thuiller W, Taberlet P, Gielly L, Rioux D, Brodbeck S, Aubert S and Manel S. 2008. Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, 1: 273-283.
- Karl I, Schmitt T and Fischer K. 2008. PGI genotype effects life history traits and temperature stress resistance in a Copper butterfly. *Functional Ecology*, 22: 887-894.
- Keller LF and Waller DM. 2002. Inbreeding effects in wild populations. *Trends in Ecology and Evolution*, 17: 230-241.
- Keller I, Nentwig W and Largiadèr CR. 2004. Recent habitat fragmentation due to roads can lead to significant genetic differentiation in an abundant flightless ground beetle. *Molecular Ecology*, 13: 2983-2994.
- Keyghobadi N. 2007. The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, 85: 1049-1064.
- Keyghobadi N, Crawford LA and Maxwell SA. 2009. Successful analysis of AFLPs from non-lethally sampled wing tissues in butterflies. *Conservation Genetics*, 10: 2021-2024.
- Keyghobadi N, Roland J and Strobeck C. 1999. Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae) *Molecular Ecology*, 8: 1481-1495.

- Keyghobadi N, Unger KP, Weintraub JD and Fonseca DM. 2006. Remnant populations of the regal fritillary (*Speyeria idalia*) in Pennsylvania: local genetic structure in a high gene flow species. *Conservation Genetics*, 7: 309-313.
- Koscinski D, Crawford LA, Keller HA and Keyghobadi N. 2011. Effects of different methods of non-lethal tissue sampling on butterflies. *Ecological Entomology*, 36: 301-308.
- Krauss J, Steffan-Dewenter I and Tschardtke T. 2004. Landscape occupancy and local population size depends on host plant distribution in the butterfly *Cupido minimus*. *Biological Conservation*, 120: 355-361.
- Krauss SL. 2000. Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology*, 9: 1241-1245.
- Lande R. 1998. Anthropogenic, ecological and genetic factors in extinction and conservation. *Researches on Population Ecology*, 40: 259-269.
- Leidner AK, Haddad, NM. 2010. Natural, not urban, barriers define population structure for a coastal endemic butterfly. *Conservation Genetics*, 11: 2311-2320.
- Lindenmayer DB and Fischer J. 2006. Habitat fragmentation and landscape change: An ecological and conservation synthesis. Island Press, Washington, DC.
- Lomolino MV, Riddle BR, Whittaker RJ and Brown JH. 2010. Biogeography, 4th edn. Sinauer Associates, Inc., Sunderland, MA.

- Louy D, Christian H, Schmitt T, Assmann T, Meyer M and Müller P. 2007. Strongly diverging population genetic patterns of three skipper species: the role of habitat fragmentation and dispersal ability. *Conservation Genetics*, 8: 671-681.
- Luikart G, England PR, Tallmon D, Jordan S and Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4: 981-994.
- Lushai G, Fjelsted W, Marcovitch O, Aagaard K, Sherratt TN, Allen JA and Maclean N. 2000. Application of molecular techniques to non-lethal tissue samples of endangered butterfly populations (*Parnassius apollo* L.) in Norway for conservation management. *Biological Conservation*, 94: 43-50.
- MacDougall-Shackleton EA, Clinchy M, Zanette L and Neff BD. 2011. Songbird genetic diversity is lower in anthropogenically versus naturally fragmented landscapes. *Conservation Genetics*, 12:1195-1203.
- Manel S, Joost S, Epperson BK, Hoderegger R, Storfer A, Rosenberg MS, Scribner KT, Bonin A and Fortin M-J. 2010. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19: 3760-3772.
- McGeoch MA. 2007. Insects and bioindication: theory and progress, in: *Insect conservation biology: proceedings of the Royal Entomological Society's 23rd Symposium*, eds. Stewart AJA, New TR and Lewis OT. CABI, Wallingford, UK.
- Meglécz E, Petenian F, Danchin E, D'Acier AC, Rasplus J-Y and Faure E. 2004. High similarity between flanking regions of different microsatellites detected within each of

- two species of Lepidoptera: *Parnassius Apollo* and *Euphydryas aurinia*. *Molecular Ecology*, 13: 1693-1700.
- Merckx T, Van Dyck H, Karlsson B, Leimar O. 2003. The evolution of movements and behaviour at boundaries in different landscapes: a common arena experiment with butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 270: 1815-1821.
- Meudt HM, Clarke AC. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12: 106-117.
- Mueller UG and Wolfenbarger LL. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution*, 14: 389-394.
- Nève G and Megléc E. 2000. Microsatellite frequencies in different taxa. *Trends in Ecology and Evolution*, 15: 376-377.
- Nève G, Barascud B, Descimon H and Baguette M. 2008. Gene flow rise with habitat fragmentation in the bog fritillary butterfly (Lepidoptera: Nymphalidae). *BMC Evolutionary Biology*, 8:84.
- New TR. 1997. *Butterfly conservation*, 2nd edn. Oxford University Press, Oxford.
- Nitepõld K, Smith AD, Osborne JL, Reynolds DR, Carreck NL, Martin AP, Marden JH, Ovaskainen O and Hanski I. 2009. Flight metabolic rate and *Pgi* genotype influence butterfly dispersal rate in the field. *Ecology*, 90: 2223-2232.

- Norberg U and Leimar O. 2002. Spatial and temporal variation in flight morphology in the butterfly *Melitaea cinxia* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, 77: 445-453.
- Paetkau D, Waits, LP, Clarkson PL, Craighead L, Vyse E, Ward R and Strobeck C. 1998. Variation in genetic diversity across the range of the North American brown bears. *Conservation Biology*, 12: 418-429.
- Parmesan C. 2003. Butterflies as bioindicators for climate change effects, in: *Butterflies: ecology and evolution taking flight*, eds. Boggs CL, Watt WB and Ehrlich PR. University of Chicago Press, London.
- Perrie LR, Brownsey PJ, Lockhart PJ and Large MF. 2003. Evidence for an allopolyploid complex in New Zealand *Polystichum* (Dryopteridaceae). *New Zealand Journal of Botany*, 41: 189-215.
- Pertoldi C, Bijlsma R and Loeschcke V. 2007. Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. *Biodiversity Conservation*, 16: 4147-4163.
- Pitra C, Suárez-Seoane S, Martín CA, Streich W-J and Alonso JC. 2011. Linking habitat quality with genetic diversity: a lesson from great bustards in Spain. *European Journal of Wildlife Research*, 57: 411-419.
- Porlier M, Bélisle M and Garant D. 2009. Non-random distribution of individual genetic diversity along an environmental gradient. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364: 1543-1554.

- Proshek B, Crawford LA, Davis CS, Desjardins S, Henderson AE and Sperling FAH. 2013. *Apodemia mormo* in Canada: population genetic data support prior conservation ranking. *Journal of Insect Conservation*, 17: 155-170.
- Prugh LR, Hodges KE, Sinclair ARE and Brashares JS. 2008. Effect of habitat area and isolation on fragmented animal populations. *Proceedings of the National Academy of Sciences in the United States of America*, 105: 20770-20775.
- Reed DH. 2010. Albatrosses, eagles and newts, oh my!: exceptions to the prevailing paradigm concerning genetic diversity and population viability? *Animal Conservation*, 13: 448-457.
- Reed DH and Frankham R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology*, 17: 230-237.
- Ronce O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Reviews in Ecology Evolution and Systematics*, 38: 231-253.
- Sala, OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Swanald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes MT, Walker BH, Walker M and Wall DH. 2000. Global biodiversity scenarios for the year 2010. *Science*, 287: 1770-1774.
- Samways MJ. 1994. *Insect conservation biology*. Chapman and Hall, London.
- Schooley RL and Branch LC. 2007. Spatial heterogeneity in habitat quality and cross-scale interactions in metapopulations. *Ecosystems*, 10: 846-853.

- Schtickzelle N, Le Boulengé E and Baguette M. 2002. Metapopulation dynamics of the bog fritillary butterfly: demographic processes in a patchy population. *Oikos*, 97: 349-360.
- Schtickzelle N, Mennechez G and Baguette M. 2006. Dispersal depression with habitat fragmentation in the bog fritillary butterfly. *Ecology*, 87: 1057-1065.
- Skrede I, Eidesen PB, Portela RP and Brochmann C. 2006. Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Molecular Ecology*, 15: 1827-1840.
- Stevens VM and Coulon A. 2012. Landscape effects on spatial dynamics: the natterjack toad as a case study, in: Dispersal ecology and evolution, eds. Clobert J, Baguette M, Benton TG and Bullock JM. Oxford University Press, Oxford.
- Stevens VM, Pavoine S and Baguette M. 2010. Variation within and between closely related species uncovers high intra-specific variability in dispersal. *PLOS ONE*, 5: e11123.
- Storz JF. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, 14: 671-688.
- Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fishman DL and Waller RW. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science*, 306: 1783-1786.

- Takami Y, Koshio C, Ishii MU, Fujii H, Hidaka T, Shimizu IM. 2004. Genetic diversity and structure of urban populations of *Pieris* butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, 13: 245-258.
- Taylor PD and Merriam G. 1995. Wing morphology of a forest damselfly is related to landscape structure. *Oikos*, 73: 43-48.
- Templeton AR, Shaw K, Routman E and Davis SK. 1990. The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*, 77: 13-27.
- Thomas JA, Bourn NAD, Clarke RT, Stewart KE, Simcox DJ, Pearman GS, Curtis R and Goodger B. 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. *Proceedings of the Royal Society of London B: Biological Sciences*, 268: 1791-1796.
- Thomas CD, Hill JK and Lewis OT. 1998. Evolutionary consequences of habitat fragmentation in a localized butterfly. *Journal of Animal Ecology*, 67: 485-497.
- Turlure C, Schtickzelle N and Baguette M. 2010. Resource grain scales mobility and adult morphology in butterflies. *Landscape Ecology*, 25: 95-108.
- Vandewoestijne S and Baguette M. 2004. Genetic population structure of the vulnerable bog fritillary butterfly. *Hereditas*, 141: 199-206.
- Van Dyck H and Baguette M. 2005. Dispersal behaviour in fragmented landscapes: routine or special movements? *Basic and Applied Ecology*, 6: 535-545.

- Vandewoestijne S and Van Dyck H. 2011. Flight morphology along a latitudinal gradient in a butterfly: do geographic clines differ between agricultural and woodland landscapes? *Ecography*, 34: 876-886.
- Vila M, Auger-Rozenberg MA, Goussard F and Lopez-Vaamonde C. 2009. Effect of non-lethal sampling on life-history traits of the protected moth *Graellsia isabelae* (Lepidoptera: Saturniidae). *Ecological Entomology*, 34: 356-362.
- Visscher PM, Smith D, Hall SJG and Willaim JL. 2001. A viable herd of genetically uniform cattle. *Nature*, 409: 303.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407-4414.
- Ye X, Skidmore AK and Wang T. 2013. Within-patch habitat quality determines the resilience of specialist species in fragmented landscapes. *Landscape Ecology*, 28: 135-147.
- Watt WB, Wheat CW, Meyer H and Martin JF. 2003. Adaptation at specific loci. VII. Natural selection, dispersal and the diversity of molecular-functional variation patterns among butterfly species complexes (*Colias*: Lepidoptera, Pieridae). *Molecular Ecology*, 12: 1265-1275.
- Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M and Powell W. 2005. Comparative analysis of population genetic structure in *Athyrium*

distenifolium (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Molecular Ecology*, 14: 1681-1695.

Zhang DX. 2004. Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology and Evolution*, 19: 507-509.

Zhang D-H and Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, 12: 583-584.

Zhivotovsky L. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8: 907-913.

Chapter 2

2 A call for more transparent reporting of error rate: The quality of AFLP data in ecological and evolutionary research¹

2.1 Introduction

Amplified fragment length polymorphism (AFLP) analysis is an established multilocus genomic fingerprinting technique commonly employed in ecological and evolutionary research in a broad range of taxa (Bensch and Åkesson 2005). The popularity of the AFLP technique (Vos et al. 1995) can be attributed to the large number of informative markers (>100) that can be developed relatively easily and inexpensively without *a priori* knowledge of the study organism's genome (Meudt and Clarke 2007). Applications of AFLPs in molecular ecology research are wide ranging and include determining levels of genetic diversity and population genetic structure, detecting hybridization, parentage analysis, detecting loci of ecological relevance, assignment of individuals, detecting markers associated with phenotype, and reconstruction of phylogenies (Bensch and Åkesson 2005; Meudt and Clarke 2007).

¹ A version of this chapter has been published and is presented here with permission from John Wiley and Sons.

Citation: Crawford LA, Kosciński D and Keyghobadi N. 2012. A call for more transparent reporting of error rates: The quality of AFLP data in ecological and evolutionary research. *Molecular Ecology*, 21: 5911-5917.

Though considered robust and reproducible, particularly in comparison to other dominant marker systems such as random amplified polymorphic DNAs (RAPDs) and inter simple sequence repeats (ISSRs), AFLPs are still susceptible to genotyping error (Mueller and Wolfenbarger 1999; Meudt and Clarke 2007). Genotyping errors are discrepancies found among multiple genotypes generated from the same sample (Bonin et al. 2004) and, regardless of the markers used, can result from causes such as low quality DNA, sample contamination, biochemical artefacts, and human error (Bonin et al. 2007). For AFLPs specifically, errors can also arise from incomplete restriction digest reactions, co-migration of non-homologous fragments (allele homoplasy), PCR plate boundary artifacts, and errors during the interpretation of AFLP profiles such as scoring background noise as real peaks (Vekemans et al. 2002; Bonin et al. 2007; Holland et al. 2008; Zhang et al. 2012). Detailed reviews of how genotyping error can be generated are provided by Bonin et al. (2004) and Pompanon et al. (2005).

In AFLP studies, reproducibility (i.e., inverse of genotyping error) often varies among loci, such that not all peaks in an AFLP profile should necessarily be retained for analysis. Rather, loci that contribute disproportionately to high error rates can be excluded and objective methods for doing so have been proposed (Whitlock et al. 2008; Herrmann et al. 2010). Thus, the overall reproducibility of AFLP datasets can vary as a result of differences in laboratory, peak-calling and locus-selection protocols (Pompanon et al. 2005). Conversely, for a given dataset, the number of loci retained for analysis can vary depending on the genotyping error rate that is accepted (e.g., Chapter 3, Crawford et al. 2011; Zhang et al. 2012).

Relatively little is known regarding the effects of AFLP error rate on downstream analyses and subsequent biological conclusions. In theory, genotyping errors could lead to inaccurate estimates of genetic diversity and population structure or false detection of selection (Vekemans et al. 2002; Koopman and Gort 2004; Bonin et al. 2007; Caballero et al. 2008). For example, Zhang et al. (2012) compared the results obtained from AFLP datasets varying in error rates (0, 1, 2, 3, 4 and > 4 %) and found that inaccurate inferences of a previously determined phylogeographic pattern were made based on datasets with > 4 % error. While it is possible that genotyping errors may not significantly bias overall conclusions (e.g., Bonin et al. 2004) such errors may contribute to higher levels of noise and reduce the power of the dataset (Meudt and Clarke 2007). Herrmann et al. (2010) showed that AFLP fragments selected using stringent criteria designed to reduce genotyping error resulted in marked differences in estimates of genetic diversity and genetic differentiation when compared with a dataset composed of randomly selected fragments.

Thus, variation in results and parameter estimates among AFLP studies may in part be related to how the AFLP profiles were analyzed and how reproducible they are. Without knowing the methodological details for how a dataset was produced and its overall reproducibility, it is virtually impossible to critically assess the results, and make meaningful comparisons among studies (Pompanon et al. 2005).

The importance of quantifying and reporting genotyping error in molecular ecology studies has been well reviewed (see Bonin et al. 2004; Pompanon et al. 2005). However, in our experience, even recently published papers rarely provide adequate information on the reliability of AFLP datasets. To quantify objectively the extent of this

problem, we assessed the quality of AFLP datasets used in current molecular ecology research. We reviewed studies published in 2010-2011 that used AFLPs in population genetic or phylogenetic research, to determine what proportion reported error rate, and of those how many appropriately followed recommended guidelines for assessing and reporting genotyping error (as per Bonin et al. 2004 and Pompanon et al. 2005). We report trends in error reporting among studies by taxonomic group, fragment scoring methods, and journal calibre (based on impact factor). Our review highlights a current gap in publication standards and we aim to encourage researchers to estimate and report AFLP genotyping error using existing guidelines, and to more transparently report how such error was quantified.

2.2 Methods

2.2.1 Literature search

We conducted a literature search using the citation database ‘Web of Science’ (Thompson Reuters; <http://www.isiknowledge.com>), restricting our search to those published articles available in-print in 2010 and 2011 (up to November 21 2011). Our intent was to examine evolutionary and ecological studies which employ AFLP data to make population genetic and phylogenetic inferences. We therefore used the *topic* keywords “AFLP” NOT “linkage” so as to exclude any research articles which used AFLPs exclusively for linkage mapping or quantitative trait analysis. We also searched using the keywords “AFLP” AND “linkage disequilibrium” in case any relevant studies had been removed by the previous search because the term linkage disequilibrium had

been mentioned. Only articles which contained “AFLP” in the title and/or abstract were retained.

We used EndNote x4 (Thompson Reuters) to search for AFLP studies which conducted research specifically in the fields of molecular ecology and/or evolution. Key terms were selected from an initial subset of articles. The most common key terms that were consistently listed in this subset were then used (*‘structure’*, *‘varia*’*, *‘differentiation’*, *‘diversity’*, *‘phylo*’*, *‘taxonomy’*, *‘scan’*, *‘gene flow’*, *‘biogeography’*, *‘adapt*’*). We recognize that this is not an exhaustive search. However, the number of papers generated by our initial search without filtering would have been logistically infeasible to characterize in detail. We used keywords which encompassed the research themes of most molecular ecology studies, and although we recognize that by using only a few keywords we may have excluded several relevant papers, we are reasonably confident that articles selected using these criteria should be representative of all molecular ecology studies that employ AFLP markers.

We read through all articles identified in this way and removed any non-relevant studies still retained by our keyword search. A study was only considered relevant if it specifically used AFLPs to conduct some form of population genetic or phylogenetic analyses, such as estimating H_e , F_{ST} , AMOVA, genetic distance, or generating a neighbour-joining tree. We used only empirical studies that reported an AFLP dataset for the first time, removing literature reviews, conference proceedings, primer notes, purely theoretical papers, etc.

2.2.2 Data collection

For each article, we collected the publication name as well as the taxonomic group studied: broadly classified as (i) *plants and red algae*, (ii) *animals*, (iii) *fungi*, (iv) *bacteria* and (v) *protists*. For all articles we recorded the total number of genotyped individuals, the number of AFLP primer combinations used, and the total number of loci and number of polymorphic loci retained in the final dataset. We also recorded the methodology used to score AFLP fragments. Typically scoring is either conducted by eye, or by using software such as GeneMapper (Applied Biosystems), GeneMarker (Softgenetics), RawGeno (Arrigo et al. 2009; available from <http://sourceforge.net/projects/rawgeno/>) or AFLPSCORE (Whitlock et al. 2008; available from www.shef.ac.uk/molecol/software/aflpscore) to analyse automatically and score AFLP datasets based on user specified parameters. As well, some studies now employ a ‘semi-automated’ approach where the positions of the marker bins used for allele calling are first determined or verified by the researcher, before the automated software is allowed to score the AFLP data (e.g., Whitlock et al. 2008; Herrmann et al. 2010). Thus, we classified the scoring approaches used by AFLP studies as either (i) *manual* (fragments scored visually), (ii) *automated* (fragments scored by a software program) or (iii) *semi-automated* (marker bins are determined or inspected manually before fragments are scored by software). Finally, we also noted the impact factor of every journal represented in our dataset using the 2010 ISI Impact Factor (Thomson Reuters).

Using replicate samples from the focal dataset (i.e., not standards) is considered the most robust way to estimate genotyping error rate (Bonin et al. 2004). It is

recommended that a substantial proportion of the total sample size be replicated (5-10%) and that original tissue samples, or at least the same DNA extract, be used to generate replicates so that estimated error rates are reflective of the entire genotyping process (restriction-ligation, pre-selective PCR, selective PCR, scoring) and include both technical and human errors (Bonin et al. 2004; Pompanon et al. 2005). We categorized all articles based on whether they reported an assessment of dataset reproducibility using replicate samples. We found that many studies claimed to use ‘reproducible’ loci based on replicates, but did not actually report a specific genotyping error rate for their dataset. Thus, articles were classified as: 1) genotyping error rate reported (GE), (2) reproducibility assessed but genotyping error rate not reported (NGE) and (3) no assessment of reproducibility reported (NR). For studies that assessed dataset reproducibility (GE and NGE) we determined the proportion of the total sample size that was replicated (i.e., number of replicate samples / total sample size), and the stage of the AFLP protocol at which replicates were generated: (i) different DNA extractions of the same tissue sample, *tissue*, (ii) from the same DNA extraction, *extraction*, (iii) different restriction-ligation reactions, *R-L*, (iv) different PCR reactions, *PCR*, (v) different sequencing runs, *sequencing* or (vi) AFLP profiles scored more than once, *scoring*.

Studies which presented more than one dataset (e.g., multiple species were examined) were represented multiple times in our analysis where a variable of interest (e.g., number of primer pairs, loci, or replicate samples) differed among the datasets. Impact factor scores, the total number of loci and the number of polymorphic loci were normalized by \log_{10} transformation. We used a non-parametric test (Kruskal-Wallis) to examine differences in the number of primer combinations used among error reporting

categories, as the former was not normalized even by transformation. All statistical analyses were conducted in SPSS v.16.0 (SPSS Inc., Chicago) and all tests were two-tailed.

2.3 Results

2.3.1 Overall trends

Following our filtering steps to select appropriate AFLP-based articles in ecology and evolution, we reviewed a total of 470 studies published in 205 different journals between 2010 and 2011. We found that 54 % of these studies did not report *any* evaluation of the reproducibility of the AFLP datasets (category NR). Of the studies which did claim to use a ‘reproducible’ dataset based on assessment of replicate samples (GE and NGE), 41 % (or 19 % of the total) did not report an associated genotyping error rate (NGE; Figure 2.1).

Overall, a large proportion of the studies examined were of plant species (plants and red algae: 60 %; animals: 22 %; fungi: 9 %; bacteria: 8 %; protists: 1 %). Interestingly however, a higher proportion of animal studies reported an AFLP error rate (35 %) than any other group (range = 10-27 %; Figure 2.2).

The number of primer combinations used by studies did not differ among the error reporting categories, GE, NGE and NR (average 5.4 - 6.8 / category; Kruskal-Wallis, $\chi^2_{(2, n = 502)} = 1.802, P = 0.406$). While the total number of loci and the number of polymorphic loci retained in the final analysis also did not differ significantly among categories (one-way ANOVA, $F_{2, 359} = 1.432, P = 0.240$; $F_{2, 363} = 1.723, P = 0.180$,

respectively), we observed a trend of fewer total loci and polymorphic loci being retained in GE studies compared to NR studies, with NGE being intermediate (Figure 2.3).

We also found variation in the scoring method used among studies: a greater proportion of GE studies (76 %) used either a semi-automatic or automatic approach than either the NGE (45 %) or NR (50 %) groups.

In addition we found that journal impact factor differed significantly among papers in the different error reporting categories (one-way ANOVA, $F_{2,440} = 32.589$, $P < 0.001$; Figure 2.4). On average the impact factor of journals of GE studies was significantly greater than that of both NR and NGE studies (post-hoc Tukey HSD test, $P < 0.001$).

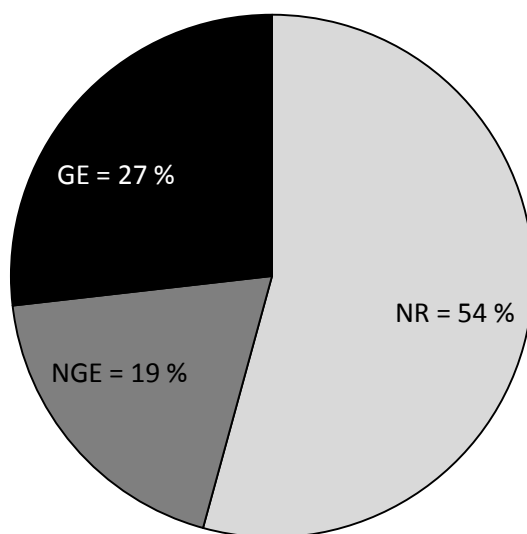


Figure 2.1 AFLP studies surveyed, categorized by reporting of dataset reproducibility. *GE*, genotyping error rate reported; *NGE*, reproducibility assessed but genotyping error rate not reported, *NR*, no assessment of reproducibility reported. Total sample size = 470 (*GE* = 126; *NGE* = 89; *NR* = 255).

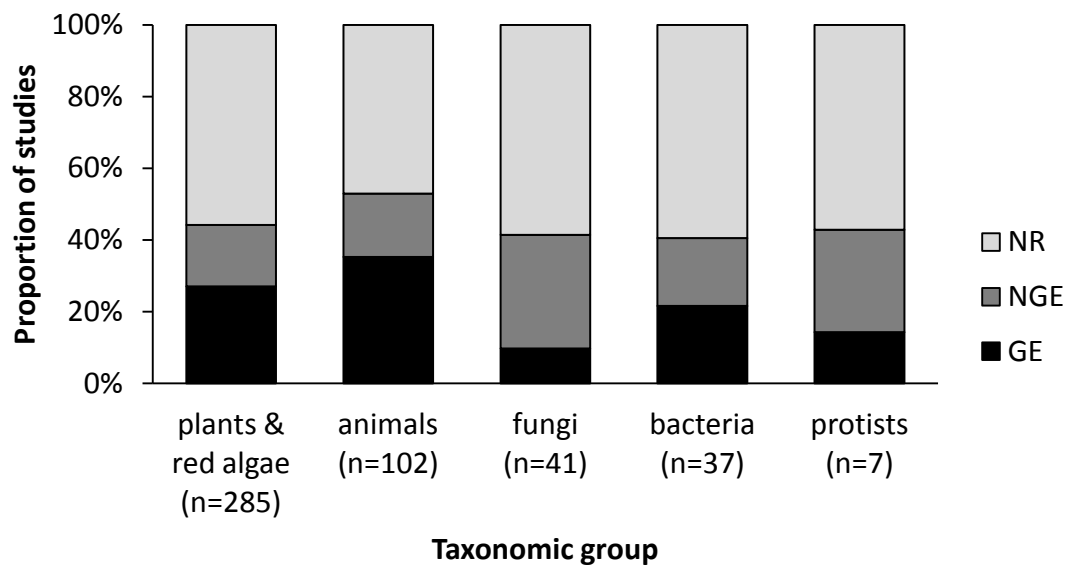


Figure 2.2 Within each taxonomic group, the proportion of surveyed AFLP studies where reproducibility was assessed and genotyping error rate was reported (*GE*), reproducibility was assessed but genotyping error rate was not reported (*NGE*) and where reproducibility was not reported at all (*NR*). Total sample size = 472.

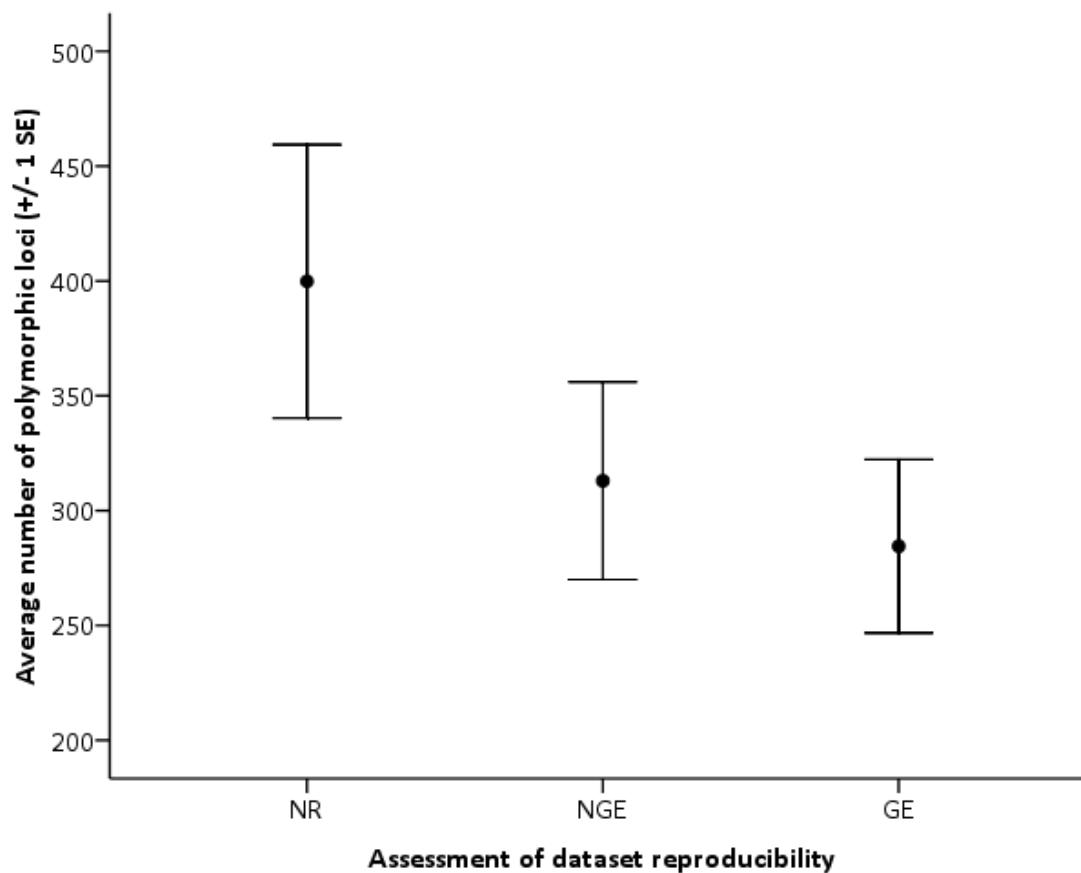


Figure 2.3 Average number of polymorphic AFLP loci retained for final analysis in studies, grouped by reporting of dataset reproducibility. Total sample size = 369 (GE = 99; NGE = 69; NR = 201). *GE*, genotyping error rate reported; *NGE*, reproducibility assessed but genotyping error rate not reported, *NR*, reproducibility was not reported.

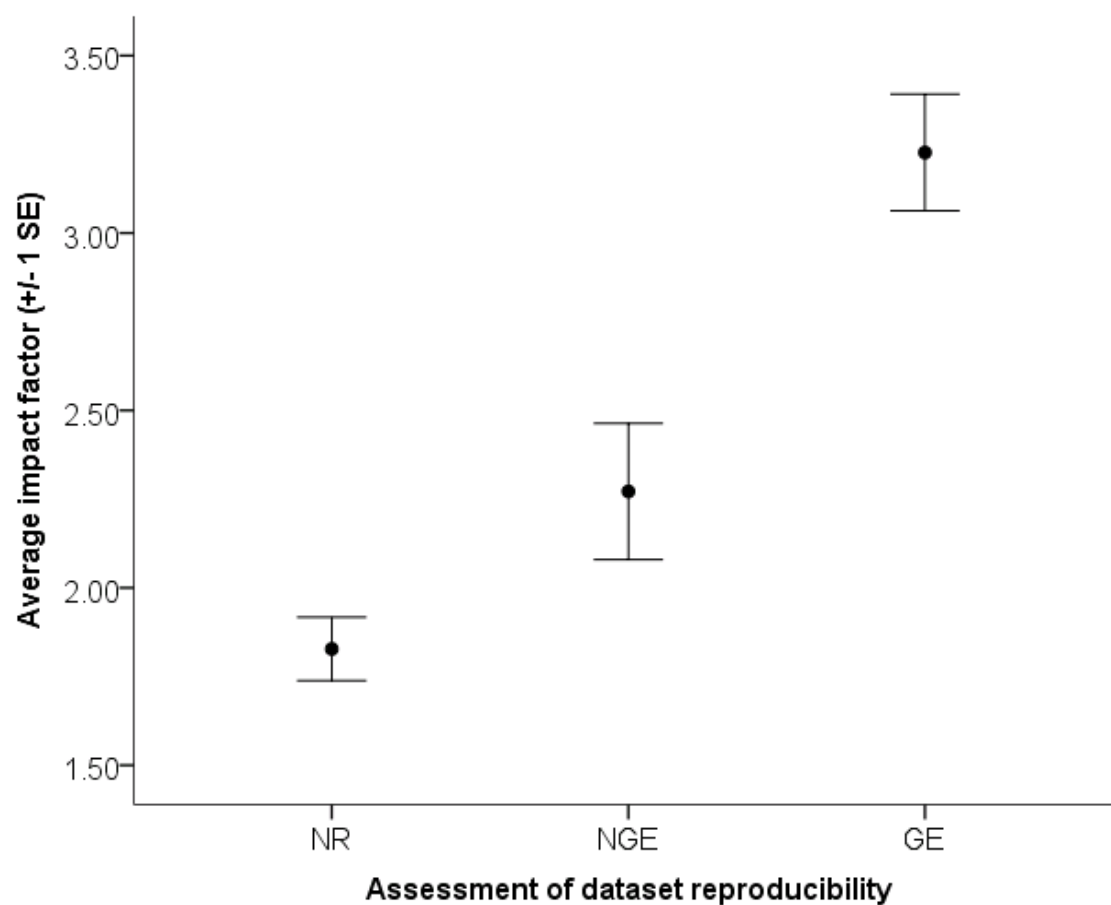


Figure 2.4 The average journal impact factor of surveyed AFLP studies within each error reporting category. Total sample size = 420 (GE = 123; NGE = 82; NR = 215). *GE*, genotyping error rate reported; *NGE*, reproducibility assessed but genotyping error rate not reported, *NR*, reproducibility was not reported.

2.3.2 Comparisons among ‘reproducible’ studies

Among all studies which reported some assessment of dataset reliability using replicate samples (GE and NGE), we observed that 35 % either did not include information on the number of samples replicated, or did not replicate a minimum of 5 % of the total sample size (Table 2.1). This proportion was greater for NGE studies than GE studies (43 and 31 %, respectively; Table 2.1). Although they represent only a small portion of all studies reviewed, it is worth noting that multiple studies (7 %) determined the reproducibility of their datasets based on only a very small number of replicate samples (< 2 %). Indeed, in some cases we recorded values as low as one replicate sample. We also noted that several studies determined reproducibility for only a portion of the final dataset (i.e. the reliability of *all* primer combinations used in the study was not examined).

Many of the studies which reported using replicate samples did generate them as recommended at either the tissue or DNA extraction stage of the AFLP protocol (37 %; Table 2.2). When we examined the studies within GE and NGE separately we found that a greater proportion of studies within GE replicated samples at the tissue and DNA extraction stages (42 and 29 %, respectively; Table 2.2). However, 10 % of studies did not replicate samples until further on in the AFLP protocol, at either the R-L, PCR, sequencing or scoring stages, and, more importantly, more than half of the total studies examined (53 %) did not report clear details for how the replicates were generated (Table 2.2). We found that the language used to describe how replicate samples were generated was often very unclear and ambiguous (e.g., “*replicates of the AFLP protocol were conducted*”, “*individuals underwent a second amplification*”). If neither of two

researchers (LAC, DK) could interpret the methodology used, the stage of replication was classified as NA (information not available; Table 2.2).

2.4 Discussion

2.4.1 Genotyping error reporting in AFLP studies

Despite much discussion on the importance of quantifying and reporting genotyping error (Bonin et al. 2004; Pompanon et al. 2005; Meudt and Clarke 2007) our review indicates that this is still not standard practice in the AFLP literature. Indeed we found that more than half of the AFLP studies we reviewed did not report any form of assessment of the reliability of their datasets. Of the studies that indicated using a ‘reproducible’ dataset, a large proportion did not report a specific error rate and did not provide specific details of how reproducibility was assessed. Moreover, many of the papers which did report an error rate failed to follow the recommended standards (Bonin et al. 2004; Pompanon et al. 2005) for quantifying error. Thus the quality of these datasets may not actually be as high as suggested by the reported error rate.

Table 2.1 Replicate sample size, as a percentage of the total number of samples in the dataset, for studies which assessed reproducibility of AFLP datasets. Results are presented as the proportion of studies within each reporting category (*GE*; *NGE*; *overall*, *GE* and *NGE* combined) that reported replicate sample sizes of < 2 %, ≥ 2 and < 5 %, ≥ 5 and < 10 %, and ≥ 10 %. Values in bold type indicate the proportion of studies which replicated a minimum of 5 % of total sample size as recommended by Bonin et al. (2004).

	Size of replicate sample (percentage of total # of samples)				
	NA	< 2 %	≥ 2 & < 5 %	≥ 5 & < 10 %	≥ 10 %
GE	0.09	0.07	0.16	0.25	0.44
NGE	0.18	0.08	0.17	0.11	0.46
Overall	0.12	0.07	0.16	0.20	0.45

Total sample size (n) = 223; GE = 139; NGE = 84

NA, information not available; *GE*, genotyping error rate reported; *NGE*, reproducibility assessed but genotyping error rate not reported

Table 2.2 The stage at which replicate samples were generated for studies which assessed reproducibility. Results are presented as proportion values, and were calculated in relation to the total number of studies within each reporting category (*GE*; *NGE*; *overall*, *GE* and *NGE* combined). Values in bold type indicate the proportion of studies which met recommended guidelines of Bonin et al. (2004).

	NA	Post- DNA extraction	DNA extraction	Tissue
GE	0.51	0.07	0.20	0.22
NGE	0.57	0.14	0.12	0.17
Overall	0.53	0.10	0.17	0.20

Total sample size = 240; GE = 148; NGE = 92

Post-DNA extraction, represents all studies classified as *R-L*, *PCR*, *sequencing* or *scoring*

NA, information not available; *GE*, genotyping error rate reported; *NGE*, reproducibility assessed but genotyping error rate not reported

We recognize that failing to report genotyping error does not necessarily mean that the data used are unreliable or that the results of such studies should be considered invalid. Certainly in many cases we found that although studies did not report an actual error value, they did report precautions taken to limit genotyping error from occurring throughout the AFLP protocol. We argue however, that reporting an associated error rate for a given dataset is an indication of a study's robustness, and it allows for appropriate evaluation of the significance of results and comparison to other studies. This is especially important for AFLPs, where the number of loci retained in the final dataset and the error rate are correlated and decisions regarding locus filtering could potentially affect the accuracy and precision of population genetic analyses (Herrmann et al. 2010; Zhang et al. 2012). In the current study, we found that AFLP articles which reported genotyping error tended to be published in journals with higher impact factors. In general, studies published in such journals are very thorough in design, so it is possible that these studies would be more likely to report genotyping error. We are not aware of variation among journals in explicit standards followed for publication of AFLP data.

More animal studies tend to report AFLP genotyping error rate compare to any other taxonomic group. Pompanon et al. (2005) documented an increase in the number of population genetics papers which dealt with genotyping error from 1989 to 2004; although their search was not exclusive to AFLPs, they also found that more non-human animal studies dealt with genotyping error than genetic studies of humans or plants. They noted that most of the animal studies addressing error had used non-invasive methods of DNA sampling, where low quality and quantity of DNA is a concern and where determining genotyping error is therefore a common practice in order to ensure dataset

quality. Thus, taxon-related differences in the assessment of dataset reliability may result from taxon-related variation in the nature of tissue sampling.

While the average number of primer combinations used did not differ among GE, NGE and NR studies, we found that the number of loci retained for analysis tended to be lower in studies which assessed reproducibility. Given that these studies also made efforts to filter out ‘error-prone’ loci, we would expect these datasets to contain on average fewer loci than unfiltered datasets. Interestingly, even though the majority of GE studies used automated sequencers to detect AFLP fragments (results not shown), and this technique is known to produce a larger number of loci than traditional gel electrophoresis for the same primer combinations and study species (Terefework et al. 2001; Papa et al. 2005; Reunova et al. 2010), we still found that the number of loci retained for analysis in these studies was less than the other reporting categories which used traditional gels more often. This suggests that in these studies the locus-filtering step is removing a large number of loci which might otherwise contribute error and hence noise to the dataset.

Throughout the articles reviewed in this study, we noted that the words ‘*reliable*’ and ‘*reproducible*’ were consistently used by authors to imply that their datasets were of high-quality. However, we found that many studies which claimed to use ‘*reproducible loci*’ did not actually report *any* details for how reproducibility was assessed and were therefore considered under the NR category. Still more concerning was our finding that even for those studies which did use replicates to assess reproducibility (GE and NGE), many did not follow the recommended guidelines entirely or did not report sufficient details of their methodology. For example, a relatively small proportion of samples were

replicated, samples were not replicated from the beginning of the protocol, and most critically genotyping error was not reported. Thus in some cases, both the terminology used and values of error reported may be very misleading to readers and indicate that the presented data are more reproducible than they actually are. Providing clear, complete details for how the dataset was generated and quantifying the associated genotyping error using a standard protocol is critical to ensure accurate interpretation of results.

Finally, we also observed a considerable amount of variation in the type of error rate which was reported among GE studies. In AFLP research, the most commonly employed metric of genotyping error is the mean error rate per locus (Pompanon et al. 2005; Bonin et al. 2007), which can be calculated as the ratio between the total number of mismatched phenotypes (band presence vs. absence) to the number of replicated phenotypes (Pompanon et al. 2005). While other acceptable methods to calculate error also exist (Pompanon et al. 2005; Holland et al. 2008) the link between these measures is not always straightforward, making them difficult to compare. Mean error rate per locus is considered a universal quality index metric in molecular research, which permits comparisons among studies and different marker systems to be made (Pompanon et al. 2005) and should therefore be considered the preferred metric of error to use in AFLP studies. While we had originally intended to examine trends in genotyping error rate values among taxonomic groups, journal impact factors, etc., we found that there was too much variation among articles in the calculations used to determine genotyping error as well as the amount of detail reported. Thus we could not easily extract this information for all studies, and conducting comparison among different groups would not have been meaningful.

2.4.2 Conclusions and recommendations

The results from this review indicate that there is currently a general lack of consistency among AFLP studies in the methods used to assess reproducibility and even in the details of methodology presented. This makes it extremely difficult to assess the quality of the AFLP datasets and to compare the results of different studies. This is an issue not only for the researchers employing these datasets, but more importantly for the integrity of AFLPs used in ecological studies. To demonstrate robustness to reviewers and readers, as well as to facilitate comparisons among studies, we encourage researchers to employ established recommended guidelines for determining genotyping error in AFLP data, and to report the following items explicitly in every publication:

- All steps taken throughout the AFLP protocol to minimize genotyping error
- Number of samples or proportion of total sample size that was replicated
- Stage in AFLP protocol at which samples were replicated
- Methods and parameters used to score and select loci for final analysis
- The specific formula which was used to calculate genotyping error, ideally reporting the commonly employed mismatch error per locus (as per Bonin et al. 2004) to allow for inter-study comparisons
- The error rate associated with the final dataset used for genetic analysis
- The initial number of loci obtained and the final number of loci retained for analysis

Following the recommended guidelines, and reporting the items listed above, will certainly improve the quality and transparency of AFLP studies for those researchers who

choose to follow them. However, it is clear that several past calls for improved reporting of genotyping error have failed to effectively promote change in how AFLP data are presented in the literature. We thus propose that journals implement publication standards for AFLP studies, and require the reporting of genotyping error rate following a specific set of guidelines (for example the list provided above) as a condition of publication, similar to requiring DNA sequence data to be publicly archived. Other molecular science disciplines, such as forensic genetics, and gene expression already follow strict reporting standards to ensure data quality (Brazma et al. 2001; Pompanon et al. 2005; Schneider 2007).

As demonstrated by our current study, even when genotyping error rates are reported, most researchers fail to comply with recommended error reporting standards. While in some cases these details may be caught during the review process, this is clearly often not happening. Standardizing the reporting of genotyping error for AFLP studies would serve several important functions including: (1) Ensure that genotyping error rates are consistently reported for all AFLP studies, and that they follow a common set of guidelines, (2) Encourage confidence in and continued use of AFLPs in population genetic research by demonstrating that the AFLP datasets used are reproducible and robust, (3) Enable accurate interpretation of results, and facilitate inter-study comparisons and meta-analysis and (4) Provide better quality control of AFLP studies, by both the researcher and the reviewers. We are not aware of any journals that currently enforce standards for reporting of genotyping error for AFLPs or any other genotype studies.

We foresee few major challenges for both researchers and journals in making the details of genotyping error rate calculations a standard requirement of AFLP studies. For

the most part, journal space limitations should not be an issue as only a few lines of text are typically required to outline the methodological details and results of calculating genotyping error (e.g., Chapter 3, Crawford et al. 2011). Should more in-depth details be required, they could be provided as supplementary information, especially when space limitations are an issue.

Over time, the use of AFLPs in molecular ecology research will likely be replaced by emerging high-resolution next-generation sequencing (NGS) techniques. However, until these technologies become more cost-effective and widely accessible, we anticipate the continued use of AFLPs in molecular ecology research. Furthermore, the problems we have highlighted in this review are not specific to just the AFLP literature alone. Indeed, a survey by Guichoux et al. (2011) of 100 microsatellite studies recently published in *Molecular Ecology* noted that only 26 % of articles reported a measure of genotyping error. As well, NGS techniques particularly those which rely on restriction enzymes for the generation of fragments (e.g., reduced-representation sequencing, restriction-site-associated DNA sequencing and multiplexed shotgun genotyping) are still susceptible to genotyping error, and require quality control (Davey et al. 2011). Therefore, a policy for reporting genotyping error should apply not only to AFLP studies, but for all genotyping studies in general. Establishing standardized guidelines (perhaps specific to different genotyping methods) that researchers should follow for publishing genotype data, will ensure the continued quality of data used in molecular ecology research, even as the technologies used change over time.

Finally, a group of ecology and evolution journals including *Molecular Ecology*, *Evolution*, *Heredity*, *The American Naturalist*, and *Journal of Evolutionary Biology*

recently joined together and introduced parallel data archiving policies (Whitlock et al. 2010). We propose that a consortium of journals could similarly implement a common set of guidelines for reporting genotyping error across all genotype studies to further ensure consistency and promote data quality among studies published in different journals.

2.5 Literature Cited

- Arrigo N, Tuszyński JW, Ehrich D, Gerdes T and Alvarez N. 2009. Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics*, 10: 33.
- Bensch S and Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.
- Bonin A, Bellemain E, Bronken E, Pompanon F, Brochmann C and Taberlet P. 2004. How to track and assess genotyping errors in population genetic studies. *Molecular Ecology*, 13: 3261-3273.
- Bonin A, Ehrich D and Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16: 3737-3758.
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J and Vingron M. 2001. Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. *Nature Genetics*, 29: 365-371.
- Caballero A, Quesada H and Rolan-Alvarez E. 2008. Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, 179: 539-554.

- Crawford LA, Desjardins S and Keyghobadi N. 2011. Fine-scale genetic structure of endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs. *Conservation Genetics*, 12: 991-1001.
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM and Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12: 499-510.
- Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F and Petit RJ. 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11: 591-611.
- Herrmann D, Poncet BN, Manel S, Rioux D, Gielly L, Taberlet P and Gugerli F. 2010. Selection criteria for scoring amplified fragment length polymorphism (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*, 53: 302-310.
- Holland BR, Clarke AC and Meudt HM. 2008. Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology*, 57: 347-366.
- Koopman WJM and Gort G. 2004. Significance tests and weighted values for AFLP similarities, based on *Arabidopsis in silico* AFLP fragment length distributions. *Genetics*, 167: 1915–1928.
- Meudt HM and Clarke AC. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12: 106-117.

- Mueller UG and Wolfenbarger LL. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution*, 14: 389-394.
- Papa R, Troglio M, Ajmone-Marsan P and Nonnis Marzano F. 2005. An improved protocol for the production of AFLP markers in complex genomes by means of capillary electrophoresis. *Journal of Animal Breeding and Genetics*, 122: 62-68.
- Pompanon F, Bonin A, Bellemain E and Taberlet P. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, 6: 847-859.
- Reunova GD, Kats IL, Muzarok TI and Zhuravlev YN. 2010. Polymorphism of RAPD, ISSR and AFLP markers of the *Panax ginseng* C.A. Meyer (Araliaceae) Genome. *Russian Journal of Genetics*, 46: 1057-1066.
- Schneider PM. 2007. Scientific standards for studies in forensic genetics. *Forensic Science International*, 165: 238-243.
- Terefework Z, Kaijalainen S and Lindstrom K. 2001. AFLP fingerprinting as a tool to study the genetic diversity of *Rhizobium galegae* isolated from *Galega orientalis* and *Galega officinalis*. *Journal of Biotechnology*, 91: 169-180.
- Vekemans X, Beauwens T, Lemaire M and Rodán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11: 139-151.

- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407-4414.
- Whitlock MC, McPeck MA, Rausher MD, Rieseberg L and Moore AJ. 2010. Data archiving. *The American Naturalist*, 175: 145-146.
- Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources* 8: 725-735.
- Zhang H and Hare MP. 2012. Identifying and reducing AFLP genotyping error: an example of tradeoffs when comparing population structure in broadcast spawning versus brooding oysters. *Heredity*, 108: 616-625.

3 Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs²

3.1 Introduction

Genetic techniques are now widely employed in conservation biology to tackle such problems as assessing extinction risk, resolving taxonomic status, detecting hybridization, and identifying sources for reintroduction (Frankham et al. 2002). To date, the management plans of many threatened species have been greatly enhanced by the knowledge gained through genetic studies (e.g., Madsen et al. 1999; Ralls et al. 2000; Wilson et al. 2000; Haig et al. 2001). Genetic markers are also powerful tools for revealing information about a species' status that may not be evident based on physical, behavioural, or demographic observations alone (Frankham et al. 2002). In particular, there are an increasing number of invertebrate animals of conservation concern for which little is known about their population structure and dispersal behaviours (New 1995). These important variables may be extremely difficult to study directly because of the small size and/or cryptic behaviour of many invertebrate species. Genetic data may thus be particularly critical in revealing key aspects of the ecology of such species.

² A version of this chapter has been published and is presented here with permission from Springer-Verlag.

The arid lowlands of the south Okanagan and Similkameen valleys of British Columbia (BC), Canada constitute a nationally rare habitat that supports approximately 15,000 invertebrate species (Cannings and Cannings 1995) and is at risk due to increasing human activity (Guppy et al. 1994). Many of the invertebrates found in this region are provincially, nationally, and even globally rare, yet relatively little is known regarding their biology including, in some cases, such basic parameters as their range and habitat requirements (Cannings and Cannings 1995). Of these rare invertebrates, several butterfly species, including the Mormon metalmark butterfly (*Apodemia mormo*; C & R Felder 1859), were among the first to be identified as threatened or endangered in this area (Guppy and Sheppard 2001).

While the Mormon metalmark in BC historically (early 1900s) occurred as far north as the Okanagan (Guppy et al. 1994), it now exists only in the Similkameen Valley. The BC population of the Mormon metalmark survives at low numbers (estimated as approximately 2000 individuals; S. Desjardins unpublished), confined to a highly restricted geographic area consisting of a small number (~15) of fragmented habitat patches. Given the population's confinement to a single valley, coupled with the butterfly's strong association with its larval food plant (snow buckwheat, *Eriogonum niveum*) and assumed sedentary habits, the BC Mormon metalmark population is considered extremely vulnerable to natural stochastic events (e.g. climatic extremes or disease outbreaks) as well as human activity (e.g., agricultural development) and was designated as endangered on the Canadian Species At Risk Act in 2003 (COSEWIC 2003). Efforts to protect critical habitat and stabilize the population since then have been hindered by a general lack of knowledge surrounding basic life history traits, habitat

requirements, dispersal capabilities and population structure of the butterfly, making it difficult to predict future population trends and develop an effective recovery strategy (COSEWIC 2003). Here, we use amplified fragment length polymorphism (AFLP) markers generated from non-lethally sampled wing tissue to assess the genetic diversity and structure of the BC population of the Mormon metalmark and provide information regarding the population's status that will be of value for future conservation and management plans.

Molecular techniques are increasingly employed in the study of endangered butterfly populations (e.g., Mountain apollo, *Parnassius apollo*: Lushai et al. 2000; Cranberry fritillary, *Boloria aquilonaris*: Vandewostejine and Baguette 2002; Karner blue, *Lycaedies melissa samuelis*: Gompert et al. 2006; Regal fritillary, *Speyeria idalia*: Keyghobadi et al. 2006; Marsh fritillary, *Euphydryas aurinia*: Sigaard et al. 2008, Crystal skipper, *Atrytonopsis* sp.: Leidner and Haddad 2010). Microsatellite markers are a popular marker of choice in conservation genetic studies due to their putative neutrality and high variability. For most Lepidoptera, however, microsatellite development is extremely challenging because of the occurrence of repetitive flanking regions (Zhang 2004) and the number of microsatellite markers used in most butterfly studies is consequently relatively low. More recently, several genetic studies of butterflies have successfully employed amplified fragment length polymorphisms (AFLPs) instead (e.g., Takami et al. 2004; Kronforst et al. 2007). However, in only a few cases have AFLPs been applied in a conservation context (Gompert et al. 2006; Collier et al. 2010; Leidner and Haddad 2010).

The AFLP technique (Vos et al. 1995) provides a large number (> 100) of informative and reliable genetic markers that can be developed quickly and inexpensively (Bensch and Akesson 2005). The markers generated are dominant and widely distributed throughout the genome, allowing an assessment of genome-wide variation (Meudt and Clarke 2007). The technique is particularly useful in non-model organisms as previous knowledge of the genome is not necessary (Bensch and Akesson 2005). Furthermore, it has been recently demonstrated that small, non-lethally sampled, pieces of butterfly wing tissue provide sufficient quantities of DNA for the analysis of AFLPs (Keyghobadi et al. 2009) and that they can be taken without lowering individual survival or affecting behaviour (Vila et al. 2009; Hamm et al. 2010; Kosciński et al. 2011; Crawford et al. 2013). Thus, the use of AFLPs for population genetic studies represents a promising tool, particularly for the study of endangered butterfly species where neither lethal sampling of individuals nor the development of microsatellites is preferable.

Best practices for AFLP marker development, scoring and analysis have been well reviewed (see Bensch and Akesson 2005; Bonin et al. 2007; Meudt and Clarke 2007). However, among studies that apply AFLPs (e.g., Lepidoptera studies; Table 3.1) there exists little consistency in the methods used or even in the details of methodology presented. In particular, the scoring parameters used, the mismatch error rate of AFLP primer combinations, as well as the total number of loci initially generated (not the total number of polymorphic loci) are often not reported. As a result, it is difficult to compare and interpret results, such as estimates of genetic diversity, among studies. These comparisons are of particular value for studies of endangered populations, where researchers are interested in examining whether levels of genetic diversity and estimates

of gene flow are lower in relation to those found in other populations or other organisms. Therefore, for the use of AFLP markers to truly be of value to conservation research it is critical for AFLP studies to provide detailed, transparent reports of the methodologies used and to only use loci with low genotype error rates. We intend for our work to act as a case study for applying AFLPs to endangered butterfly populations in general, and thus we provide detailed accounts of the entire AFLP fragment scoring process, the methods used to calculate genetic diversity measures, and we report mismatch error rates so that our results may be more meaningfully interpreted.

3.2 Methods

3.2.1 Study species

The Mormon metalmark is a small butterfly (wingspan of 25-35mm) in the family Riodinidae that is closely associated with disturbed, arid regions that contain its larval host plant (wild buckwheats, *Eriogonum sp.*). The range of *Apodemia mormo*, *sensu lato* is widespread in Mexico, and the western United States, while in Canada, the species occurs only in southern British Columbia and southwestern Saskatchewan (SK; Layberry 1998). While morphologically similar, the two Canadian populations of the Mormon metalmark in southern BC and southern SK are considered discrete, nationally significant populations, as they are geographically isolated from each other, as well as from other populations in the core of the species' distribution to the south (but see

Table 3.1 Summary of representative studies using AFLPs to measure genetic diversity in natural populations of Lepidoptera.

Species	Sample size	Fragment analysis method	AFLP scoring program(s) used	AFLP scoring parameters	Assessment of genotyping error	Initial no. loci	Polymorphic loci (%)	H _e	Source
<i>Vanessa atalanta</i>	277	6% polyacrylamide gel	NA	Bands clear and easily scored in all gels	NA	199	84.9	NA	Brattsröm et al. (2010)
<i>Theclinesthes albocincta</i>	248	ABI Prism 3730 automated sequencer ^a	GENEMAPPER v3.7 ^a	60-350bp; >50-70rfu; 1.2bp bin width	NA	363	13.8-41.9	0.066-0.139	Collier et al. (2010)
<i>Trichoplusia ni</i>	1082	LI-COR 4200 automated sequencer ^b	SAGA v2.0 ^b	NA	5.4% error rate (laboratory and scoring)	204	82.8	0.165-0.303	Franklin et al. (2010)
<i>Atrytonopsis new species 1</i>	98	LI-COR 4200 and 4300 sequencers ^b	Quantar v1.08 ^c	NA	1.7 % error rate	NA	68.5-100.0	0.274-0.416	Leidner and Haddad (2010)
<i>Ostrinia nubilalis</i>	180	LI-COR 4200 automated sequencer ^b	SAGA v3.2 ^b	NA	NA	164	72.0-94.0	0.237-0.376	Krumm et al. (2008)
<i>Tortrix viridana</i>	401	ALF express II automated sequencer ^d	ALF win Fragmentanalyser v1.02 ^d	NA	NA	74	62.7-77.9	0.130-0.160	Schroeder and Degen (2008)
<i>Grapholita molesta</i>	87	6% polyacrylamide gel	NA	Only scored reproducible fragments	NA	236	58.2-99.4	0.186-0.100	Timm et al. (2008)
<i>Lycoides melissa melissa</i>	79	ABI Prism 377 automated sequencer ^a	GENESCAN ^a	≥150rfu; excluded bands present in <5% individuals	95.5% scored bands detected in replicate samples	143	90.9	NA	Gompert et al. (2006)
<i>Lycoides melissa samuelis</i>	111	ABI Prism 377 automated sequencer ^a	GENESCAN ^a	≥150rfu; excluded bands present in <5% individuals	95.5% scored bands detected in replicate samples	143	86.7	NA	Gompert et al. (2006)
<i>Cydia pomonella</i>	128	6% polyacrylamide gel	NA	NA	NA	214	35.2-98.9	0.060-0.180	Timm et al. (2006)
<i>Pieris rapae;</i>	626	ABI Prism 373 automated sequencer ^a	GENESCAN v3.1.2, GENOTYPER v2.5 ^a	≥90bp	NA	484	10.1-24.0	0.044-0.064	Takami et al. (2004)
<i>Pieris melete</i>	235	ABI Prism 373 automated sequencer ^a	GENESCAN v3.1.2, GENOTYPER v2.5 ^a	≥90bp	NA	484	6.4-16.5	0.031-0.056	Takami et al. (2004)

H_e, expected heterozygosity; NA, information not reported; * Individuals collected from greenhouse and field populations; ^aApplied Biosystems; ^bLI-COR Biosciences; ^cKeyGene; ^dAmersham Biosciences

Proshek et al. 2013). The BC and SK populations also differ in their choice of host plant: *Eriogonum niveum* (snow buckwheat) and *Eriogonum pauciflorum* (fewflower buckwheat) respectively (COSEWIC 2003).

3.2.2 Sample collection

We sampled adult butterflies using hand nets from fourteen separate habitat patches during 2006-2008 (mid-August to early-September) within the Similkameen River Valley, British Columbia, Canada (49°12.469' N, 119°49.295' W; Figure 3.1). These sites represent all currently known BC populations that occur on Crown land or privately owned property. We used fine iris scissors and forceps to remove a small piece of tissue (3mm x 3mm, or ~1-5% of the hind wing area) from the hind wings of each individual. These 'wing clips' were immediately stored in absolute ethanol. All butterflies were released from the location of their initial capture.

3.2.3 DNA extraction and AFLP analysis

Genomic DNA was extracted from the wing tissue samples using the QIAgen DNeasy® Blood and Tissue Kit, and then concentrated using a standard ethanol precipitation (as in Keyghobadi et al. 2009). For AFLP analysis we used the commercial AFLP Plant Mapping Kit (Applied Biosystems) following a modified version of the manufacturer's protocol (Keyghobadi et al. 2009). Based on relative polymorphism and reproducibility, the following five selective primer combinations were used to obtain AFLP profiles for *A. mormo*: *EcoRI*-AAC/*MseI*-CAC, *EcoRI*-AGC/*MseI*-CAC, *EcoRI*-AGC/*MseI*-CAT,

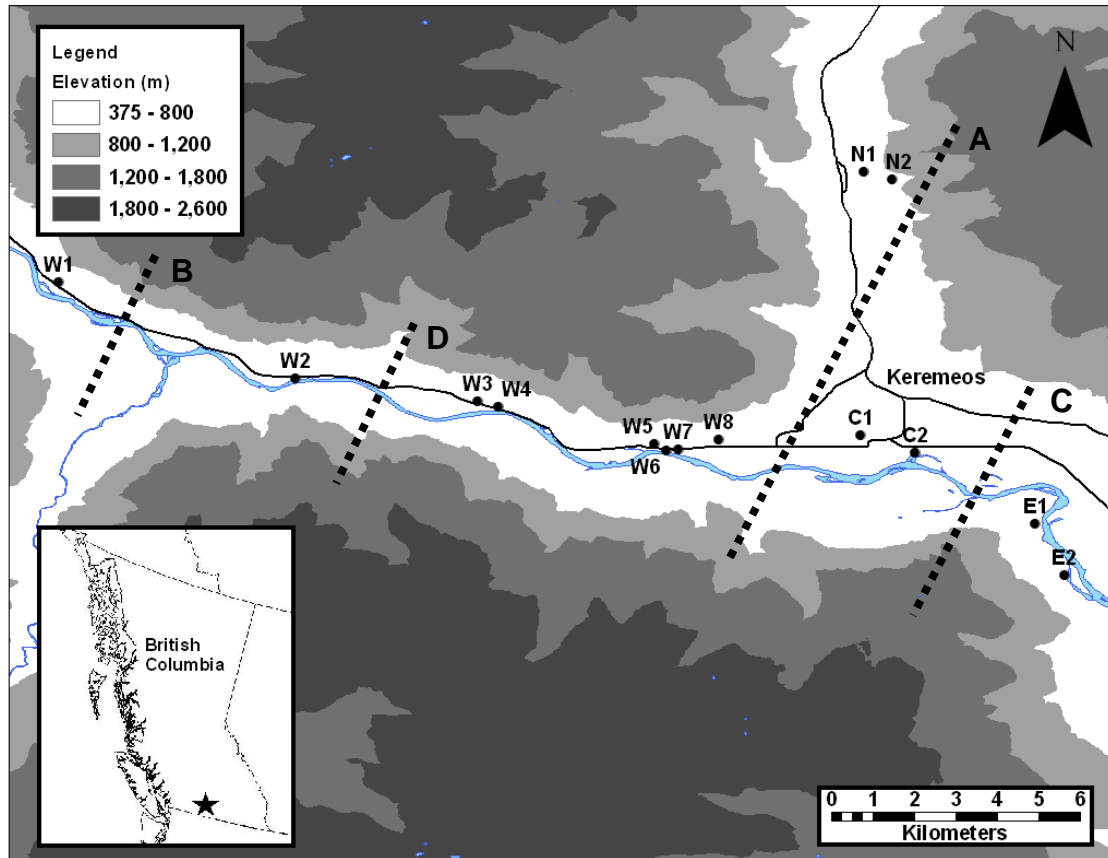


Figure 3.1 Locations of the 14 studied sub-populations of *Apodemia mormo* in the Similkameen River Valley, British Columbia, Canada.

Site codes correspond to the general orientation of each sub-population in relation to the town of Keremeos (W = west; C = central; E = east; N = north). The insert shows the location of the studied population in British Columbia. The Similkameen River is depicted in blue, major roads are depicted in black, and elevation is depicted in grey and measured as meters above sea level. Areas of high genetic differentiation within the British Columbian population of *Apodemia mormo*, as identified by BARRIER v2.2 (Manni et al. 2004), are indicated by dashed black lines. Barriers which possessed high bootstrap support (>50 %) are labelled in order of importance (A, B, C, and D).

EcoRI-AGC/*MseI*-CTT, and *EcoRI*-AAG/*MseI*-CAA. To ensure that no contamination occurred we included negative controls (i.e., Milli-Q purified water in place of DNA template and PCR products) at each step of the protocol. Fluorescently labelled AFLP fragments were separated and sized using a 3730S Genetic Analyzer (Applied Biosystems).

We determined AFLP fragment sizes and peak heights in GENEMAPPER v.4.0 (Applied Biosystems) using a semi-automated approach. Specifically, we allowed GENEMAPPER to identify automatically AFLP loci (bins) between 100 and 500 base pairs (bp) in size. We ignored fragments smaller than 100 bp to reduce the incidence of size homoplasy (Vekemans et al. 2002). All bins were set to a width of one base pair, and those fragments with peak heights below 50 relative fluorescence units (rfu) were assumed to represent instrument noise (Keyghobadi et al. 2009) and were not scored. To ensure that bin positions were assigned accurately, all bins were then checked manually. Any bins possessing fragments that overlapped with adjacent bins were removed. As well, we adjusted bins assigned off-centre of any peak distributions. Finally, all AFLP profiles were checked manually to ensure successful amplification and were either re-run or removed from analysis if the fingerprint failed to amplify or appeared to possess many unique fragments.

We left the AFLP peak-height data un-normalized in GENEMAPPER, and then normalized and scored the data in AFLPScore v.1.3b (Whitlock et al. 2008) using the data filter and absolute phenotype-calling threshold settings. Based on mismatch error analysis of replicate samples, AFLPScore is an objective method of AFLP phenotype scoring that establishes optimal scoring parameters while minimizing genotyping error.

Briefly, a locus selection threshold based on mean peak-height is applied, and loci with peak-heights equal or above this threshold are retained for further analysis. For each AFLP profile, a phenotype-calling threshold is then used to identify fragments as either present or absent. We estimated mismatch error rate separately for each primer combination using 36 replicate individuals. Due to the small size of the wing-clips used in this study and the small amounts of DNA thus available to us, replicate samples were generated by subjecting two aliquots of the same DNA extraction to the entire AFLP protocol independently.

We initially tested various combinations of locus and phenotype scoring thresholds in AFLPSCORE, ranging from 60-250 rfu each, to determine the optimal scoring parameters. We found that error rate did not decrease appreciably above values of 100 rfu for both thresholds, and therefore we chose to use locus and phenotype scoring thresholds of 100 rfu to ensure the reliability of the loci we retained. Upon conducting downstream genetic analyses, however, we found low estimates of population genetic variability and were concerned that we had used overly stringent scoring criteria that may have removed the more variable peaks. Therefore, we compared estimates of variability and population structure obtained using a less stringent criterion of 75 rfu (for both locus and phenotype-calling) which slightly increased error rate but retained more loci compared to using the 100 rfu thresholds. We did not include an additional comparison to threshold values of 60 rfu due to concerns that background instrument noise could affect peak calling.

In addition, for each set of threshold values (75 rfu and 100 rfu) we also compared the results obtained from genetic analyses using a dataset with all retained loci (dataset

A) versus using a dataset from which we further removed all loci (peaks) that were present in only a single individual (dataset B). The removal of such singleton peaks is often performed in AFLP analyses, since singleton loci may represent false markers (e.g., as a result of incomplete digestion) which can upwardly bias estimates of homozygosity (following a 5% polymorphism criterion; Milot et al. 2007). All subsequent population genetic analyses were thus conducted separately using the four described datasets (75-A, 75-B, 100-A, 100-B) with the exception of the measures of inbreeding (F_{AFLP}) which were only calculated using dataset 100-B.

3.2.4 Data analysis

3.2.4.1 Genetic diversity

We first estimated allele frequencies for each sample (hereafter referred to as a ‘sub-population’) using the program AFLP-SURV v.1.0 (Vekemans et al. 2002). We applied a Bayesian method with a non-uniform prior distribution of allele frequencies which is robust to minor departures from Hardy-Weinberg equilibrium (Zhivotovsky 1999). We then measured levels of genetic diversity within each sub-population by calculating the number and the proportion of loci that were polymorphic at the 5% level, as well as unbiased estimates of expected heterozygosity (or Nei’s genetic diversity) (following Lynch and Milligan 1994). We also calculated pairwise F_{ST} values, with significance tests based on 10,000 permutations. Input files for AFLP-SURV were prepared using the R script AFLPDAT (Ehrich 2006).

An AFLP-based measure of individual inbreeding coefficient, F_{AFLP} , was estimated for each individual sampled using a simulation approach in the program

FAFLPcalc (Dasmahapatra et al. 2008). Using raw AFLP counts, FAFLPcalc estimates band frequencies, which are then employed to produce simulated data assuming a wide range of inbreeding (f) values. The optimal fit between the observed distribution of f values and one of the simulated data sets is then determined. This method assumes that at least half of the individuals surveyed are outbred ($f = 0$), and that inbred individuals are more homozygous and will possess more null phenotypes than under random expectations. FAFLPcalc is a Visual Basic Macro for use in Excel, written with the capacity to handle up to 250 loci (less than our current dataset). To calculate F_{AFLP} values representative of the entire AFLP dataset we ran the simulation analysis in FAFLPcalc using a dataset consisting of 250 randomly selected loci, selected using a random number generator in Excel. This procedure was repeated a total of 20 times and the results were then averaged.

3.2.4.2 Population genetic structure

Using estimated allele frequencies and following Lynch and Milligan (1994) we calculated in AFLP-SURV the following unbiased estimates of population genetic structure: total gene diversity (expected heterozygosity or gene diversity in the overall sample), average gene diversity within sub-populations, average gene diversity among populations in excess of that observed within populations and F_{ST} (the ratio of between-site relative to within-site genetic variation).

We used Monmonier's (1973) maximum difference algorithm implemented in the program BARRIER v2.2 (Manni et al. 2004) to identify whether any significant genetic discontinuities existed among sites in the BC population. In brief, a Voronoï tessellation

is produced using geographic coordinates of the sampling locations, and from this a Delaunay triangulation is obtained. Using this geographic network, along with a corresponding genetic distance matrix, the Monmonier's algorithm then identifies barriers or zones along the edges of the triangulation where genetic distance values are greatest, based on the number of barriers defined by the user. The significance (or robustness) of the identified barriers is then assessed by means of bootstrap matrix analysis, and typically decrease with rank. We tested a series of barrier numbers, ranging from 1-14, and used 100 pairwise F_{ST} matrices (generated in AFLP-SURV) to assess the robustness of the computed barriers.

We also examined patterns of genetic differentiation among sub-populations by applying principal coordinates analysis (PCoA) to a genetic distance matrix of pairwise F_{ST} estimates using the data standardization option in GenAlEx v.6.0 (Peakall and Smouse 2006).

Finally, we tested for patterns of isolation-by-distance (IBD) by plotting measures of genetic dissimilarity between pairs of habitat patches (pairwise F_{ST}) against geographic distance (km). We determined the statistical significance of these correlations using the Mantel test (Mantel 1967) implemented in GenAlEx based on 9,999 permutations. Both F_{ST} and geographic distance were linearly transformed ($F_{ST}/(1-F_{ST})$, $\ln(\text{geographic distance})$), respectively; Rousset 1997). For IBD analysis we used two measures of geographic distance: (1) straight-line distance between sites, and (2) distance between sites along valley bottoms, which reflects the more likely flight route of individuals. The straight-line distance between all pairs of sites was calculated based on their geographical coordinates using GenAlEx. To estimate the pairwise distance between sites, constraining

movement of individuals to the Similkameen valley (minimizing movement over topographical barriers) we used the ruler function in Google Earth v. 5.1 (Google Inc.).

3.3 Results

3.3.1 AFLP analysis and phenotype scoring

Out of a starting sample size of 479, the AFLP phenotypes of 467 individuals were successfully scored. For six individuals, selective amplification appeared to fail; thus these individuals were removed from further analyses. As well, we excluded six individuals exhibiting many unique peaks not present in any other individuals in our study. Based on the 100-rfu scoring criterion, the five selective primer combinations generated a total of 484 AFLP loci with a mean mismatch error rate of 1.55% (dataset 100-A; Table 3.2). In comparison, the 75-rfu scoring criterion produced 526 loci with a mean mismatch error rate of 1.59% (dataset 75-A; Table C.1). When singletons were removed from both datasets, 326 (dataset 100-B) and 484 (dataset 75-B) loci were retained.

A comparison of all statistical analyses performed using the four datasets is provided in Appendix C. Briefly, we found that the estimates of genetic diversity were slightly higher for the 100-rfu datasets compared to the 75-rfu datasets, and for B datasets compared to A datasets (Table C.2). All population genetic structure analyses produced similar results (Table C.3-C.6). However, for the principal coordinates analysis more variation was explained by the first two coordinates for the 100-B dataset than the other datasets (Table C.5). As well, in the IBD analysis both 100-rfu datasets possessed higher

r^2 values and lower P values than their 75-rfu counterparts, while both B datasets possessed higher r^2 values but identical P values when compared to their A counterparts (Table C.6). Given its apparently higher explanatory power in these analyses, all further results reported are based on the 100-B dataset.

3.3.2 Genetic diversity

In general, levels of polymorphism were low within all BC sub-populations (PPL range = 0.172 – 0.236; H_e range = 0.068 – 0.090; Table 3.3). The proportion of polymorphism did not differ greatly among sites, with the most polymorphic site differing from the least polymorphic site by 6.4 % (or 21 loci). Similarly, expected heterozygosity was low and of similar value across all sub-populations (greatest difference between sites: 2.2 % points; Table 3.3).

Although not significantly different from zero, estimates of mean individual inbreeding within sub-populations, measured as average F_{AFLP} /sub-population, appeared higher in the two eastern sites (E1, E2) and the most western site (W1), than in sites W4, W5, W6 and W7 (Table 3.3).

3.3.3 Population genetic structure

Total gene diversity of the BC population was estimated at 0.082, with average gene diversity within and among sub-populations estimated as 0.079 (\pm 0.002 SE) and 0.004 (\pm 0.0004 SE), respectively. The estimate of differentiation among sub-populations (F_{ST}) was found to be statistically significant (0.043 ± 0.105 SE, $P < 0.00001$).

BARRIER identified four regions of high genetic differentiation within the BC population that exhibited high bootstrap support (bootstrap support >50 %; Figure 3.1) when four or more barriers were defined (Table C.4). The first and strongest barrier to be identified separated the western and northern sites from the central and eastern sites. The second barrier separated the most western site (W1) from the rest of the population, suggesting that this site is genetically isolated. The third barrier separated the central and eastern sites indicating that despite being geographically proximate, the eastern sites are genetically distinct from the central sites. Finally, the fourth barrier separated the second most western (W2) site from the rest of the population suggesting that, similar to W1, it is also genetically isolated. When we removed the most geographically isolated sites (W1, N1 and N2) from the barrier detection analysis the same barriers, in the same order, were detected as before (excluding the second barrier which separated W1).

Principal coordinates analysis clustered sites into four main groups (Figure 3.2). These clusters are consistent with results of pairwise F_{ST} comparisons between sites (results not shown) as well as the groups identified by BARRIER analysis. Together, the first two coordinates explained 69.54 % of the variation present within the data set.

We found a significant isolation by distance relationship among the BC sub-populations, and this relationship was stronger when we used geographic distance along valleys ($r^2 = 0.247$, $P < 0.0001$; Figure 3.3b) compared to Euclidean geographic distance ($r^2 = 0.182$, $P = 0.002$; Figure 3.3a).

Table 3.2 Summary of the AFLP phenotype scoring results for all selective primer combinations. Mismatch error rates generated by setting both locus and phenotype thresholds to 100rfu.

Selective primer combination	Mismatch error rate%	Initial number of loci	Number of loci retained
<i>EcoRI</i> -AAC/ <i>MseI</i> -CAC	1.43	112	94
<i>EcoRI</i> -AGC/ <i>MseI</i> -CAC	1.74	107	93
<i>EcoRI</i> -AGC/ <i>MseI</i> -CAT	1.68	135	112
<i>EcoRI</i> -AGC/ <i>MseI</i> -CTT	1.28	109	80
<i>EcoRI</i> -AAG/ <i>MseI</i> -CAA	1.63	129	109
	(Mean 1.55)	(Total 592)	(Total 488)

Table 3.3 Within sub-population and global genetic diversity for the British Columbia population of *Apodemia mormo*. Global measures are reported as mean values.

Site	N	PPL	H _e (±SE)	F _{AFLP} (±SE) **
W1	45	0.224	0.088 (0.008)	0.477 (0.105)
W2	39	0.221	0.077 (0.008)	0.197 (0.079)
W3	19*	0.221	0.068 (0.008)	0.118 (0.128)
W4	38	0.224	0.072 (0.007)	-0.079 (0.077)
W5	44	0.181	0.071 (0.008)	0.074 (0.074)
W6	41	0.181	0.068 (0.007)	-0.048 (0.083)
W7	15*	0.227	0.083 (0.008)	0.017 (0.130)
W8	38	0.230	0.082(0.008)	0.196 (0.107)
C1	38	0.218	0.070 (0.007)	0.098 (0.066)
C2	40	0.172	0.070 (0.008)	0.143 (0.077)
E1	40	0.236	0.090 (0.008)	0.511 (0.140)
E2	29	0.227	0.090 (0.008)	0.435 (0.120)
N1	28	0.209	0.085 (0.008)	0.290 (0.124)
N2	13*	0.227	0.087 (0.008)	0.263 (0.195)
All	467	0.214	0.079 (0.008)	

N, number of analysed samples; PPL, proportion of polymorphic loci; H_e, expected heterozygosity; F_{AFLP}, individual inbreeding coefficient, averaged across individuals per sub-population.

* The minimum recommended sample size for detecting polymorphism using a 5% criterion is 20 individuals, thus caution should be employed when interpreting the results for these sub-populations.

**Average of 20 runs, each consisting of 250 loci which were randomly selected from a possible 326 loci using a random number generator in Excel. Standard error values were calculated separately for each run and then averaged.

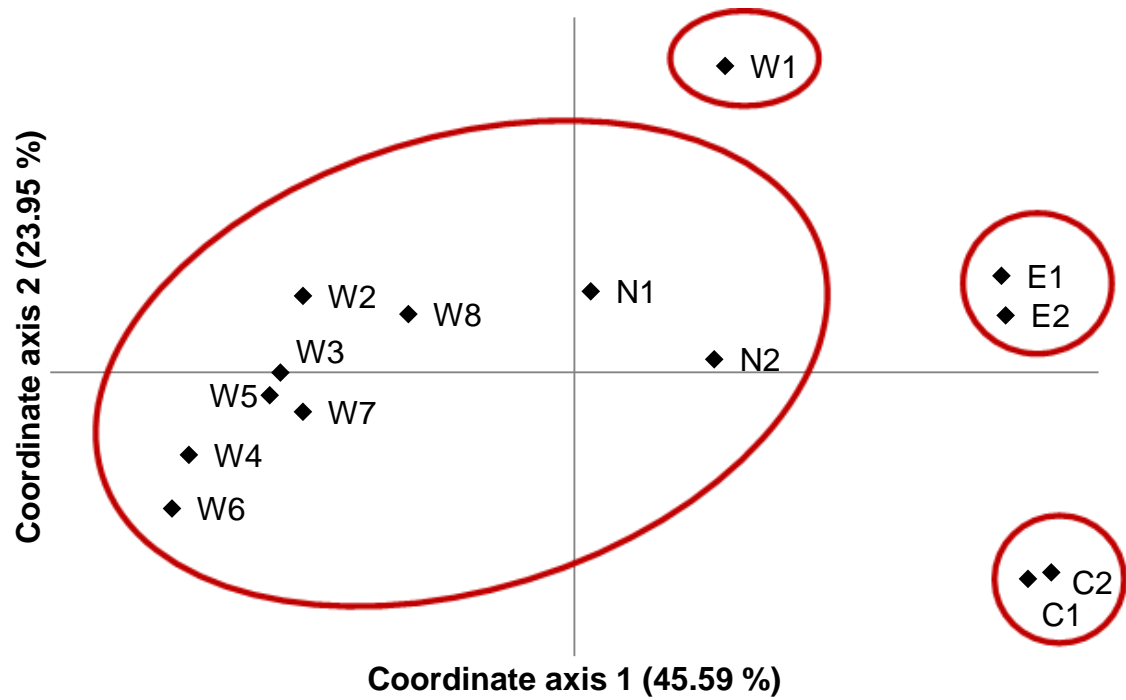


Figure 3.2 Principal coordinate analysis (PCoA) of all 14 sub-populations of *Apodemia mormo* based on a genetic distance matrix (F_{ST}). The genetically differentiated regions previously identified by BARRIER are indicated by red circles.

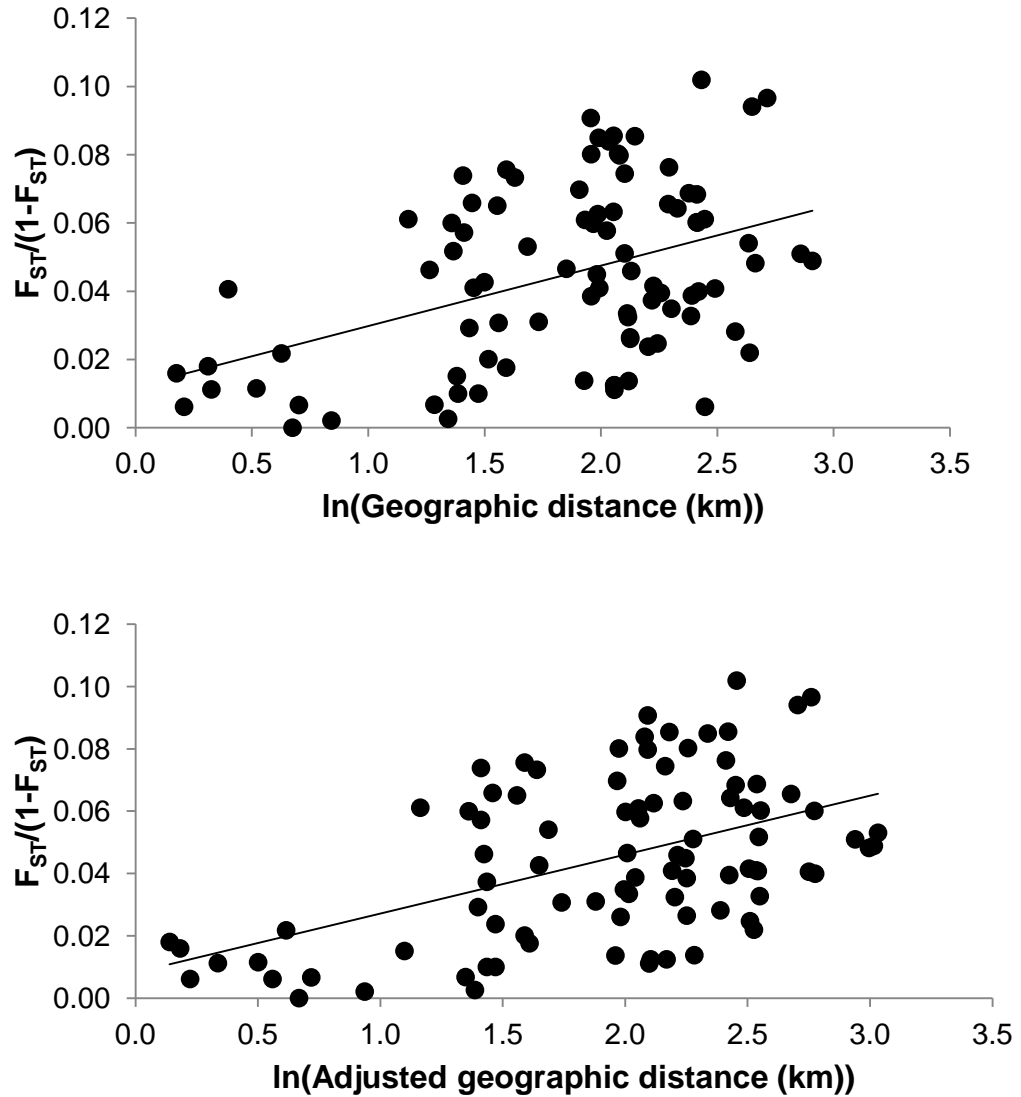


Figure 3.3 Relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (ln transformed) among all studied sub-populations.

Geographic distance was measured as (a) straight line distance between sites, or (b) 'adjusted' distance between sites (based on constraining movement to valley bottoms). Each point represents a pair of sub-populations. The linear regression lines are shown to indicate the underlying trends ($r^2 = 0.182$; 0.247 respectively).

3.4 Discussion

3.4.1 AFLP analysis and phenotype scoring

We successfully applied AFLPs generated from non-lethally sampled wing tissue to examine the genetic structure and diversity of an endangered butterfly population. Our study is among only a few to apply AFLPs to an endangered butterfly (Gompert et al. 2006; Leidner and Haddad 2010) and is the first to do so using non-lethal tissue samples. Given the difficulties often associated with developing and scoring microsatellite markers for many Lepidoptera species (Zhang 2004; Megléc et al. 2007) our results indicate that AFLPs represent a suitable alternative molecular marker for conducting conservation genetic studies of butterflies.

Our final method of scoring and phenotyping AFLP variation is a conservative one compared to most other AFLP studies (Table 3.1). We removed questionable peaks contributing to phenotype error rate, used scoring parameters to select loci (100-500bp, 100rfu) and call phenotypes (100 rfu), and removed singleton loci from the final AFLP dataset. Our stringent approach to locus selection resulted in a lower number of loci retained for genetic analyses, but led to lower error rates and ultimately, a more powerful dataset for inferring population structure. In comparing four datasets generated using different threshold values and in the inclusion or exclusion of singleton loci, we found that the most stringent dataset (thresholds of 100rfu and singleton loci removed), despite retaining the fewest number of loci, produced the strongest signals of population structure as indicated by lower variability in estimates, higher r^2 values and higher levels of statistical significance (Appendix C). In combination with a very low mismatch error rate, these results indicate that the loci retained under the more conservative selection

criteria were highly repeatable and informative, generating the dataset with the highest ‘signal to noise’ ratio. Given these results, it is clear that increasing the number of loci in AFLP datasets by lowering selection stringency does not necessarily compensate for the decreased reliability of the additional loci (e.g., Herrmann et al. 2010). To obtain a sufficient number of AFLP loci, we strongly recommend the maintenance of more conservative locus selection criteria coupled with the use of additional selective primer combinations rather than applying loose selection criteria across a small number of selective primer sets (as discussed in Whitlock et al. 2008). Furthermore, as indicated in Table 3.1, locus selection criteria are often not reported in many studies using AFLPs, nor are error rates. To enable comparisons among studies, as well as adequate evaluation of the data, it is therefore equally important that details of the analysis methods, and associated error rates, be provided.

3.4.2 Population genetic structure of the Mormon metalmark

Despite the small geographic scale of the studied region, our analyses indicate that a high degree of spatial genetic structure exists among the BC Mormon metalmark sub-populations. Our results thus suggest that many sub-populations may be experiencing low levels of gene flow and/or high levels of genetic drift. Similar findings have previously been reported for other endangered butterfly species inhabiting fragmented landscapes. For example, Vandewoestijne and Baguette (2004) observed high genetic differentiation in a vulnerable Belgian population of the Bog fritillary using RAPD (randomly amplified polymorphic DNA) markers. Limited dispersal and gene flow among populations, particularly among the most isolated populations that showed the greatest differentiation,

were thought to have contributed to population sub-structuring. Similarly, Sigaard et al. (2008) attributed the patterns of restricted gene flow observed among isolated Danish populations of the Marsh fritillary, detected using microsatellites, to increasing habitat fragmentation. As well, a very isolated population of the Regal fritillary, typically considered to be a high gene flow species, was shown to exhibit high genetic differentiation and restricted gene flow at a small spatial scale (<10km) using microsatellite and mitochondrial markers (Keyghobadi et al. 2006). Restricted gene flow in the Mormon metalmark may in part be due to the natural patchiness of the host-plant distribution, but habitat fragmentation has most likely been further intensified as a result of agriculture and urban development.

Overall, sub-populations of the Mormon metalmark displayed a significant pattern of isolation by distance, and PcoA analysis and pairwise F_{ST} comparisons also suggested that gene flow may be highest among sites that are geographically most proximate (e.g., W2-W8). The two northern sites (N) are an exception to this pattern however, as they appear to be more genetically similar to the western sites (W) than the central sites (C) despite being geographically closer to the latter. The northern sites may therefore not be as isolated as was previously thought. Although much of the landscape separating the two regions is currently dominated by agriculture, along the base of the mountain ranges there may be suitable, but not yet recognized, habitat patches through which dispersal can occur. A thorough survey of this region has not been performed as much of the land is privately owned.

BARRIER analysis identified four areas of high genetic differentiation (with high bootstrap support) within the BC population. The second and fourth barriers identified

(B, D) are likely a result of geographic distance isolating the two most western sites. The distances between these sites and all others exceed the suggested dispersal capability of the butterfly based on mark-recapture data (600 m estimated by Arnold and Powell 1983; 4 km estimated by S. Desjardins, unpublished), with very few suitable intervening habitat patches available. In contrast, the first and third barriers to be identified likely represent regions of high genetic differentiation that are not simply a result of geographic isolation. The first barrier to be identified (A) separated the central and eastern sites from the northern and western sites, despite the fact that the northern sites are geographically the most isolated. Furthermore, when we removed the northern sites from the BARRIER analyses we still observed the same barrier between the central and western regions, thus providing further evidence that the observed barrier is not an artefact of isolation by distance. As both central sites are located within the town of Kermeos, this strongly suggests that the town is acting as an important barrier preventing individuals from dispersing to the nearest western sites. While there is urban development separating the two central sites from one-another, C1 is a newly established site which occurs within a man-made gravel pit and is thought to have been colonized by individuals from the nearby C2 site (COSEWIC 2003). This recent colonization (approximately 5-11 years ago; COSEWIC 2003) may account for why these two sites appear genetically similar.

Urban areas have been shown to significantly hinder dispersal and gene flow in a variety of taxa including insects, amphibians, reptiles and birds (e.g., Noel 2007; Vandergast 2009; Delaney et al. 2010). In particular, several butterfly species inhabiting urban areas have been found to exhibit limited gene flow among local populations and reduced genetic diversity in comparison to populations inhabiting rural areas (Wood and

Pullin 2002; Takami et al. 2004). However, not all species are hindered by urbanization, and some may even thrive in such areas. Leidner and Haddad (2010) found that urbanization posed no physical barrier to dispersal and gene flow for the highly mobile Crystal skipper, and in fact, may have facilitated the movement of individuals over short distances.

The third barrier to be identified in our study (C) separated the central sites from the eastern sites. Geographically, these sites are not exceptionally distant. For example, 2.5 km separate C2 from E1. In comparison, a similar distance separates W4 from W5, which were not found to be genetically dissimilar from each other. Therefore, a barrier to gene flow other than simply distance also separates the central and eastern sub-populations, however the exact nature of that barrier is not apparent. The eastern sites are the only sites to be located on the south side of the Similkameen river, but are also separated from the nearest neighbouring sub-populations by a large amount of cultivated land and urban development. While agricultural and urban areas have been shown to limit butterfly dispersal (e.g., Dover 1991; Ries and Debinski 2001; Wood and Pullin 2002; Schtickzelle and Baguette 2003; Takami et al. 2004), the river itself may also be a factor. There are few examples where water bodies have been identified as limiting gene flow in butterfly species. However, Leidner and Haddad (2010) identified ocean inlets separating barrier islands to be barriers to gene flow in the Crystal skipper.

3.4.3 Genetic variation

Through comparison to other AFLP studies of Lepidoptera (Table 3.1), the BC population of the Mormon metalmark appears to exhibit low genetic diversity. However, differences in AFLP scoring methodology may influence these diversity estimates, making direct comparisons among species difficult. It should also be noted, that as the BC population is located at the northern boundary of the Mormon metalmark's distribution, low genetic diversity within this population may simply be a result of its peripheral location (as reviewed in Eckert et al. 2008).

While no estimates of mean individual inbreeding coefficient (F_{AFLP}) within sites were statistically different from zero, variation in their magnitude among sites suggests that some of the more geographically isolated sites (E1, E2 and W1) may be more homozygous than some of the more connected sites (W4-W7). This is concordant with our observation that the geographically isolated sites also appear to be genetically isolated. Likewise, the lowest inbreeding estimates were observed for the western sub-populations W4-W7, previously suggested to be the largest sub-populations (in both habitat availability and population numbers; COSEWIC 2003).

3.4.4 Future management and recommendations

As a host-plant specialist with limited dispersal capabilities, the survival of the BC population of the Mormon metalmark will be highly dependent on habitat availability. We found significant genetic structure within this population, indicating limitations to dispersal and potentially high levels of drift in some sub-populations. Thus, our results

underline the importance of maintaining habitat patches of sufficient size and connectivity for the persistence of this population.

Our results also suggest that small habitat patches may act as important stepping-stones between sub-populations. Specifically, such stepping-stones may be facilitating the connectivity of the geographically distant northern sub-populations to the western ones. Thus, ideally as many habitat patches as possible should be conserved, regardless of size.

Finally, we identified the town of Keremeos to be an important genetic barrier, which was not intuitive based on the geographic location of known sub-populations. Previous studies of various butterfly species inhabiting urban areas have reported habitat availability to be a critical factor in determining the likelihood of population persistence in such environments, particularly when dispersal ability is limited (Maes and Van Dyck 2001; Wood and Pullin 2002). Consequently, future management should include establishing stepping-stone habitat sites within the town of Keremeos which would facilitate movement of individuals within the town itself as well as from the west to the east side of the Similkameen valley.

3.5 Literature cited

- Arnold RA and Powell JA. 1983. *Apodemia mormo langei*, in: Ecological studies of six endangered butterflies (Lepidoptera, Lycaenidae): Island biogeography, patch dynamics and design of habitat preserves. University of California Press, Berkeley, CA.
- Bensch S and Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.
- Bonin A, Ehrich D and Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16: 3737-3758.
- Brattström O, Åkesson S and Bensch S. 2010. AFLP reveals cryptic population structure in migratory European red admirals (*Vanessa atalanta*). *Ecological Entomology*, 35: 248-252.
- Cannings S and Cannings R. 1995. Rare invertebrates of the south Okanagan. Province of British Columbia, Ministry of Environment, Lands and Parks. Available from: <http://www.env.gov.bc.ca>.
- Collier N, Gardner M, Adams M, McMahon CR, Benkendorff K and Mackay DA. 2010. Contemporary habitat loss reduces genetic diversity in an ecologically specialized butterfly. *Journal of Biogeography*, 37: 1277-1287.

- Committee on the Status of Endangered Wildlife in Canada (COSEWIC). 2003. COSEWIC assessment and updated status report on the Mormon Metalmark *Apodemia mormo* in Canada. Ottawa, Canada.
- Crawford LA, Kosciński D, Watt KM, McNeil JN and Keyghobadi N. 2013. Mating and oviposition success of a butterfly are not affected by non-lethal tissue sampling. *Journal of Insect Conservation*, doi: 10.1007/s10841-013-9566-8.
- Dasmahapatra KK, Lacy RC and Amos W. 2008. Estimating levels of inbreeding using AFLP markers. *Heredity*, 10: 286-295.
- Delaney KS, Riley SPD and Fisher RN. 2010. A rapid, strong, and convergent genetic response to urban habitat fragmentation in four divergent and widespread vertebrates. *PLOS ONE*, 5: e12767.
- Dover JW. 1991. The conservation of insects on arable farmland, in: The conservation of insects and their habitats, eds. Collins NM and Thomas JA. Academic Press, London.
- Eckert CG, Samis KE and Loughheed SC. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, 17: 1170-1188.
- Ehrich D. 2006. AFLPDAT: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*, 6: 603-604.
- Frankham R, Ballou JD and Briscoe DA. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

- Franklin MT, Ritland CE and Myers JH. 2010. Spatial and temporal changes in genetic structure of greenhouse and field populations of cabbage looper, *Trichoplusia ni*. *Molecular Ecology*, 19: 1122-1133.
- Gompert Z, Nice CC, Fordyce JA, Forister ML and Shapiro AM. 2006. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. *Molecular Ecology*, 15: 1759-1768.
- Guppy CS, Shepard JH and Kondla NG. 1994. Butterflies and skippers of conservation concern in British Columbia. *Canadian Field Naturalist*, 108: 31-40.
- Guppy CA and Shepard JH. 2001. Butterflies of British Columbia: Including Western Alberta, Southern Yukon, the Alaska Panhandle, Washington, Northern Oregon, Northern Idaho, and Northwestern Montana. University of British Columbia Press, Vancouver, Canada.
- Haig SM, Wagner RS, Forsman ED and Mullins TD. 2001. Geographic variation and genetic structure in spotted owls. *Conservation Genetics*, 2: 25-40.
- Hamm CA, Aggarwal D and Landis DA. 2010. Evaluating the impact of non-lethal DNA sampling on two butterflies, *Vanessa cardui* and *Satyrodes eurydice*. *Journal of Insect Conservation*, 14: 11-18.
- Herrmann D, Poncet BN, Manel S, Rioux D, Gielly L, Taberlet P and Gugerli F. 2010. Selection criteria for scoring amplified fragment length polymorphism (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*, 53: 302-310.

- Keyghobadi N, Unger KP, Weintraub JD and Fonseca DM. 2006. Remnant populations of the regal fritillary (*Speyeria idalia*) in Pennsylvania: local genetic structure in a high gene flow species. *Conservation Genetics*, **7**: 309-313.
- Keyghobadi N, Crawford, LA and Maxwell SA. 2009. Successful analysis of AFLPs from non-lethally sampled wing tissues in butterflies. *Conservation Genetics*, **10**: 2021-2024.
- Koscinski D, Crawford LA, Keller HA and Keyghobadi N. 2011. Effects of different methods of non-lethal tissue sampling on butterflies. *Ecological Entomology*, **36**: 301-308.
- Kronforst MR, Salazar C, Linares M and Gilbert LE. 2007. No genomic mosaicism in a putative hybrid butterfly species. *Proceedings of the Royal Society B: Biological Sciences*, **274**: 1255-1264.
- Krumm JT, Hunt TE, Skoda SR, Hein GL, Lee, DJ, Clark PL and Foster JE. 2008. Genetic variability of the European corn borer, *Ostrinia nubilalis*, suggests gene flow between populations in the Midwestern United States. *Journal of Insect Science*, **8**: 1-12.
- Layberry RA, Hall PW and Lafontaine JD. 1998. The butterflies of Canada. University of Toronto Press, Toronto.
- Leidner AK and Haddad NM. 2010. Natural, not urban, barriers define population structure for a coastal endemic butterfly. *Conservation Genetics*, **11**: 2311-2320.

- Lushai G, Fjelsted W, Marcovitch O, Aagaard K, Sherratt TN, Allen JA and Maclean N. 2000. Application of molecular techniques to non-lethal tissue samples of endangered butterfly populations (*Parnassius apollo* L.) in Norway for conservation management. *Biological Conservation*, 94: 43-50.
- Lynch M and Milligan BG. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, 3: 91-99.
- Madsen T, Shine R, Olsson M and Wittzell H. 1999. Restoration of an inbred adder population. *Nature*, 402: 34-35.
- Maes D and Van Dyck H. 2001. Butterfly diversity loss in Flanders (north Belgium): Europe's worst case scenario? *Biological Conservation*, 99: 263-276.
- Manni F, Guérard E and Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". *Human Biology*, 76: 173-190.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209-220.
- Meglécz E, Anderson SJ, Bourget D, Butcher R, Caldas A, Cassel-Lundhagen A, d'Acier AC, Dawson DA, Faure N, Fauvelot C, Franck P, Harper G, Keyghobadi N, Kluetsch C, Muthulakshmi M, Nagaraju J, Patt A, Péténian F, Silvain JF and Wilcock HR. 2007. Microsatellite flanking region similarities among different loci within insect species. *Insect Molecular Biology*, 16: 175-185.

- Millot E, Weimerskirch H, Duchesne P and Bernatchez L. 2007. Surviving with low genetic diversity: the case of albatrosses. *Proceedings of the Royal Society B: Biological Sciences*, 274: 779-787.
- Meudt HM and Clarke AC. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12: 106-117.
- Monmonier M. 1973. Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis*, 3: 245-261.
- New TR. 1995. An introduction to invertebrate conservation biology. Oxford University Press, New York.
- Noel S, Ouellet M, Galois P and Lapointe F. 2007. Impact of urban fragmentation on the genetic structure of eastern red-backed salamander. *Conservation Genetics*, 8:599-606.
- Peakall R and Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6: 288-295.
- Proshek B, Crawford LA, Davis CS, Desjardins S, Henderson AE and Sperling FAH. 2013. *Apodemia mormo* in Canada: Population genetic data support prior conservation ranking. *Journal of Insect Conservation*, 17: 155-170.
- Ralls K, Ballou JD, Rideout BA and Frankham R. 2000. Genetic management of chondrodystrophy in the California condor. *Animal Conservation*, 3: 145-153.
- Ries L and Debinski DM. 2001. Butterfly responses to habitat edges in the highly fragmented prairies of central Iowa. *Journal of Animal Ecology*, 70:840-852.

- Rousset F. 1997. Genetic differentiation and estimation of gene flow from F -statistics under isolation by distance. *Genetics*, 145: 1219-1228.
- Schroeder H and Degen B. 2008. Spatial genetic structure in populations of the green oak leaf roller, *Tortrix viridana* L. (Lepidoptera, Tortricidae). *European Journal of Forest Research*, 127: 447-453.
- Schtickzelle N and Baguette M. 2003. Behavioural responses to habitat patch boundaries restrict dispersal and generate emigration-patch area relationships in fragmented landscapes. *Journal of Animal Ecology*, 72: 533-545.
- Sigaard P, Pertoldi C, Madsen AB, Sogaard B and Loeschcke V. 2008. Patterns of genetic variation in isolated populations of the endangered butterfly *Euphydryas aurinia*. *Biological Journal of the Linnean Society*, 95: 677-687.
- Takami Y, Koshio C, Ishii MU, Fujii H, Hidaka T and Shimizu IM. 2004. Genetic diversity and structure of urban populations of Pieris butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, 13: 245-258.
- Timm AE, Geertsema H and Warnich L. 2006. Gene flow among *Cydia pomonella* (Lepidoptera: Tortricidae) geographic and host populations in South Africa. *Journal of Economic Entomology*, 99: 341-348.
- Timm AE, Geertsema H and Warnich L. 2008. Population genetic structure of *Grapholita molesta* (Lepidoptera: Tortricidae) in South Africa. *Annals of the Entomological Society of America*, 101: 197-203.

- Vandewoestijne S and Baguette M. 2004. Genetic population structure of the vulnerable bog fritillary butterfly. *Hereditas*, 141: 199-206.
- Vandergast AG, Lewallen EA, Deas J, Bohonak AJ, Weissman DB and Fisher RN. 2009. Loss of genetic connectivity and diversity in urban microreserves in a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus* n. sp. "santa monica") *Journal of Insect Conservation*, 13: 329-345.
- Vekemans X, Beauwens T, Lemaire M and Rodán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11: 139-151.
- Vila M, Auger-Rozenberg MA, Goussard F and Lopez-Vaamonde C. 2009. Effect of non-lethal sampling on life-history traits of the protected moth *Graellsia isabelae* (Lepidoptera: Saturniidae). *Ecological Entomology*, 34: 356-362.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407-4414.
- Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resource*, 8: 725-735.
- Wilson PJ, Grewal S, Lawford ID, Heal JNM, Granacki AG, Pennick D, Theberges MT, Voigt DR, Waddell W, Chambers RE, Paquet PC, Goulet G, Cluff D and White BN.

2000. DNA profiles of the eastern Canadian wolf and the red wolf provide evidence for a common evolutionary history independent of the gray wolf. *Canadian Journal of Zoology*, 78: 2156-2166.

Wood BC and Pullin AS. 2002. Persistence of species in a fragmented urban landscape: the importance of dispersal ability and habitat availability for grassland butterflies. *Biodiversity and Conservation*, 11:1451-1468.

Zhang DX. 2004. Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology and Evolution*, 19: 507-509.

Zhivotovsky L.1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8: 907-913.

Chapter 4

4 Local patch characteristics determine patterns of genetic diversity in a glacial relict, peatland specialist butterfly

4.1 Introduction

The conservation of populations and communities within fragmented landscapes is often focused on preserving large, well-connected areas of habitat (Fahrig 2001; Baguette 2004). Within a conservation genetics framework, this can maintain a high level of genetic variability since, in theory, smaller and more isolated habitat patches support small populations that are genetically impoverished due to reduced gene flow and increased genetic drift and inbreeding (Keyghobadi 2007). Maintenance of genetic diversity, in turn, is important as genetic variation is often correlated with individual fitness as well as population viability and persistence (e.g., Saccheri et al. 1998; Hansson and Westerberg 2002; Reed and Frankham 2003; Vandewoestijne et al. 2008; but see Reed 2010). However, habitat quality can also significantly influence population dynamics (Thomas et al. 2001; Baguette et al. 2011), and higher quality habitat patches generally support larger populations which are more genetically diverse (de Vere et al. 2009; Pitra et al. 2011). Thus, in addition to area and isolation, the overall quality of a habitat patch can be an important determinant of genetic variability and long-term population persistence.

Some habitats, such as glacial relict habitats (e.g., temperate peatlands) are naturally fragmented due to abiotic and/or biotic factors (Hampe and Jump 2011). The species that occur in these habitats are often highly specialized and adapted to specific environmental conditions (Habel and Schmitt 2012) which constrain their occurrence to isolated populations (Spitzer and Danks 2006; Hampe and Jump 2011). Given this geographic restriction, such specialist species usually experience reduced gene flow, which results in lower intrapopulation genetic diversity and increased interpopulation genetic differentiation (Habel and Schmitt 2012). Despite a lack of genetic exchange, highly isolated populations may nonetheless persist because of purging of deleterious alleles over multiple generations (Habel and Zachos 2012), making them less susceptible to the negative consequences of genetic bottlenecks or inbreeding (Crnokrak and Barrett 2002; Reed 2010). As biota inhabiting naturally highly fragmented habitats appear to be adapted to reduced dispersal and gene flow, genetic variation and population viability are predicted to be less sensitive to changes in the surrounding landscape, such as increased isolation and loss of potential habitat (Habel and Schmitt 2012; Habel and Zachos 2012). For these specialists, local patch characteristics (e.g., habitat patch size and quality) may be the more critical factors influencing genetic diversity and population persistence; however, to our knowledge few empirical tests of this prediction currently exist.

Across much of Europe and the northeastern United States, relict peatland habitats have become increasingly degraded and isolated due to human activities (Giberson and Hardwick 1999; Spitzer and Danks 2006; Savage et al. 2011). Consequently, many of the specialized plant and animal species endemic to these unique habitats are now of high conservation concern (van Sway et al. 2006; Turlure et al. 2009). Furthermore, peatland

ecosystems are sensitive to increases in atmospheric nitrogen deposition (Tomassen et al. 2003) as well as changes in precipitation and temperature regimes (Breeuwer et al. 2010). Although many plant and animal species have responded to current climate change by shifting their phenology or distribution ranges (Parmesan et al. 1999; Parmesan 2006), glacial relict species cannot easily disperse from sites that become unsuitable because they often have limited dispersal abilities and rely on resources that only occur under very specific environmental conditions (Hampe and Petit 2005; Franco et al. 2006; Goffart et al. 2010; Habel et al. 2011). Therefore, determining to what extent landscape structure versus local patch characteristics influence patterns of population genetic structure and diversity in peatland specialist species may allow us to better predict how they will respond to increasing habitat fragmentation and degradation. Such an understanding will also aid in assessing vulnerability of populations to environmental change and guide appropriate mitigation and management plans to conserve populations that are most at risk. Patterns of gene flow and population genetic structure have been described for several glacial relict species inhabiting western Europe (e.g., Vandewoestijne and Baguette 2004; Nève et al. 2008; Finger et al. 2009; Habel et al. 2010; Drees et al. 2011), while comparatively little research has been conducted for such species in North America.

In this study, we examined patterns of population genetic structure and diversity, and their relationships with local patch characteristics and surrounding landscape structure, in the bog copper (*Lycaena epixanthe*, Boisduval and Le Conte 1835), a glacial relict butterfly endemic to temperate Nearctic peatlands. We examined local populations within a landscape in central Ontario, Canada, where peatland habitat remains relatively

untouched by human activities. The objectives of our study are: (1) to establish whether local populations of the bog copper exhibit the genetic characteristics of a habitat specialist (i.e., low genetic variability within populations and high genetic differentiation among populations) and (2) to test the hypothesis that for a habitat specialist, differences in intrapopulation genetic diversity are better explained by local patch characteristics (area, quality) than by the surrounding landscape structure (isolation, composition). By addressing these questions in an undisturbed landscape, we aim to assess the dispersal potential and natural levels of gene flow among populations of the bog copper, as well as identify which landscape features may be most influential for the long-term persistence of this species. The results of this study will have important implications for the conservation of the bog copper as well as other relict habitat specialists.

4.2 Methods

4.2.1 Study species

The bog copper is a very small (wingspan: 17-22 mm), univoltine butterfly found in eastern North America. It is considered a relict species whose distribution was more widespread before the Pleistocene glaciation (Wright 1983). In Canada it ranges in a band from Riding Mountain National Park, Manitoba east to St. John's, Newfoundland, and it also occurs along the eastern United States as far south as Virginia (Layberry et al. 1998). Both the larvae and adults feed exclusively on cranberry plants (*Vaccinium macrocarpum* and *V. oxycoccos*) which typically grow in acid peatlands, and thus the species is considered a strict habitat specialist (Wright 1983; Cech and Tudor 2005). While

cranberries can grow in less saturated substrates, the bog copper is only known to occur in open wetland habitats with permanently wet, sunny substrates (Opler and Malikul 1992; Layberry et al. 1998). The adult's flight behaviour is slow and low to the ground, and butterflies typically occur in discrete colonies in which they are believed to carry out their entire lifecycle (Wright 1983; Cech and Tudor 2005). Within suitable habitat patches, this species is often locally abundant and can reach high population densities (e.g., mean relative abundance > 200 (individuals/hour), Swengel and Swengel 2011). Given the species' sedentary flight behaviour and strict habitat requirements it is expected to be a relatively poor disperser whose movement is constrained by the amount of surrounding forest and open water habitat. While this species' global status is *apparently secure* (G4, NatureServe 2013), in the United States bog coppers are listed as imperiled in Pennsylvania and critically imperiled in West Virginia and Maryland, primarily due to habitat loss and fragmentation (Cech and Tudor 2005).

4.2.2 Study area

The study was conducted within the southern region of Algonquin Provincial Park, Ontario, Canada (UTM: 17N 692550E 5049669N; Figure 4.1). At an elevation of 443 m above sea level, the climate of this region is generally cool and moist, with a mean annual temperature of 4 °C and a mean annual precipitation of 1182.8 mm (Environment Canada; Canadian Climate Normals 1971-2000; station Dwight, Ontario, <http://www.climate.weatheroffice.gc.ca>). Algonquin Provincial Park is situated within the Great Lakes-St. Lawrence forest region of Ontario, a transition zone between the southern deciduous forest of eastern North America and the northern coniferous boreal

forest. In this region coniferous tree species such as eastern red cedar (*Juniperus virginiana*), eastern hemlock (*Tsuga canadensis*), eastern white pine (*Pinus strobus*), and red pine (*P. resinosa*) occur amongst deciduous broad-leaved species such as yellow birch (*Betula alleghaniensis*), northern red oak (*Quercus rubra*), sugar maple (*Acer saccharum*), and red maple (*A. rubrum*). As well, several boreal species including white birch (*Betula papyrifera*), tamarack (*Larix laricina*), white spruce (*Picea glauca*), and black spruce (*P. mariana*) also occur. The landscape is predominated by forest (81%), freshwater lakes (10%) and wetland habitats (8%) including acid bogs and poor fens (collectively referred to as 'peatlands'). Peatlands are classified as permanently saturated wetlands exhibiting an accumulation of un-decomposed organic matter (peat). Water is contained within these habitats via either a high water table, an underlying impervious stratum or by climatic conditions that limit evaporation (Johnson 1985; Marshall et al. 1999). The rate of decomposition of plant material is very slow in these habitats due to a lack of inflowing nutrients, poor drainage and a buildup of acids (Gore 1983; Tiner 1999). Bogs are extremely acidic (e.g., pH < 4.2) and nutrient-poor peatlands, whose only water input is through precipitation (Mitsch et al. 2009). In contrast, fens are less acidic and nutrient-poor than bogs as they are also fed by slow-moving ground water (Gore 1983; Spitzer and Danks 2006). Depending on groundwater flow and chemistry, the pH of fens can range from mildly acidic 'poor fens' (e.g., pH = 4.1-5.8), to extremely alkaline, 'rich fens' (e.g., pH > 6.7; Mitsch et al. 2009). Bogs and poor fens are inhabited by a unique set of flora which are highly adapted to living under waterlogged and acidic conditions (Johnson 1985) including sphagnum mosses (*Sphagnum* spp.), low-growing heaths (e.g. bog laurel, *Kalmia polifolia*; cranberry, *Vaccinium* spp.; labrador tea, *Ledum*

groenlandicum; and leatherleaf, *Chamaedaphne calyculata*) and carnivorous plants (e.g. northern pitcher plant, *Sarracenia purpurea*; and sundews, *Drosera* spp.). Due to increased nutrient availability, poor fens can also support a wider array of tree and shrub species than bogs (Johnson 1985; Tiner 1999).

4.2.3 Sample collection

Tissue from adult butterflies was non-lethally sampled in 13 separate peatlands in July 2009 (Table 4.1, Figure 4.1). Within five of those peatlands we collected individuals from two discrete regions (Table 4.1). Because bog coppers are thought to be extremely sedentary, we were interested in whether individuals from different sampling locations within a peatland represented discrete colonies, and we initially conducted population genetic analyses for these sub-samples separately.

Individuals were captured using hand-nets, and a small piece of tissue or ‘wing-clip’ ($\sim 0.1 \text{ cm}^2$) was removed from both hind-wings using iris scissors and immediately stored in absolute ethanol. All butterflies were then marked, to prevent resampling, and released from the location of their initial capture. We have demonstrated previously for other butterfly species that small pieces of wing tissue can be sampled non-lethally without lowering individual survival, or affecting short-term flight behaviour, mating, and oviposition (Koscinski et al. 2011; Crawford et al. 2013).

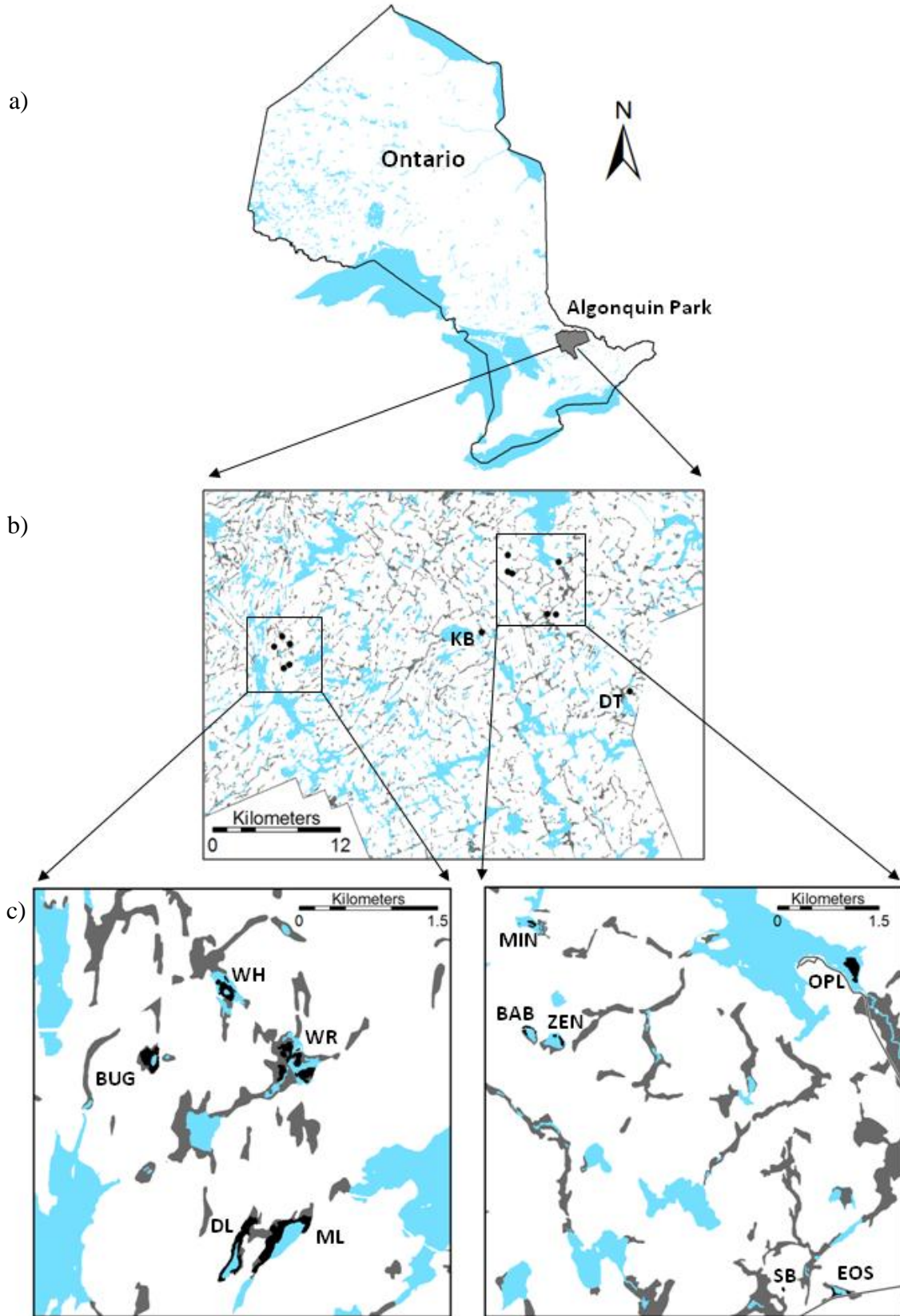


Figure 4.1 Map of the study system and sampling locations for *Lycaena epixanthe* outlining the location of Algonquin Provincial Park in Ontario, Canada (a), the study region within Algonquin Provincial Park (b) and the locations of the 13 studied populations (b, c). Forest, the dominant land cover is depicted in white, open water in blue, wetlands in dark gray and the studied peatlands in black.

Table 4.1 Collection record of the 13 studied populations of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada, with coordinates measured as the centroid of butterfly capture in each peatland or survey region, the number of individuals initially collected ($n_{\text{collected}}$) and the number of individuals successfully amplified and phenotyped for AFLP analyses (n_{AFLP}).

Code	Peatland	Region	UTM Coordinates (17N)		$n_{\text{collected}}$	n_{AFLP}
			Easting	Northing		
BUG	'Buggy'	-	679515.01	5049067.38	25	25
WH	Wolf Howl Pond	-	680263.51	5049909.77	41	41
DL	Dizzy Lake	a	680239.50	5046830.66	20	20
		b	680420.74	5047193.93	20	20
ML	Mizzy Lake	a	681141.86	5047379.28	19	18
		b	680892.62	5047327.68	24	24
WR	West Rose Lake	-	680935.89	5049256.81	38	38
KB	'Kearney'	-	698978.67	5050431.11	39	37
MIN	Minor Lake	a	701460.15	5057456.41	25	25
		b	701446.35	5057410.10	15	15
BAB	Bab Lake	-	701436.90	5055911.00	40	40
ZEN	Zenobia Lake	-	701883.80	5055742.06	39	38
SB	Spruce Bog	-	705183.73	5052048.05	16	16
EOS	Eos Lake	a	706133.08	5052006.34	11	11
		b	705978.98	5051977.83	29	29
OPL	Opeongo Lake	-	706267.58	5056808.38	41	41
DT	D. Thompson Lake	a	712915.21	5044885.79	20	20
		b	713048.59	5044791.02	20	19

Throughout tissue collection in each surveyed peatland, we kept a record of the total time required to capture all sampled individuals (Table D.1). This allowed us to determine the catch per unit effort (number of individuals / person-hour), a coarse index of population density, in each peatland. Weather conditions were similar across all sampling sites (i.e., we only captured butterflies on warm and sunny days).

4.2.4 DNA extraction and AFLP marker selection

Wing-clips were removed from ethanol, allowed to air dry for several minutes, and genomic DNA was extracted using the DNeasy® Blood and Tissue Kit (QIAGEN, Germantown, MD) following the manufacturer's protocol. The DNA was eluted in two volumes of 200 μ L (total = 400 μ L) to maximize total yield and DNA concentrations ranged from ~ 50-100 ng/ μ L.

Population genetic data were obtained using amplified fragment length polymorphism (AFLP) molecular markers, a multilocus genomic fingerprinting technique widely used in ecological and evolutionary research on non-model species (Bensch and Åkesson 2005). We have shown previously that small pieces of butterfly wing tissue provide sufficient quantities of DNA for the development of AFLPs (Keyghobadi et al. 2009; Crawford et al. 2011). AFLP profiles for the bog copper were generated following a protocol adapted from two standard fluorescent AFLP methods, that of Clarke and Meudt (2005) and the commercial AFLP Plant Mapping Kit (Applied Biosystems, Foster City, CA), the details of which are outlined in Appendix A. After initially screening 64 selective primer combinations (Table A.1), the following five combinations were selected based on relative polymorphism and reproducibility of fragments: *Eco*RI-ACA/*Mse*I-

CTT, *EcoRI*-AAC/*MseI*-CAC, *EcoRI*-AAC/*MseI*-CTC, *EcoRI*-AAG/*MseI*-CAA, and *EcoRI*-ACC/*MseI*-CAC. Negative controls were included in each step of the protocol to detect potential DNA contamination (i.e., Milli-Q purified water was used in place of DNA template, and restriction-ligation and PCR products). Fluorescently labelled AFLP fragments were separated and sized using a 3730S Genetic Analyzer (Applied Biosystems).

4.2.5 AFLP Analysis

AFLP fragment sizes and non-normalized peak heights were determined using GENEMAPPER v.4.0 (Applied Biosystems) based on comparison to a size standard ladder (LIZ-500) loaded with each sample. To reduce the incidence of size homoplasy (Vekemans et al. 2002) and artefactual peaks caused by instrument noise, we used GENEMAPPER to identify initially those AFLP loci (i.e., fragments that fall within user-defined size ‘bins’) between 100 and 500 base pairs (bp) in size, and above 100 relative fluorescence units (rfu) in height (see Chapter 3). We set bin width to one bp, and visually checked the position of all bins to ensure that they were accurately assigned by GENEMAPPER. Specifically, we adjusted bins if they were off-centre of a peak distribution, and removed bins if they contained fragments that overlapped with an adjacent bin. As well, we manually checked every AFLP profile to confirm successful amplification. If a sample failed to amplify or appeared to possess many unique fragments it was run through the AFLP protocol again using a new aliquot of the same DNA extraction, and if the amplification failed a second time then the sample was removed from analysis.

The AFLP peak-height data were then normalized and scored in AFLPSCORE v.1.3b (Whitlock et al. 2008) using the data filter and absolute phenotype-calling threshold settings. AFLPSCORE applies an objective approach to AFLP phenotype scoring and uses mismatch error rate analysis of replicate samples to ascertain optimal phenotype scoring parameters that both minimize mismatch error and maximize the number of loci retained. First a locus selection threshold is determined based on mean peak height, and only loci whose peak-heights are equal or above this threshold are retained. Then a phenotype-calling threshold is applied for each AFLP profile to score loci as either present or absent. We used 32 replicate individuals (~ 7 % of total sample size; see Chapter 2) to estimate mismatch error rate for each primer combination. All of the collected wing tissue was required for DNA extraction, thus replicate samples represent two aliquots of the same DNA extraction which were independently subjected to the entire AFLP protocol.

For each primer combination we determined the optimal scoring parameters by testing multiple combinations of locus and phenotype scoring thresholds (from 100-3000 and 100-1000 rfu, respectively) and we identified those thresholds which provided the maximal number of retained loci while maintaining a mismatch genotyping error rate of less than 5 % (acceptable error rates for AFLPs; Bonin et al. 2004). Using the appropriate threshold values we then created four increasingly inclusive datasets with decreasing stringency regarding mean mismatch error rates (< 5, < 4, < 3 and < 2 %) across all primer combinations, herein after referred to as datasets 5A, 4A, 3A and 2A. Estimates of homozygosity (following a 5% criterion) can be upwardly biased by unique or 'singleton' loci which may not represent true AFLP loci (e.g., as a result of incomplete

digestion; Milot et al. 2007). Thus, for each dataset we created a second more stringent dataset where we removed all loci that were present in only a single individual, here after referred to as datasets 5B, 4B 3B and 2B.

Using the eight described datasets (5A, 5B, 4A, 4B, 3A, 3B, 2A and 2B) we calculated estimates of population genetic structure and genetic variability (described below) and examined the trade-offs between reduced genotyping error and the number of retained loci on our ability to detect population genetic structure and variation in genetic diversity estimates among populations. Based on these comparisons we identified the optimal data set for our study system, which we subsequently used to perform all population genetic analyses.

4.2.6 Landscape evaluation

4.2.6.1 Surrounding Landscape Structure

We assessed landscape structure in the area surrounding each surveyed peatland, by evaluating a circular sector of 1 km radius surrounding the centroid of butterfly capture at each site. Given the bog copper's presumed limited dispersal capabilities, close association with its host-plant, and based on preliminary genetic analyses which suggested pronounced genetic differentiation among populations located less than 500 m apart, a 1 km radius represents a suitable scale at which we would expect the bog copper to be influenced by landscape structure. We confirmed this by initially testing for landscape effects on intrapopulation genetic diversity (details below) using radii of 100 m, 250 m, 500 m and 1 km. Indeed we found that the 1 km scale explained the most

variation in all genetic diversity measures among populations (results not shown), thus we present the results of the landscape analysis using data extracted from the 1km scale. Some surveyed peatlands were geographically proximate which resulted in considerable overlap of radii at larger spatial scales. Thus, to avoid pseudoreplication we did not test radii larger than 1km.

We used a digital raster-based land cover classification map (Ontario Ministry of Natural Resources, Forest Resources Inventory, 2005) with a grain size of 5 m² to determine, for the circular sector surrounding each peatland: the area proportion and an index of proximity of potentially suitable habitat (wetland), and the area proportion of forest and open water habitat (Table B.1). Based on field observations, we considered all classified wetland types in our study area (open muskeg, treed muskeg and brush and alder), as potentially suitable habitat for the bog copper. We assumed that all forest stands, regardless of species composition, would serve as significant barriers to dispersal for the bog copper. Thus we classified the land cover types in our study area generally as either *forests*, *wetlands*, *open water*, *rock* or *unclassified land*. The proximity index provides a measure of both the degree of isolation and the degree of fragmentation of a particular habitat type within a defined neighborhood or search radius (McGarigal et al. 2002). It was calculated as the sum of the area of all potentially suitable wetland habitat whose edges fall within a 500 m radius of the focal patch, divided by the squared distance between the focal patch and all other wetland patches. Thus, all else being equal, a patch within a neighbourhood containing more wetland habitat will have a larger mean proximity index. Likewise, a patch will have a larger mean proximity index when surrounding wetland habitat occurs as larger, more contiguous, and/or proximate patches

(McGarigal et al. 2002). Other studies have previously demonstrated open water to be a barrier to dispersal in butterflies (Leidner and Haddad 2010) and therefore we considered open water as an aspect of surrounding landscape structure predicted to be negatively related to genetic diversity. However, given the bog copper's reliance on a generally high water table, the amount of water in the surrounding landscape could also affect local patch quality, in which case it would be predicted to have a positive relationship with genetic diversity. All measures of landscape spatial structure were performed using ArcGIS v.10.0 (Esri, Redlands, CA) and FRAGSTATS v.4.0 (McGarigal et al. 2002).

4.2.6.2 Local Patch Characteristics

We determined the geographic area (ha) of each sampled peatland (Table B.1) using a combination of high resolution aerial imagery (Ontario Ministry of Natural Resources, Forest Resource Inventory 2006, 40 cm accuracy), a vector-based version of the same land cover classification map described above, and our own ground truthing.

Within each surveyed peatland we also estimated the abundance and distribution of host-plant available for the bog copper (Table B.1). In July 2010 we recorded cranberry cover (as a proportion) within 1 m² quadrats spaced 15 m apart along continuous transects. Transects were parallel, separated by 20 m and covered the length of each peatland. The total number of surveyed quadrats varied among peatlands in proportion to total area, and we recorded their locations using high-accuracy GPS (Trimble GeoHX). Mean host-plant density was calculated as the mean cranberry cover

for all quadrats in a peatland. We then calculated an index of total host-plant abundance for each peatland by multiplying mean host-plant density by total peatland area.

4.2.7 Data analysis

4.2.7.1 Population genetic structure

Unbiased estimates of population genetic structure based on estimated allele frequencies were calculated in AFLP-SURV v.1.0 (Vekemans et al. 2002). Allele frequencies are estimated using a Bayesian approach with a non-uniform prior distribution which is robust to minor departures from Hardy-Weinberg equilibrium (Zhivotovsky 1999). We first calculated pairwise F_{ST} values between all populations and sampling locations within populations ($N = 18$ sampling locations; Table 4.1). We evaluated the statistical significance of pairwise F_{ST} based on 10,000 random permutations, and only considered values greater than the 99th percentile of the null distribution to be statistically significant (Vekemans et al. 2002). Within DL, DT and EOS sub-samples were not significantly differentiated ($P > 0.05$, Table E.2), thus all individuals within each of these peatlands were considered to represent a single population and were pooled in subsequent analyses. Sub-samples from MIN and ML did however exhibit significant differentiation ($P < 0.01$, Table E.2) and to be conservative we considered these samples as separate sub-populations in subsequent analyses (i.e., $N = 15$).

Patterns of genetic differentiation among populations were examined using principal coordinates analysis (PCoA) of a covariance matrix of pairwise F_{ST} estimates in GenAlEx v.6.0 following the data standardization option (Peakall and Smouse 2006). As well, we assessed the degree of between-population relative to within-population genetic

variation using analysis of molecular variance (AMOVA) with significance of the variance components tested using 10,000 random permutations in the program Arlequin v.3.1 (Excoffier et al. 2005). Based on estimated allele frequencies we also used AFLP-SURV to calculate global F_{ST} .

We tested for patterns of isolation-by-distance (IBD) by plotting measures of genetic dissimilarity between pairs of populations (pairwise F_{ST}) against geographic distance, and determined the statistical significance of these correlations using the Mantel test (Mantel 1967) in GenAlEx based on 9,999 permutations. Both F_{ST} and geographic distance were linearly transformed ($F_{ST}/(1-F_{ST})$ and $\log_{10}(1 + \text{geographic distance in km})$, respectively). Geographic distances among populations were calculated as straight-line distances between the centroids of butterfly capture using ArcGIS.

4.2.7.2 Genetic diversity

Measures of genetic diversity based on estimated allele frequencies were determined for each population using AFLP-SURV. For each population, levels of genetic diversity were measured as the proportion of loci that were polymorphic at the 5% level (PPL), and as unbiased estimates of expected heterozygosity (H_e) or Nei's genetic diversity (following Lynch and Milligan 1994). For each population we also report the mean number of loci (n_m), the number of private loci or loci unique to a single population (n_p), and the number of fixed loci (n_f) as determined by FAMDV.1.25 (Schlüter and Harris 2006).

4.2.7.3 Landscape effects on genetic diversity

We used multiple linear regressions to investigate the effects of surrounding landscape structure (proportion of wetland, forest, and water habitat, and mean proximity of wetland habitat) and local patch characteristics (peatland size, mean host-plant density and host-plant abundance) on intrapopulation genetic diversity measures (PPL, H_e). Because we found evidence of very limited gene flow, even for nearby peatlands (see Results), each peatland could be treated as an independent replicate with respect to genetic diversity measures. Prior to analysis, proportion variables (PPL, H_e , % cover) were logit transformed (Warton and Hui 2011) and all predictors centered as is recommended for multiple regressions (Aiken and West 1991). We then confirmed normality for all variables (Kolmogorov-Smirnov tests; all $P > 0.05$) and evaluated collinearity among predictors based on pairwise Pearson's correlation coefficient values. Area and host-plant predictor variables were highly correlated (Pearson's correlation; all $r > 0.75$, all $P < 0.05$), as was wetland and forest cover (Pearson's correlation; all $r > 0.75$, all $P < 0.05$). Thus we retained the four predictors which did not exhibit significant collinearity (i.e., peatland area, proportion of forest cover, proportion of open water and mean proximity of wetland habitat) for analyses. Starting from a full-model (all main factors and two-way interaction terms), we removed non-significant terms with P -values above 0.10 following a stepwise mixed selection method. To avoid pseudoreplication of the two peatlands with differentiated sub-populations (MIN and ML) the analysis only included the sub-population with the larger sample from each of those peatlands (i.e. MINa and MLb).

To assess how host-plant availability varies with peatland patch size, we used linear regressions to examine the relationships between patch size and mean cranberry density, the coefficient of variation in cranberry density, and the index of total cranberry abundance. We assessed variability in host-plant density within each peatland by calculating the coefficient of variation in cranberry density among all quadrats. High coefficient of variation values indicate a heterogeneous range of cranberry densities, whereas low values indicate that cranberry density is fairly homogenous. We also used linear regression to explore whether variation in our estimates of population density among surveyed peatlands could be explained by variation in cranberry density. All statistical analyses were performed using IBM SPSS v.20 (IBM Corp, New York) and JMP v.8 (SAS Institute Inc, Cary, NC).

4.3 Results

4.3.1 AFLP analysis and phenotype scoring

The AFLP phenotypes of 477 individuals (of 482 sampled) were successfully scored (Table 4.1). The selective amplifications of one individual failed, and that individual was removed from further analyses. Five other individuals were excluded because they possessed many unique peaks not present in any other individuals in our study. Using different phenotype scoring criteria which allowed for varying levels of mean mismatch error, the number of AFLP loci generated by the five selective primer combinations ranged from 126-319 (Table E.3). Overall, increasing genotyping error tolerated resulted in a greater number of loci to be retained. After removing singleton loci from each dataset, the final number of loci retained was reduced to 112-294 (Table E.3).

A comparison of all statistical analyses performed using the eight datasets is provided in Appendix B (Tables E.4-E.7). Briefly, for the estimates of overall differentiation (global F_{ST} ; Table E.4) and the proportion of molecular variance among populations (AMOVA; Table E.5) we observed very little effect of increasing mismatch error rate, and observed only limited differences between the original datasets (A) and the datasets from which singletons were omitted (B). As well, comparison of the PCoA results revealed that the variation explained by the first two coordinates was not greatly improved by increasing genotyping error tolerance, nor by removing singletons (Table E.6). However, in the IBD analysis, increasing mismatch error tolerance resulted in higher r^2 and lower P -values (Table E.7). We found that mean estimates of genetic diversity (PPL, H_e) remained relatively unchanged, but that the range of values estimated for each population generally decreased with increasing genotyping tolerance. As well, the mean genetic diversity estimates and the range of population values generated by Dataset B were greater than those of Dataset A (Table E.4). Overall, the singleton datasets with moderate levels (< 3- 4 %) of genotyping error appear to be most informative for genetic studies of the surveyed bog copper populations, as in comparison to the other datasets, they provided both good explanatory power to detect population structure and an ability to resolve variation in genetic diversity among populations. Thus, we report the results of all further analyses based on the 4B dataset which has a mean mismatch error of 3.62% (Table 4.2) and after removing singletons consists of 190 loci (Table E.3).

Table 4.2 Summary of the AFLP phenotype scoring results for *Lycaena epixanthe* for all selective primer combinations following a < 4% mismatch error rate criterion.

Selective primer combination	Scoring Threshold (rfu)		Mismatch error rate %	Initial no. of loci	No. of loci retained
	Locus	Phenotype			
<i>EcoRI-ACA/MseI-CTT</i>	900	300	3.92	161	37
<i>EcoRI-AAC/MseI-CAC</i>	2000	500	3.98	177	34
<i>EcoRI-AAC/MseI-CTC</i>	900	300	3.09	122	49
<i>EcoRI-AAG/MseI-CAA</i>	500	100	3.23	123	62
<i>EcoRI-ACC/MseI-CAC</i>	2500	1200	3.89	177	42
			3.62 (mean)	760 (total)	224 (total)

4.3.2 Population structure

All pairwise F_{ST} values, except for that between EOS and OPL, suggested significant genetic differentiation between populations ($P < 0.01$, Table 3.3). Overall estimated differentiation among populations (global F_{ST}) was statistically significant (0.095 ± 0.175 SE, $P < 0.001$), and AMOVA results indicated a relatively large proportion of variance was partitioned among populations (% of variance among and within populations: 9.11%, 90.89 %, respectively, $P < 0.001$).

The first two coordinates of the PCoA explained 68.57 % of the variation present within the dataset (Figure 4.2). Consistent with the pairwise F_{ST} comparisons, the PCoA indicated no obvious spatial pattern of population structure. Interestingly, the separation of populations along the first axis, which explains more than half of the variation (51.54 %), appears to correspond to levels of genetic diversity (see below) rather than to geographic location. For example, DT and BAB, the peatlands with the highest genetic diversity estimates (Table 4.4), were closely grouped together despite being separated by over 16 km.

Across all Algonquin populations, we found a significant pattern of IBD ($r^2=0.182$, $P < 0.001$; Figure 4.3a). Given the high genetic differentiation indicated by the F_{ST} and AMOVA analyses we also tested for IBD among nearby populations (i.e., separated by less than 7 km) located within the western and eastern halves of the study area (western: BUG, DL, ML, WH, WR; eastern: BAB, EOS, MIN, OPL, SB, ZEN; Figure 4.1). Within each of these regions we did not find a significant relationship (western sites: $r^2 = 0.057$, $P = 0.191$; eastern sites: $r^2 = 0.039$, $P = 0.158$; Figure 4.3b, c).

4.3.3 Genetic diversity

Levels of genetic diversity were relatively low within the Algonquin population of the bog copper in comparison to other butterfly species (see Table 3.1); however there was considerable variation in estimates of diversity among populations (Table 3.4).

Specifically, n_f , the number of fixed loci, ranged from 3 to 73; PPL, the proportion of polymorphic loci varied from 0.237 to 0.632; H_e , expected heterozygosity, ranged from 0.100 to 0.287. All estimates of genetic diversity were highly correlated with one another (Pearson's $r > 0.90$, $P < 0.05$ in all cases).

4.3.4 Effects of landscape

For both dependent variables (PPL, H_e), multiple linear regression results indicated that all land cover composition predictors (% forest and open water) and the index of proximity of wetland habitat had no influence on genetic diversity estimates (Table 4.5).

There was a strong negative relationship between genetic diversity and total peatland patch size; however, the predictor was only statistically significant in the model which included % water cover (Table 4.5). Examining the relationship between total patch size and genetic diversity, site ML appeared to be an outlier that deviated considerably from the overall trend. By removing this site, the explanatory power of all predictor variables increased, and a significant proportion of the variation in genetic diversity estimates among the studied sites could be explained by both the total peatland patch size (negative relationship) and the proportion of open water in the surrounding landscape (positive relationship; Table 4.5). Mean proximity and the proportion of forest habitat still had no significant effect on genetic diversity estimates.

Among our studied sites, mean cranberry density was negatively related to peatland patch size ($r^2 = 0.481$, $P = 0.009$; Figure 4.4a), while overall cranberry abundance and the coefficient of variation in mean cranberry density were positively related to peatland size ($r^2 = 0.526$, $P = 0.005$, and $r^2 = 0.484$, $P = 0.008$, respectively; Figure 4.4b, c). Estimates of population density of the bog copper in each peatland, derived from catch by unit effort, were positively associated with mean cranberry density ($r^2 = 0.142$, $P = 0.205$; Figure 4.5a) and negatively associated with total peatland size ($r^2 = 0.024$, $P = 0.614$; Figure 4.5b), although these relationships were not significant. Two populations, SB and WR, were clear outliers to this relationship, and their removal resulted in a highly significant positive linear relationship with mean cranberry density ($r^2 = 0.897$, $P < 0.001$). The explanatory power of peatland size on population density also improved considerably, but the relationship was still not significant ($r^2 = 0.256$, $P = 0.113$).

Table 4.3 Pairwise F_{ST} values for the surveyed populations and sub-populations (N = 15) of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada.

	BUG	WH	DL	MLa	MLb	WR	KB	MINa	MINb	BAB	ZEN	SB	EOS	OPL	DT
BUG	-														
WH	0.083	-													
DL	0.093	0.074	-												
MLa	0.109	0.066	0.101	-											
MLb	0.098	0.049	0.094	0.049	-										
WR	0.101	0.044	0.096	0.078	0.072	-									
KB	0.156	0.103	0.179	0.106	0.055	0.109	-								
MINa	0.171	0.111	0.188	0.120	0.064	0.114	0.038	-							
MINb	0.140	0.107	0.181	0.107	0.070	0.085	0.073	0.034	-						
BAB	0.200	0.124	0.198	0.122	0.071	0.162	0.034	0.062	0.118	-					
ZEN	0.093	0.049	0.111	0.080	0.045	0.077	0.064	0.058	0.073	0.065	-				
SB	0.118	0.056	0.116	0.051	0.040	0.085	0.054	0.072	0.095	0.071	0.042	-			
EOS	0.120	0.061	0.130	0.079	0.040	0.086	0.058	0.062	0.068	0.058	0.048	0.045	-		
OPL	0.146	0.087	0.160	0.107	0.064	0.110	0.070	0.063	0.073	0.068	0.055	0.060	0.021*	-	
DT	0.237	0.158	0.242	0.167	0.108	0.189	0.056	0.091	0.148	0.026	0.116	0.106	0.080	0.099	-

* populations that are not significantly differentiated, $\alpha = 0.01$ level

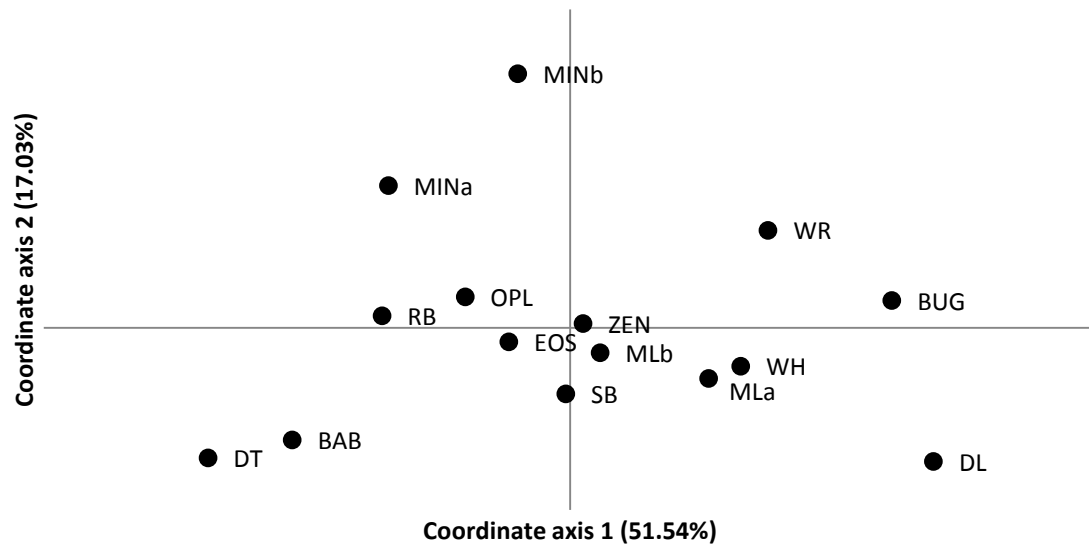
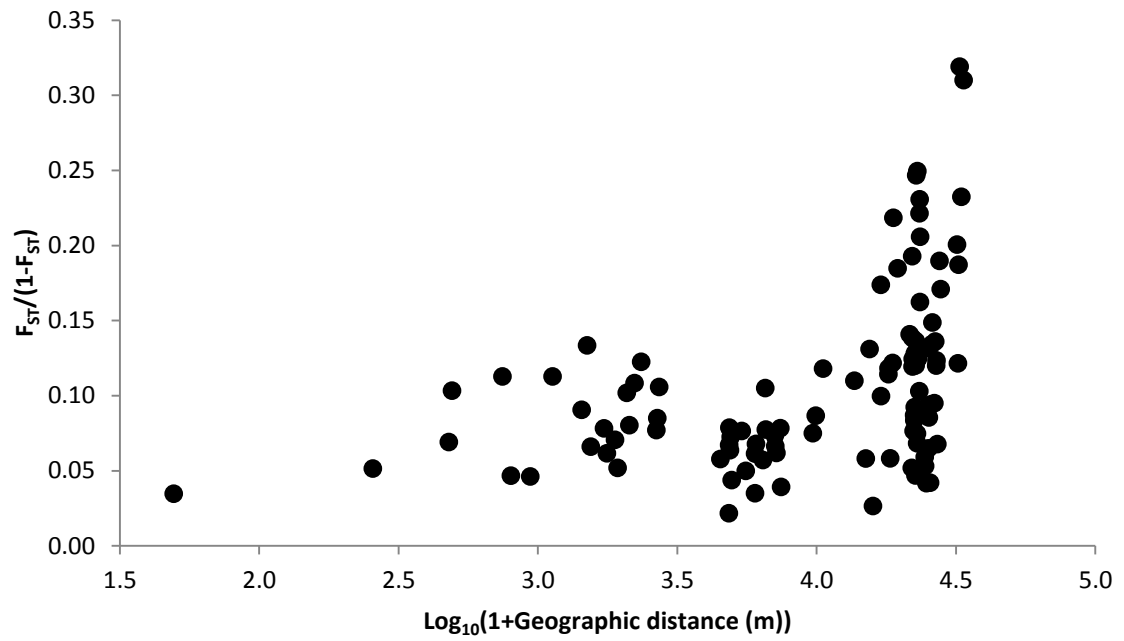
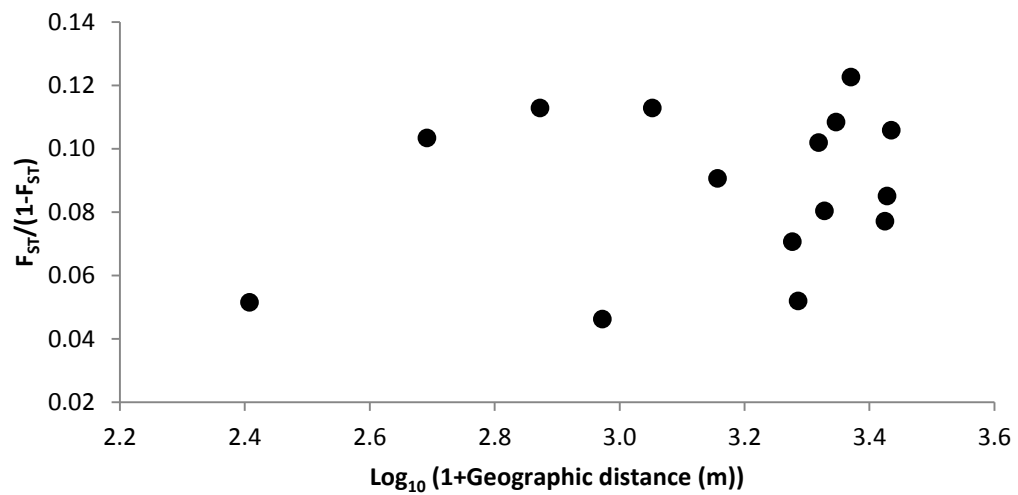


Figure 4.2 Principal coordinate analysis of all surveyed *Lycaena epixanthe* populations in Algonquin Provincial Park based on a covariance matrix of pairwise F_{ST} values.

a)



b)



c)

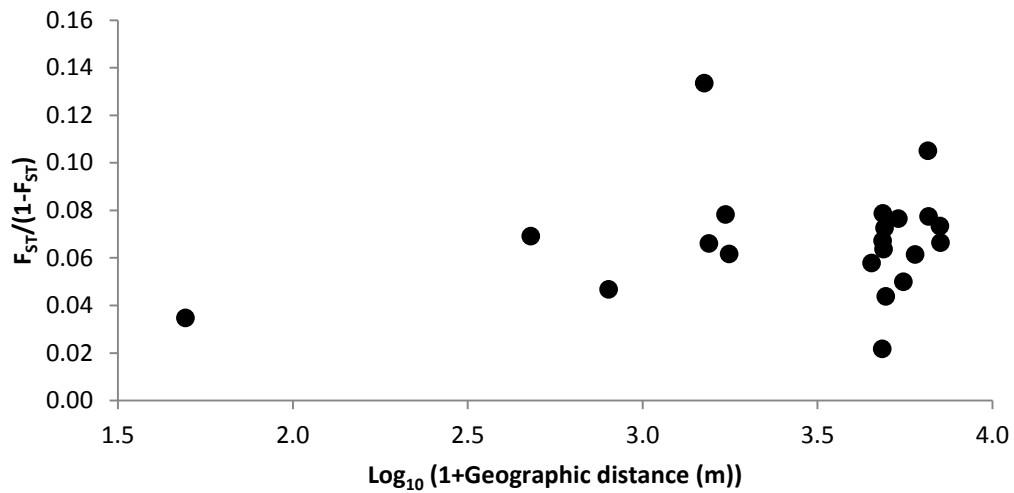


Figure 4.3 Relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (\log_{10} transformed) among a) all 15 studied populations and sub-populations, b) western populations (BUG, DL, MLa, MLb, WH and WR) and c) eastern populations (BAB, EOS, MINa, MINb, OPL, SB and ZEN).

Table 4.4 Mean population genetic diversity for *Lycaena epixanthe* in Algonquin Provincial Park, Canada. For peatlands with differentiated sub-populations (ML and MIN), estimates are presented for each sub-population separately, and with all individuals pooled.

Peatland	Region	n_m (\pm SE)	n_f	n_p	PPL	H_e (\pm SE)
BUG		108.880 (0.821)	72	0	0.237	0.107 (0.011)
WH		107.707 (0.809)	45	0	0.389	0.152 (0.012)
DL		109.300 (0.473)	73	0	0.237	0.100 (0.011)
ML	a	107.500 (1.138)	64	1	0.305	0.155 (0.013)
	b*	100.958 (2.017)	41	0	0.389	0.184 (0.014)
	pooled	103.762 (1.318)	30	1	0.453	0.183 (0.013)
WR		107.842 (0.547)	55	0	0.321	0.130 (0.012)
KB		100.135 (1.682)	15	1	0.542	0.231 (0.014)
MIN	a*	101.440 (1.591)	31	0	0.453	0.211 (0.014)
	b	103.867(1.133)	59	0	0.316	0.165 (0.014)
	pooled	102.350 (1.087)	25	0	0.495	0.201 (0.014)
BAB		95.125 (1.941)	5	0	0.595	0.262 (0.015)
ZEN		106.026 (0.913)	28	1	0.479	0.190 (0.013)
SB		103.313 (1.632)	46	0	0.374	0.188 (0.014)
EOS		102.000 (1.077)	20	0	0.505	0.198 (0.014)
OPL		101.439 (1.080)	28	0	0.458	0.185 (0.013)
DT		93.154 (2.519)	3	0	0.632	0.287 (0.016)

*sub-population used in multiple linear regression analyses

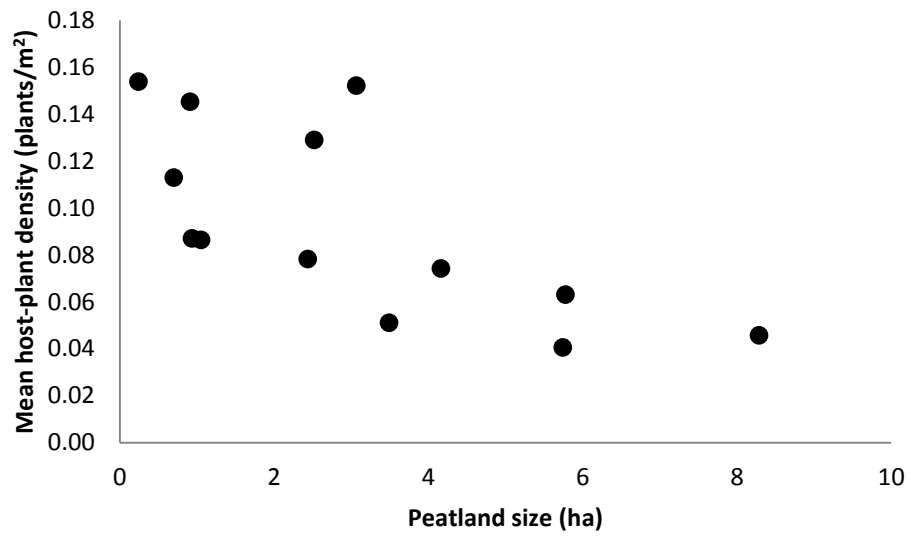
n_m , mean number of loci; n_f , number of fixed loci; n_p , number of private loci; PPL, proportion of polymorphic loci using a 5% criterion; H_e , expected heterozygosity.

Table 4.5 Effect of landscape and patch predictors on estimates of genetic diversity in *Lycaena epixanthe* for a) all 13 surveyed peatland sites, and b) for all sites except Mizzy Lake (ML).

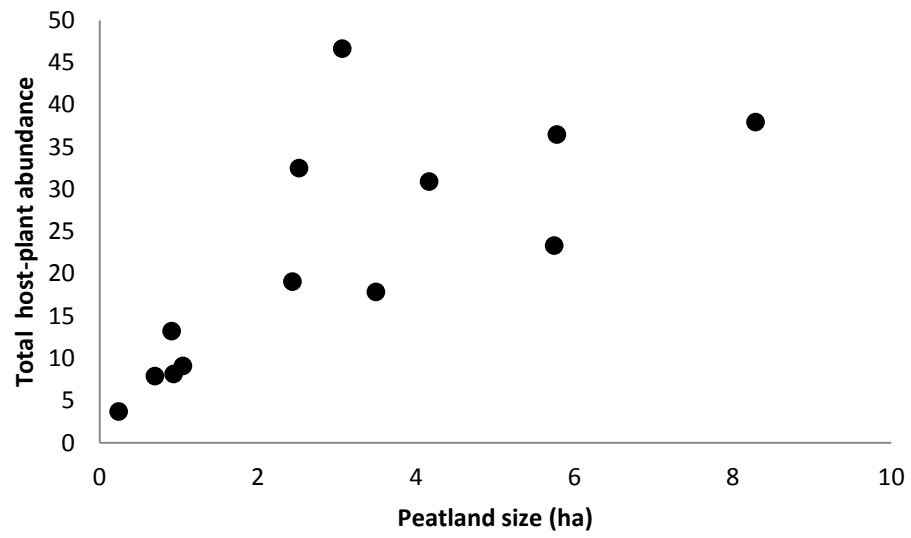
Model	adj. r^2	df error	df total	F	P-value	Variable	Coefficient (\pm SE)	t	P-value for t-test
a) All sites									
PPL	0.297	10	12	3.529	0.069	Intercept	0.441 (0.071)	6.17	<0.001
						Water	0.437 (0.223)	1.96	0.079
						TotArea	-0.075 (0.032)	-2.32	0.043
H_e	0.275	10	12	3.279	0.080	Intercept	-0.102 (0.031)	-3.33	0.008
						Water	-0.169 (0.097)	1.75	0.110
						TotArea	-0.032 (0.014)	-2.32	0.043
b) ML removed									
PPL	0.422	9	11	5.013	0.034	Intercept	0.400 (0.072)	5.55	<0.001
						Water	0.524 (0.217)	2.41	0.039
						TotArea	-0.118 (0.041)	-2.85	0.019
H_e	0.611	9	11	9.641	0.006	Intercept	-0.129 (0.025)	-5.09	<0.001
						Water	0.226 (0.076)	2.96	0.016
						TotArea	-0.060 (0.015)	-4.16	0.003

PPL, proportion of polymorphic loci; H_e , expected heterozygosity

a)



b)



c)

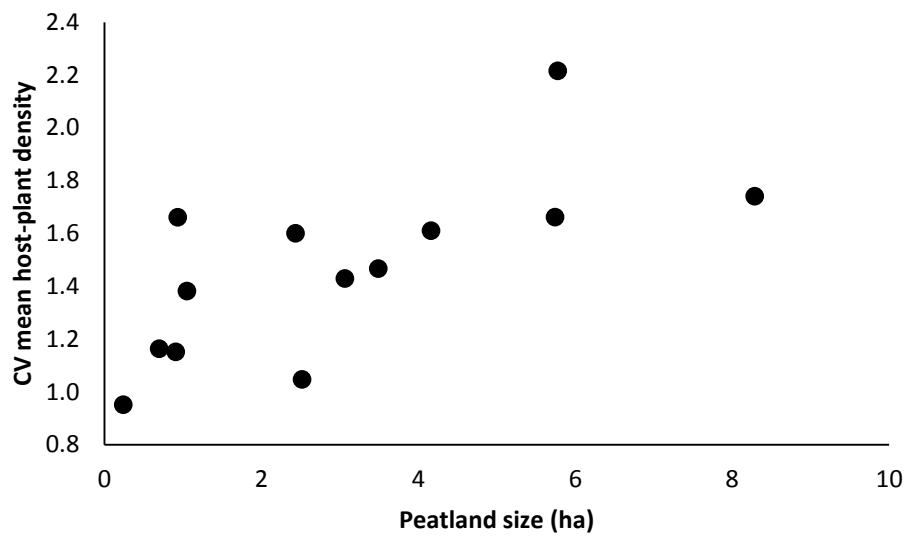
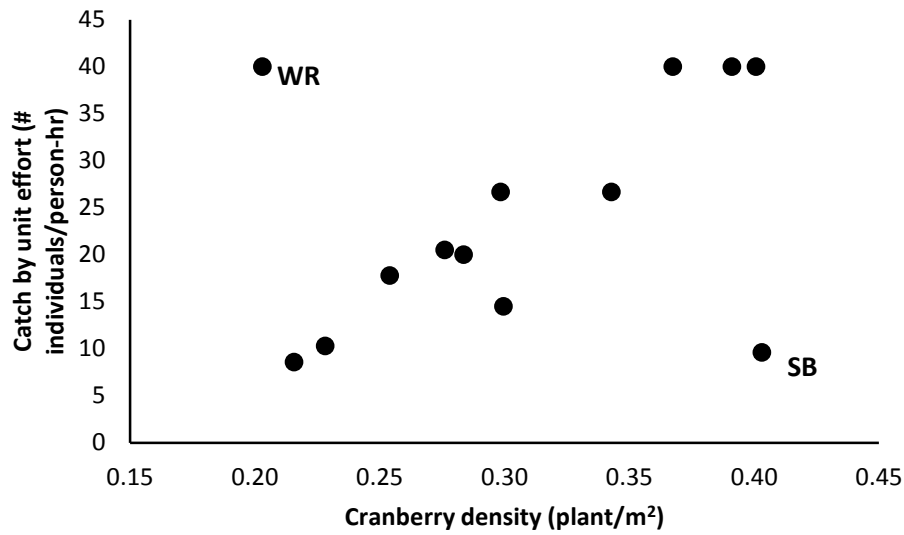


Figure 4.4 Relationships between peatland size (ha) and a) mean cranberry density, b) cranberry abundance, and c) the coefficient of variation in cranberry density, for peatlands in Algonquin Provincial Park, Canada.

a)



b)

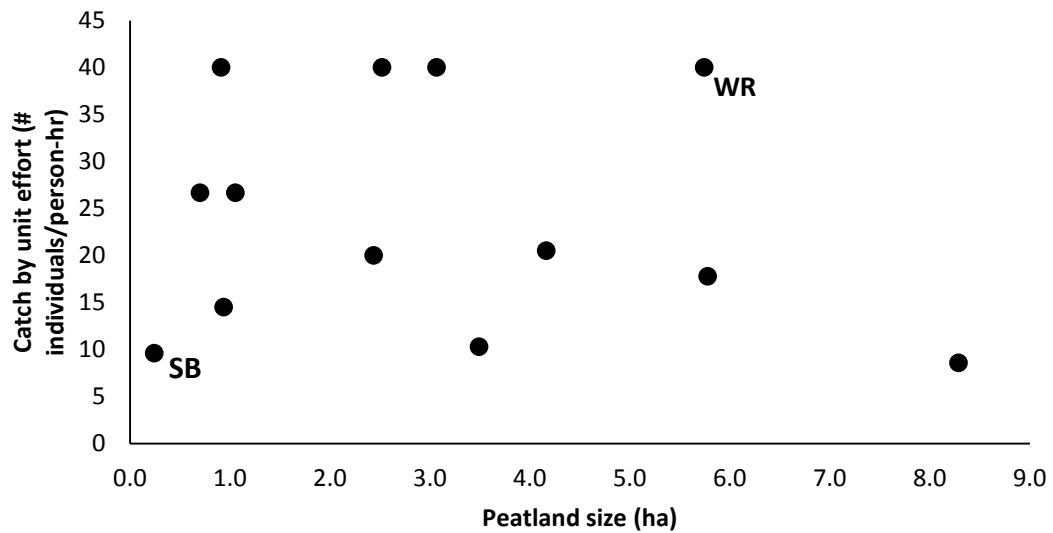


Figure 4.5 Relationship between catch per unit effort (number of individuals / person-hour) of *Lycaena epixanthe* and a) mean cranberry density and b) peatland size. Two outlier populations, SB and WR, are indicated.

4.4 Discussion

4.4.1 AFLP analysis and phenotype scoring

Using a modified version of standard AFLP protocols, we successfully amplified AFLPs generated from non-lethally sampled wing tissue in the bog copper butterfly. In contrast to the majority of current AFLP studies (see Chapter 2, Crawford et al. 2012; Chapter 3, Crawford et al. 2011), we applied a rigorous approach to select only the most repeatable and highly informative AFLP loci for use in population genetic analyses. Comparisons of results obtained from datasets varying in genotyping error rate indicate a trade-off between mean mismatch error and dataset resolution. The AFLP datasets with a moderate level of genotyping error tolerance (< 3-4%), but still within the acceptable range for AFLP studies (Bonin et al. 2004), provided the best compromise for detecting both genetic divergence among populations and genetic variability within populations. Our results suggest that using overly stringent locus selection criteria aimed at essentially removing all genotyping error (< 2%) may limit the information content of the resulting dataset by excessively reducing the number of retained loci. In a recent comparative study examining the effects of reduced AFLP genotyping error on population genetic analysis, Zhang and Hare (2012) similarly concluded that for their study system, an intermediate level of genotyping error tolerance (3-4%) was optimal for making population genetic inferences. That being said, overall we detected the same patterns of population genetic structure and diversity among populations regardless of which dataset was examined. Thus, for this particular study, datasets varying in mismatch error rates from < 2% to < 5% did not result in a dramatic gain or loss of population genetic signal.

4.4.2 Population genetic structure

Our results support the prediction that the bog copper butterfly exhibits a pattern of population genetic structure typical of a relict habitat specialist. Even at a small geographic scale (approximately 35 km), and within a relatively undisturbed natural landscape, we observed a high degree of spatial genetic structure among populations of the bog copper in Algonquin Provincial Park (Table 4.3). Recent analyses using neutral microsatellite molecular markers (E. Winkler, L. Crawford and N. Keyghobadi unpublished) and variable wing-spot pattern traits (C.Y. Kim, L. Crawford and N. Keyghobadi unpublished) also suggest significant divergence among the studied populations. High genetic differentiation has been reported for other specialist insect species inhabiting naturally fragmented landscapes (e.g., Finger et al. 2009; Ortego et al. 2010). Among locally situated populations of the bog copper, there was no detectable pattern of isolation by distance (Figure 4.3b, c), which in combination with high differentiation suggests that gene flow is highly restricted in this species even among geographically proximate populations (Hutchison and Templeton 1999). Indeed, our results indicate that genetically distinct colonies may exist even within the same peatland, a pattern which has also been thought to occur in *Agonum ericeti*, a peatland specialist ground beetle (Drees et al. 2011). Unlike other butterfly species which have demonstrated increasing genetic differentiation as a result of recent anthropogenic habitat fragmentation and isolation (e.g., Vandewoestijne and Baguette 2004; Sigaard et al. 2008), the high genetic structure of the bog copper in this landscape is likely a result of an extremely low dispersal ability, and the natural patchiness of their habitat.

The significant pattern of isolation by distance (IBD) observed at the larger scale (all populations pooled) is counterintuitive given the marked genetic structure indicated by the F_{ST} and AMOVA analyses, and the lack of IBD at smaller scales. It is unlikely that this pattern reflects an equilibrium between ongoing gene flow and genetic drift (Hutchinson and Templeton 1999). Interestingly, the larger-scale pattern of IBD appears to be linked to differences in levels of genetic diversity among populations. Lower genetic divergence (pairwise F_{ST}) appears to be related to similar estimates of genetic diversity of populations, as seen in the PCoA analysis where populations were grouped along the first coordinate axis according to differences in observed levels of genetic diversity (Figure 4.2). Several populations which share similar estimates of genetic diversity and thus lower measures of genetic divergence, also happen to be geographically proximate to each other (e.g., BUG and DL; Figure 4.1, Table 4.4). This may be driving the pattern of IBD observed at the larger scale. The large-scale IBD could also reflect historical relationships among populations related to post-glacial colonization patterns. However, analysis of the same samples using microsatellites found no evidence of IBD at this spatial scale (E. Winkler, L. Crawford and N. Keyghobadi unpublished). Thus, the larger-scale IBD pattern observed here is more likely an artefact of using a dominant marker with only two alleles per locus, where populations with low diversity (i.e. many invariant loci) are more likely to display similar allele frequencies.

4.4.3 Genetic variation

Levels of genetic diversity in the Algonquin Park populations of the bog copper appear to be moderate to low in comparison to other AFLP studies of Lepidoptera (Table 3.1

Chapter 3; Crawford et al. 2011). However, as Chapter 2 (Crawford et al. 2012) discusses in more detail, differences in laboratory protocol and the methods used to score AFLP phenotypes can generate datasets that greatly vary in the number of loci retained, potentially influencing estimates of genetic diversity and population differentiation (e.g., Herrmann et al. 2010; Zhang and Hare 2012). Thus, direct comparisons among species and studies should be made with some caution. Nevertheless, we detected considerable variation in genetic diversity measures among different populations within our study. All measures of genetic diversity varied considerably among the 13 populations analysed (Table 4.4), and demonstrated a larger range of values than was previously observed for the Mormon metalmark butterfly (Table 3.3 in Chapter 3; Crawford et al. 2011). This phenomenon (i.e., high among population variation in genetic diversity) has similarly been reported for other glacial relict species using different molecular techniques (peat moss, *Polytrichum commune*, Wilson and Provan 2002; violet copper butterfly, *Lycaena helle*, Finger et al. 2009; pitcher plant midge, *Metriocnemus knabi*, Rasic and Keyghobadi 2012).

4.4.4 Landscape effects on genetic diversity

Our results support the hypothesis that local habitat patch characteristics are more important predictors of population genetic diversity in the bog copper than features of the surrounding landscape related to isolation. Multiple linear regression analyses suggest that total peatland patch size and the proportion of water in the surrounding landscape were the factors that best explained variation among populations in genetic diversity measures, while the index of proximity (geographic isolation) and the proportion of forest

(composition) in the surrounding landscape did not influence genetic variation. We initially considered that proportion of open water could be an isolation variable as open water has been shown to be a major barrier to dispersal in other butterfly species (e.g., Leidner and Haddad 2010), in which case it would have a negative effect on population genetic diversity in the bog copper. However, our results revealed a strong positive relationship between proportion of open water and genetic diversity, which suggests that in fact proportion of water in the surrounding landscape is indicative of local peatland conditions (as outlined below) and should actually be considered as a patch quality characteristic. Geographic isolation and habitat connectivity have been shown to be good predictors of genetic diversity for other insect species inhabiting fragmented landscapes (e.g., Lange et al. 2010; Ortego et al. 2012). However, our results suggest that for the glacial relict bog copper, the geographic distance separating habitat patches and the amount of nearby habitat have little influence on local population genetic diversity. This is consistent with our population genetic structure analyses which showed no pattern of isolation by distance at the local scale, and high genetic differentiation among all surveyed populations indicated by the F_{ST} estimates and AMOVA.

It has been demonstrated by many genetic studies, including in peatland specialist species (Drees et al. 2011), that habitat patch size is positively correlated with genetic diversity, and this effect is likely mediated by higher population size in larger patches (reviewed in Frankham 1996; Keyghobadi 2007). Interestingly, in our study, we detected a significant negative relationship between genetic diversity and total peatland patch size. We hypothesize that this inverse relationship results in part from differences in the distribution of cranberry among different sized peatlands; specifically, the host-plant

density is greater and its distribution less spatially variable in smaller peatlands compared to larger ones (Figure 4.4). Not only is the plant the exclusive larval and adult food source, but cranberry plants are the preferred location used by adults for basking, perching, and courtship and pairing activities (Wright 1983). Thus, for this weak-flying species, a high host-plant density may concentrate adults (e.g. Turlure et al. 2010a) providing access to both essential resources and to mating partners which may increase mean reproductive success and local population density. This results in a situation where bog copper population density is actually highest in the smallest peatland patches, and population size does not increase in proportion to increasing patch size. In this respect, our results are consistent with a population abundance study conducted in Northern Wisconsin, USA, that noted the highest abundances of bog coppers in some of the smallest peatlands examined (Swengel and Swengel 2011). Indeed the positive association between mean cranberry density and our estimates of bog copper population density suggest that in our study area the butterfly's abundance is positively linked to cranberry density rather than total bog size. Furthermore, in populations inhabiting larger peatlands where the host-plant density is lower and less evenly distributed, individuals are more sparsely distributed across the peatland and may thus be more likely to experience skewed success in mate-finding and reproduction. The combination of higher population density and more even distribution of reproductive success would result in higher effective population size, and thus higher genetic diversity, in the smaller peatlands with high concentrations of host-plants.

In butterflies, variation in microclimatic conditions (e.g., moisture, light intensity, and temperature) can affect female oviposition behaviours, egg and larval survival, and

overall habitat use (e.g., Merrill et al. 2008; Ashton et al. 2009; Gibbs et al. 2012; Krämer et al. 2012) and likely contributed to differences in habitat quality among our study sites. Wright (1983) observed that female bog coppers specifically target cranberry plants located on *Sphagnum* hummocks for oviposition or they lay their eggs along the perimeter of the peatland on cranberry located at the base of sedges. The *Sphagnum* carpet on the peatland floor is an important thermoregulator (Spitzer and Danks 2006), which provides a cool, humid environment that helps to reduce desiccation of both the eggs and larvae of the bog copper (Wright 1983). Similarly, other Palaearctic peatland butterflies (bog fritillary, *Procllossia eunomia*; cranberry fritillary; *Boloria aquilonaris*; and violet copper, *Lycaena helle*) have been found to preferentially lay their eggs in *Sphagnum* hummocks or in areas with high moisture and cool temperatures (Turlure et al. 2009; 2010a). Thus, these microenvironmental conditions may be important non-consumable resources that can significantly impact fitness and ultimately population dynamics. As well, the geometry of different sized peatlands may also affect the overall habitat quality for the bog copper, as smaller peatlands tend to have a higher perimeter-to-area ratio which could potentially provide a larger area for oviposition.

Many glacial relict insects are well-adapted to cool and humid environments (Addo-Bediako et al. 2002; Spitzer and Danks 2006). Indeed, Turlure et al. (2010a) experimentally demonstrated greater larval survival of the cranberry fritillary caterpillar at lower temperatures. As well, they also used a field study of two Belgian peatlands to determine that higher caterpillar densities were particularly associated with *Sphagnum* hummocks located within early successional or humid areas. In these high-quality areas, the interior of hummocks provide a stable, relatively cool environment, which buffers

against fluctuating air temperatures and likely serves as a cold thermal refuge for caterpillars improving their survival. Accordingly Turlure et al. (2010a) found fewer caterpillars in drier and later successional areas of the peatland where the thermal buffering ability of the hummocks was reduced. They also noted that among their two study sites, the density of adult butterflies differed six-fold. They attributed the varying densities to the quality and quantity of adult and caterpillar resources present in each site which may have been affected by oviposition behaviour and larval survival.

While neither the presence or distribution of hummocks, nor the successional state of peatlands, were explicitly recorded in this study, the proportion of open water in the surrounding landscape may be indicative of the relative moisture content (patch characteristic) of different peatlands. Lower water tables are associated with increased shrub cover and a reduction of *Sphagnum* hummocks (Turlure et al. 2010a) and thus would result in lower quality habitat for the bog copper. In addition, water bodies function as thermoregulators (Caissie 2006) so it may be that the spring air temperatures in peatlands occurring in areas with more open water (higher water tables) are reduced which provides a more favourable thermal microclimate for larval growth and survival. Further study is required in order to determine the exact influence of the amount of open water on habitat quality for the bog coppers. Alterations in hydrology in the surrounding landscape through anthropogenic activities (e.g., drainage, de-forestation) or via climate change could indirectly affect the natural dynamics of peatland habitats, which in turn, may have dire consequences for their inhabitants.

4.4.5 Management and Conservation Implications

The current focus of many landscape genetic studies is to examine the influence of the intervening landscape on patterns of population genetic structure and diversity (Holderegger and Wagner 2008; Storfer et al. 2010). The results of our study, however, emphasize that local patch characteristics can also have important effects on population genetics. Particularly for highly specialized species, such as the bog copper, intervening landscape configuration and composition may have little influence on genetic structure and diversity, as these species are adapted to living under highly isolated conditions. For genetic studies of specialist species, patch characteristics, including descriptors of habitat quality, need to be considered in landscape analyses.

Recent ecological research has shown that an understanding of the functional habitat of a butterfly species (i.e., host-plant availability, microclimatic conditions) can significantly improve conservation efforts and population viability analyses (Dennis et al. 2006; Grundel and Pavlovic 2007; Turlure et al. 2010b). While many studies have empirically demonstrated habitat quality to be an important predictor of patch occupancy, and population abundance (e.g., Dennis and Eales 1997; Thomas et al. 2001; Fleishman et al. 2002; Krauss et al. 2004), relatively little research has focused on examining the genetic signatures of habitat quality across different populations (but see Porlier et al. 2009; Pitra et al. 2011; Alda et al. 2013). A more comprehensive understanding of the genetic consequences of changes in habitat quality, will contribute to developing more effective conservation and restoration strategies for at-risk populations, especially for those species such as the bog copper that are naturally highly fragmented.

In addition to being very closely tied to the distribution of its host-plant, the bog copper may also be strongly affected by local microclimatic conditions as have been found for other glacial relict butterflies (Turlure et al. 2009; 2010a; Goffart et al. 2010; Habel et al. 2011). While humidity, temperature, successional stage and the presence of *Sphagnum* hummocks were all found to influence caterpillar survival and density in the cranberry fritillary (Turlure et al. 2010a), different species inhabiting the same the environment often require different ecological resources (e.g., bog fritillary and violet copper; Turlure et al. 2009). Thus, species-specific detailed investigations are required to determine the functional adult and larval resources required for the bog copper. Increasing habitat fragmentation may not directly impact isolated populations of this species; however degradation of peatland habitat through direct destruction or through alterations to the hydrology of the surrounding landscape may indirectly change local environmental conditions which are important to ensuring long-term population persistence.

4.5 Literature Cited

- Addo-Bediako A, Chown SL and Gaston KJ. 2002. Metabolic cold adaptation in insects: a large-scale perspective. *Functional Ecology*, 16: 332-338.
- Aiken LS and West SG. 1991. Multiple regression: testing and interpreting interactions. Sage Publications, Newbury Park, CA.
- Alda F, Garcia J, Garcia JT and Suarez-Seoane S. 2013. Local genetic structure on breeding grounds of a long-distance migrant passerine: the bluethroat (*Luscinia svecica*) in Spain. *Journal of Heredity*, 104: 36-46.
- Ashton S, Gutiérrez D and Wilson RJ. 2009. Effects of temperature and elevation on habitat use by a rare mountain butterfly: implications for species responses to climate change. *Ecological Entomology*, 34: 437-446.
- Baguette M. 2004. The classical metapopulation theory and the real, natural world: a critical appraisal. *Basic and Applied Ecology*, 5: 213-224.
- Baguette M, Clobert J and Schtickzelle N. 2011. Metapopulation dynamics of the bog fritillary butterfly: experimental changes in habitat quality induced negative density-dependent dispersal. *Ecography*, 34: 170-176.
- Bensch S and Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.

Bonin A, Bellemain E, Bronken E, Pompanon F, Brochmann C and Taberlet P. 2004.

How to track and assess genotyping errors in population genetic studies. *Molecular Ecology*, 13: 3261-3273.

Breeuwer A, Heijmans M, Robroek B and Berendse F. 2010. Field simulation of global

change: transplanting northern bog mesocosms southward. *Ecosystems*, 13: 712-726.

Cassie D. 2006. The thermal regime of rivers: a review. *Freshwater Biology*. 51: 1389-

1406.

Cech R and Tudor G. 2005. Butterflies of the East Coast. Princeton University Press,

Princeton, NJ.

Clarke A and Meudt H. 2005. AFLP (amplified fragment length polymorphism) for

multilocus genomic fingerprinting. Alan Wilson Centre for Ecology and Evolution,

Massey University, New Zealand. Available at: <http://www.clarkeresearch.org>.

Crawford LA, Desjardins S and Keyghobadi N. 2011. Fine-scale genetic structure of an

endangered population of the Mormon metalmark butterfly (*Apodemia mormo*)

revealed using AFLPs. *Conservation Genetics*, 12: 991-1001.

Crawford LA, Kosciński D and Keyghobadi N. 2012. A call for more transparent

reporting of error rates: the quality of AFLP data in ecological and evolutionary

research. *Molecular Ecology*, 21: 5911-5917.

- Crawford LA, Koscinski D, Watt KM, McNeil JN and Keyghobadi N. 2013. Mating and oviposition success of a butterfly are not affected by non-lethal tissue sampling. *Journal of Insect Conservation*. doi: 10.1007/s10841-013-9566-8.
- Crokrak P and Barrett SCH. 2002. Purging the genetic load: a review of the experimental evidence. *Evolution*, 56: 2347-2358.
- de Vere N, Jongejans E, Plowman A and Williams E. 2009. Population size and habitat quality affect genetic diversity and fitness in the clonal herb *Cirsium dissectum*. *Oecologia*, 159: 59-68.
- Dennis RLH and Eales HT. 1997. Patch occupancy in *Coenonympha tullia* (Müller, 1976) (Lepidoptera: Satyrinae): habitat quality matters as much as patch size and isolation. *Journal of Insect Conservation*, 1: 167-176.
- Dennis RLH, Shreeve TG and Van Dyck H. 2006. Habitats and resources: the need for a resource-based definition to conserve butterflies. *Biodiversity and Conservation*, 15: 1943-1966.
- Drees C, Zumstein P, Thorsten B-D, Härdtle W, Martern A, Meyer H, von Oheimb G and Assmann T. 2011. Genetic erosion in habitat specialist shows need to protect large peat bogs. *Conservation Genetics*, 12: 1651-1656.
- Excoffier L, Laval G and Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, 1:47-50.
- Fahrig L. 2001. How much habitat is enough? *Biological Conservation*, 100: 65-74.

- Finger A, Schmitt T, Zachos FE, Meyer M, Assmann T and JC Habel. 2009. The genetic status of the violet copper *Lycaena helle* – a relict of the cold past in times of global warming. *Ecography*, 32: 382-390.
- Fleishman E, Ray C, Sjögren-Gulve P, Boggs CL and Murphy DD. 2002. Assessing the roles of patch quality, area, and isolation in predicting metapopulation dynamics. *Conservation Biology*, 16: 706-716.
- Franco AMA, Hill JK, Kitschke C, Collingham YC, Roy DB, Fox R, Huntley B and Thomas CD. 2006. Impacts of climate warming and habitat loss on extinctions at species' low-latitude range boundaries. *Global Change Biology*, 12: 1545-1553.
- Frankham R. 1996. Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10: 1500-1508.
- Gibbs M, van Dyck H and Breuker CJ. 2012. Development on drought-stressed host plants affects life history, flight morphology and reproductive output relative to landscape structure. *Evolutionary Applications*, 5: 66-75.
- Giberson D and Hardwick ML. 1999. Pitcher plants (*Sarracenia purpurea*) in eastern Canadian peatlands: ecology and conservation of the invertebrate inquilines, in: *Invertebrates in freshwater wetlands of North America*, eds. Batzer DP, Rader RB and Wissinger SA. John Wiley and Sons, Inc., New York.
- Goffart P, Schtickzelle N and Turlure C. 2010. Conservation and management of the habitats of two relict butterflies in the Belgian Ardenne: *Proclissiana eunomia* and

- Lycaena helle*, in: Relict species: phylogeography and conservation biology, eds. Habel JC and Assmann T. Springer, Heidelberg, Germany.
- Gore AJP. 1983. Ecosystems of the world Vol 4A: Analytical studies. Mires: swamp, bog, fen, and moor. Elsevier, Amsterdam.
- Grundel R and Pavlovic NB. 2007. Resource availability, matrix quality, microclimate, and spatial pattern as predictors of patch use by the Karner blue butterfly. *Biological Conservation*, 135: 135-144.
- Habel JC, Finger A, Schmitt T and Neve G. 2010. Survival of the endangered butterfly *Lycaena helle* in a fragmented environment: genetic analyses over 15 years. *Journal of Zoological Systematics and Evolutionary Research*, 49: 25-31.
- Habel JC, Rödder D, Schmitt T and Nève G. 2011. Global warming will affect the genetic diversity and uniqueness of *Lycaena helle* populations. *Global Change Biology*, 17: 194-205.
- Habel JC and Schmitt T. 2012. The burden of genetic diversity. *Biological Conservation*, 147: 270-274.
- Habel JC and Zachos FE. 2012. Habitat fragmentation versus fragmented habitats. *Biodiversity and Conservation*, 21: 2987-2990.
- Hampe A and Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics*, 42: 313-333.

- Hampe A and Petit RJ. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, 8: 461-467.
- Hansson B and Westerberg L. 2002. On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology*, 11: 2467-2474.
- Herrmann D, Poncet BN, Manel S, Rioux D, Gielly L, Taberlet P and Gugerli F. 2010. Selection criteria for scoring amplified fragment length polymorphism (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*, 53: 302-310.
- Holderegger R and Wagner HH. 2008. Landscape genetics. *BioScience*, 58: 199-207.
- Hutchinson DW and Templeton AR. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, 53: 1898-1914.
- Johannesen J, Veith M and Seitz A. 1996. Population genetic structure of the butterfly *Melitaea didyma* (Nymphalidae) along a northern distribution range border. *Molecular Ecology*, 5: 259-267.
- Johnson CW. 1985. Bogs of the northeast. University Press of New England, Hanover, NH.
- Keyghobadi N, Crawford L and Maxwell S. 2009. Successful analysis of AFLPs from non-lethally sampled wing tissues in butterflies. *Conservation Genetics*, 10: 2021-1024.

- Keyghobadi N. 2007. The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, 85: 1049-1064.
- Koscinski D, Crawford LA, Keller HA and Keyghobadi N. 2011. Effects of different methods of non-lethal tissue sampling on butterflies. *Ecological Entomology*, 36: 301-308.
- Krämer B, Kämpf I, Enderle J, Poniatowski D and Fartmann T. 2012. Microhabitat selection in a grassland butterfly: a trade-off between microclimate and food availability. *Journal of Insect Conservation*, 16: 857-865.
- Krauss J, Steffan-Dewenter I and Tschardt T. 2004. Landscape occupancy and local population size depends on host plant distribution in the butterfly *Cupido minimus*. *Biological Conservation*, 120: 355-361.
- Lange R, Durka W, Holzhauser SIJ, Wolters V and Diekötter T. 2010. Differential threshold effects of habitat fragmentation on gene flow in two widespread species of bush crickets. *Molecular Ecology*, 19: 4936-4948.
- Layberry RA, Hall PW and Lafontaine JD. 1998. The Butterflies of Canada. University of Toronto Press, Toronto.
- Leidner AK and Haddad NM. 2010. Natural, not urban, barriers define population structure for a coastal endemic butterfly. *Conservation Genetics*, 11; 2311-2320.
- Lynch M and Milligan BG. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, 3: 91-99.

- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209-220.
- Marshall SA, Finnamore AT and Blades DCA. 1999. Canadian peatlands: diversity and habitat specialization of the Arthropod fauna, in: *Invertebrates in freshwater wetlands of North America*, eds. Batzer DP, Rader RB and Wissinger SA. John Wiley and Sons, Inc., New York.
- McGarigal K, Cushman SA, Neel MC and Ene E. 2002. FRAGSTATS: Spatial pattern analysis program for categorical maps. Available at:
<http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- Merrill RM, Gutiérrez D, Lewis OT, Gutiérrez J, Díez SB and Wilson RJ. 2008. Combined effects of climate and biotic interactions on the elevational range of a phytophagous insect. *Journal of Animal Ecology*, 77: 145-155.
- Milot E, Weimerskirch H, Duchesne P and Bernatchez L. 2007. Surviving with low genetic diversity: the case of albatrosses. *Proceedings of the Royal Society B*, 274: 779-787.
- Mitsch WJ, Gosselink JG, Anderson CJ and Zhang L. 2009. *Wetland Ecosystems*. John Wiley and Sons, Inc., Hoboken, NJ.
- NatureServe. 2013. NatureServe Explorer: An online encyclopedia of life. NatureServe, Arlington, VA. Available from: <http://www.natureserve.org/explorer>.

- Nève G, Barascud B, Descimon H and Baguette M. 2008. Gene flow rise with habitat fragmentation in the bog fritillary butterfly (Lepidoptera: Nymphalidae). *BMC Evolutionary Biology*, 8: 84.
- Opler PA and Malikkul V. 1992. Eastern Butterflies. Peterson Field Guide. Houghton Mifflin Company, Boston.
- Ortego J, Aguirre MP, Cordero PJ. 2010. Population genetics of *Mioscirtus wagneri*, a grasshopper showing a highly fragmented distribution. *Molecular Ecology*, 19: 472-483.
- Ortego J, Aguirre MP and Cordero PJ. 2012. Landscape genetics of a specialized grasshopper inhabiting highly fragmented habitats: a role for spatial scale. *Diversity and Distributions*, 18: 481-492.
- Parmesan C, Ryrholm N, Stefanescu C, Hill JK, Thomas CD, Descimon H, Huntley B, Kaila L, Kullberg J, Tammaru T, Tennent WJ, Thomas JA and Warren M. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, 399: 579-583.
- Parmesan C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution and Systematics*, 37: 637-669.
- Peakall R and Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.

- Pitra C, Suárez-Seoane S, Martín CA, Streich W-J and Alonso JC. 2011. Linking habitat quality with genetic diversity: a lesson from great bustards in Spain. *European Journal of Wildlife Research*, 57: 411-419.
- Porlier M, Bélisle M and Garant D. 2009. Non-random distribution of individual genetic diversity along an environmental gradient. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364: 1543-1554.
- Rasic G and Keyghobadi N. 2012. From broadscale patterns to fine-scale processes: habitat structure influences genetic differentiation in the pitcher plant midge across multiple spatial scales. *Molecular Ecology*, 21: 223-236.
- Reed DH and Frankham R. 2003. The correlation between population fitness and genetic diversity. *Conservation Biology*, 17: 230-237.
- Reed DH. 2010. Albatrosses, eagles and newts, oh my!: exceptions to the prevailing paradigm concerning genetic diversity and population viability? *Animal Conservation*, 13: 448-457.
- Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W and Hanski I. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature*, 392: 491-494.
- Savage J, Wheeler TA, Moores AMA and Tallefer AG. 2011. Effects of habitat size, vegetation cover, and surrounding land use on Diptera diversity in temperate Nearctic bogs. *Wetlands*, 31: 125-134.

- Schlüter PM and Harris SA. 2006. Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes*, 6: 569-572.
- Sigaard P, Pertoldi C, Madsen AB, Sjøgaard B and Loeschcke V. 2008. Patterns of genetic variation in isolated populations of the endangered butterfly *Euphydryas aurinia*. *Biological Journal of the Linnean Society*, 95: 677-687.
- Spitzer K and Danks HV. 2006. Insect biodiversity of boreal peat bogs. *Annual Review of Entomology*, 51: 137-161.
- Storfer A, Murphy MA, Spear SF, Holderegger R and Waits LP. 2010. Landscape genetics: where are we now? *Molecular Ecology*, 19: 3496-3514.
- Swengel AB and Swengel SR. 2011. High and dry or sunk and dunked: lessons for tallgrass prairies from quaking bogs. *Journal of Insect Conservation*, 15: 165-178.
- Tiner RW 1999. Wetland indicators: a guide to wetland identification, delineation, classification, and mapping. Lewis Publishers, Boca Raton.
- Thomas JA, Bourn NAD, Clarke RT, Stewart KE, Simcox DJ, Pearman GS, Curtis R and Goodger B. 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. *Proceedings of the Royal Society of London B: Biological Sciences*, 268: 1791-1796.
- Tomassen HBM, Smolders AJP, Lamers LPM and Roelofs JGM. 2003. Stimulated growth of *Betula pubescens* and *Molinia caerulea* on ombrotrophic bogs: role of high levels of atmospheric nitrogen deposition. *Journal of Ecology*, 91: 357-370.

- Turlure C, Choutt J, Baguette M and Van Dyck H. 2010a. Microclimate buffering and resource-based habitat in a glacial relict butterfly: significance for conservation under climate change. *Global Change Biology*, 16: 1883-1893.
- Turlure C, Choutt J, Van Dyck H, Baguette M and Schtickzelle N. 2010b. Functional habitat area as a reliable proxy for population size: case study using two butterfly species of conservation concern. *Journal of Insect Conservation*, 14: 379-388.
- Turlure C, Van Dyck H, Schtickzelle N and Baguette M. 2009. Resource-based habitat definition, niche overlap and conservation of two sympatric glacial relict butterflies. *Oikos*, 118: 950-960.
- van Sway CA, Warren MS and Loïs G. 2006. Biotope use and trends of European butterflies. *Journal of Insect Conservation*, 10: 189-209.
- Vandewoestijne S and Baguette M. 2004. Genetic population structure of the vulnerable bog fritillary butterfly. *Hereditas*, 141: 199-206.
- Vandewoestijne S, Schtickzelle N and Baguette M. 2008. Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology*, 6: 46.
- Vekemans X, Beauwens T, Lemaire M and Rodán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11: 139-151.

- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407-4414.
- Warton DI and Hui FKC. 2011. The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92: 3-10.
- Wilson PJ and Provan J. 2003. Effect of habitat fragmentation on levels and patterns of genetic diversity in natural populations of the peat moss *Polytrichum commune*. *Proceedings of the Royal Society B: Biological Sciences*. 270: 881-886.
- Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources* 8: 725-735.
- Wright DM. 1983. Life history and morphology of the immature stages of the Bog Copper butterfly *Lycaena epixanthe* (Bsd. & Le C.) (Lepidoptera: Lycaenidae). *The Journal of Research on the Lepidoptera*, 22: 47-100.
- Zhang H and Hare MP. 2012. Identifying and reducing AFLP genotyping error: an example of tradeoffs when comparing population structure in broadcast spawning versus brooding oysters. *Heredity*, 108: 616-625.
- Zhivotovsky L. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8: 907-913.

Chapter 5

5 Flight morphology corresponds to both broad- and fine-scale landscape structure in a highly specialized glacial relict butterfly (*Lycaena epixanthe*)

5.1 Introduction

Movement of individuals is a key process that mediates the response of population dynamics, population genetics and local adaptation to landscape structure (Clobert et al. 2001; Hanski et al. 2004). However, movement ability itself may change with shifts in selection pressures accompanying ecological and landscape change. With increasing habitat fragmentation, and the associated extinction risks, there is considerable interest in understanding factors that affect movement ability or propensity in mobile organisms (Van Dyck and Baguette 2005; Baguette and Van Dyck 2007; Baguette et al. 2012).

Morphological traits associated with flight in insects, such as relative thorax mass and wing loading (total body mass/wing area), have been shown to be reliable proxies for movement across many different species (e.g., Chai and Srygley 1990; Dudley 1990; Kuusaari et al. 1996; Berwaert et al. 2002; Turlure et al. 2010b). As well, several within-species comparisons have revealed that these flight-morphological traits have a heritable basis and can thus evolve in response to environmental change (Thomas et al. 1998; Hill et al. 1999a,b; Roff and Fairbairn 2001; Merxck et al. 2003). Consequently, flight-morphological traits including thorax mass, wing area and wing loading are commonly used to study how the mobility of insects responds to different ecological conditions, particularly in the context of habitat fragmentation (e.g., Taylor and Merriam 1995;

Thomas et al. 1998; Berwaerts et al. 1998; Hill et al. 1999a, b; Van Dyck and Matthysen 1999; Norberg and Leimar 2002).

In insects, the thorax and abdomen together comprise over 80% of the total insect body mass (Thomas et al. 1998). The thorax consists primarily of flight muscles (for example, in adult butterflies it consists of 90 % flight muscle by mass; Dudley 1991), and thus thorax size is considered a dependable approximation of the overall flight muscle allocation of an individual (Wickman 1992). Likewise, because the abdomen contains the reproductive organs, its size can be used as an indication of overall fecundity (i.e., spermatophore or egg production; Hill et al. 1999a, b). Thus the relative allocation to thoracic muscle versus abdomen is expected to reflect differences in selection pressures on flight versus reproduction (Marden and Chai 1991; Zera and Denno 1997). Flight types involving rapid acceleration and manoeuvrability generally require a larger investment into flight muscles and thorax mass relative to abdomen size (Chai and Srygley 1990; Srygley and Dudley 1993; Marden 2000). Wing loading is positively related to flight speed and wing-beat frequency in many butterfly species, but also contributes to an increased cost of flight as a higher wing-beat frequency requires a higher body temperature and is energetically expensive (Bartholomew and Casey 1978). Particularly for small butterflies, individuals with a higher wing loading must spend relatively more time basking in order to increase their body temperature (Gilbert 1984; Heinrich 1986). Thus for many species, prolonged flight is often associated with larger wing area and lower wing loading (Dudley 2000).

Insects use flight for many purposes including dispersal, defined here as the movement of individuals between discrete habitat patches (Bowler and Benton 2005;

Ronce 2007). With increasing habitat fragmentation, the costs and benefits associated with dispersal for a given species may be altered, leading to evolutionary changes in flight-related morphology (Van Dyck and Matthysen 1999; Baguette et al. 2012). The strength and direction of selection imposed by habitat fragmentation on flight-morphology is highly species-specific, however, and largely depends on the spatial configuration of required resources (Van Dyck and Matthysen 1999). When all necessary resources are concentrated within a single habitat patch, increased isolation is expected to select against more mobile phenotypes due to an increased cost of dispersal out of the natal habitat patch (e.g., Dempster 1991; Schtickzelle et al. 2006). In contrast, when complementary resources become spatially segregated across different habitat patches due to habitat fragmentation such that an individual cannot obtain all required resources in a single patch, selection is expected to favour individuals with higher mobility (e.g., Taylor and Merriam 1995). Moreover, beyond its role in moving among habitat patches, flight is also important in insects for many daily tasks known as ‘station keeping’ or ‘routine movements’, including mate-location, foraging and oviposition site selection, which generally occur within a single habitat patch (Van Dyck and Baguette 2005). Changes in local conditions within habitat patches that alter the routine movements of individuals may thus also affect flight-morphology. For example, variation in the spatial heterogeneity of adult nectar resources was found to correspond with flight-morphology in four butterfly species (*Boloria aquilonaris*, *Clossiana selene*, *Lycaena hippothoe* and *Procllossiana eunomia*) across two peatlands (Turlure et al. 2010b). Where nectar resources were widespread, individuals exhibited lower wing loading and larger wing

areas, as well as an increase in female abdomen size and investment into fecundity, as inferred by the ratio between thorax and abdomen volumes.

Flight in insects is a multifaceted trait that is likely influenced by many factors including surrounding landscape structure, and resource distribution and quality. Current research on the evolution of flight morphology in butterflies has primarily focused on the effects of landscape structure (e.g., Thomas et al. 1998; Hill et al. 1999a; Hill et al. 1999b; Norberg and Leimar 2002; Merckx et al. 2003; Vandewoestijne and Van Dyck 2011). For some species, however, fine-scale variation in local habitat conditions may exert more influence on morphological traits (as a result of routine movements) than broad-scale landscape structure, and empirical studies examining this relationship are needed (Norberg and Leimar 2002; Turlure et al. 2010b). For example, glacial relict species, such as those that live in temperate peatland habitats, are highly specialized and adapted to live under very particular environmental conditions (Spitzer and Danks 2006; Hampe and Jump 2011). This close habitat association naturally restricts the spatial distribution of such species, and dispersal and gene flow among populations are often low (Schtickzelle et al. 2006). Thus, variation in flight-related morphology among populations may reflect adaptation to local ecological conditions rather than to surrounding landscape structure.

Currently, many animal species endemic to relict peatland habitats are listed as species of high conservation concern (van Sway et al. 2006; Turlure et al. 2009). This is primarily due to the increasing degradation and isolation of peatland habitats as a result of human land-use activities (Giberson and Hardwick 1999; Spitzer and Danks 2006; Savage et al. 2011). As well, these species are predicted to be extremely sensitive to

changes in local environmental conditions brought on by global climate change (e.g., changes in precipitation and temperature regimes; Turlure et al. 2010a; Habel et al. 2011). Understanding which factors influence mobility in such species is important for predicting future responses to landscape and environmental change.

Here, we use a glacial relict butterfly (bog copper, *Lycaena epixanthe* Boisduval and Le Conte 1835) to investigate the relative influence of surrounding landscape structure and local habitat characteristics on inter-population variation in flight morphology. The bog copper is a strict habitat specialist, endemic to temperate Nearctic peatlands where its adult and larval food plant (bog cranberry, *Vaccinium macrocarpum* and *V. oxycoccos*) typically grows (Wright 1983; Cech and Tudor 2005). We conducted our study within a relatively undisturbed landscape in central Ontario, Canada, where peatland habitat occurs as naturally fragmented networks. Previous population genetic analyses in this study system (Chapter 4) indicated that local populations are highly genetically differentiated, suggesting limited dispersal among habitat patches even at a small spatial scale. If the surrounding landscape structure (composition and configuration) limits dispersal in this species, and given that all necessary resources are found within individual habitat patches, we predict that individuals will invest less into flight ability as habitat patch connectivity decreases, due to increased costs of dispersal. Such individuals should thus exhibit a reduced relative thorax mass and wing area, and an increased relative abdomen mass and wing load, relative to individuals in landscapes with higher among connectivity. On the other hand, because populations are potentially highly genetically isolated, functional flight morphology may be adapted to local (within-patch) habitat conditions, such as within-

patch variation in resource distribution. In Chapter 4 we found that host plant density decreased and was less evenly distributed with increasing peatland size. In larger peatlands then, adult butterflies may require higher mobility to access an adequate quantity of this critical resource. Thus, if fine-scale habitat conditions influence individual mobility over evolutionary time, we predict more mobile flight types (larger relative thorax mass and wing area, and smaller relative abdomen and wing loading) in larger peatlands.

5.2 Methods

5.2.1 Study species

The bog copper is the smallest of the North American coppers (Lycaeninae) with a wingspan of 17-22 mm (Layberry et al. 1998). Adult flight behaviour is typically slow and low to the ground, and males have been observed to follow a ‘perching’ mate-location strategy: they establish territories on cranberry plants and sit and wait, flying out to intercept passing females (Wright 1983). Females lay eggs singly on the underside of host plant leaves, apparently selecting sites near the water’s edge and/or in regions where the peatland substrate (*Sphagnum* moss spp.) is well-saturated (Wright 1983). Adults nectar almost exclusively on *Vaccinium* spp. and feed on drops of dew (Wright 1983).

The bog copper’s distribution ranges across central and eastern Canada, in a band from Riding Mountain National Park, Manitoba east to St. John’s Newfoundland, and south along the eastern United States to Virginia (Layberry et al. 1998). It is a common and often locally abundant species (Ehrlich 1984; Swengel and Swengel 2011), however,

due to increasing habitat loss and fragmentation the species is now listed as imperiled in Pennsylvania, and critically imperiled in West Virginia and Maryland (Cech and Tudor 2005).

5.2.2 Study area

We conducted this study within the Great Lake-St. Lawrence forest region of Ontario, Canada, in Algonquin Provincial Park (UTM: 17N 692550E, 5049669N; Figure 5.1). This region is a transition zone between northern boreal coniferous forest and southern deciduous forest and spatially isolated patches of acid bog and fen habitat (collectively referred to as ‘peatlands’) are prevalent across the landscape. Forested habitats which predominate the landscape (81% of land cover) include the deciduous broad-leaved species yellow birch (*Betula alleghaniensis*), northern red oak (*Quercus rubra*), sugar maple (*Acer saccharum*), and red maple (*A. rubrum*), as well as the conifers eastern red cedar (*Juniperus virginiana*), eastern hemlock (*Tsuga canadensis*), eastern white pine (*Pinus strobus*), and red pine (*P. resinosa*). Boreal species common to the area include white birch (*Betula papyrifera*), tamarack (*Larix laricina*), white spruce (*Picea glauca*), and black spruce (*P. mariana*). The remaining land cover consists primarily of freshwater lakes (10%) and wetland habitats including peatlands (8%). Peatlands, also known as ‘organic wetlands’ are highly saturated ecosystems, characterized by an excess accumulation of plant detritus (Marshall et al. 1999; Charman 2002). Waterlogging and oxygen-poor conditions encourage the formation and accumulation of peat in these habitats (Charman 2002). Bogs are particularly acidic, nutrient poor peatlands, which receive water solely through precipitation (Gore 1983; Tiner 1999). They are home to an

endemic plant community specially adapted to living under nutrient poor, waterlogged and acidic conditions (Johnson 1985) including sphagnum mosses (*Sphagnum spp.*), low-growing heaths (e.g., bog laurel, *Kalmia polifolia*; cranberry, *Vaccinum spp.*; labrador tea, *Ledum groenlandicum*; and leatherleaf, *Chamaedaphne calyculata*) and carnivorous plants (e.g., northern pitcher plant, *Sarracenia purpurea*; and sundews, *Drasera spp.*). Peatlands which also receive water from additional ground water sources tend to be less acidic and nutrient poor, and can support a wider community of plant species (Johnson 1985; Tiner 1999). These ecosystems are generally classified as fens, however depending on groundwater flow and chemistry, pH levels can vary dramatically from ‘poor fens’ which are slightly acidic to ‘rich fens’ which are strongly alkaline (Mitsch et al. 2009).

5.2.3 Specimen collection and preparation

Adult male and female bog coppers were collected from eight different peatlands in Algonquin Provincial Park, Ontario, Canada in July 2011 (Figure 5.1; Table 5.1). Butterflies were captured using hand nets, placed in glassine envelopes, and subsequently stored at -20 °C. In an effort to survey a representative sample of the morphological variation present within each population, we collected butterflies from each peatland on at least two separate occasions, allowing two to three days to pass between each collection. We collected only newly emerged individuals, which we identified based on their degree of wing wear following a four category scale (1, very fresh and intact, to 4, heavily damaged wings; Vandewoestijne and Van Dyck 2011). We identified the sex of each individual based on the spot patterning of the dorsal forewings which differs between the sexes (Cech and Tudor 2005). Working within a protected provincial park,

we limited the number of individuals collected per site so as to minimize the effects of sampling on each population.

5.2.4 Measurement of morphological characters

Adults were thawed to room temperature, and body parts (head, thorax, abdomen, legs and wings) were carefully dissected and dried to a constant mass at 50 °C for 24h. Body parts were individually weighed on a high-precision microbalance (sensitivity ± 0.001 mg, Mettler Toledo MX5). Wings were then mounted between glass microscope slides, and digitally photographed under standardized light conditions using a Nikon D3X series camera (105 mm macro lens). We used ImageJ software v.1.5m (available from <http://image.nih.gov/ij/>) to measure the area of the ventral forewing of each individual, randomly selecting the right or left side for measurement. Wing area was calculated using a macro which automated the outlining and area measurement of the wings in ImageJ. Using a sub-set of samples, the area measurements as determined by the macro were found to be highly correlated with the mean of three area measurements obtained manually in ImageJ (Pearson's $r = 0.959$, $P < 0.001$, $N = 20$). Wing loading was calculated as total mass/forewing area (Van Dyck and Wiklund 2002; Vandewoestijne and Van Dyck 2011).

Table 5.1 Collection records for adult male and female butterflies of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada. UTM coordinates (17N) represent the centroid of butterfly capture.

Site	Peatland	Easting	Northing	N_{males}	N_{females}
WH	Wolf Howl Pond	680263.51	5049909.77	23	11
WR	West Rose Lake	680935.89	5049256.81	19	13
MIN	Minor Lake	701460.15	5057456.41	18	16
BAB	Bab Lake	701436.90	5055911.00	15	17
SUN	Sunday Creek	705574.81	5051097.38	17	18
EOS	Eos Lake	706133.08	5052006.34	19	14
COS	Costello Creek	707014.15	5054251.22	17	13
OPL	Opeongo Lake	706267.58	5056808.38	17	14

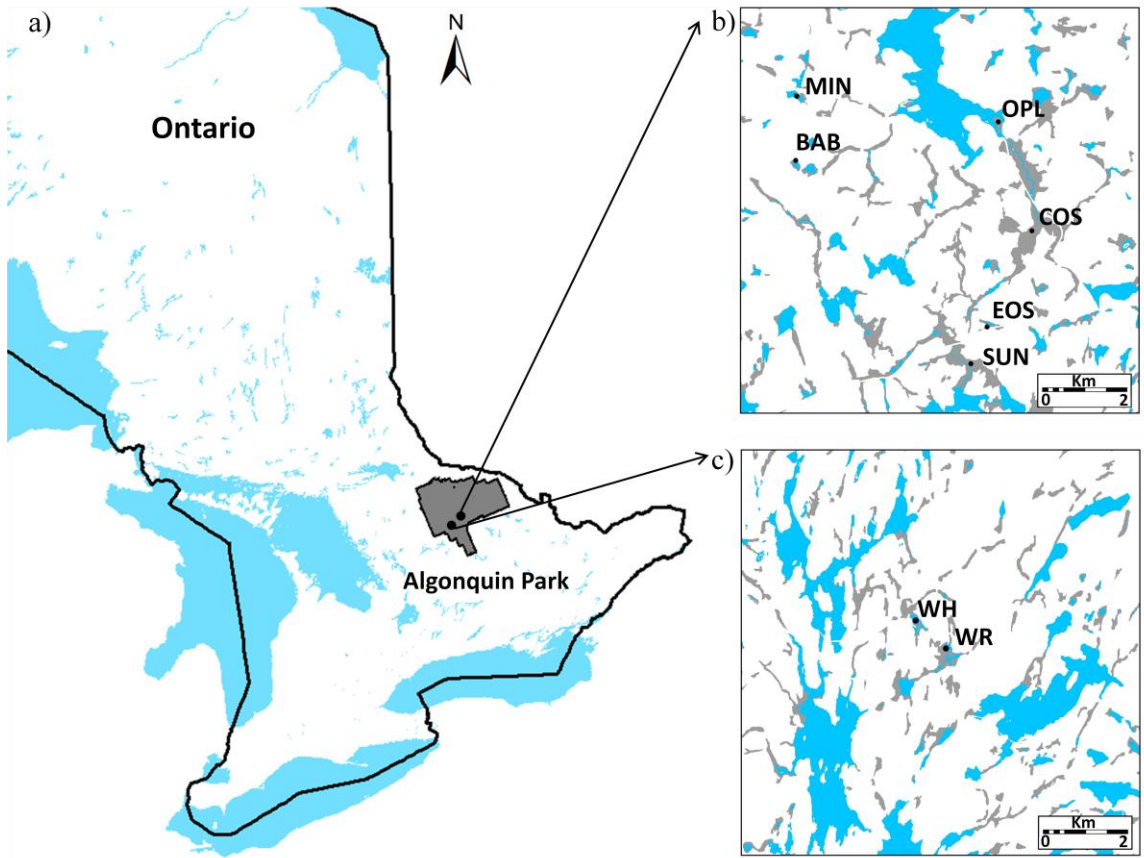


Figure 5.1 Map of the study system and sampling locations of *Lycaena epixanthe*: (a) location of Algonquin Provincial Park in Ontario, Canada; (b, c) location of the eight studied populations. Within Algonquin Park, forest habitat is indicated in white, open water in blue, and wetlands in gray.

5.2.5 Landscape evaluation

The geographic area of each sampled peatland (Table B.1) was determined as previously described in Chapter 3 through use of high-resolution orthophotography (Ontario Ministry of Natural Resources, Forest Resources Inventory 2006, 40 cm accuracy), a digital vector-based land cover classification map of Algonquin Park (Ontario Ministry of Natural Resources, Forest Resource Inventory, 2005) and personal ground truthing. We established in Chapter 4, through extensive transect surveys of 13 peatland habitats located within the same study system (six of which are included in the current study), that peatland size is a reliable indicator of the spatial dimension of host-plant resources. There was a significant negative relationship between peatland size and the density of cranberry cover ($r^2 = 0.481$, $P < 0.009$; Figure 4.4a) and conversely, a significant positive relationship between peatland size and the coefficient of variation in cranberry density ($r^2 = 0.484$, $P < 0.008$; Figure 4.4c). Thus, the density of cranberry was higher and its distribution less variable in smaller peatlands compared to larger ones. In the present study we therefore use peatland size as a proxy for the spatial dimension of host-plant resources. Two of the surveyed peatlands in this study (COS and SUN) were extremely large in size (> 22 ha) and it was logistically infeasible to conduct extensive surveys of cranberry cover to the same resolution as used in Chapter 4.

Following the same methods and justifications as described in Chapter 4, we also assessed landscape structure within a 1 km buffer radius around the centroid of butterfly capture in each surveyed peatland (Table B.1). Based on a rasterized version of the land cover classification map described above (grain size = 5 m^2), we determined the area proportion of wetland (potentially suitable habitat), forest and open water habitat. In

addition, we also determined the mean proximity of wetland habitat as an index of the relative isolation and fragmentation of wetland habitat within each buffer (McGarigal et al. 2002). Extraction of all landscape variables were performed using ArcGIS v.10.0 (ESRI, Redlands, California) and FRAGSTATS v.4.0 (McGarigal et al. 2002).

5.2.6 Statistical analyses

All morphological traits were first \log_{10} transformed (Hill et al. 1999b) and we used the Kolmogorov-Smirnov test to confirm that they did not significantly deviate from normality (in all cases $P > 0.05$). We then confirmed whether morphological traits differed between males and females; since trait variances differed between the sexes, we used the Mann-Whitney U test to evaluate sex differences for each morphological character. To assess whether variation in local mobility (i.e., flight-morphological phenotypes) among populations was related to surrounding landscape structure, and/or to within-patch habitat size and resource distribution, we used residual values of each morphological character (i.e. thorax mass, abdomen mass, wing area and wing loading) that accounted for body size and thus represented the relative investment to different body parts (Hill et al. 1999a, b). Because males and females may differ in their allocation patterns, and response to landscape and local conditions, we conducted all analyses separately for each sex.

We confirmed that the slopes of the relationships between total mass and each of thorax mass, abdomen mass, wing area, and wing loading were consistent among populations using ANCOVAs with site as a fixed factor and total mass as a covariate. In

both sexes, all traits were significantly correlated with total mass ($P < 0.001$), and for all traits there were no significant interaction effects of site*total mass ($P > 0.05$). Thus, for each sex separately, we conducted one linear regression of each of thorax mass, abdomen mass, wing area and wing loading on total mass, and used the residual values for each individual in further analyses.

To determine whether the relative investment into different morphological characters was associated with local patch size and/or surrounding landscape structure, we applied linear mixed-effect models using the R package lme4 v.0.999999-2 (D. Bates, M. Maechler and B. Bolker, available at <http://mumin.r-forge.r-project.org/>). Peatland area and mean proximity variables were \log_{10} transformed to improve normality, and proportion variables (wetland, forest and open water habitat) were logit transformed (Warton and Hui 2011). Tests for collinearity among all landscape variables had identified strong correlations (i.e., Pearson's $r > 0.70$) between proportion of wetland habitat and peatland area and between wetland habitat and proportion of forest habitat. Thus we retained only uncorrelated landscape measures (i.e., peatland area, proportion of forest habitat, proportion of open water habitat, and mean proximity) as explanatory variables. Also, individuals were nested within each peatland as a random effect, since individuals captured within the same peatland represent non-independent replicates. All predictor variables were standardized by subtracting the mean from each observation and dividing by the standard deviation.

To determine the best model and the relative contribution of each predictor in explaining variation in the response variables, we employed a multimodel inference approach (Burnham and Anderson 2002) using the R packages MuMIn v.1.9.5 (K Bartoń,

available at <http://mumin.r-forge.r-project.org>) and AICcmodavg v.1.30 (MJ Mazerolle, available at <http://cran.r-project.org/package=AICcmodavg>). We generated a candidate set of models, consisting of all possible combinations of predictor variables (i.e., 16 models). Models were then ranked based on second order Akaike information criterion values (AICc), which corrects for small sample sizes (Burnham and Anderson 2002). The top model in the set was determined as the model exhibiting the lowest AICc value, and we considered any additional models within 2 AICc values of the top model ($\Delta AICc$ values < 2) to be equally well supported (Burnham and Anderson 2002). Support for the top-ranked model was evaluated based on Akaike model weights (w_i). Assuming that the true model has been included within the candidate model set, w_i can be considered as the probability that model M_i represents the true model explaining variation in a given morphological trait. Strong support for a top-ranked model is evidenced by a weight (w_i) of 0.9 or greater, and an AIC_c value that is at least four units smaller than the second-ranked model (Burnham and Anderson 2002).

In this study we found in all cases that more than one model was equally well supported (ΔAIC_c values < 2) so we used model averaging of the top candidate models ($AIC_c < 2$) to determine the relative importance of each explanatory variable ($w_{+}(i)$) in explaining the response, as well as to obtain parameter estimates and their unconditional standard errors. We calculated relative importance for each predictor by summing the AICc weights of that predictor across those top models in which it was included. Model averaged parameter estimates and unconditional standard errors for each predictor were determined using the weighted average of the parameter estimates across the same top models. All statistical analyses were performed using IBM SPSS v.20 (IBM Corp,

Armonk, NY) and R version 3.0.0 (R Development Core Team 2013; R Foundation for Statistical Computing, Vienna).

5.3 Results

As expected, values of all morphological traits differed significantly between the sexes (Table 5.2). The overall body size of females was larger than males, as were absolute measures of thorax size, abdomen size, wing area and wing loading (Mann-Whitney U, all P -values < 0.001). After accounting for body size, there were no significant differences observed between the sexes in relative investment into different body parts (Mann-Whitney U, all P -values > 0.05).

AIC_c model selection results indicated that the morphological traits of both males and females were associated with multiple landscape variables (Table 5.3). Overall, the relative importance of each landscape predictor varied depending on the morphological trait examined. The estimated effects of predictors also differed between males and females, however, the direction of responses were fairly consistent (Table 5.4). In males, the results of the linear mixed model regression analyses indicated that our data provided very strong support for a negative relationship between relative thorax mass (i.e., residual of thorax mass on total body mass) and both peatland size and forest cover ($w_+(i) = 1.00$ and 1.00 , respectively; Table 5.4). For relative abdomen mass we found strong support for a positive relationship with peatland size and water cover ($w_+(i) = 1.00$ and 0.711 , respectively; Table 5.4), and we found good support for water cover negatively

influencing both relative wing area and relative wing loading ($w_+(i) = 0.682$ and 0.615 , respectively; Table 5.4).

In females, our analyses suggest a very strong negative influence of the proportion of forest cover on relative thorax mass ($w_+(i) = 1.00$). Relative thorax mass was also negatively related to peatland size and proportion of open water with good support ($w_+(i) = 0.604$ and 0.572 , respectively; Table 5.4). Although forest cover, peatland size and open water all demonstrated the expected positive relationship with abdomen mass, forest cover was the only predictor with considerable weight ($w_+(i) = 0.714$; Table 5.4). Relative wing area and wing loading were both best explained by mean proximity of wetland habitat, but with only limited support ($w_+(i) = 0.429$ and 0.466 , respectively; Table 5.4).

Table 5.2 Absolute mean morphology measurements (\pm SE) of adult male and female *Lycaena expixanthe* collected from eight peatlands in Algonquin Provincial Park, Ontario, Canada. Total mass, thorax mass and abdomen mass (mg) represent dry weights. Wing loading was calculated as total dry mass / wing area.

Site	Total mass (mg)	Thorax mass (mg)	Abdomen mass (mg)	Wing area (mm²)	Wing loading
<i>Males</i>					
WH	4.898 (0.120)	1.831 (0.040)	1.242 (0.051)	0.546 (0.007)	8.971 (0.174)
WR	5.028 (0.114)	1.847 (0.023)	1.374 (0.063)	0.549 (0.007)	9.203 (0.213)
MIN	4.838 (0.876)	1.795 (0.265)	1.300 (0.064)	0.527 (0.007)	0.406 (0.003)
BAB	4.858 (0.094)	1.800 (0.035)	1.300 (0.043)	0.544 (0.010)	8.959 (0.169)
SUN	5.226 (0.941)	1.881 (0.034)	1.613 (0.047)	0.535 (0.008)	9.779 (0.175)
EOS	5.501 (0.119)	2.01 (0.032)	1.554 (0.058)	0.564 (0.007)	9.746 (0.188)
COS	5.422 (0.176)	1.952 (0.039)	1.589 (0.102)	0.556 (0.009)	9.740 (0.231)
OPL	4.746 (0.108)	1.819 (0.031)	1.284 (0.053)	0.530 (0.007)	8.955 (0.149)
<i>Females</i>					
WH	7.466 (0.235)	1.960 (0.046)	3.798 (0.159)	0.546 (0.0132)	13.693 (0.342)
WR	7.914 (0.258)	2.021 (0.055)	4.115 (0.182)	0.582 (0.011)	13.596 (0.332)
MIN	7.184 (0.270)	1.927 (0.049)	3.536 (0.204)	0.547 (0.009)	13.145 (0.456)
BAB	7.748 (0.234)	1.976 (0.048)	3.988 (0.171)	0.588 (0.006)	13.194 (0.394)
SUN	8.316 (0.204)	2.171 (0.037)	4.276 (0.144)	0.590 (0.009)	14.113 (0.319)
EOS	7.216 (0.199)	2.110 (0.032)	3.318 (0.154)	0.591 (0.012)	12.354 (0.417)
COS	7.633 (0.415)	2.176 (0.083)	3.624 (0.299)	0.596 (0.010)	12.813 (0.668)
OPL	7.736 (0.308)	2.113 (0.053)	3.812 (0.246)	0.589 (0.009)	13.024 (0.526)

Table 5.3 Summary of model selection results for morphological traits in male and female *Lycaena epixanthe*. Models with $\Delta_i AIC_c < 4$ are presented, and are ranked according to AIC_c .

Model	logLik	AIC_c	$\Delta_i AIC_c$	w_i	Model	logLik	AIC_c	$\Delta_i AIC_c$	w_i
FEMALE					MALE				
<i>Relative thorax mass</i>					<i>Relative thorax mass</i>				
Area+For+Wat	59.712	-130.2	0.00	0.244	Area+For	142.919	-298.7	0.00	0.378
For	63.013	-129.0	1.18	0.135	Area+For+Wat	139.211	-297.3	1.45	0.183
Area+For	61.576	-128.9	1.26	0.130	Area+For+Prox	139.455	-296.8	1.91	0.145
For+Wat	60.851	-128.6	1.59	0.110	Area+For+Wat+Prox	135.751	-295.1	3.64	0.061
Prox	62.918	-127.9	2.29	0.078					
For+Prox	61.189	-127.8	2.41	0.073					
Area+For+Prox	59.510	-126.7	3.42	0.044					
Area+For+Wat+Prox	57.459	-126.6	3.59	0.040					
<i>Relative abdomen mass</i>					<i>Relative abdomen mass</i>				
Prox	12.816	-26.1	0.00	0.139	Area+Wat	80.006	-169.4	0.00	0.298
For	12.493	-26.1	0.02	0.138	Area	81.981	-168.3	1.11	0.172
Area+For+Wat	9.797	-26.0	0.16	0.128	Area+For+Wat	77.607	-167.6	1.76	0.123
Area+For	11.439	-25.8	0.35	0.117	Area+Wat+Prox	77.215	-166.7	2.68	0.078
For+Wat	10.651	-25.2	0.89	0.089	Area+Prox	79.503	-166.6	2.77	0.075
Intercept	13.366	-25.0	1.16	0.078	Area+For	79.484	-166.0	3.39	0.055
For+Prox	11.088	-24.7	1.38	0.070					
Wat+Prox	10.482	-23.6	2.56	0.039					
Area+Prox	10.566	-23.5	2.61	0.038					
Area+For+Prox	9.825	-23.5	2.62	0.037					
Area	11.399	-23.2	2.94	0.032					
Wat	11.286	-23.1	3.05	0.030					
For+Wat+Prox	9.072	-22.6	3.51	0.024					
Area+For+Wat+Prox	8.084	-22.4	3.76	0.021					
<i>Relative wing area</i>					<i>Relative wing area</i>				
Prox	209.063	-426.0	0.00	0.227	Wat	301.300	-615.0	0.00	0.194
Intercept	211.822	-425.7	0.34	0.192	Area+Wat	296.907	-614.4	0.56	0.146

For	208.193	-424.6	1.44	0.111	Intercept	305.138	-614.0	1.01	0.117
Area	207.934	-423.9	2.11	0.079	For+Wat	296.456	-613.5	1.51	0.091
For+Prox	205.105	-423.5	2.56	0.063	Area	300.833	-613.3	1.67	0.084
Area+Prox	204.960	-423.5	2.57	0.063	For	300.574	-612.6	2.35	0.060
Wat+Prox	204.852	-423.4	2.60	0.062	Wat+Prox	296.619	-612.2	2.82	0.047
Wat	207.677	-423.4	2.62	0.061	Area+Prox	296.629	-612.0	3.01	0.043
For+Wat	204.131	-422.3	3.76	0.035	For+Prox	296.765	-611.9	3.03	0.042
					Prox	300.576	-611.9	3.10	0.041
					Area+Wat+Prox	292.396	-611.7	3.31	0.037
					Area+For+Wat	292.486	-611.5	3.51	0.033
<i>Relative wing loading</i>					<i>Relative wing loading</i>				
Prox	-139.187	282.9	0.00	0.228	Wat	-101.892	202.8	0.00	0.184
Intercept	-139.529	283.2	0.32	0.194	Area+Wat	-103.311	203.0	0.29	0.160
For	-140.081	284.4	1.50	0.108	Intercept	-100.759	203.6	0.82	0.122
Area	-140.304	285.0	2.08	0.081	Area	-102.140	204.1	1.36	0.093
Area+Prox	-140.181	285.4	2.56	0.063	For+Wat	-103.720	205.0	2.28	0.059
For+Prox	-140.045	285.5	2.58	0.063	For	-102.515	205.1	2.33	0.057
Wat+Prox	-140.287	285.5	2.59	0.062	Area+Prox	-103.424	205.3	2.56	0.051
Wat	-140.575	285.5	2.62	0.062	Wat+Prox	-103.589	205.6	2.85	0.044
For+Wat	-140.046	286.8	3.87	0.033	Prox	-102.431	205.6	2.88	0.044
					Area+For+Wat	-104.714	205.8	3.07	0.040
					Area+Wat+Prox	-104.857	205.9	3.09	0.039
					For+Prox	-103.472	205.9	3.13	0.039
					Area+For	-103.640	206.6	3.81	0.027

Log likelihood statistics (logLik), second order Akaike information criterion (AIC_c), $\Delta_i AIC_c$, and Akaike weights (w_i) are derived from linear-mixed model regressions. *Area*, total peatland area; *For*, proportion of forest habitat; *Wat*, proportion of open water habitat; *Prox*, mean proximity of potentially suitable wetland habitat.

Table 5.4 Effect of landscape variables on morphological traits in male and female *Lycaena epixanthe*. Model-averaged Akaike weights ($w_+(i)$), parameter estimates ($\widehat{\beta}_j$), and standard errors (SE) are derived from linear mixed-model regressions after model selection.

Variable	Relative thorax mass			Relative abdomen mass			Relative wing area			Relative wing loading		
	$w_+(i)$	$\widehat{\beta}_j$	SE	$w_+(i)$	$\widehat{\beta}_j$	SE	$w_+(i)$	$\widehat{\beta}_j$	SE	$w_+(i)$	$\widehat{\beta}_j$	SE
<i>Males</i>												
Peatland area	1.000	-0.033	0.012	1.000	0.042	0.017	0.364	-0.003	0.003	0.452	0.058	0.047
Forest cover	1.000	-0.033	0.012	0.208	0.021	0.027	0.144	0.002	0.002	-	-	-
Water cover	0.260	-0.006	0.007	0.711	0.025	0.015	0.682	-0.004	0.002	0.615	0.071	0.038
Proximity	0.205	-0.006	0.011	-	-	-	-	-	-	-	-	-
<i>Females</i>												
Peatland area	0.604	-0.053	0.038	0.323	0.080	0.062	-	-	-	0.161	-0.083	0.150
Forest cover	1.000	-0.059	0.028	0.714	0.073	0.046	0.209	-0.006	0.006	0.228	0.115	0.157
Water cover	0.572	-0.027	0.021	0.287	0.039	0.034	-	-	-	0.175	0.022	0.143
Proximity	-	-	-	0.275	-0.092	0.053	0.429	0.012	0.008	0.466	-0.272	0.191

5.4 Discussion

In this study, we examined variation in flight-related morphological traits among local populations of the bog copper butterfly. Using an AIC_c inferential approach our results provide evidence for effects of both the surrounding landscape structure, and the distribution of local host-plant resources, on flight morphology. This suggests that for naturally fragmented species like the bog copper, local conditions in addition to landscape structure may be important in influencing mobility, and thus should be considered in long-term species management plans.

Given the bog copper's small size and thus inherent limited dispersal ability (Sekar 2012), coupled with its dependency on cranberry as the sole adult and larval food source, the cost of long, continuous flights through an inhospitable matrix of forest and open water habitat is presumed to be high. As well, previous population genetic analyses of the bog copper detected a high degree of genetic differentiation among local populations (Chapter 4). Thus, we hypothesized that increasing habitat isolation should increase the cost of dispersal among habitat patches and select against mobile phenotypes. Indeed we found strong evidence that in both sexes, thorax allocation decreased with increasing amount of forest in the surrounding landscape. Following colonization of peatland habitats, long-distance movements in the bog copper (as inferred by relative thorax mass) may have been selected against in response to increasing forest cover over time (Hanski et al. 2004). Changes in dispersal propensity and allocation to thorax muscles in relation to the degree of habitat isolation have been demonstrated in other insect species experiencing contemporary habitat fragmentation (e.g., Dempster 1991; Heidinger et al. 2010). The populations examined here however have likely been

isolated over relatively long time scales, and may have become adapted to local conditions, which could overrule the effects of landscape structure (e.g., Turlure et al. 2009). Our results however, indicate that at least for the bog copper butterfly, habitat isolation (as measured by the amount of forest in the surrounding landscape) appears to negatively affect morphological traits associated with mobility.

In addition to dense forest cover, open water has been shown to be a significant barrier to movement in other butterfly species (e.g., Leidner and Haddad 2010), and thus we predicted that we might observe less mobile phenotypes with increasing amounts of open water in the surrounding landscape. Our results showed that in males, relative abdomen size and wing loading were larger and relative wing area smaller in habitat patches surrounded by more open water. While these results support our hypothesis of water as a dispersal barrier, we did not find any influence of proportion of open water on male relative thorax mass. Furthermore, although proportion of forest in the landscape strongly influenced flight morphology in females, our analyses did not provide strong support for any effect of open water on female morphology. In Chapter 4 we found that estimates of genetic diversity increased when populations were surrounded by a higher proportion of open water habitat, indicating that the amount of open water may be linked to differences in habitat quality among sites, including differences in water and nutrient availability. Thus it is possible that the proportion of open water may indirectly lead to changes in male morphology as a result of varying local habitat conditions, rather than by directly hindering mobility through the surrounding landscape. Mounting evidence suggests that variation in microclimatic conditions and host-plant quality, via changes in water and nutrient availability, can indirectly influence morphological as well as

behavioural and life-history traits in butterflies (Talloen et al. 2004; Turlure et al. 2010a; Gibbs et al. 2012; Turlure et al. 2013; Van Velde et al. 2013). For example, male *Pararge aegeria* butterflies reared on drought-stressed plants in which leaf nitrogen, carbon and water concentrations were reduced were found to invest less into reproduction (spermatophore size) compared to control individuals (Van Velde et al. 2013). In a separate experiment on the same species, females reared on drought-stressed plants exhibited lower wing loading and reproductive output (reduced fecundity). The results of these studies highlight that the functional morphology of butterflies can also be influenced by fine-scale, local heterogeneity in ecological conditions in addition to broader landscape features. Thus the connection between proportion of open water in the landscape, male flight morphology, and habitat quality identified in this study warrants further exploration.

It is now well-established that the abundance and spatial distribution of nectar and larval host-plants can influence the movement of butterflies across fragmented landscapes (Brommer and Fred 2001). While few empirical examples currently exist, it is also thought that fine-scale resource heterogeneity within habitat patches can impact mobility (Baguette and Van Dyck 2007; Turlure et al. 2010b). While we predicted investment in flight to increase with increasing peatland size due to more heterogeneous host-plant coverage, our results instead showed that allocation to the thorax in both sexes increased with declining peatland size, although the effect was not significant in females. One explanation for the relatively large thoraxes observed in smaller peatlands, is that smaller habitat patches are more likely to experience local extinctions and have subsequently been re-colonised by dispersive individuals with larger thoraxes, as shown in some other

butterfly species (Kuussaari et al. 1996; Hanski et al. 2002; 2004; Hill et al. 1999a). However, we have established that host-plant density is greater and more homogenous in smaller peatlands in our study system (Chapter 4), and previous genetic analyses (Chapter 4) and population abundance surveys of the bog copper (Swengel and Swengel 2011) both indicate that small habitat patches may have the carrying capacity to support relatively large, stable populations. Thus it seems unlikely that populations in small habitat patches are more likely to have recently undergone extinction. Given that the effects of peatland area on thorax and abdomen mass were considerably stronger in males than females, it is possible that peatland size affects male morphology by influencing mate location behaviours. Bog coppers have been classified as a perching species (Wright 1983) and this is likely to be the case in small peatlands with dense cranberry cover, where males need to establish territories in order to ensure access to mates. In larger peatlands where cranberry cover is less dense and more heterogeneous however, males may instead switch to a more ‘patrol’ type of mate location behaviour, more actively flying around in order to locate mates (Scott 1986). Typically, perchers have relatively large thoraxes compared to patrollers, which allows for more quick, rapid and powerful flight (Betts and Wootton 1988; Dudley 1990; Wickman 1992). A similar explanation was proposed by Thomas et al. (1998) to account for the negative relationship between relative thorax mass and habitat patch area observed in a metapopulation of the lycaenid *Plebejus argus* occupying fragmented patches of limestone grassland. Furthermore, males of the butterfly species *P. aegeria* have been shown to exhibit different mate-location strategies and corresponding differences in flight morphological design in relation to habitat structure (Shreeve 1984; Van Dyck and Matthysen 1999; Van Dyck

2003). Whether males of the bog copper similarly adopt alternative mate-location strategies depending on resource distribution is unknown and requires further study.

We found that males and females exhibited sexual dimorphism, as is common across most butterfly species, with females being generally larger than males (Wickman 1992; Layberry et al. 1998; Scott 1986; Turlure et al. 2010b). While the morphological traits of both sexes generally responded in the same direction to landscape predictors, the relative importance and statistical significance of explanatory variables differed. This may reflect sex-specific differences in ecology whereby investment into flight-related morphology is shaped by different factors (Norberg and Leimar 2002; Heidinger et al. 2010). As males spend most of their active time searching for mates (Shreeve 1992), their morphology is likely to be highly influenced by local conditions that affect mate location (i.e., mobility at the local scale) rather than long-distance dispersal (mobility at the landscape scale; Wickham 1992; Hanski et al. 2004). In butterflies, females are often the more dispersive sex (Baker 1984; Scott 1986) and in contrast to males their morphology may be more reflective of factors that impede or facilitate movement at the landscape scale. Additionally, trade-offs between flight ability and fecundity have been shown to occur in females of many insect species (Zera and Denno 1997; Zera and Harshman 2001; Hughes et al. 2003, but see Hanski et al. 2006), therefore it is also possible that the morphology of female bog coppers may be responding to factors influencing reproduction rather than movement per se (Hanski et al. 2004).

Flight morphology characters have been demonstrated for many insect species to be reliable proxies of flight ability (Chai and Srygley 1990; Kuusaari et al. 1996; Berwaert et al. 2002; Turlure et al. 2010b). Thus, we interpret the variation in flight

morphology observed here among local populations of the bog copper butterfly to reflect real differences in movement ability. As the functional relationship between flight morphology and flight performance has not been evaluated specifically for the bog copper, however, we caution that future work is still needed to confirm this relationship. Our correlative approach using direct measurements of wild-caught adults allows us to examine patterns of phenotypic variation among populations; however it does not allow us to identify whether the underlying mechanism responsible for these differences is local adaptation or phenotypic plasticity. The life-history of the bog copper makes it extremely challenging to rear under laboratory conditions (Wright 1983; L. Crawford unpublished), and thus conduct a common garden experiment which would allow us to identify whether phenotypic variation among sites is due to genetic differences. Regardless, we observed a considerable amount of morphological variation among local populations of the bog copper suggesting that the flight morphology of these butterflies is responding on some timescale to both local habitat conditions and landscape structure.

5.5 Literature Cited

- Baguette M, Legrand D, Freville H, Van Dyck H and Ducatez S. 2012. Evolutionary ecology of dispersal in fragmented landscapes, in: Dispersal ecology and evolution, eds. Clobert J, Baguette M, Benton TG and Bullock JM. Oxford University Press, Oxford.
- Baguette M and Van Dyck H. 2007. Landscape connectivity and animal behaviour: functional grain as a key determinant for dispersal. *Landscape Ecology*, 22: 1117-1129.
- Baker RR. 1984. The dilemma: When and how to go or stay, in: The biology of butterflies, eds. Vane-Wright RI and Ackery PR. Academic Press, London.
- Bartholomew GA and Casey TM. 1978. Oxygen consumption of moths during rest, pre-flight warm-up and flight in relation to body size and wing morphology. *Journal of Experimental Biology*, 76: 11-25.
- Berwaerts K, Van Dyck H and Aerts P. 2002. Does flight morphology relate to flight performance? An experimental test with the butterfly *Pararge aegeria*. *Functional Ecology*, 16: 484-491.
- Berwaerts K, Van Dyck H, Van Dongen S and Matthysen E. 1998. Morphological and genetic variation in the speckled wood butterfly (*Pararge aegeria* L.) among differently fragmented landscapes. *Netherlands Journal of Zoology*, 48: 241-253.
- Betts CR and Wootton RJ. 1988. Wing shape and flight behaviour in butterflies (Lepidoptera: Papilionoidea and Hesperioidea): A preliminary analysis. *Journal of Experimental Biology*, 138: 271-288.

- Bowler DE and Benton TG. 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80: 205-225.
- Brommer JE and Fred MS. 2001. Movement of the Apollo butterfly *Parnassius apollo* related to host plant and nectar plant patches. *Ecological Entomology*, 24: 125-131.
- Burnham KP and Anderson DR. 2002. Model selection and multimodel inference: A practical information-theoretic approach, 2nd edn. Springer-Verlag, New York.
- Cech R and Tudor G. 2005. Butterflies of the East Coast. Princeton University Press, Princeton, NJ.
- Chai P and Srygley RB. 1990. Predation and the flight, morphology, and temperature of Neotropical rain-forest butterflies. *The American Naturalist*, 135: 748-765.
- Charman D. 2002. Peatlands and environmental change. John Wiley and Sons Ltd, London.
- Clobert J, Ims RA and Rousset F. 2004. Causes, mechanisms and consequences of dispersal, in: Ecology, genetics, and evolution of metapopulations, eds. Hanski I and Gaggiotti OE. Academic Press Inc., London.
- Dempster JP. 1991. Fragmentation, isolation and mobility of insect populations, in: Conservation of insects and their habitats, eds. Collins NM and Thomas JA. Academic Press Inc., San Diego, CA.
- Dudley R. 1990. Biomechanics of flight in Neotropical butterflies: morphometrics and kinematics. *Journal of Experimental Biology*, 150: 37-53.

Dudley R. 1991. Biomechanics of flight in Neotropical butterflies: aerodynamics and mechanical power requirements. *Journal of Experimental Biology*, 150: 335-357.

Ehrlich PR. 1984. The structure and dynamics of butterfly populations, in: The biology of butterflies, eds. Vane-Wright RI and Ackery PR. Academic Press, London.

Gibbs M, Van Dyck H and Breuker CJ. 2012 . Development on drought-stressed host plants affects life history, flight morphology and reproductive output relative to landscape structure. *Evolutionary Applications*, 5: 66-75.

Giberson D and Hardwick ML. 1999. Pitcher plants (*Sarracenia purpurea*) in eastern Canadian peatlands: ecology and conservation of the invertebrate inquilines, in: Invertebrates in freshwater wetlands of North America, eds. Batzer DP, Rader RB and Wissinger SA. John Wiley and Sons, Inc., New York.

Gilbert FS. 1984. Thermoregulation and the structure of swarms in *Syrphus ribesii* (Syrphidae). *Oikos*, 42: 249-255.

Gore AJP. 1983. Ecosystems of the world Vol 4A: Analytical studies. Mires: swamp, bog, fen, and moor. Elsevier, Amsterdam.

Habel JC, Rödder D, Schmitt T and Nève G. 2011. Global warming will affect the genetic diversity and uniqueness of *Lycaena helle* populations. *Global Change Biology*, 17: 194-205.

Hampe A and Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics*, 42: 313-333.

- Hanski I, Breuker CJ, Schöps K, Setchfield R and Nieminen M. 2002. Population history and life history influence the migration rate of female Glanville fritillary butterflies. *Oikos* 98: 87-97.
- Hanski I, Erälahti C, Kankare M, Ovaskainen O and Sirén H. 2004. Variation in migration propensity among individuals maintained by landscape structure. *Ecology Letters*, 7: 958-966.
- Hanski I, Saastamoinen M and Ovaskainen O. 2006. Dispersal-related life-history trade-offs in a butterfly metapopulation. *Journal of Animal Ecology*, 7: 91-100.
- Heidinger IMM, Hein S and Bonte D. 2010. Patch connectivity and sand dynamics affect dispersal-related morphology of the blue-winged grasshopper *Oedipoda caerulescens* in coastal grey dunes. *Insect Conservation and Diversity*, 3: 205-212.
- Hill JK, Thomas CD and Blakeley DS. 1999a. Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121: 165-170.
- Hill JK, Thomas CD and Lewis OT. 1999b. Flight morphology in fragmented populations of a rare British butterfly, *Hesperia comma*. *Biological Conservation*, 87: 277-283.
- Hughes CL, Hill JK and Dytham C. 2003. Evolutionary trade-offs between reproduction and dispersal in populations at expanding range boundaries. *Proceedings of the Royal Society B*, 270: S147-S150.
- Johnson CW. 1985. Bogs of the northeast. University Press of New England, Hanover, NH.
- Kuussaari M, Nieminen M and Hanski I. 1996. An experimental study of migration in the Glanville fritillary butterfly *Melitaea cinxia*. *Journal of Animal Ecology*, 65: 791-801.

- Layberry RA, Hall PW and Lafontaine JD. 1998. *The Butterflies of Canada*. University of Toronto Press, Toronto.
- Leidner AK and Haddad NM. 2010. Natural, not urban, barriers define population structure for a coastal endemic butterfly. *Conservation Genetics*, 11; 2311-2320.
- Marden JH. 2000. Variability in the size, composition and function of insect flight muscles. *Annual Review of Physiology*, 62: 157-178.
- Marden JH and Chai P. 1991. Aerial predation and butterfly design: How palatability, mimicry, and the need for evasive flight constrain mass allocation. *American Naturalist*, 138: 15-36.
- Marshall SA, Finnamore AT and Blades DCA. 1999. Canadian peatlands: diversity and habitat specialization of the Arthropod fauna, in: *Invertebrates in freshwater wetlands of North America*, eds. Batzer DP, Rader RB and Wissinger SA. John Wiley and Sons, Inc., New York.
- McGarigal K, Cushman SA, Neel MC and Ene E. 2002. FRAGSTATS: Spatial pattern analysis program for categorical maps. Available at:
<http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- Merckx T, Van Dyck H, Karlsson B, Leimar O. 2003. The evolution of movements and behaviour at boundaries in different landscapes: a common arena experiment with butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 270: 1815-1821.
- Mitsch WJ, Gosselink JG, Anderson CJ and Zhang L. 2009. *Wetland Ecosystems*. John Wiley and Sons, Inc., Hoboken, NJ.

- Norberg U and Leimar O. 2002. Spatial and temporal variation in flight morphology in the butterfly *Melitaea cinxia* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, 77: 445-453.
- Roff DA and Fairbairn DJ. 2001. The genetic basis of dispersal and migration, and its consequences for the evolution of correlated traits, in: Dispersal, eds. Clobert J, Danchin E, Dhondt AA and Nichols JD. Oxford University Press, Oxford.
- Ronce O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Reviews in Ecology Evolution and Systematics*, 38: 231-253.
- Savage J, Wheeler TA, Moores AMA and Taillefer AG. 2011. Effects of habitat size, vegetation cover, and surrounding land use on Diptera diversity in temperate Nearctic bogs. *Wetlands*, 31: 125-134.
- Schtickzelle N, Mennechez G and Baguette M. 2006. Dispersal depression with habitat fragmentation in the bog fritillary butterfly. *Ecology*, 87: 1057-1065.
- Sekar S. 2012. A meta-analysis of the traits affecting dispersal ability in butterflies: can wingspan be used as a proxy? *Journal of Animal Ecology*, 81: 174-184.
- Scott JA. 1986. The butterflies of North America: a natural history and field guide. Stanford University Press, Stanford, CA.
- Shreeve TG. 1984. Habitat selection, mate-location, and micro-climatic constraints on the activity of the speckled wood butterfly *Pararge aegeria*. *Oikos*, 42: 371-377.
- Shreeve TG. 1992. Adult behaviour, in: The ecology of butterflies in Britain, ed. Dennis RLH. Oxford University Press, Oxford.

- Spitzer K and Danks HV. 2006. Insect biodiversity of boreal peat bogs. *Annual Review of Entomology*, 51: 137-161.
- Swengel AB and Swengel SR. 2011. High and dry or sunk and dunked: lessons for tallgrass prairies from quaking bogs. *Journal of Insect Conservation*, 15: 165-178.
- Talloe W, Van Dyck H and Lens L. 2004. The cost of melanisation: butterfly wing coloration under environmental stress. *Evolution*, 58: 360-366.
- Taylor PD and Merriam G. 1995. Wing morphology of a forest damselfly is related to landscape structure. *Oikos*, 73: 43-48.
- Thomas CD, Hill JK and Lewis OT. 1998. Evolutionary consequences of habitat fragmentation in a localized butterfly. *Journal of Animal Ecology*, 67: 485-497.
- Tiner RW 1999. Wetland indicators: a guide to wetland identification, delineation, classification, and mapping. Lewis Publishers, Boca Raton.
- Turlure C, Chouff J, Baguette M and Van Dyck H. 2010a. Microclimate buffering and resource-based habitat in a glacial relict butterfly: significance for conservation under climate change. *Global Change Biology*, 16: 1883-1893.
- Turlure C, Radchuk V, Baguette M, Meijrink M, van den Burg A, WallisDeVries M and van Duinen G-J. 2013. Plant quality and local adaptation undermine relocation in a bog specialist butterfly. *Ecology and Evolution*, 3: 244-254.
- Turlure C, Schtickzelle N and Baguette M. 2010b. Resource grain scales mobility and adult morphology in butterflies. *Landscape Ecology*, 25: 95-108.

- Turlure C, Van Dyck H, Schtickzelle N and Baguette M. 2009. Resource-based habitat definition, niche overlap and conservation of two sympatric glacial relict butterflies. *Oikos*, 118: 950-960.
- Van Dyck H. 2003. Mate location: a matter of design? Adaptive morphological variation in the speckled wood butterfly, in: *Butterflies: Ecology and evolution taking flight*, eds. Boggs CL, Watt WB and Ehrlich PR. University of Chicago Press, London.
- Van Dyck H and Baguette M. 2005. Dispersal behaviour in fragmented landscapes: routine or special movements? *Basic and Applied Ecology*, 6: 535-545.
- Van Dyck H and Baguette M. 2012. Dispersal and habitat fragmentation in invertebrates- examples from widespread and localized butterflies, eds. Clobert J, Baguette M, Benton TG and Bullock JM. Oxford University Press, Oxford.
- Van Dyck H and Matthysen E. 1999. Habitat fragmentation and insect flight: a changing 'design' in a changing landscape? *Trends in Ecology and Evolution*, 5: 172-174.
- Van Dyck H and Wiklund C. 2002. Seasonal butterfly design: morphological plasticity among three developmental pathways relative to sex, flight and thermoregulation. *Journal of Evolutionary Biology*, 15: 216-225.
- van Sway CA, Warren MS and Loïs G. 2006. Biotope use and trends of European butterflies. *Journal of Insect Conservation*, 10: 189-209.
- Vande Velde V, Schtickzelle N and Van Dyck H. 2013. Effect of larval food stress on male adult behaviour, morphology and reproductive investment in the butterfly *Pararge aegeria*. *Evolutionary Ecology*, 27: 221-234.

- Vandewoestijne S and Van Dyck H. 2011. Flight morphology along a latitudinal gradient in a butterfly: do geographic clines differ between agricultural and woodland landscapes? *Ecography*, 34: 876-886.
- Warton DI and Hui FKC. 2011. The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92: 3-10.
- Wickman P-O. 1992. Sexual selection and butterfly design – a comparative study. *Evolution*, 5: 1525-1536.
- Wright DM. 1983. Life history and morphology of the immature stages of the Bog Copper butterfly *Lycaena epixanthe* (Bsd. & Le C.) (Lepidoptera: Lycaenidae). *The Journal of Research on the Lepidoptera*, 22: 47-100.
- Zera AJ and Denno RF. 1997. Physiology and ecology of dispersal polymorphism in insects. *Annual Review of Entomology*, 42: 207-230.
- Zera AJ and Harshman LG. 2001. The physiology of life history tradeoffs in animals. *Annual Review of Ecology and Systematics*, 32: 95-126.

Chapter 6

6 Molecular signatures of selection associated with fine-scale landscape heterogeneity in a relict butterfly, *Lycaena epixanthe*

6.1 Introduction

Humans have induced wide-spread environmental change, such as land-use and climate change as well as the introduction of invasive species and environmental contaminants, which has placed new selective pressures on species (Reusch and Wood 2007), altering dispersal patterns, the timing of life-history events, population dynamics and overall fitness (Lovejoy and Hannah 2005; Parmesan 2006). While broad-scale shifts in geographic distributions in response to environmental change have been predicted fairly accurately for many species (Thuiller et al. 2008; Pereira et al. 2010), current distribution models do not possess the spatial and biological resolution necessary to forecast the evolutionary response of populations at a local level (Hampe and Jump 2011). Local adaptation to fine-scale landscape heterogeneity (i.e., variation in habitat size, quality and isolation) has the potential to buffer populations against larger-scale climatic and landscape changes (Thuiller et al. 2008; Willis and Baghwat 2009). Thus, an understanding of how current local adaptations have been shaped by fine-scale landscape heterogeneity may allow us to more accurately predict the future geographic distributions and survival of organisms (Manel et al. 2010a; Hampe and Jump 2011).

The emerging field of landscape genomics provides a framework for studying the effects of the environment/landscape on adaptive genetic variation in natural populations (Holderegger et al. 2008; Joost et al. 2007; Manel et al. 2010a; Schoville et al. 2012). Through the integration of high-resolution genomic data, environmental datasets and spatial statistical methods, landscape genomic studies seek to link the spatial distribution of alleles potentially under selection to environmental variables (Schoville et al. 2012). Putative adaptive loci are identified using genome-scan methods which screen genetic markers such as amplified fragment length polymorphisms (AFLPs), single nucleotide polymorphisms (SNPs) and microsatellites (SSRs), and apply either a population-based outlier locus detection method or a correlative landscape genetics approach to find loci exhibiting signatures of natural selection (Manel et al. 2010a). Outlier locus detection methods identify potentially adaptive loci as those exhibiting higher genetic differentiation (e.g., F_{ST}) among populations than expected under a neutral model (Vitalis et al. 2001; Beaumont 2005). In contrast, the landscape genetics approach identifies candidate loci based on the correlation of allele frequencies with clinal ecological data (Joost et al. 2007; Holderegger et al. 2010). In most cases, these identified molecular markers are closely linked to genomic regions under selection rather than directly experiencing selection themselves (Maynard Smith and Haigh 1974; Holderegger et al. 2008).

Particularly in non-model species for which little or no genomic information is available, genome scans using AFLP molecular markers have proven to be a useful tool for detecting signatures of selection (Meudt and Clarke 2007). Compared to other marker systems, the AFLP protocol is relatively inexpensive, and can easily screen hundreds to

thousands of reliable polymorphic loci distributed across the entire genome (Bensch and Åkesson 2005). Indeed an increasing number of landscape genomic studies on non-model species have employed AFLP-based genome scans in parallel with environmental data to identify genomic regions potentially under selection as well as the corresponding ecological factors acting as drivers of selection (reviewed in Holderegger et al. 2008; Schoville et al. 2012). To date, the focus of many of these AFLP-based studies has been to identify broad-scale environmental variation (e.g. gradients of latitude, altitude, temperature, and precipitation) associated with patterns of adaptive genetic variation surveyed across relatively large spatial scales (e.g., Bonin et al. 2006; Joost et al. 2007; Manel et al. 2010b; Poncet et al. 2010; Keller et al. 2012; Bothwell et al. 2013) or to identify loci of adaptive relevance associated with contrasting habitat types (e.g., Campbell and Bernatchez 2004; Collin and Fumagalli 2011; Buckley et al. 2012). In contrast, relatively little research has sought to explore how fine-scale landscape heterogeneity can contribute to patterns of adaptive genetic variation.

Relict species represent a natural laboratory for investigating the influence of landscape heterogeneity on local adaptive processes. Most relict species are habitat specialists, and possess limited dispersal abilities, thus, they typically occur as discrete populations in small and isolated habitat patches which may experience largely independent evolutionary processes and trajectories (Spitzer and Danks 2006; Habel et al. 2010; Hampe and Jump 2011). Landscape heterogeneity, defined as variation in the size, quality and connectivity of suitable habitat patches, exposes these isolated populations to different ecological selection pressures through which local adaptations may evolve (Kawecki and Ebert 2004). Strong ecological selection generally acts more rapidly in

small, isolated populations (Maynard Smith 1976; Habel et al. 2010); however, if selective pressures are weak, the negative effects of genetic drift associated with small, isolated populations may overwhelm selection and result in an overall loss of adaptive genetic variation (Nei et al. 1975). Furthermore, small populations which experience fluctuating demographics are also likely to lose adaptive genetic variation over time through genetic drift (Kawecki and Ebert 2004). Thus, signatures of selection may be less evident in relict populations as a result of small population demography. Nevertheless, relict populations are well-known for exhibiting adaptations to local ecological conditions (Kawecki and Ebert 2004) and thus represent a candidate system to study the effects of landscape heterogeneity on patterns of adaptive genetic variation at a fine spatial scale.

Here we use the bog copper butterfly (*Lycaena epixanthe*, Boisduval and Le Conte 1835), a relict species endemic to temperate Nearctic peatlands, as a model to explore the relationship between fine-scale landscape heterogeneity and adaptive genetic variation. While the bog copper's distribution is thought to have been much more widespread prior to the Pleistocene glaciation (Wright 1983), today the butterfly occurs across eastern North America, from Manitoba east to St. John's Newfoundland in Canada, and as far south as Virginia in the United States, in areas with relatively cool and humid microclimatic conditions (Layberry et al. 1998; Hampe and Jump 2011). The bog copper is a strict habitat specialist as both the larvae and adults feed exclusively on the bog cranberry (*Vaccinium macrocarpum* and *V. oxycoccos*) which typically only grows in acidic peatlands (Wright 1983; Cech and Tudor 2005). Adult butterflies are very small in size (wingspan = 17- 22 mm) and exhibit limited flight capabilities (Wright 1983; Cech and Tudor 2005), which, together with their close habitat association spatially restricts

the distribution of the bog copper to discrete habitat patches. Even at a small spatial scale, however, the ecological conditions within different peatland habitats (e.g., host-plant availability and quality, moisture content, and exposure) and the composition of the landscape surrounding each habitat patch can be extremely diverse, and populations occupying different peatlands are likely exposed to different selection pressures as a result of this heterogeneity.

In a previous population genetic study of the bog copper butterfly using AFLP molecular markers (Chapter 4), we demonstrated that nearby populations were genetically differentiated (significant pairwise F_{ST} values and no pattern of isolation by distance), suggesting limited gene flow among populations at a very small spatial scale. We also found that population genetic diversity was negatively associated with peatland size, and positively with the amount of open water surrounding the studied peatland. We hypothesized that these landscape variables may reflect differences in microclimatic conditions among habitat patches (e.g., host-plant distribution, water availability and thermal regimes) which are known to affect population viability in other butterfly species (e.g., Turlure et al. 2010a; 2013). In another study (Chapter 5) we examined morphological variation associated with flight ability among local populations of the bog copper and found that individual phenotypes varied in relation to both the surrounding landscape structure and local patch characteristics. In particular, increasing forest cover was associated with decreased investment into flight in both males and females. As well, the relative allocation to flight in males corresponded positively to peatland size and amount of open water in the surrounding area, suggesting that male mate-location strategies may differ among populations as a result of differences in local ecological

conditions. Although we were unable to disentangle the underlying mechanism responsible for the observed morphological variation (i.e., genetic adaptation vs. phenotypic plasticity), these results as well as those of the former study lend support to the hypothesis that fine-scale landscape heterogeneity (i.e., both local patch characteristics and surrounding landscape structure) may be imposing selective pressures on populations of this relict butterfly species.

In the present study, we used an AFLP-based genome scan to identify putative candidate outlier loci under divergent selection. We then used a multimodel inference approach to test for associations between allele frequencies of these candidate loci and the fine-scale landscape variables previously identified as important predictors through population genetic and morphological analyses (Chapters 3 and 4). This allowed us to evaluate the relative influence of each landscape variable as a potential selective force shaping patterns of genetic differentiation across the genome of the bog copper.

6.2 Methods

6.2.1 Study sites and data collection

We surveyed a total of 551 bog copper butterflies (16 - 43 individuals from each site) collected from 15 discrete peatland sites within Algonquin Provincial Park, Ontario, Canada during July 2009 and 2010 (Table 6.1; Figure 6.1). The samples from 13 of these locations were evaluated in Chapter 4 for AFLP-based population genetic analyses (Table 4.1 in Chapter 4). We chose peatland sites that varied in total size and relative isolation. For each study site, we used four uncorrelated landscape variables previously described

in Chapters 4 and 5 which encompass measurements of local habitat patch area and quality, as well as surrounding landscape composition and isolation: (1) the geographic area of each sampled peatland, which can also be considered as a proxy for host-plant density and distribution; (2) the area proportion of forest; and (3) open water habitat; and (4) the mean proximity of wetland habitat. Landscape variables 2-4 were evaluated within a 1 km buffer radius of the centroid of butterfly capture in each surveyed peatland. As outlined in Chapters 4 and 5 the extraction of all landscape variables was performed using ArcGIS v.10.0 (ESRI, Redlands, California) and FRAGSTATS v.4.0 (McGarigal et al. 2002).

As described in Chapters 3 (Crawford et al. 2011) and 4 we employed a non-lethal method (Koscinski et al. 2011; Crawford et al. 2013) to collect tissue samples for genetic analyses. Adult butterflies were captured using hand-nets and a small piece of tissue ($\sim 0.1 \text{ cm}^2$) was removed from both hind-wings using fine iris scissors. The wing tissue ('wing-clip') was immediately stored in absolute ethanol, and the butterfly marked and released from its initial point of capture.

DNA was extracted from each wing-clip sample using the DNeasy® Blood and Tissue Kit (QIAGEN, Germantown, MD), following the methods described in Chapter 4. We then used a modified AFLP protocol (outlined in Chapter 4 and Appendix A) to generate unique AFLP profiles for each bog copper individual based on five selective primer combinations. Non-normalized AFLP profiles were visualized and sized using GENEMAPPER v.4.0 (Applied Biosystems) based on a size standard ladder (LIZ-500). As before, we excluded any AFLP fragments falling within user-defined size 'bins' that were less than 100 base pairs in size and/or less than 100 relative fluorescent units in

height, in order to minimize the occurrence of size homoplasy (Vekemans et al. 2002) and artefactual instrument noise in our dataset. The locations of all bins assigned by GENEMAPPER were manually confirmed and adjusted if necessary (outlined in Chapter 4) and we manually reviewed the AFLP profile for every individual and removed any profiles that harboured multiple unique fragments or failed to properly amplify. Following the same procedure described in Chapter 4 we then used AFLPScore v.1.3b (Whitlock et al. 2008) to normalize and score the peak-height data for the AFLP profiles of all individuals surveyed across the 15 peatlands (i.e., we conducted an independent AFLPScore analysis on this updated AFLP dataset which included sites COS and SUN). The mismatch genotyping error rate of the AFLP dataset was determined by AFLPScore based on 35 replicate samples (~6 % of total sample size). Replicate individuals were selected at random, and represent two aliquots of the same DNA extraction which have separately undergone the whole AFLP procedure.

Previously in Chapter 4, we compared the results of population genetic analyses performed using analogous AFLP datasets varying in overall genotyping error rate (2 – 5 %) and established that a mismatch genotyping error rate of approximately 4 % was optimal for our study system. Therefore, in the current study, we used AFLPScore to identify the optimal locus and phenotype scoring thresholds which would allow us to achieve a mismatch error rate of approximately 4 % while retaining the maximal number of loci possible. To limit the occurrence of false positives in the outlier detection analyses, we also removed all monomorphic loci as well as any loci with a minor allele frequency less than 0.05 (as recommended by Foll and Gaggiotti 2008).

Table 6.1 Collection record of the 15 studied populations of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada. Individuals were surveyed in July 2009 and 2010, and the coordinates for each peatland are measured as the centroid of butterfly capture. Both the number of individuals initially collected ($n_{\text{collected}}$) and the number of individuals successfully amplified and phenotyped for AFLP analyses (n_{AFLP}) are presented.

Code	Peatland	Year	UTM Coordinates (17N)		$n_{\text{collected}}$	n_{AFLP}
			Easting	Northing		
BUG	'Buggy'	2009	679515.01	5049067.38	25	25
WH	Wolf Howl Pond	2009	680263.51	5049909.77	41	41
DL	Dizzy Lake	2009	680239.50	5046830.66	40	40
ML	Mizzy Lake	2009	681141.86	5047379.28	43	42
WR	West Rose Lake	2009	680935.89	5049256.81	38	38
KB	'Kearney'	2009	698978.67	5050431.11	39	37
MIN	Minor Lake	2009	701460.15	5057456.41	40	40
BAB	Bab Lake	2009	701436.90	5055911.00	40	40
ZEN	Zenobia Lake	2009	701883.80	5055742.06	39	38
SUN	Sunday Creek	2009	705574.82	5051097.38	41	41
SB	Spruce Bog	2009	705183.73	5052048.05	16	16
EOS	Eos Lake	2009	706133.08	5052006.34	40	40
COS	Costello Creek	2010	707014.15	5054251.22	27	27
OPL	Opeongo Lake	2009	706267.58	5056808.38	41	41
DT	D. Thompson Lake	2009	712915.21	5044885.79	40	39

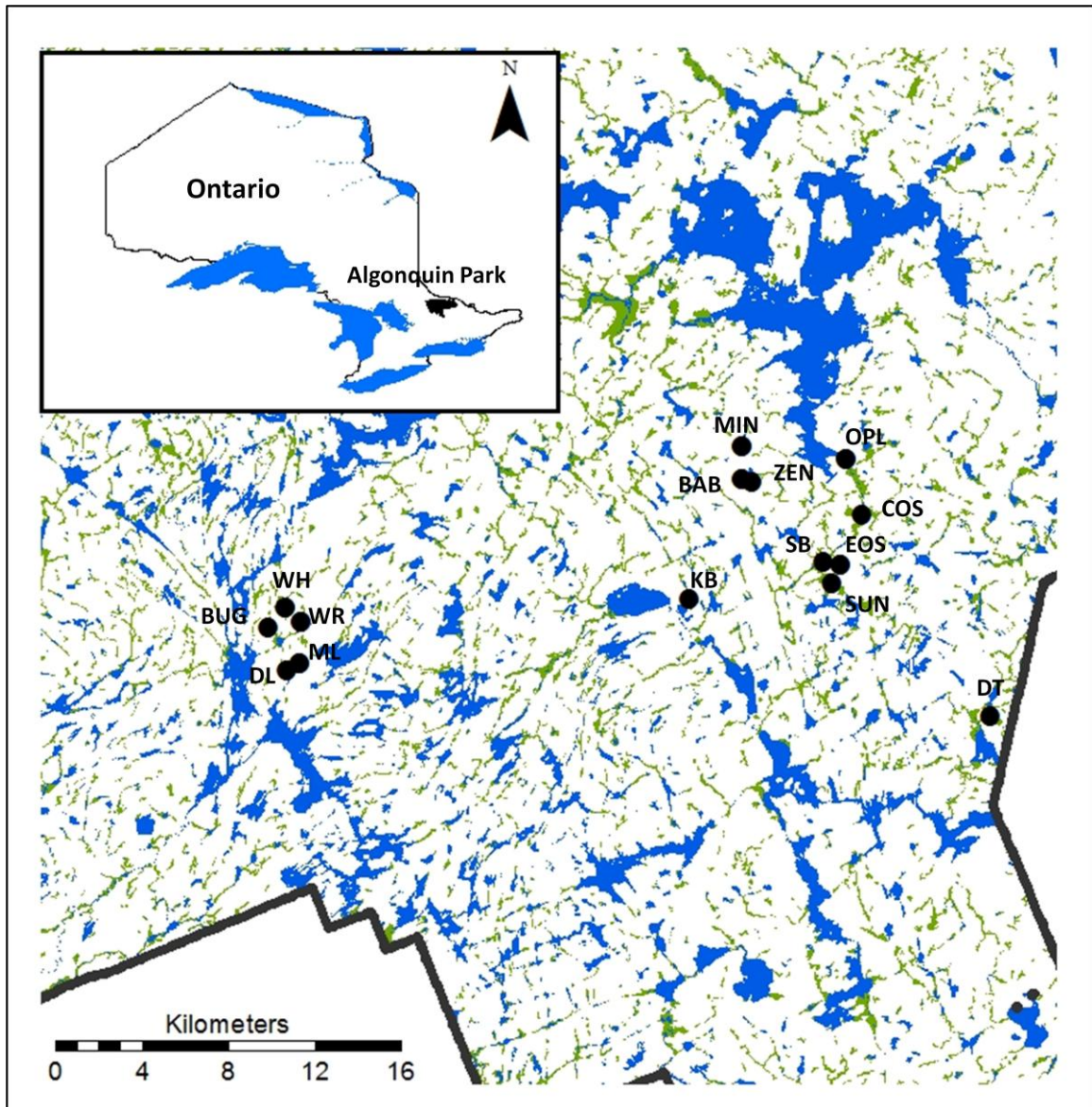


Figure 6.1 Sampling locations of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada. The predominant land cover types within the study system are indicated: forest, white; open water, blue; and wetlands, green. The inset shows the location of Algonquin Provincial Park in Ontario.

6.2.2 Detecting potentially adaptive loci

Previous genetic structure analysis of the bog copper within the same study system (Chapter 4) established that individuals occupying different peatland habitat patches were essentially genetically isolated. Thus we considered each peatland as an independent population in outlier detection analyses. To identify AFLP loci exhibiting signatures of divergent selection (greater differentiation among populations than expected under a neutral model of evolution) we used two F_{ST} outlier-based programs: (1) DFDIST (Beaumont and Balding 2004), as implemented in the workbench Mcheza (Antao and Beaumont 2011) which uses a frequentist inference method, and (2) BayeScan v.2.1 (Foll and Gaggiotti 2008) which employs a Bayesian inference method. Identifying loci as putative outliers via two different algorithms allows for a more rigorous test for signatures of selection (Perez-Figueroa et al. 2010; Narum and Hess 2011).

DFDIST compares observed population differentiation (F_{ST}) coefficients at each locus to a null distribution generated by coalescent simulations under a classical island model (i.e., drift-migration equilibrium) and identifies loci which display comparatively high levels of differentiation. We used 50 000 loci to model the null distribution, with a mean neutral F_{ST} calculated from a trimmed version of the observed dataset with putative outlier loci removed. Employing a 5% significance level, we considered those AFLP loci exhibiting observed F_{ST} values greater than the 95% upper quantile of the null distribution as putatively adaptive loci experiencing divergent selection. To correct for multiple testing we also set a false discovery rate (FDR) of 10% (Benjamin and Hochberg 1995). The simulated neutral distribution generated by DFDIST has been shown to be robust to mild departures from the simple island model (Beaumont 2005).

BayeScan, in contrast, considers all of the data simultaneously and directly estimates the probability of a locus being under selection. BayeScan is considered a more conservative outlier test that is less likely to detect false positives, because it does not assume a simple island model, allowing effective population sizes and the amount of genetic drift between populations to vary (Perez-Figueroa et al. 2010), and because it directly accounts for the issue of multiple testing (Foll and Gaggiotti 2008). Two alternative models of differentiation are defined for each locus, a selection-based model and a neutral-based model. A reversible jump Markov Chain Monte Carlo (RJ-MCMC) approach is then used to estimate the posterior probability for each model. As recommended by Foll and Gaggiotti (2008), we employed the default values for the RJ-MCMC algorithm parameters: 5000 iterations conducted for 20 pilot runs (i.e., total iterations = 100 000), and to reduce the occurrence of false positives we used conservative prior odds of 10:1 in favour of a neutral model. In Bayesian statistics posterior probabilities are used for model choice decision, although they cannot be directly interpreted or compared to classical *P*-values such as those calculated in DFDIST (Foll and Gaggiotti 2008). Instead, we used a posterior probability threshold > 0.76 to identify putative candidate outlier loci demonstrating ‘substantial’ evidence for selection according to Jeffreys’ scale (Jeffreys 1961; Foll and Gaggiotti 2008).

6.2.3 Identifying landscape variables associated with putative outlier loci

The loci identified to be potentially under directional selection were used in separate multiple linear regression analyses to explore whether variation in allele frequencies at

each locus could be explained by the landscape factors examined. The frequency of individuals scored for the presence allele at a given candidate locus in each population was used as the response variable and was logit transformed. As predictor variables we used log-transformed measures of peatland size and mean proximity of wetland habitat, and logit transformed measures of proportion of forest habitat and open water habitat. Prior to analyses, all explanatory variables were standardized by subtracting the mean from each observation and dividing by the standard deviation. We applied a multimodel inference approach (Burnham and Anderson 2002) which allowed us to determine the best model explaining allele frequency variation among populations, as well as identify which landscape predictors (if any) were most influential for each locus. Candidate models composed of all 16 possible combinations of explanatory variables were ranked according to Akaike information criterion values corrected for small sample size (AICc; Burnham and Anderson 2002). All models within 2 AICc values of the top-ranked model (ΔAICc values < 2) were considered to be equally well supported, top candidate models (Burnham and Anderson 2002). The relative importance of each explanatory variable (w_{+i}) then was evaluated based on its Akaike model weight (w_i) calculated through model averaging of the top candidate models. Model selection and averaging was performed using the R package MuMIn v.1.9.5 (K Bartoń, available at <http://mumin.r-forge.r-project.org>), and all statistical analyses were conducted using R statistical software version 3.0.0 (R Development Core Team 2013; R Foundation for Statistical Computing, Vienna, Austria).

6.3 Results

6.3.1 AFLP analysis and phenotype scoring

We successfully scored the AFLP phenotypes for a total of 545 individuals representing 15 bog copper populations (Table 6.1). Six individuals repeatedly failed to amplify or had many unique alleles and were therefore excluded from further analyses. Following genotyping error analysis in AFLPSCORE, the resulting dataset consisted of 226 repeatable AFLP loci with a mean mismatch error rate of 3.76 % across the five selective primer combinations (Table 6.2). To limit the occurrence of false positives, only polymorphic loci with a minor allele frequency ≥ 0.05 were retained for outlier detection analyses. Thus the final AFLP dataset consisted of 130 loci (Table 6.2).

6.3.2 Outlier locus detection

Applying two different F_{ST} outlier-based approaches to our AFLP dataset, we detected multiple loci exhibiting notably high levels of differentiation among populations suggestive of divergent selection. DFDIST identified eight AFLP loci as putative outliers above the 95 % upper quantile of the neutral distribution (Table 6.3). However, after controlling for multitest correction (FDR = 10 %) none of the loci remained significant. BayeScan identified five loci exhibiting substantial support (posterior probabilities > 0.76) for divergent selection according to Jeffrey's scale (Jeffreys 1961; Table 6.3). The corresponding FDR (q -value) for all loci was equal or lower than 11 % (Table 6.3). This q -value threshold is considerably more rigorous than a classic P -value threshold of 11% (Foll and Gaggiotti 2008). Overall, five different candidate loci were consistently

detected by both BayeScan and DFDIST (before multi-test correction): loci 34, 54, 87, 94 and 100 (Table 6.3). We considered these five loci to be potentially under divergent selection and used each in a separate multiple regression analysis to test for associations with landscape variables.

6.3.3 Associations between landscape variables and putative outlier loci

Using an AIC_c inferential model selection approach our results suggest that all of the landscape variables were associated with at least one putative outlier locus (Table 6.4). Allele frequencies of four of the five loci were associated with variation in peatland size, while proportion of open water also explained allele frequency variation for three of the five loci (Table 6.5). As well, our data provide substantial support for an influence of proportion of forest cover and moderate support for an influence of wetland proximity on allele frequencies at locus 54 (Table 6.5).

Table 6.2 Details of the five selective primer combinations used and their contribution to the final AFLP dataset. The results of the phenotype scoring and mismatch error rate analysis performed using AFLPScore are presented, with the scoring parameters (locus and phenotype scoring thresholds) used for each primer combination indicated. We present the number of AFLP loci initially identified using GENEMAPPER, the number of loci retained by AFLPScore, and the final number of loci used in outlier detection analyses following the removal of loci with minor allele frequencies < 0.05.

Selective primer combination	Scoring Threshold		Mismatch error rate %	Number of AFLP loci		
	Locus	Phenotype		GENEMAPPER	AFLPScore	Final dataset
<i>EcoRI-ACA/MseI-CTT</i>	600	100	3.74	161	39	22
<i>EcoRI-AAC/MseI-CAC</i>	1400	100	3.95	176	35	16
<i>EcoRI-AAC/MseI-CTC</i>	1100	400	3.94	121	43	21
<i>EcoRI-AAG/MseI-CAA</i>	600	200	3.51	123	69	50
<i>EcoRI-ACC/MseI-CAC</i>	2800	1300	3.68	123	40	21
			3.76 (mean)	704 (total)	226 (total)	130 (total)

Table 6.3 AFLP loci identified as putatively under divergent selection for *Lycaena expixanthe*. The results from DFDIST (P -value, F_{ST}) and BayeScan (posterior probability, q -value, F_{ST}) are presented for each candidate locus.

Locus no.	DFDIST*		BayeScan**		
	P -value	F_{ST}	Posterior probability	q -value	F_{ST}
34	0.990	0.143	0.866	0.094	0.155
54	0.994	0.163	0.803	0.111	0.142
62	0.983	0.129	0.484	0.327	0.108
87	0.995	0.163	0.914	0.061	0.152
91	0.979	0.143	0.405	0.351	0.104
94	0.994	0.162	0.964	0.036	0.169
100	0.970	0.129	0.877	0.084	0.155
130	0.998	0.190	0.073	0.525	0.071

We used an AFLP dataset consisting of 130 loci to detect putative candidate outlier loci (locus no.). Bolded values indicate loci which met criteria for outlier detection: DFDIST = $P < 0.05$; BayeScan = posterior probability > 0.76 .

*After controlling for multitest correction (false discovery rate = 10 %) all loci initially identified by DFDIST were no longer significant.

**Note that q -value thresholds calculated in BayeScan are considerably more stringent than the equivalent P -value threshold used in classical statistics.

Table 6.4 Summary of model selection results for the five candidate outlier loci identified to be under divergent selection by DFDIST and BayeScan. Models with $\Delta_i\text{AIC}_c < 4$ are presented, and are ranked according to AIC_c values.

Locus no.	Model	logLik	AIC_c	$\Delta_i\text{AIC}_c$	w_i
34	Intercept	-24.309	53.6	0.00	0.429
	Area	-23.828	55.8	2.22	0.142
	Wat	-24.033	56.2	2.63	0.115
	Prox	-24.226	56.6	3.01	0.095
	For	-24.309	56.8	3.18	0.088
54	Area+For+Prox	18.671 5	4.0	0.00	0.260
	For	23.113 5	4.4	0.40	0.213
	Area+For	21.719 5	5.4	1.43	0.127
	For+Prox	21.742 5	5.5	1.48	0.124
	Intercept	25.359 5	5.7	1.71	0.111
87	Area+Wat	-18.609	49.2	0.00	0.284
	Intercept	-22.624	50.2	1.03	0.169
	Wat	-21.086	50.4	1.13	0.161
	Area	-21.418	51.0	1.80	0.115
	For+Wat	-20.469	52.9	3.72	0.044
	For	-22.486	53.2	3.94	0.040
94	Area+Wat	20.712 5	3.4	0.00	0.28
	Wat	22.827 5	3.8	0.41	0.228
	Intercept	25.239 5	5.5	2.05	0.100
	Area+Wat+Prox	19.410 5	5.5	2.06	0.100
	Wat+Prox	22.027 5	6.1	2.63	0.075
	Area	24.472 5	7.1	3.70	0.044
	For+Wat	22.665 5	7.3	3.91	0.040
100	Wat	-22.204	52.6	0.00	0.292
	Area+Wat	-20.341	52.7	0.09	0.279
	Intercept	-24.763	54.5	1.94	0.111
	For+Wat	-21.591	55.2	2.59	0.080
	Wat+Prox	-22.114	56.2	3.64	0.047
	Area	-24.137	56.5	3.87	0.042

Log likelihood statistics ($\log\text{Lik}$), second order Akaike information criterion (AIC_c), $\Delta_i\text{AIC}_c$, and Akaike weights (w_i) are derived from linear multiple regressions. *Area*, total peatland area; *For*, proportion of forest habitat, *Wat*, proportion of open water habitat; *Prox*, mean proximity of potentially suitable wetland habitat.

Table 6.5 The relative importance of landscape variables acting as potential drivers of genetic variation in the five candidate outlier loci for *Lycaena epixanthe*. For each locus, model-averaged AIC_c weights ($w_{+}(i)$) derived from multiple linear regressions after model selection are presented.

Locus no.	Peatland size	Forest cover	Water cover	Proximity
34				
54	0.464	0.868		0.460
87	0.547		0.609	
94	0.551		1.000	
100	0.409		0.837	

6.4 Discussion

We applied an AFLP-based genome scan approach to detect signatures of selection among local populations of a relict species, the bog copper butterfly, and used AIC_c model selection to assess the relationship between landscape variables and candidate outlier loci. While most landscape genomic studies to date have examined patterns of adaptive genetic variation surveyed across much larger spatial scales (e.g., Bonin et al. 2006; Joost et al. 2007; Manel et al. 2010b; Poncet et al. 2010; Keller et al. 2012; Bothwell et al. 2013) or between contrasting habitat types (e.g., Campbell and Bernatchez 2004; Collin and Fumagalli 2011; Buckley et al. 2012), our study provides evidence for natural selection influencing populations at a relatively fine spatial scale due to landscape heterogeneity. Using two different F_{ST} -based outlier detection programs we identified five different loci potentially under divergent selection, representing 3.8 % of the total loci surveyed. Allele frequency variation at four of these five loci was associated with landscape variables, particularly peatland size and the proportion of open water habitat in the surrounding landscape. Despite the potential for small population size to overwhelm the effects of selection, our results provide molecular evidence of local adaptation in the bog copper, and suggest that patterns of adaptive genetic differentiation among populations are indeed influenced by fine-scale landscape heterogeneity. Our study thus demonstrates the utility of relict species as a model system for investigating the influence of landscape heterogeneity on local adaptive processes.

6.4.1 Detecting potentially adaptive loci using an F_{ST} -based outlier approach

Using a conservatively generated AFLP dataset, consisting of only highly repeatable, polymorphic loci, we made efforts where possible to control for factors that could potentially increase the probability of type I error in the statistical outlier detection tests. In particular, during our initial selection of AFLP loci, we limited the occurrence of AFLP size homoplasy (Caballero et al. 2008), we evaluated population genetic structure prior to outlier analyses (Excoffier et al. 2009) and we accounted for multiple comparisons (Perez-Figueroa et al. 2010). Outlier loci were then classified using DFDIST and BayeScan following stringent criteria, and we used the results of these two programs in combination to further reduce the risk of using false outlier loci. All five of the putative adaptive loci identified by BayeScan were also identified by DFDIST (before multi-test correction), indicating considerable concordance between the two programs. When compared under a range of different simulated scenarios (i.e., varying values of mean neutral F_{ST} , mean selection coefficients and proportion of true selective loci) BayeScan has been shown to perform more efficiently than DFDIST after multi-test correction, particularly in fully neutral and low selection scenarios (Pérez-Figueroa et al. 2010). However, false signatures of evolutionary divergence can be generated by spatial and historical effects (Kawecki and Ebert 2004; Schoville et al. 2012), and it is recommended that the results of any outlier detection analyses be cautiously evaluated with this issue in mind (Manel et al. 2010a; Perez-Figueroa et al. 2010; Narum and Hess 2011). For example, if gene flow is spatially restricted among populations (i.e., isolation by distance, IBD), genetic drift could cause neutral alleles to change in frequency (Wright 1943). However, the effects of drift should be apparent across the entire genome

rather than at specific loci. As well, spatial genetic structure among populations can result in correlated allele frequencies, which if hidden and/or not accounted for, can generate a large number of false positive loci in outlier tests (Excoffier et al. 2009). Furthermore, past demographic events such as population bottlenecks could also increase genetic drift and population differentiation resulting in false outliers (Holderegger et al. 2008).

BayeScan takes demographic effects into account when modelling population divergence (Foll and Gaggiotti 2008), and we accounted for population structure and the presence of IBD prior to the population-level outlier analyses. Thus, we can be confident that the five outliers identified by both BayeScan and DFDIST represent differentiation caused by natural selection rather than neutral evolutionary processes.

6.4.2 Identifying landscape variables affecting potentially adaptive loci

AIC_c analyses indicated that the proportion of open water habitat, as well as peatland size, most strongly influenced the allele frequencies of the candidate loci. The proportion of forest cover and mean proximity of wetland habitat, both indicators of habitat fragmentation and isolation, were found to associate with one of the selected loci (locus 54). Interestingly both peatland size and proportion of open water were also found to be the best predictors of variation in genetic diversity, and measurements of male flight-morphology among populations of the bog copper (Chapters 4 and 5). They likely reflect variation in local ecological conditions (e.g. host-plant availability, moisture content, and exposure) within each peatland habitat patch which could be responsible for the putative adaptive genetic variation observed here.

Differences in host-plant distribution and microclimatic conditions are known to affect female oviposition behaviours, male mate-location strategies and overall habitat use in butterflies (e.g., Merrill et al. 2008; Ashton et al. 2009; Krämer et al. 2012) which could induce evolutionary changes in traits such as flight ability, thermal regulatory ability and fecundity (Shreeve 1986; Berwearts et al. 1998; Vandewoestijne Van Dyck 2011; Gibbs et al. 2012; Vande Velde et al. 2013). As well, most peatland-associated species are thought to be cold-adapted (Addo-Bediako et al. 2002; Spitzer and Danks 2006), and indeed a recent experimental study of the cranberry fritillary butterfly (*Boloria aquilonaris*), a Palearctic peatland specialist, found that larval survival was greater under cooler environmental conditions (Turlure et al. 2010a). *Sphagnum* moss which carpets peatland floors is an important thermoregulator (Spitzer and Danks 2006) and particularly *Sphagnum* hummocks which tend to be located within peatlands in early successional or humid areas (Turlure et al. 2010a, b), provide cool, moist environments which buffer against fluctuating air temperatures (Turlure et al. 2010a). Differences among sites in the availability and quality of these thermal refuges may contribute to different thermal tolerances among populations. Furthermore, these hummocks may also serve as important oviposition sites, as both the eggs and larvae of the bog copper are known to be sensitive to desiccation (Wright 1983) and thus individuals may be locally adapted to differences in water availability.

In Chapter 5, we found evidence for a negative relationship between proportion of forest cover and relative thorax mass in male and female bog coppers, suggesting a reduction in flight capability in response to increasing habitat isolation. Changes in flight morphology (indicative of a loss in dispersal propensity) in response to habitat

fragmentation have similarly been documented for other butterfly species (Dempster 1991; Schtickzelle et al. 2006), and based on common garden experiments several lines of evidence now suggest a genetic basis for these morphological differences (Thomas et al. 1998; Hill et al. 1999a, b; Merxck et al. 2003). In this study we found substantial support for forest cover influencing genetic variation at locus 54, suggesting that indeed the proportion of forest in the surrounding landscape may be an important selective force affecting populations of the bog copper. It is important to note that we screened a relatively small proportion of the genome (i.e., < 150 loci) here, which certainly limited the number of candidate outliers we were able to detect. Screening hundreds or thousands of loci would potentially allow us to identify additional outlier markers also associated with surrounding landscape structure.

6.4.3 Conclusion and future perspectives

Many relict habitat specialists, including the bog copper butterfly, are currently under threat throughout much of their range and are at risk of extinction due to ongoing anthropogenic induced environmental change. Understanding how local landscape heterogeneity contributes to the evolution of populations will allow us to better predict how species may respond to future habitat change. In this study, we used candidate AFLP markers identified using an outlier-based genome scan approach, to characterize potential local adaptation among isolated populations of a non-model species and identify several measurements of landscape heterogeneity which may be acting as causal agents of selection. Using anonymous AFLP markers of unknown sequence content does not provide conclusive evidence that the genetic variation observed at these loci affords a

selective advantage (Nunes et al. 2012). Thus, in order to confirm the role of these candidate loci in local adaptation, the functional mechanisms operating at these loci need to be determined (Reusch and Wood 2007; Holderegger et al. 2010). This can be accomplished through a combination of molecular characterisation and hypothesis-driven selection experiments (e.g., common garden; Holderegger et al. 2008). Determining the functional significance associated with genetic variation at outlier loci would allow us to examine how tolerable different bog copper populations are to changes in ecological conditions. Such information is pertinent for ensuring the long-term viability of this highly specialized species, especially for those populations experiencing habitat change. As well, the ability to accurately identify populations harbouring distinct signatures of local adaptation will be invaluable for the selection of source populations for reintroduction or supplementation of threatened populations (Thomas 2011; Turlure et al. 2013). This may be particularly challenging for relict species where populations are often locally adapted. Indeed, our results highlight that even at a relatively small spatial scale, local populations exhibited evidence of genetic adaptations potentially due to habitat patch-specific ecological conditions.

6.5 Literature Cited

- Addo-Bediako A, Chown SL and Gaston KJ. 2002. Metabolic cold adaptation in insects: a large-scale perspective. *Functional Ecology*, 16: 332-338.
- Antao T and Beaumont MA. 2011. Mcheza: a workbench to detect selection using dominant markers. *Bioinformatics*, 27: 1717-1718.
- Ashton S, Gutiérrez D and Wilson RJ. 2009. Effects of temperature and elevation on habitat use by a rare mountain butterfly: implications for species responses to climate change. *Ecological Entomology*, 34: 437-446.
- Beaumont MA. 2005. Adaptation and speciation: what can F_{ST} tell us? *Trends in Ecology and Evolution*, 20: 435-440.
- Beaumont MA and Balding DJ. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13: 969-980.
- Benjamin Y and Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57: 289-300.
- Bensch S and Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.
- Berwaerts K, Van Dyck H, Van Dongen S and Matthysen E. 1998. Morphological and genetic variation in the speckled wood butterfly (*Pararge aegeria* L.) among differently fragmented landscapes. *Netherlands Journal of Zoology*, 48: 241-253.

- Bothwell H, Bisbing S, Overgaard Therkildsen N, Crawford L, Alvarez N, Holderegger R and Manel S. Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis* L.), using a landscape genetic approach. *Conservation Genetics*, 14: 467-481.
- Buckley J, Butlin RK and Bridle JR. 2012. Evidence for evolutionary change associated with the recent range expansion of the British butterfly, *Aricia agestis*, in response to climate change. *Molecular Ecology*, 21: 267-280.
- Burnham KP and Anderson DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. Springer-Verlag, New York.
- Caballero A, Quesada H, Rolan-Alvarez E. 2008. Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, 179: 539-554.
- Campbell D and Bernatchez L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, 21: 945-956.
- Cech R and Tudor G. 2005. Butterflies of the East Coast. Princeton University Press, Princeton, NJ.
- Collin H and Fumagalli L. 2011. Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (*Phoxinus phoxinus*, Cyprinidae). *Molecular Ecology*, 20: 4490-4502.

- Crawford LA, Desjardins S and Keyghobadi N. 2011. Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs. *Conservation Genetics*, 12: 991-1001.
- Crawford LA, Kosciński D, Watt KM, McNeil JN and Keyghobadi N. 2013. Mating and oviposition success of a butterfly are not affected by non-lethal tissue sampling. *Journal of Insect Conservation*. doi: 10.1007/s10841-013-9566-8.
- Dempster JP. 1991. Fragmentation, isolation and mobility of insect populations, in: Conservation of insects and their habitats, eds. Collins NM and Thomas JA. Academic Press Inc., San Diego, CA.
- Excoffier L, Hofer T and Foll M. 2009. Detecting loci under selection in a hierarchically structure population. *Heredity*, 103: 285-298.
- Foll M and Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180: 977-993.
- Gibbs M, Van Dyck H and Breuker CJ. 2012. Development on drought-stressed host plants affects life history, flight morphology and reproductive output relative to landscape structure. *Evolutionary Applications*, 5: 66-75.
- Habel JC, Schmitt T and Assmann T. 2010. Relict species research: some concluding remarks, in: Relict species: phylogeography and conservation biology, eds. Habel JC and Assmann T. Springer-Verlag, Berlin.

- Hampe A and Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics*, 42: 313-333.
- Hill JK, Thomas CD and Blakeley DS. 1999a. Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121: 165-170.
- Hill JK, Thomas CD and Lewis OT. 1999b. Flight morphology in fragmented populations of a rare British butterfly, *Hesperia comma*. *Biological Conservation*, 87: 277-283.
- Holderegger R, Herrmann D, Poncet B, Gugerli F, Thuiller W, Taberlet P, Gielly L, Rioux D, Brodbeck S, Aubert S and Manel S. 2008. Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, 1: 273-283.
- Holderegger R, Bühler D, Gugerli F, Manel S. 2010. Landscape genetics of plants. *Trends in Plant Science*, 15: 675-683.
- Jeffreys H. 1961. The theory of probability, 3rd edn. Oxford University Press, New York.
- Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G and Taberlet P. 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16: 3955-3969.
- Kawecki TJ and Ebert D. 2004. Conceptual issues in local adaptation. *Ecology Letters*, 7: 1225-1241.

- Keller I, Schuler J, Bezault E and Seehausen O. 2012. Parallel divergent adaptation along replicated altitudinal gradients in Alpine trout. *BMC Evolutionary Biology*, 12: 210.
- Koscinski D, Crawford LA, Keller HA and Keyghobadi N. 2011. Effects of different methods of non-lethal tissue sampling on butterflies. *Ecological Entomology*, 36: 301-308.
- Krämer B, Kämpf I, Enderle J, Poniatowski D and Fartmann T. 2012. Microhabitat selection in a grassland butterfly: a trade-off between microclimate and food availability. *Journal of Insect Conservation*, 16: 857-865.
- Layberry RA, Hall PW and Lafontaine JD. 1998. The Butterflies of Canada. University of Toronto Press, Toronto.
- Lovejoy TE and Hannah L. 2005. Climate change and biodiversity. Yale University Press, New Haven, CT.
- Luikart G, England PR, Tallmon D, Jordan S and Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4: 981-994.
- Manel S, Joost S, Epperson BK, Hoderegger R, Storfer A, Rosenberg MS, Scribner KT, Bonin A and Fortin M-J. 2010a. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19: 3760-3772.

- Manel S, Poncet BN, Legendre P, Gugerli F and Holderegger R. 2010b. Common factors drive adaptive genetic variation at different spatial scale in *Arabidopsis alpina*. *Molecular Ecology*, 19: 3824-3835.
- Maynard Smith J. 1976. What determines the rate of evolution? *The American Naturalist*, 110: 331-338.
- Maynard Smith J and Haigh J. 1974. The hitchhiking effect of a favourable gene. *Genetic Research Cambridge*, 23: 23-35.
- Meudt HM and Clarke AC. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12: 106-117.
- Merrill RM, Gutiérrez D, Lewis OT, Gutiérrez J, Díez SB and Wilson RJ. 2008. Combined effects of climate and biotic interactions on the elevational range of a phytophagous insect. *Journal of Animal Ecology*, 77: 145-155.
- Merckx T, Van Dyck H, Karlsson B, Leimar O. 2003. The evolution of movements and behaviour at boundaries in different landscapes: a common arena experiment with butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 270: 1815-1821.
- Narum SR and Hess JE. 2011. Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11: 184-194.
- Nei M, Maruyama T and Chakraborty R. 1975. Bottleneck effect and genetic variability in populations. *Evolution*, 29: 1-10.

- Nunes VL, Beaumont MA, Butlin RK and Paulo OS. 2012. Challenges and pitfalls in the characterization of anonymous outlier AFLP markers in non-model species: lessons from an ocellated lizard genome scan. *Heredity*, 109: 340-348.
- Parmesan C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution and Systematics*, 37: 637-669.
- Pereira HM, Leadley PW, Proença V, Alkemade R, Scharlemann JPW, Fernandez-Manjarres JF, Araújo MB, Balvanera P, Biggs R, Cheung WWL, Chini L, Cooper HD, Gilman EL, Guénette S, Hurtt GC, Huntington HP, Mace GM, Oberdorff T, Revenga C, Rodrigues P, Scholes RJ, Sumaila UR and Walpole M. 2010. Scenarios for global biodiversity in the 21st century. *Science*, 330: 1496-1501.
- Pérez-Figueroa A, Garcia-Pereira MJ, Saura M, Rolan-Alvarez E and Caballero A. 2010. Comparing three different methods to detect selective loci using dominant marker. *Journal of Evolutionary Biology*, 23: 2267-2276.
- Poncet B, Herrmann D, Gugerli F. 2010, Taberlet P, Holderegger R, Gielly L, Rioux D, Thuiller W, Aubert S and Manel S. Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. *Molecular Ecology*, 19: 2896-2907.
- Reusch TBH and Wood TE. 2007. Molecular ecology of global change. *Molecular Ecology*, 16: 3973-3992.

- Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C and Manel S. 2012. Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution and Systematics*, 43: 23-43.
- Schtickzelle N, Mennechez G and Baguette M. 2006. Dispersal depression with habitat fragmentation in the bog fritillary butterfly. *Ecology*, 87: 1057-1065.
- Shreeve TG. 1986. Egg laying by the Speckled Wood Butterfly (*Pararge aegeria*) - The role of female behaviour, host plant abundance and temperature. *Ecological Entomology*, 11: 229-236.
- Spitzer K and Danks HV. 2006. Insect biodiversity of boreal peat bogs. *Annual Review of Entomology*, 51: 137-161.
- Storz JF. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, 14: 671-688.
- Thomas CD. 2011. Translocation of species, climate change and the end of trying to recreate past ecological communities. *Trends in Ecology and Evolution*, 26: 216-221.
- Thomas CD, Hill JK and Lewis OT. 1998. Evolutionary consequences of habitat fragmentation in a localized butterfly. *Journal of Animal Ecology*, 67: 485-497.
- Thuiller W, Albert C, Araújo MB, Berry PM, Cabeza M, Guisan A, Hickler T, Midgley GF, Paterson J, Schurr FM, Sykes MT and Zimmermann NE. 2008. Predicting global change impacts on plants species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, 9: 137-152.

- Turlure C, Choutt J, Baguette M and Van Dyck H. 2010a. Microclimate buffering and resource-based habitat in a glacial relict butterfly: significance for conservation under climate change. *Global Change Biology*, 16: 1883-1893.
- Turlure C, Choutt J, Van Dyck H, Baguette M and Schtickzelle N. 2010b. Functional habitat area as a reliable proxy for population size: case study using two butterfly species of conservation concern. *Journal of Insect Conservation*, 14: 379-388.
- Turlure C, Radchuk V, Baguette M, Meijrink M, van den Burg A, WallisDeVries M and van Duinen G-J. 2013. Plant quality and local adaptation undermine relocation in a bog specialist butterfly. *Ecology and Evolution*, 3: 244-254.
- Vande Velde V, Schtickzelle N and Van Dyck H. 2013. Effect of larval food stress on male adult behaviour, morphology and reproductive investment in the butterfly *Pararge aegeria*. *Evolutionary Ecology*, 27: 221-234.
- Vandewoestijne S and Van Dyck H. 2011. Flight morphology along a latitudinal gradient in a butterfly: do geographic clines differ between agricultural and woodland landscapes? *Ecography*, 34: 876-886.
- Vekemans X, Beauwens T, Lemaire M and Rodán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11: 139-151.
- Vitalis R, Dawson K and Boursot P. 2001. Interpretation of variation across marker loci as evidence of selection. *Genetics*, 158: 1811-1823.

- Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources* 8: 725-735.
- Willis KJ and Baghwat SA. 2009. Biodiversity and climate change. *Science*, 326: 806-807.
- Wright DM. 1983. Life history and morphology of the immature stages of the Bog Copper butterfly *Lycaena epixanthe* (Bsd. & Le C.) (Lepidoptera: Lycaenidae). *The Journal of Research on the Lepidoptera*, 22: 47-100.
- Wright S. 1943. Isolation by distance. *Genetics*, 28: 114-138.

Chapter 7

7 General Discussion

In heterogeneous landscapes, spatial variation in biotic and abiotic factors influence local and regional population dynamics, genetics and evolutionary adaptation (Stevens and Coulon 2012). For many species, recent anthropogenic land-use change has led to sudden and dramatic increases in landscape heterogeneity by fragmenting natural habitats into smaller, more isolated patches over a relatively short time period (Habel and Zachos 2012). Understanding the implications of increased habitat fragmentation for ecological and evolutionary dynamics of populations is one of the most pressing concerns in conservation research (Debinski and Holt 2000). Relict populations created via both anthropogenic and natural fragmentation processes are currently of high conservation concern (Habel et al. 2010), and determining which landscape components are most important for maintaining population viability in these two types of relict species is important for future species management.

While most ecological and genetic studies to date have considered the negative impacts of habitat fragmentation on recently isolated populations, comparatively little research has explored whether naturally fragmented populations exhibit the same response (Habel and Zachos 2012). Using butterflies as a model system, the overarching goal of my dissertation was to evaluate, for populations inhabiting fragmented landscapes, how patterns of genetic and morphological variation are spatially partitioned in relation to features of the landscape and local habitat. To this end, I studied relict populations which have been recently fragmented due to anthropogenic activities as well

as relict populations which are fragmented primarily due to natural processes. Taken together, the chapters of my dissertation provide valuable insight into how landscape history and species ecology can influence the genetics of relict populations.

7.1 AFLPs: an alternative marker system for assessments of neutral and adaptive genetic variation in butterflies

I successfully generated and analysed AFLP profiles for two different butterfly species using, for the first time, DNA extracted from non-lethal samples of butterfly wing tissue. Following two different AFLP protocols, a proprietary commercial protocol (ABI AFLP Plant Mapping Kit; Applied Biosystems, Foster City, CA; Chapter 3), as well as a modified AFLP protocol using standard PCR reagents (Chapters 4 and 6), I demonstrated that small pieces of wing tissue can provide sufficient quantities of DNA for AFLP generation. For both butterfly species, I surveyed populations within a relatively small geographic range (< 20 km, Mormon metalmark; <35 km, bog copper), and using AFLP datasets consisting of a few hundred loci was able to elucidate genetic signatures of population differentiation (Chapter 3 and 4). In addition, in Chapter 6, I successfully applied a genome-scan approach to the bog copper AFLP dataset and identified 5 outlier loci demonstrating unusually high genetic differentiation, providing evidence for divergent selection. These results demonstrate the utility of AFLPs as an alternative molecular marker system for genetic studies of butterflies and other non-model organisms. By generating a large number of multilocus markers which provided an estimate of genome-wide genetic variation (Meudt and Clarke 2007), my datasets were sufficiently powerful to resolve genetic differences among populations at a relatively fine

spatial scale. As well, by screening a large portion of the genome the likelihood of finding several markers linked to genomic regions under selection was high (Bensch and Åkesson 2005). Thus AFLPs are an extremely informative and versatile marker system which can be applied to address research questions in both population genetics and population genomics.

My review of the current AFLP literature in Chapter 2 revealed that published results for the majority of studies using AFLPs, are based on datasets in which the reproducibility of the AFLP loci has not been adequately assessed. As well, I found that a pervasive lack of consistency exists among AFLP studies in both the methods used to assess reproducibility and in the details of methodology presented. This is of concern because loci which exhibit high genotyping error are likely to contribute noise to a dataset, thereby reducing the dataset's resolution and power (Meudt and Clarke 2007). Furthermore, it has been theorized that genetic analyses generated using such inconsistent datasets may result in erroneous estimates of genetic diversity and population structure, or false detection of signatures of selection (Vekemans et al. 2002; Koopman and Gort 2004; Bonin et al. 2007; Caballero et al. 2008). For this reason it has been previously recommended that non-reliable loci (i.e., exhibiting relatively high genotyping error rates) be removed from the dataset prior to downstream analyses (Bonin et al. 2004; Pompanon et al. 2005). This in part ensures that the loci retained for genetic analyses are highly reproducible, and also allows the results among studies to be meaningfully evaluated and compared (Pompanon et al. 2005). My review of the literature has thus highlighted an important gap in the publishing standards of AFLP data in ecological and evolutionary research.

Given the problems with analysis and reporting of AFLP error which I identified in the literature (Chapter 2), I provided detailed accounts of the methodology used to generate, score and analyse the AFLP data presented in Chapters 3, 4 and 6. I followed the recommended guidelines for assessing and quantifying genotyping error (Bonin et al. 2004; Pompanon et al. 2005), and used an objective method (Whitlock et al. 2008) to identify those loci which contributed disproportionately to high error rates. Filtering out these ‘error-prone’ loci lowered the total number of loci retained for use in final genetic analyses, although arguably this step removed a considerable amount of noise from the datasets which provided for greater power to detect population structure. Indeed my ability to identify genetic differences among populations at relatively small spatial scales (Chapters 3 and 4) suggests that the information content of the AFLP datasets was extremely high. By transparently reporting the methodological details followed and the genotyping error associated with each AFLP dataset used, my data chapters can serve as models for other AFLP studies to follow.

7.2 Genetics of relict populations in fragmented landscapes

7.2.1 Relict populations in anthropogenically fragmented habitats

Contemporary habitat fragmentation as a result of anthropogenic activities has been shown to significantly hinder dispersal and gene flow among populations in a variety of taxa (reviewed in Saunders et al. 1991; Fahrig 2003). A reduction in gene flow can be detrimental to the long-term persistence of populations as it is gene flow which maintains genetic variation within populations by opposing the force of genetic drift, and

introducing potentially adaptive alleles (Segelbacher et al. 2010). Thus, one major goal of many species conservation plans is to maintain functional connectivity among remnant populations in order to promote gene flow and sustain levels of genetic diversity (Van Dyck and Baguette 2005). Evaluating patterns of neutral genetic variation within and among populations inhabiting fragmented landscapes, through the use of molecular genetic markers, allows us to infer patterns of gene flow and connectivity (e.g., Vandewoestijne and Baguette 2004; Schwartz and Karl 2005). These results can provide valuable insights into the genetic diversity and structure of fragmented populations which can be useful for conservation purposes, for example by identifying populations particularly vulnerable to extinction or, conversely, integral for the maintenance of the larger population network and therefore a top priority for protection (Keyghobadi 2007).

In Chapter 3, I found that sub-populations of the Mormon metalmark in the Similkameen Valley in British Columbia (BC), Canada, exhibited patterns of genetic diversity and structure highly characteristic of anthropogenically fragmented populations. Levels of genetic diversity were generally low within all sub-populations, and in particular, the most geographically isolated sites exhibited some evidence of decreased homozygosity in relation to the other sites. This was concordant with estimates of sub-population genetic differentiation, in which I found that the more geographically isolated sites were also the most genetically isolated (higher differentiation). Despite a small geographic range (<20 km), sub-populations exhibited a high degree of spatial genetic structure overall, indicating limited dispersal and restricted gene flow, particularly among sub-populations separated by urban development. These results suggest that the functional connectivity among sites is very low, and that increased gene flow is necessary

to ensure the long-term persistence of this population. Thus future management plans should target preserving existing habitat patches as well as creating or restoring additional patches which will facilitate the movement of individuals across the landscape.

It is important to note that based on my assessment of genome-wide genetic variation, I cannot infer whether the low levels of genetic diversity observed in small, isolated sub-populations of the Mormon metalmark correspond to adverse consequences for fitness (i.e., decreased survival and/or reproductive success) and long-term population viability. Ideally, a causal link between neutral or genome-wide genetic diversity and fitness could be demonstrated through common-garden breeding experiments examining key life-history traits such as female fecundity and offspring growth and survival. However, for endangered populations like the Mormon metalmark, such manipulations are not feasible because they would require the removal of a considerable proportion of the overall population. As an alternative, a correlative approach could be used to relate estimates of genetic variation with life history traits indicative of fitness (Fjerdingstad et al. 2007). For example, to demonstrate a relationship between neutral genetic diversity and fitness in a metapopulation of the chalk-blue butterfly (*Polyommatus coridon*), Vandewoestijne et al. (2008) used an estimate of adult lifetime expectancy as a measure of individual fitness. For many butterfly species, life span determines lifetime reproductive success and thus can be considered an important fitness component (Thornhill and Alcock 1983). Despite the inherent limitations of using a correlational approach, for threatened populations such as the BC population of the Mormon metalmark, examining relationships between genetic diversity and readily measured

fitness surrogates is a feasible approach that can provide insight into whether populations with lower genetic diversity suffer a loss in fitness.

7.2.2 Relict populations in naturally fragmented habitats –the importance of local habitat patch characteristics

Most studies examining the genetic implications of habitat fragmentation consider species which normally occupy interconnected habitats (Keyghobadi 2007), and focus on understanding the influence of patch size and isolation, and the characteristics of the intervening landscape, on patterns of population genetic diversity and structure (Holderegger and Wagner 2008; Storfer et al. 2010). While increasing habitat loss and isolation are expected to have negative consequences for many species (Fahrig 2003), those which occur in naturally patchy habitats may show a contrasting response to contemporary anthropogenic habitat fragmentation (Habel and Zachos 2012). Because these habitat specialists are thought to be adapted to persist under isolated conditions, changes in the surrounding landscape structure as a result of anthropogenic activities (i.e., increased habitat isolation and loss of suitable habitat) may have little impact on populations (Habel and Schmitt 2012; Habel and Zachos 2012). Instead, the dynamics and long-term persistence of populations are more likely to be dependent upon local habitat characteristics (e.g., habitat patch size and quality). In Chapter 3, I provide empirical evidence in support of this prediction. Multiple linear regression analyses revealed that a significant proportion of variation in genetic diversity within local populations of the bog copper butterfly was explained by two habitat patch descriptors: total peatland patch size and proportion of water in the surrounding landscape, likely

indicators of local habitat quality (i.e., host-plant quality and distribution, temperature and moisture regimes). In contrast, I found descriptors of habitat isolation (mean proximity of wetland habitat) and landscape composition (proportion of forest habitat) to have no significant effect on diversity estimates. These results indicate that for the bog copper butterfly, local habitat conditions rather than surrounding landscape structure may be the main determinants of population dynamics and genetics. My study highlights the importance of considering species ecology and habitat history when predicting population viability in fragmented landscapes (Ye et al. 2013). For specialist species which are naturally fragmented, the most effective strategy to ensure long-term persistence may be the management and restoration of remaining suitable habitat patches (Habel and Zachos 2012). The genetic consequences of changes in habitat quality have largely been underexplored to date (but see Porlier et al. 2009; Pitra et al. 2011; Alda et al. 2013). However, as my results indicate, for specialist species it is particularly critical that future work be conducted in order to understand the relationship between habitat quality, and genetic diversity and fitness.

While naturally fragmented species may be somewhat resistant to the negative impacts of increasing habitat isolation, their typically low levels of genetic diversity may nonetheless limit adaptive potential and the ability to respond to changing environmental conditions. Changes in habitat quality and suitability due to global change (e.g., changes in climate, and nutrient and hydrological cycles, etc.) and also via local human intervention (e.g., deforestation, drainage, etc.), are prevalent worldwide. Naturally fragmented species often have limited dispersal capabilities and habitat tolerances, which limit their ability to move across the landscape should their habitat patch become

unsuitable (Ye et al. 2013). Thus, the ability to respond via genetic adaptation to changes in local habitat conditions is essential for long-term persistence of such species (Reusch and Wood 2007). In Chapter 6 I used an AFLP-based genome scan approach to search for molecular evidence of selection among local populations of the bog copper butterfly, a relict habitat specialist. I identified five candidate outlier loci exhibiting unusually high levels of genetic differentiation, which is indicative of divergent selection across sample sites. Several measurements reflective of local landscape heterogeneity, both local habitat patch conditions and surrounding landscape structure, were identified as potential agents of selection. Molecular evidence for local adaptation in response to varying ecological conditions has been demonstrated in organisms distributed across large geographic ranges (e.g., European Alps, Manel et al. 2010b; Bothwell et al. 2013) and among contrasting habitat types (e.g., benthic vs. pelagic zone; Campbell and Bernatchez 2004). However, my study is one of the first to document patterns of adaptive genetic variation in relation to landscape heterogeneity at a relatively fine spatial scale. Understanding how current local adaptations in populations have been influenced by fine-scale landscape heterogeneity is important for making accurate predictions regarding the future geographic distributions and survival of organisms experiencing environmental change (Manel et al. 2010a; Hampe and Jump 2011). Thus, my study provides new insights into the landscape factors influencing patterns of adaptive genetic variation in a naturally fragmented species.

7.3 Mobility in naturally fragmented relict populations

Dispersal ability or tendency is a heritable trait which has been shown to respond rapidly to ecological and landscape change (Thomas et al. 1998; Hill et al. 1999a, b; Merxck et al. 2003). Most research to date has focused on understanding the effects of habitat fragmentation (i.e., habitat loss and isolation) on mobility and dispersal, and the resulting consequences for local population genetics and dynamics (e.g., Norberg and Leimar 2002; Vandewoestijne and Van Dyck 2011). However, spatial heterogeneity within habitat patches may also affect dispersal ability by altering patterns of routine movement (Turlure et al. 2010). Such an effect may be particularly prominent in naturally fragmented species, where dispersal of individuals among populations is generally low due to their close habitat association (Hampe and Jump 2011). Thus, it is important to understand how both landscape structure and local ecological conditions within habitat patches may influence mobility in order to accurately predict how the dynamics and genetics of local populations may respond to future landscape change. This is an area of research where few empirical examples currently exist.

In Chapter 5, I evaluated whether local habitat patch characteristics and/or measurements of surrounding landscape structure were responsible for variation in potential mobility observed among local populations of the bog copper butterfly. To assess movement ability, I used morphological measurements associated with flight (e.g., thorax mass and wing loading) which have been demonstrated to be reliable indicators of flight ability in other butterfly species (Chai and Srygley 1990; Kuusaari et al. 1996; Berwaert et al. 2002; Turlure et al. 2010). My results provided evidence for an effect of both local habitat conditions and landscape structure on flight-related morphological

traits. For example, increasing amounts of forest in the surrounding landscape, indicative of increased habitat patch isolation, appeared to correspond with less mobile phenotypes in both males and females. However in males, we also found that investment in flight was greatest in smaller peatlands in which host-plant density is higher and more homogeneously distributed. A potential explanation for this finding is that individuals may be using different mate location strategies (i.e., perching vs. patrolling), which require different flight designs (Betts and Wootton 1988; Dudley 1990; Wickman 1992), in response to differences in population and host-plant density among peatlands. This study highlights that for a naturally fragmented species, morphological traits associated with mobility may be responding to both local habitat patch characteristics and surrounding landscape structure. It also supports the hypothesis that local habitat conditions contribute to morphological variation in butterflies, and should thus be considered when predicting the response of population dynamics and genetics to landscape change.

7.4 Conclusion

My doctoral research demonstrates the utility and power of AFLP analysis to study non-model and threatened organisms, and has identified current gaps in error rate reporting amongst published AFLP literature. I established best practices for AFLP analysis and applied them to study population connectivity and structure in an anthropogenically fragmented butterfly species, as well as a naturally fragmented butterfly species. The findings of my dissertation demonstrate that for populations inhabiting fragmented landscapes, both landscape-scale and within-patch characteristics can contribute to

patterns of genetic and phenotypic variation among populations. With increasing anthropogenic fragmentation of natural habitats, such information and tools will be valuable for predicting future population trends and preserving evolutionary potential of relict species and populations, particularly those with specialized habitat requirements.

7.5 Literature Cited

- Alda F, Garcia J, Garcia JT and Suarez-Seoane S. 2013. Local genetic structure on breeding grounds of a long-distance migrant passerine: the bluethroat (*Luscinia svecica*) in Spain. *Journal of Heredity*, 104: 36-46.
- Bensch S and Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14: 2899-2914.
- Bonin A, Bellemain E, Bronken E, Pompanon F, Brochmann C and Taberlet P. 2004. How to track and assess genotyping errors in population genetic studies. *Molecular Ecology*, 13: 3261-3273.
- Bonin A, Ehrich D and Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16: 3737-3758.
- Bothwell H, Bisbing S, Overgaard Therkildsen N, Crawford L, Alvarez N, Holderegger R and Manel S. Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis* L.), using a landscape genetic approach. *Conservation Genetics*, 14: 467-481.
- Caballero A, Quesada H and Rolan-Alvarez E. 2008. Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, 179: 539-554.

- Campbell D and Bernatchez L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, 21: 945-956.
- Debinski DM and Holt RD. 2000. A survey and overview of habitat fragmentation experiments. *Conservation Biology*, 14: 342-355.
- Habel JC, Assmann T, Schmitt T and Avise JC. 2010. Relict species: from past to future, in: Relict species: phylogeography and conservation biology, eds. Habel JC and Assmann T. Springer, Heidelberg, Germany.
- Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Reviews of Ecology, Evolution and Systematics*, 34: 487-515.
- Fjordingstad EJ, Schtickzelle N, Manhes P, Gutierrez A and Clobert J. 2007. Evolution of dispersal and life history strategies – *Tetrahymena ciliates*. *BMC Evolutionary Biology*, 7: 133.
- Habel JC and Schmitt T. 2012. The burden of genetic diversity. *Biological Conservation*, 147: 270-274.
- Habel JC and Zachos FE. 2012. Habitat fragmentation versus fragmented habitats. *Biodiversity and Conservation*, 21: 2987-2990.
- Hampe A and Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics*, 42: 313-333.

- Hill JK, Thomas CD and Blakeley DS. 1999a. Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121: 165-170.
- Hill JK, Thomas CD and Lewis OT. 1999b. Flight morphology in fragmented populations of a rare British butterfly, *Hesperia comma*. *Biological Conservation*, 87: 277-283.
- Holderegger R and Wagner HH. 2008. Landscape genetics. *BioScience*, 58: 199-207.
- Keyghobadi N. 2007. The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, 85: 1049-1064.
- Koopman WJM and Gort G. 2004. Significance tests and weighted values for AFLP similarities, based on Arabidopsis *in silico* AFLP fragment length distributions. *Genetics*, 167: 1915–1928.
- Manel S, Joost S, Epperson BK, Holderegger R, Storfer A, Rosenberg MS, Scribner KT, Bonin A and Fortin M-J. 2010a. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19: 3760-3772.
- Manel S, Poncet BN, Legendre P, Gugerli F and Holderegger R. 2010b. Common factors drive adaptive genetic variation at different spatial scale in *Arabidopsis alpina*. *Molecular Ecology*, 19: 3824-3835.
- Merckx T, Van Dyck H, Karlsson B, Leimar O. 2003. The evolution of movements and behaviour at boundaries in different landscapes: a common arena experiment with butterflies. *Proceedings of the Royal Society B: Biological Sciences* 270: 1815-1821.

- Meudt HM and Clarke AC. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12: 106-117.
- Norberg U and Leimar O. 2002. Spatial and temporal variation in flight morphology in the butterfly *Melitaea cinxia* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, 77: 445-453.
- Pitra C, Suárez-Seoane S, Martín CA, Streich W-J and Alonso JC. 2011. Linking habitat quality with genetic diversity: a lesson from great bustards in Spain. *European Journal of Wildlife Research*, 57: 411-419.
- Pompanon F, Bonin A, Bellemain E and Taberlet P. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, 6: 847-859.
- Porlier M, Bélisle M and Garant D. 2009. Non-random distribution of individual genetic diversity along an environmental gradient. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364: 1543-1554.
- Reusch TBH and Wood TE. 2007. Molecular ecology of global change. *Molecular Ecology*, 16: 3973-3992.
- Saunders DA, Hobbs RJ and Margules CR. 1991. Biological consequences of ecosystem fragmentation – a review. *Conservation Biology*, 5: 18-32.
- Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy OJ, Holderegger R, Taberlet P, Waits LP and Manel S. 2010. Applications of landscape

- genetics in conservation biology: concepts and challenges. *Conservation Genetics*, 11: 375-385.
- Schwartz TS and Karl SA. 2005. Population and conservation genetics of the gopher tortoise (*Gopherus polyphemus*). *Conservation Genetics*, 6: 917-928.
- Storfer A, Murphy MA, Spear SF, Holderegger R and Waits LP. 2010. Landscape genetics: where are we now? *Molecular Ecology*, 19: 3496-3514.
- Thomas CD, Hill JK and Lewis OT. 1998. Evolutionary consequences of habitat fragmentation in a localized butterfly. *Journal of Animal Ecology*, 67: 485-497.
- Thornhill R and Alcock J. 1983. *The Evolution of Insect Mating Systems*. Harvard University Press, Cambridge, MA.
- Turlure C, Schtickzelle N and Baguette M. 2010. Resource grain scales mobility and adult morphology in butterflies. *Landscape Ecology*, 25: 95-108.
- Vandewoestijne S and Baguette M. 2004. Genetic population structure of the vulnerable bog fritillary butterfly. *Hereditas*, 141: 199-206.
- Vandewoestijne S, Schtickzelle N and Baguette M. 2008. Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology*, 6: 46.
- Vandewoestijne S and Van Dyck H. 2011. Flight morphology along a latitudinal gradient in a butterfly: do geographic clines differ between agricultural and woodland landscapes? *Ecography*, 34: 876-886.

- Van Dyck H and Baguette M. 2005. Dispersal behaviour in fragmented landscapes: routine or special movements? *Basic and Applied Ecology*, 6: 535-545.
- Vekemans X, Beauwens T, Lemaire M, Rodán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11: 139-151.
- Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, 8: 725-735.
- Ye X, Skidmore AK and Wang T. 2013. Within-patch habitat quality determines the resilience of specialist species in fragmented landscapes. *Landscape Ecology*, 28: 135-147.

Appendix A: Laboratory protocol used to optimize AFLPs for *Lycaena epixanthe*

Amplified fragment length polymorphism (AFLP) profiles for *Lycaena epixanthe* were generated following a protocol adapted from the standard AFLP methods of Clarke and Meudt (2005) and the commercial AFLP Plant Mapping Kit (Applied Biosystems, Foster City, CA):

- 1) Genomic DNA was digested with the restriction enzymes *EcoRI* and *MseI*, and synthetic DNA adaptors (Table A.1) with complementary sticky ends were simultaneously ligated to the resulting overhangs created by the restriction enzymes (i.e., a ‘restriction-ligation’ (R-L) reaction). Five and a half microlitres of template DNA (approximately 250 ng) were added to 1.0 μL of enzyme master mix and 4.5 μL of R-L master mix to a total reaction volume of 11.0 μL (see Table A.2 for adaptor annealing protocol and Table A.3 for detailed R-L protocol). The R-L reactions were incubated overnight at 24 °C, subsequently diluted with 89.0 μL of TE_{0.1} buffer and stored at -20°C.
- 2) Diluted R-L fragments were then amplified through PCR using two ‘pre-selective’ primers (*Eco-A* and *Mse-C*) that are complementary to the adaptor sequences and have one additional nucleotide base at the 3’ end (Table A.1). Only R-L fragments which contain bases complementary to these additional bases at the 3’ end of the primers, immediately next to the restriction sites, will be amplified. Thus, the number of fragments is reduced by approximately 1/16 (Bensch and Åkesson 2005). In this pre-selective PCR reaction, 4.0 μL of diluted R-L DNA was added to 16.0 μL of pre-selective master mix for a total reaction volume of 20.0 μL (see Table A.4 for detailed pre-selective PCR

protocol). Ten microlitres of pre-selective PCR product was subsequently diluted with 25.0 μL TE_{0.1} and stored at -20 °C.

3) A sub-set of the pre-selective fragments was then amplified through a second PCR using two ‘selective’ primers (*Eco*-Axx and *Mse*-Cxx) that have an additional two nucleotide bases at the 3’ end (Table A.5), further reducing the number of fragments by 1/256 (Bensch and Åkesson 2005). All *Eco*-Axx primers were labelled using fluorescent dyes (Table A.1) to allow for visualization and scoring on an automated sequencer. For the selective PCR reaction, 3.0 μL of the diluted pre-selective PCR product was added to 17.0 μL of selective master mix to a total volume of 20.0 μL (see Table A.5 for detailed selective PCR protocol).

Different subsets of fragments are produced by using selective primers that vary in the last two nucleotide bases. We initially trialed all different pairwise combinations of eight *Eco*+Axx primers and eight *Mse*+Cxx primers (total of 64 combinations; Table A.1) and selected the following five selective primer combinations based on relative polymorphism and reproducibility of fragments: *Eco*RI-ACA/*Mse*I-CTT, *Eco*RI-AAC/*Mse*I-CAC, *Eco*RI-AAC/*Mse*I-CTC, *Eco*RI-AAG/*Mse*I-CAA, and *Eco*RI-ACC/*Mse*I-CAC. Negative controls were included in each step of the protocol to detect any contamination, and all PCR reactions were performed using a Bio-Rad DNA Engine, Peltier Thermal Cycler. Fluorescently labelled AFLP fragments were separated and sized using a 3730S Genetic Analyzer (Applied Biosystems) at the Nucleic Acid Protein Services Unit in the Michael Smith Laboratories at the University of British Columbia.

Table A.1 Oligonucleotides used for the AFLP analysis of *Lycaena epixanthe*. PCR primer sequences are indicated in the 5' to 3' orientation, and the lengths of the oligonucleotides in base pairs (bp). Note that the two *EcoRI* and *MseI* adaptors must be separately annealed prior to use in the restriction-ligation reaction (Table A.2). Bolded type denotes selective sequences.

Protocol Stage	Primer	Sequence (5'-3')	Length (bp)
Restriction-Ligation	<i>EcoRI</i> Adaptor:		
	<i>EcoRI</i> Adaptor 1	CTCGTAGACTGCGTACC	17
	<i>EcoRI</i> Adaptor 2	AATTGGTACGCAGTCTAC	18
	<i>MseI</i> Adaptor:		
	<i>MseI</i> Adaptor 1	GACGATGAGTCCTGAG	16
	<i>MseI</i> Adaptor 2	TACTCAGGACTCAT	14
Pre-Selective PCR	<i>EcoRI</i> Primer:		
	<i>Eco+A</i>	GACTGCGTACCAATTCA	17
	<i>MseI</i> Primer:		
	<i>Mse+C</i>	GATGAGTCCTGAGTAAC	17
Selective PCR	<i>EcoRI</i> Primers:		
	* <i>Eco+ANN</i>	^F GACTGCGTACCAATTCA AC	19
		^F GACTGCGTACCAATTCA AG	19
		^F GACTGCGTACCAATTCA ACA	19
		^V GACTGCGTACCAATTCA ACC	19
		^V GACTGCGTACCAATTCA ACG	19
		^V GACTGCGTACCAATTCA ACT	19
		^N GACTGCGTACCAATTCA AGC	19
		^N GACTGCGTACCAATTCA AGG	19
	<i>MseI</i> Primers:		
	<i>Mse+CNN</i>	GATGAGTCCTGAGTAAC AA	19
		GATGAGTCCTGAGTAAC AC	19
		GATGAGTCCTGAGTAAC AG	19
		GATGAGTCCTGAGTAAC AT	19
		GATGAGTCCTGAGTAAC TA	19
	GATGAGTCCTGAGTAAC TC	19	
	GATGAGTCCTGAGTAAC TG	19	
	GATGAGTCCTGAGTAAC TT	19	

*Fluorescently labelled dye; ^F = 6 FAM; ^V = VIC, ^N = NED

Table A.2 Protocol for annealing a) *EcoRI* and b) *MseI* adaptors: for each reaction component, the initial concentration and volume are indicated. *EcoRI* and *MseI* adaptor reactions were independently prepared.

Rxn components	Initial concentration	Volume (μL)
a) <i>EcoRI</i> Adaptor Master Mix		
<i>EcoRI</i> Adaptor 1	1 nmol/ μL	1.00
<i>EcoRI</i> Adaptor 2	1 nmol/ μL	1.00
Milli-Q water	-	108.00
T ₁₀ E ₁ *	-	90.00
Total		200.00
b) <i>MseI</i> Adaptor Master Mix		
<i>MseI</i> Adaptor 1	1 nmol/ μL	10.0
<i>MseI</i> Adaptor 2	1 nmol/ μL	10.0
Milli-Q water	-	90.0
T ₁₀ E ₁ *	-	90.0
Total		200.00

* T₁₀E₁ = 10 mM Tris, 1 mM EDTA

Adaptor Master Mixes were heated for 8 minutes at 93 °C in a dry heating block, then cooled at room temperature for 10 minutes, centrifuged for 10 seconds at 1400 g and stored at -20 °C.

Table A.3 Restriction-ligation (R-L) protocol outlining for a) Enzyme Master Mix, b) R-L Master Mix, and c) R-L Reaction: the components, initial concentration and volume.

Components	Initial concentration	Volume (μL)
a) Enzyme Master Mix		
Milli-Q water	-	0.55
T4-DNA ligase buffer (with ATP)	10 x	0.10
NaCl	0.5 M	0.10
BSA	1.0 mg/ μL	0.05
<i>MseI</i>	100 U	0.10
<i>EcoRI</i>	500 U	0.05
T4 DNA Ligase	100 Weiss U	0.05
Total		1.00
b) R-L Master Mix		
T4 DNA ligase buffer (with ATP)	10x	1.00
NaCl	0.5 M	1.00
BSA	1.0 mg/ μL	0.50
<i>MseI</i> Adaptor	1nm/ μL	1.00
<i>EcoRI</i> Adaptor	1nm/ μL	1.00
Total		4.50
c) R-L Reaction		
<i>Enzyme Master Mix</i>	-	1.00
<i>R-L Master Mix</i>	-	4.50
<i>Genomic DNA</i>	~250 ng	5.50
Total		11.00

Table A.4 PCR protocol for Pre-Selective AFLP amplification, outlining for each reaction component the initial concentration and volume.

Rxn components	Initial concentration	Volume (μL)
Milli-Q Water	-	3.05
Betaine	3M	7.00
dNTPs	10 mM	0.50
MgCl ₂	25 mM	1.25
PCR Buffer (no MgCl ₂)	10x	2.00
<i>Eco</i> +A Primer	10 pmol/ μL	1.00
<i>Mse</i> +C Primer	10 pmol/ μL	1.00
<i>Taq</i> Polymerase	5 U/ μL	0.20
Diluted R-L DNA	-	4.00
Total		20.00

Cycling parameters were: one cycle of 72°C for 2 minutes; 25 cycles of 94°C for 20 seconds, 56°C for 30 seconds, and 72°C for 2 minutes; and a final cycle of 60°C for 30 minutes. Ramp speed was set to 1.0°C / second.

Table A.5 PCR protocol for Selective AFLP amplification outlining for each reaction component the initial concentration and volume.

Rxn components	Initial concentration	Volume (μL)
Milli-Q Water	-	8.80
dNTPs	10 mM	0.50
MgCl ₂	25 mM	3.50
PCR Buffer (no MgCl ₂)	10x	2.00
<i>Eco</i> +Axx Primer	10 pmol/ μL	1.00
<i>Mse</i> +Cxx Primer	10 pmol/ μL	1.00
<i>Taq</i> Polymerase	5 U/ μL	0.20
Pre-Selective PCR Product	-	3.00
Total		20.00

Cycling parameters were: one cycle of 94°C for 2 minutes; 11 touchdown cycles of 94°C for 20 seconds, 66-57°C for 30 seconds, and 72°C for 2 minutes; 22 cycles of 94 °C for 20 seconds, 56°C for 30 seconds, and 72°C for 2 minutes; and a final cycle of 60°C for 30 minutes. Ramp speed was set to 1.0 °C / second.

Literature Cited

Clarke A and Meudt H. 2005. AFLP (amplified fragment length polymorphism) for multilocus genomic fingerprinting. Alan Wilson Centre for Ecology and Evolution, Massey University, New Zealand. Available from: <http://www.clarkeresearch.org>.

Appendix B: Landscape data reported in Chapters 3-5

Table B.1 Summary of the landscape variables collected for the 15 Algonquin Provincial Park populations of *Lycaena epixanthe*. Surrounding landscape structure measures i) proportion of wetland, ii) forest and iii) water habitat, and iv) mean proximity of wetland habitat, were assessed within a 1 km buffer radius surrounding the centroid of butterfly capture in each surveyed peatland. Local patch characteristics v) peatland size (ha), vi) mean host-plant density (/m²), vii) the coefficient of variation in mean host-plant density (CV) and viii) total host-plant abundance (/m²) reflect particular attributes of each surveyed site.

Peatland	Proportion land cover			Mean proximity	Peatland size	Host-plant distribution	
	Wetland	Forest	Water			Mean density (CV)	Total abundance
BUG	0.11	0.83	0.04	18.77	3.49	0.08 (0.98)	16.50
WH	0.16	0.82	0.02	34.36	2.52	0.14 (1.00)	32.81
DL	0.08	0.86	0.04	0.90	5.78	0.15 (1.22)	32.56
ML	0.10	0.85	0.05	4.27	8.29	0.09 (1.01)	36.49
WR	0.16	0.81	0.04	383.96	5.74	0.08 (0.93)	25.36
KB	0.09	0.70	0.15	27.92	3.06	0.28 (0.81)	49.55
MIN	0.06	0.89	0.04	38.64	0.91	0.19 (0.87)	13.74
BAB	0.04	0.92	0.03	19.24	1.05	0.13 (0.93)	9.92
ZEN	0.06	0.90	0.03	18.60	0.70	0.14 (0.94)	8.37
SUN	0.21	0.65	0.09	66.23	47.65	-	-
SB	0.23	0.72	0.02	80.24	0.24	0.15 (0.95)	3.70
EOS	0.12	0.81	0.02	245.43	0.93	0.21 (0.75)	7.41
COS	0.29	0.67	0.01	321.83	22.3	-	-
OPL	0.09	0.69	0.19	109.43	4.16	0.12 (1.09)	30.05
DT	0.23	0.66	0.09	59.00	2.44	0.12 (1.13)	21.21

Peatland size represents the overall area of the sampled habitat patch, mean host-plant density was calculated as the mean cranberry cover for all quadrats in a surveyed peatland, and total host-plant abundance was determined by multiplying mean host-plant density by total peatland area.

Appendix C: Chapter 3 Supplementary Material

AFLP datasets were generated by setting both loci and phenotype scoring thresholds to either 75- or 100-rfu in AFLPScore v1.3b (Whitlock et al. 2008). ‘A’ indicates the original dataset produced from AFLPScore consisting of all retained loci; ‘B’ indicates the original dataset with singleton loci removed. The values reported in the body of the text are indicated in bold.

Table C.1 Summary of the AFLP phenotype scoring results for all selective primer combinations.

Selective primer combination	Scoring Threshold		Mismatch error rate %	Initial number of loci	Number of loci retained
	Locus	Phenotype			
<i>EcoRI</i> -AAC/ <i>MseI</i> -CAC	75	75	1.46	112	99
	100	100	1.43	112	94
<i>EcoRI</i> -AGC/ <i>MseI</i> -CAC	75	75	1.77	107	101
	100	100	1.74	107	93
<i>EcoRI</i> -AGC/ <i>MseI</i> -CAT	75	75	1.71	135	117
	100	100	1.68	135	112
<i>EcoRI</i> -AGC/ <i>MseI</i> -CTT	75	75	1.30	109	95
	100	100	1.28	109	80
<i>EcoRI</i> -AAG/ <i>MseI</i> -CAA	75	75	1.71	129	117
	100	100	1.63	129	109
	75	75	1.59 (mean)	592 (total)	529 (total)
	100	100	1.55 (mean)	592 (total)	488 (total)

**Table C.2 Within sub-population and global genetic diversity for the British Columbia population of *Apodemia mormo*.
Global measures are reported as mean values.**

Site	N	75 rfu thresholds				100 rfu thresholds			
		A (526 loci)		B (380 loci)		A (484 loci)		B (326 loci)	
		PPL	H _e (±SE)	PPL	H _e (±SE)	PPL	H _e (±SE)	PPL	H _e (±SE)
W1	45	0.154	0.061 (0.006)	0.213	0.083 (0.007)	0.151	0.061 (0.006)	0.224	0.088 (0.008)
W2	39	0.144	0.050 (0.005)	0.200	0.069 (0.007)	0.149	0.052 (0.005)	0.221	0.077 (0.008)
W3	19	0.141	0.044 (0.005)	0.195	0.060 (0.007)	0.149	0.046 (0.005)	0.221	0.068 (0.008)
W4	38	0.150	0.050 (0.005)	0.208	0.069 (0.006)	0.151	0.049 (0.005)	0.224	0.072 (0.007)
W5	44	0.135	0.047 (0.005)	0.187	0.064 (0.007)	0.122	0.048 (0.005)	0.181	0.071 (0.008)
W6	41	0.144	0.044 (0.005)	0.200	0.061 (0.006)	0.122	0.046 (0.005)	0.181	0.068 (0.007)
W7	15*	0.152	0.055 (0.006)	0.211	0.075 (0.007)	0.153	0.057 (0.006)	0.227	0.083 (0.008)
W8	38	0.144	0.053 (0.005)	0.200	0.073 (0.007)	0.155	0.056 (0.006)	0.230	0.082 (0.008)
C1	38	0.143	0.047 (0.005)	0.197	0.064 (0.006)	0.147	0.048 (0.005)	0.218	0.070 (0.007)
C2	40	0.143	0.047 (0.005)	0.197	0.064 (0.007)	0.116	0.048 (0.005)	0.172	0.070 (0.008)
E1	40	0.171	0.063 (0.006)	0.237	0.086 (0.008)	0.159	0.062 (0.006)	0.236	0.090 (0.008)
E2	29	0.156	0.061 (0.006)	0.216	0.083 (0.007)	0.153	0.062 (0.006)	0.227	0.090 (0.008)
N1	28	0.135	0.056 (0.005)	0.187	0.077 (0.007)	0.140	0.058 (0.006)	0.209	0.085 (0.008)
N2	13*	0.146	0.060 (0.006)	0.203	0.081 (0.008)	0.153	0.061 (0.006)	0.227	0.087 (0.008)
All	467	0.147	0.053 (0.005)	0.204	0.072 (0.007)	0.144	0.054 (0.009)	0.214	0.079 (0.008)

N, number of analysed samples; PPL, proportion of polymorphic loci; H_e, expected heterozygosity.

Table C.3 Genetic structure summary for the British Columbian population of *Apodemia mormo*.

		H_t	H_w (±SE)	H_b (±SE)	F_{ST} (±SE)
75	A	0.055	0.053 (0.002)	0.002 (0.0003)	0.039 (0.113)*
	B	0.075	0.072 (0.002)	0.003 (0.0003)	0.039 (0.112)*
100	A	0.056	0.054 (0.002)	0.002 (0.0003)	0.042 (0.107)*
	B	0.082	0.079 (0.002)	0.004 (0.0004)	0.043 (0.105)*

*p < 0.00001

H_t, total gene diversity; H_w, average gene diversity within populations; H_b, average gene diversity among populations in excess of that observed within populations; F_{ST}, proportion of the total gene diversity that occurs among as opposed to within populations.

Table C.4 Robustness of the genetic barriers identified by BARRIER based on 100 permuted genetic distance matrices (F_{ST}). Barriers are listed in order of importance (A, B, C, etc.), and are considered to possess high bootstrap support when bootstrap values are greater than 50.

Number of barriers tested for:	Scoring criteria			
	75-A	75-B	100-A	100-B
<i>1 barrier</i>	75-76	72	74-75	70-87
<i>2 barriers</i>				
A	91-92	91	87-89	88-91
B	65	62	71	70
<i>3 barriers</i>				
A	97-99	95-96	94-96	94-97
B	96	86	92	90
C	88	92	81	83
<i>4 barriers</i>				
A	98-100	98-99	98-100	97-100
B	100	96	100	99
C	96	99	98	100
D	71	74	84	87
<i>5 barriers</i>				
A	98-100	99-100	99-100	98-100
B	100	98	100	100
C	99	100	100	100
D	90	90	97	96
E	45	45	45	39-48

Table C.5 Summary of the principal coordinate analysis (PCoA) of all studied sub-populations, based on a genetic distance matrix (F_{ST}). The amount of variability that is accounted for by the first two coordinates is reported.

Scoring criteria	% of variation explained	
	Coordinate 1	Coordinate 2
75	A	45.26
	B	45.11
100	A	45.78
	B	45.59

Table C.6 Summary of the mantel tests results examining the relationship between genetic differentiation (F_{ST} / (1- F_{ST}) and geographical distance (ln transformed) among all studied sub-populations. Geographic distance was measured as both straight line distance between sites (Euclidean), and ‘adjusted’ distance between sites (based on constraining movement to valley bottoms).

Geographic distance measure	Scoring criteria	r^2	P
Euclidean	75	A	0.144
		B	0.147
	100	A	0.179
		B	0.182
Adjusted	75	A	0.206
		B	0.209
	100	A	0.242
		B	0.247

$\alpha = 0.05$

Appendix D: Chapter 4 Supplementary Material

Table D.1 Capture record for individuals of *Lycaena epixanthe* surveyed in 13 peatlands in Algonquin Provincial Park. An index of relative butterfly density (number of individuals captured per hour) was determined for each peatland by dividing the total number of individuals captured by the capture duration, and then standardized by the number of butterfly catchers (capture effort).

Peatland	# individuals captured	Time (hr)	Capture effort	Relative butterfly density (# indivs. /hr)
BUG	18	1.75	1	10.29
WH	40	1.00	1	40.00
DL	40	2.25	1	17.78
ML	30	1.75	2	8.57
WR	40	1.00	1	40.00
KB	40	1.00	1	40.00
MIN	40	1.00	1	40.00
BAB	40	1.50	1	26.67
ZEN	40	1.50	1	26.67
SB	12	1.25	1	9.60
EOS	29	1.00	2	14.50
OPL	41	2.00	1	20.50
DT	40	2.00	1	20.00

Table D.2 Pairwise F_{ST} values for 13 surveyed populations of *Lycaena epixanthe* in Algonquin Provincial Park, Ontario, Canada. Sub-samples from the five bogs where individuals were collected from two discrete regions (DL, ML, MIN, EOS and DT) are analysed as separate sub-populations. Values in bold indicate populations which were not significantly differentiated ($\alpha = 0.01$).

	BUG	WH	DLa	DLb	MLa	MLb	WR	KB	MINa	MINb	BAB	ZEN	SB	EOSa	EOSb	OPL	DTa	DTb
BUG	-																	
WH	0.083	-																
DLa	0.108	0.093	-															
DLb	0.090	0.066	0.025	-														
MLa	0.109	0.066	0.130	0.090	-													
MLb	0.098	0.049	0.120	0.082	0.049	-												
WR	0.101	0.044	0.117	0.085	0.078	0.072	-											
KB	0.156	0.103	0.196	0.175	0.106	0.055	0.109	-										
MINa	0.171	0.111	0.208	0.179	0.120	0.064	0.114	0.038	-									
MINb*	0.140	0.107	0.213	0.161	0.107	0.070	0.085	0.073	0.034	-								
BAB	0.200	0.124	0.219	0.192	0.122	0.071	0.162	0.034	0.062	0.118	-							
ZEN	0.093	0.049	0.129	0.100	0.080	0.045	0.077	0.064	0.058	0.073	0.065	-						
SB	0.118	0.056	0.137	0.110	0.051	0.040	0.085	0.054	0.072	0.095	0.071	0.042	-					
EOSa	0.104	0.046	0.138	0.096	0.077	0.042	0.066	0.085	0.088	0.088	0.103	0.059	0.050	-				
EOSb	0.119	0.063	0.155	0.121	0.077	0.040	0.089	0.055	0.057	0.063	0.050	0.045	0.045	0.009	-			
OPL	0.146	0.087	0.183	0.154	0.107	0.064	0.110	0.070	0.063	0.073	0.068	0.055	0.060	0.037	0.019	-		
DTa	0.180	0.105	0.203	0.181	0.120	0.062	0.138	0.040	0.068	0.110	0.016	0.071	0.069	0.087	0.043	0.062	-	
DTb	0.280	0.199	0.301	0.279	0.205	0.147	0.229	0.076	0.116	0.181	0.043	0.154	0.139	0.162	0.104	0.134	0.019	-

Comparison of AFLP datasets varying in genotyping error rate

AFLP datasets were generated using AFLPSCORE v1.3b (Whitlock et al. 2008) using the data filter and absolute phenotype-calling threshold settings. ‘Dataset A’ indicates the original dataset produced from AFLPSCORE consisting of all retained loci; ‘Dataset B’ indicates the original dataset with singleton loci removed.

Table D.3 Summary of the optimal AFLP phenotype scoring parameters and associated mismatch error rates and number of retained loci for all selective primer combinations. The values reported in Chapter 3 are indicated in bold.

Selective primer combination	Scoring Threshold		Mismatch error rate %	Number of loci		
	Locus	Phenotype		Initial	Dataset A	Dataset B
<i>EcoRI-ACA/MseI-CTT</i>	1300	200	1.75	161	23	21
	1200	200	2.51	161	26	23
	900	300	3.92	161	36	30
	500	100	4.78	161	52	46
<i>EcoRI-AAC/MseI-CAC</i>	3000	100	2.42	177	16	16
	2800	500	3.08	177	22	22
	2000	500	3.98	177	34	29
	1300	500	4.75	177	57	52
<i>EcoRI-AAC/MseI-CTC</i>	1800	400	1.66	122	30	24
	1000	300	2.89	122	41	32
	900	300	3.09	122	48	37
	400	100	4.72	122	70	65
<i>EcoRI-AAG/MseI-CAA</i>	1000	100	1.94	123	28	28
	800	200	2.97	123	51	49
	500	100	3.23	123	62	60
	300	100	4.59	123	94	93
<i>EcoRI-ACC/MseI-CAC</i>	2100	600	1.95	177	29	25
	2300	800	2.61	177	37	31
	2500	1200	3.89	177	41	34
	2000	800	4.49	177	46	38
		<2%	1.94 (mean)	760 (total)	126 (total)	112 (total)
		<3%	2.81 (mean)	760 (total)	177 (total)	157 (total)
		<4%	3.62 (mean)	760 (total)	221 (total)	190 (total)
		<5%	4.67 (mean)	760 (total)	319 (total)	294 (total)

Table D.4 Summary of the overall genetic structure and intrapopulation genetic diversity results for *Lycaena epixanthe* in Algonquin Provincial Park (sub-samples within bogs are not pooled; N = 18). For each AFLP dataset we present global F_{ST} (\pm standard error), and proportion of polymorphic loci (PPL) and expected heterozygosity (H_j) averaged across sub-samples, with the range indicated in parentheses.

Scoring criteria	Dataset	F_{ST} (\pm SE)	PPL	H_j
<2%	A	0.112 (0.215)*	0.363 (0.571)	0.157 (0.251)
	B	0.111 (0.216)*	0.408 (0.643)	0.177 (0.282)
<3%	A	0.107 (0.186)*	0.355 (0.474)	0.157 (0.207)
	B	0.107 (0.187)*	0.400 (0.535)	0.176 (0.232)
<4%	A	0.103 (0.170)*	0.355 (0.434)	0.158 (0.184)
	B	0.103 (0.171)*	0.410 (0.490)	0.183 (0.198)
<5%	A	0.094 (0.161)*	0.352 (0.383)	0.158 (0.153)
	B	0.094 (0.162)*	0.382 (0.415)	0.171 (0.166)

* $P < 0.001$

Table D.5 Summary of analysis of molecular variance (AMOVA) results for the Algonquin Provincial Park populations of *Lycaena epixanthe* (sub-samples within bogs were not pooled; N =18).

Scoring criteria	Dataset	Source of variation	d.f.	Sum of squares	Variance component	% of variance
<2%	A	Among populations	17	261.539	0.403	7.75
		Within populations	459	2200.909	4.795	92.25
	B	Among populations	17	260.997	0.403	7.79
		Within populations	459	2187.535	4.766	92.21
<3%	A	Among populations	17	449.845	0.725	8.90
		Within populations	459	3402.843	7.414	91.10
	B	Among populations	17	448.797	0.724	8.96
		Within populations	459	3377.168	7.358	91.04
<4%	A	Among populations	17	619.233	1.005	9.12
		Within populations	459	4595.079	10.011	90.88
	B	Among populations	17	618.157	1.005	9.18
		Within populations	459	4565.338	9.946	90.82
<5%	A	Among populations	17	894.526	1.419	8.48
		Within populations	459	7027.832	15.310	91.52
	B	Among populations	17	893.533	1.419	8.51
		Within populations	459	7001.455	15.254	91.49

Table D.6 Summary of the principal coordinate analysis (PCoA) for all studied populations of *Lycaena epixanthe* (sub-samples within bogs were not pooled; N = 18). Based on a covariance matrix of pairwise F_{ST} values, the amount of variability that is accounted for by the first two coordinates is reported.

Scoring criteria	Dataset	% of variation explained	
		Coordinate 1	Coordinate 2
<2%	A	53.55	18.40
	B	53.72	18.28
<3%	A	56.48	17.10
	B	65.29	13.42
<4%	A	55.25	16.02
	B	56.11	15.94
<5%	A	60.27	13.77
	B	54.42	15.74

Table D.7 Summary of the mantel test results examining the relationship between genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distance (log transformed) among all studied populations of *Lycaena epixanthe* (sub-samples within bogs were not pooled; N = 18).

Scoring criteria	Dataset	r^2	Significance (P -value)
<2%	A	0.132	<0.001
	B	0.132	<0.001
<3%	A	0.168	<0.001
	B	0.148	<0.001
<4%	A	0.190	<0.001
	B	0.191	<0.001
<5%	A	0.177	<0.001
	B	0.158	<0.001

$\alpha = 0.05$

Literature Cited

Whitlock R, Hipperson H, Mannarelli M, Butlin RK and Burke T. 2008. An objective, rapid, and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, 8: 725-735.

Appendix E: Permission to reproduce published material

A version of Chapter 2 was previously published in *Molecular Ecology*. Permission to reproduce this work in my PhD thesis has been granted by John Wiley and Sons.

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Jul 09, 2013

This is a License Agreement between Lindsay Crawford ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3177371469004
License date	Jun 27, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Molecular Ecology
Licensed content title	A call for more transparent reporting of error rates: the quality of AFLP data in ecological and evolutionary research
Licensed copyright line	© 2012 Blackwell Publishing Ltd
Licensed content author	Lindsay A. Crawford,Daria Koscinski,Nusha Keyghobadi
Licensed content date	Nov 5, 2012
Start page	5911
End page	5917
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Total	0.00 USD

A version of Chapter 3 was previously published in *Conservation Genetics*. Permission to reproduce this work in my PhD thesis has been granted by Springer-Verlag.

This is a License Agreement between Lindsay Crawford ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3177621458210
License date	Jun 28, 2013
Licensed content publisher	Springer
Licensed content publication	Conservation Genetics
Licensed content title	Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (<i>Apodemia mormo</i>) revealed using AFLPs
Licensed content author	Lindsay A. Crawford
Licensed content date	Jan 1, 2011
Volume number	12
Issue number	4
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of your thesis / dissertation	Genetic and morphological variation of butterflies in relict habitats
Expected completion date	Sep 2013
Estimated size(pages)	250
Total	0.00 USD

Curriculum Vitae

Name: Lindsay A. Crawford

Post-secondary Education and Degrees: The University of Western Ontario
London, Ontario, Canada
2007-2013 PhD Biology (with Environment & Sustainability)

The University of Western Ontario
London, Ontario, Canada
2003-2007 BSc Honours specialization in Conservation Biology

Honours and Awards: National Sciences and Engineering Research Council of Canada
Post-Graduate Scholarship (NSERC-PGS) – 2010-2013

Malcolm Ferguson Award in Life Science – 2012, 2013

Collaborative Graduate Program in Environment & Sustainability
Travel Award – 2010, 2012

The Xerces Society for Invertebrate Conservation
Joan Mosenthal Dewind Bursary – 2010-2011

Province of Ontario Graduate Scholarship (OGS) – 2008-2009

Biology Graduate Travel Award – 2009

Collaborative Graduate Program in Environment & Sustainability
Entrance Scholarship – 2008

University of Western Ontario, Gold Medal Award – 2007

University of Western Ontario, Scholarship of Distinction – 2003

Related Work Experience Graduate Teaching Assistant
The University of Western Ontario
2007-2012

Research Assistant
The University of Western Ontario
2007-2008

Peer-Reviewed Publications:

- Crawford LA***, Koscinski D*, Watt KM, McNeil JN and Keyghobadi N. 2013. Mating and oviposition success of a butterfly are not affected by non-lethal tissue sampling. *Journal of Insect Conservation*, doi:10.1007/s10841-013-9566-8. *indicates equal first authorship
- Bothwell H, Bisbing S, Therkildsen NO, **Crawford LA**, Alvarez N, Holderegger R and Manel S. 2013. Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis* L.), using a landscape genetic approach. *Conservation Genetics*, 14: 467-481.
- Proshek B, **Crawford LA**, Davis CS, Desjardins S, Henderson AE and Sperling FAH. 2013. *Apodemia mormo* in Canada: Population genetic data support prior conservation ranking. *Journal of Insect Conservation*, 17: 155-170.
- Crawford LA**, Koscinski D and Keyghobadi N. 2012. A call for more transparent reporting of error rates: the quality of AFLP data in ecological and evolutionary research. *Molecular Ecology*, 21: 5911-5917.
- Crawford LA**, Desjardins S and Keyghobadi N. 2011. Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs. *Conservation Genetics*, 12: 991-1001.
- Koscinski, D*, **Crawford LA***, Keller HA, Keyghobadi N. 2011. Effects of different methods of non-lethal tissue sampling on butterflies. *Ecological Entomology*, 36**:
301-308. *indicates equal first authorship; **Editor's Choice Article, Volume 36, Issue 3
- Keyghobadi N, **Crawford LA** and Maxwell SA. 2009. Successful analysis of AFLPs from non-lethally sampled wing tissue in butterflies. *Conservation Genetics*, 10: 201-202