From species to genes: ecological and evolutionary mechanisms structuring diversity in space and time

by

Zachary Grant MacDonald

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

Doctor of Philosophy

in

Conservation Biology

Department of Renewable Resources
University of Alberta

Abstract

My Ph.D. thesis addresses three foundational questions in conservation biology: i) what is biodiversity and how is it best measured? ii) how does variation in habitat configuration, habitat composition, and environmental conditions affect emergent patterns of species diversity? and iii) how do these same factors relate to genomic variation within single species?

Our analyses of temporal patterns of butterfly species diversity resolved that negative relationships between species richness (total number of species) and species evenness (relative abundances of species) may compromise the efficacy of many diversity indices. As environmental conditions become more favorable, richness often increases while evenness decreases, meaning indices that conflate these two components of diversity (e.g., information entropies) show little change through time. Based on these findings, we outline analytical recommendations for citizen-science and long-term monitoring programmes.

Building on this work, we established a series of research projects that utilized lake islands as a naturally fragmented landscape for investigating effects of habitat configuration and composition on the diversity of butterflies and vascular plants. Overall, species richness was generally unrelated to degree of fragmentation, supporting stochastic assembly of species consistent with the sample-area effect. However, by developing a novel modelling framework that addresses abundances and occurrences of individual species, we were able to resolve that, after controlling for the sample-area effect, variation in island area and isolation disproportionately affect smaller, less-mobile butterfly species. Importantly, these analyses clearly demonstrate how emergent patterns of species richness can obscure important, species-specific responses to fragmentation. Our findings question previous, richness-based support for the recently proposed and widely debated habitat amount hypothesis, which posits that

conservation efforts should focus solely on preserving the maximum amount of habitat irrespective of its degree of fragmentation. Additionally, we carried out a series of experimental releases of butterflies in the lake-island matrix. Tracking movements of released individuals suggested there is significant disparity in species' ability to navigate fragmented landscapes and that visual senses play a primary role in habitat detection.

The last section of my thesis addresses gene flow and climate-associated genomic variation within Dod's Old World swallowtail butterfly, *Papilio machaon dodi*, throughout its Canadian range. Using a combination of genomic analyses and habitat suitability models, we identified two distinct evolutionary lineages (north *vs* south) that are genetically and ecologically divergent, maintained by local adaptation to climatic conditions. Based on climate change projections, we predicted that the northern lineage is likely to be extirpated and displaced by the southern lineage within 50 years. After controlling for climate-associated genetic variation, configurations of suitable habitat were unrelated to genetic connectivity within *P. m. dodi*. This result challenges a foundational method in ecology: the use of habitat suitability models to infer patterns of connectivity between isolated populations when genetic data are unavailable.

The combination of these thesis projects demonstrates clear utility for integrating biogeography, landscape ecology, and population genomics to address cumulative effects of habitat fragmentation, habitat loss, and climate change on the ecology and evolution of species. Much of my thesis work suggests that an autecological approach, addressing responses of individual species to geographic and environmental factors, may be most sensible from a theoretical perspective and most effective from a conservation perspective. However, continued work and consilience among biogeography, landscape ecology, and population genomics are

required to resolve whether generalizations across taxa are viable and applicable to conservation practice.

Preface

This dissertation is an original work by Zachary G. MacDonald. Except for changes to formatting and the amelioration of a small number of typos, all texts, tables, and figures for the six data chapters are unaltered from their published versions. Supplemental information, supporting information, and appendices are available online (journal publication) for each data chapter; links are imbedded within the text of this thesis.

Chapter 2: this chapter is published in Biodiversity and Conservation:

MacDonald, Z. G., Nielsen, S. E., Acorn, J. H. (2017) Negative relationships between species richness and evenness render common diversity indices inadequate for assessing long-term trends in butterfly diversity. Biodiversity and Conservation, 26(3), 617-629. https://doi.org/10.1007/s10531-016-1261-0

Chapter 3: this chapter is published in Oecologia, was featured on the journal's cover, and was awarded the *Hanksi Prize* for best student-led paper in animal ecology (2018):

MacDonald, Z. G., Anderson, I. D., Acorn, J. H., Nielsen, S. E. (2018) Decoupling habitat fragmentation from habitat loss: butterfly species mobility obscures fragmentation effects in a naturally fragmented landscape of lake islands. Oecologia, 186(1), 11-27. https://doi.org/10.1007/s00442-017-4005-2

Chapter 4: this chapter is published in Journal of Biogeography:

MacDonald, Z. G., Anderson, I. D., Acorn, J. H., Nielsen, S. E. (2018) The theory of island biogeography, the sample-area effect, and the habitat diversity hypothesis: complementarity in a naturally fragmented landscape of lake islands. Journal of Biogeography, 45, 2730-2743.

https://doi.org/10.1111/jbi.13460

Chapter 5: this chapter is published in Ecography:

MacDonald, Z. G., Deane, D. C., Lamb, C. T., He, F., Acorn, J. H., Nielsen, S. E. (2021) Distinguishing effects of area *per se* and isolation from the sample-area effect for true islands and habitat fragments. Ecography, In Press. https://doi.org/10.1111/ecog.05563

Chapter 6: this chapter is published in Journal of Insect Science:

MacDonald, Z. G., Acorn, J. H., Zhang, J., Nielsen, S. E. (2019) Perceptual range, perceptual ability, and visual habitat detection by greater fritillary butterflies. Journal of Insect Science, 19(4), 1-10. https://doi.org/10.1093/jisesa/iez060

Chapter 7: this chapter is published in Molecular Ecology:

MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, J. H., Nielsen, S. E., Sperling, F. A. H. (2020) Gene flow and climate-associated genetic variation in a vagile habitat specialist. Molecular Ecology, 29(20), 3889-3906. https://doi.org/10.1111/mec.15604

Dedication

It's probably unconventional for a Ph.D. candidate to dedicate their thesis to one of their supervisors. But, this one goes to John Acorn. Appealing to the late Stephen J. Gould's wacky application of contingencies to evolutionary theory, one can't help but apply similar logic to their own life. As I stomach this libertarian (*c.f.* deterministic) perspective, playback the tape of my own life, and try to pinpoint contingencies that culminated in my current circumstances, meeting John is the most obvious. His passion and knowledge of biodiversity, science, and philosophy are contagious, and that's why I'm here.

Acknowledgements

This section should be longer than the thesis itself. But, alas, that would be weird and liable to raise some bureaucratic eyebrows. So where do I begin? Should I rattle off an exhaustive list of names, to which the owners I owe an unpayable debt? Probably not. Thus, to be (relatively) brief, I'll impose some structure on this section, cherry-picking the thematic sequence of an enigmatic painting addressing the nature of life: Paul Gauguin's *Where Do We Come From? What Are We? Where Are We Going?* However, as this section addresses the nature of my thesis (which I hope is not described in future work as enigmatic), my adaptation will take on a narcissistic flavour, focusing on my own past, present, and trajectory, followed by an acknowledgement of funding sources.

Where Did I Come From? I attribute my biophilic tendencies to a very supportive family. Growing up, my parents not only tolerated a large collection of fish, reptiles, and amphibians, but encouraged it. Mom would make two trips to the family cabin each summer; the first to drop my sister and I off, and the second to go back for the dozens of critters and their enclosures. Dad was similarly encouraging, spending weeks building a small pond in my childhood room and dealing with hundreds of escaped crickets (food for the critters) that would inevitably clog our house's furnace filters. Jess, my sister, never minded an escaped lizard or snake frequenting her room, and has always been a positive influence. My grandparents, Coreen and Paul Evashkevich (maternal) and Sylvia and Jim Horning (paternal), played a central role in my development as a person interested in all things wild, and I thank them for their unwavering emotional support and taxidermy skills. My friends have always been supportive and encouraging; enjoying a plethora of backcountry activities together helped us maintain strong connections with local landscapes. Towards the end of my undergraduate degree, I was fortunate to connect with John Acorn, who

would become one of my most influential mentors. It was John that originally turned me onto the ecology and evolution of local butterflies and birds, as well as philosophy of science, setting the stage for our future research.

What am I? I'm an evolutionary ecologist. In 2015, I started a M.Sc. under the supervision of Scott Nielsen and John Acorn. Scott is among the most comprehensive thinkers I have had the pleasure of interacting with. He has been an incredibly supportive mentor and many of his perspectives have given shape to the way I think about science today. Scott's breadth of research interests and abilities have helped me to develop and broaden my own. Throughout my graduate programme, John continued to be a central figure in my academic and personal development in all possible dimensions. His passion and ostensible knowledge of biodiversity and philosophy continue to be a guiding light in the existential void of science.

Rolling up from a M.Sc. to a Ph.D., I had the pleasure of engaging in additional work with my committee members, Fangliang He and Felix Sperling. Fangliang is the sharpest theoretician I have had the pleasure of interacting with. Working and conversing with him and his former student, David Deane, has been an incredibly uplifting experience. Throughout the second half of my Ph.D., working with my other committee member, Felix, and his lab became of increasing importance to both my personal and academic development. Felix and I began collaborating on a number of projects that involved me retooling as a population geneticist—one of these culminated in Chapter 7 of this thesis. His unwavering support has been nothing short of invaluable to what is now materializing as my career trajectory.

Scott, John, Fangliang, and Felix have each been incredibly supportive of my academic wonderings, which may have instead been perceived as a lack of focus. For example, we often remind ourselves of a jovial Ralph Waldo Emerson quote: "A foolish consistency is the

hobgoblin of little minds, adored by little statesmen and philosophers and divines. With consistency a great soul has simply nothing to do. He may as well concern himself with his shadow on the wall. Speak what you think now in hard words, and tomorrow speak what tomorrow thinks in hard words again, though it contradict everything you said today." I hope my thesis is understood in this spirit, and that hints of thematic discontinuity signify our adaptive development as practitioners of an evolving science.

Throughout my graduate programme, additional collaborations and conversations with Julian Dupuis, Janet Sperling, Erin Campbell, Amanda Roe, Corey Davis, Iraleigh Anderson, Tyler Nelson, David Deane, Andreas Hamman, and Jian Zhang have greatly helped to broaden my understandings of ecology and evolutionary biology. I consider each of them to be very important mentors. I also had the pleasure of spending extensive time in the field with Iraleigh Anderson, Victoria Masquillier, and Tyler Nelson, and have learned immeasurable things from each of them. In addition to these folks, I have had many important other mentors throughout my graduate studies. Mom and Del Dorscheid, thanks for all the last-minute feedback on grant applications, advice on career paths, and for the emotional support. Nana and Dad, thanks for your help with field preparations, emotional support, and all the "treats, taxes, and taxidermy." Bike rides with two University of Alberta faculty members, Derek MacKenzie and Rene Belland, have been as educational as they have been therapeutic. Conversations and birding with Lu Carbyn have similarly helped me find my way. Interacting with my peers, particularly Federico Riva, with whom I shared an office (dubbed the "Butt Cave"), helped to broaden my understandings of butterfly ecology and how we can model it to death. I also owe a great debt to many other peers, who, over the past five years, have become my best friends. Many nights, spent drinking uncountable beers at "Philosophy Pints" and building vivariums for dendrobatid

frogs with the "Viv Fam," helped to positively blur the lines between scientific and personal dimensions of life.

For simplicity, I have written many sections of this thesis in a personal, active voice: e.g., "I resolved…" or "I concluded…". While this is indeed simple, it is also misleading. I haven't stood on the shoulders of giants—I have been purposefully held up and propelled by them. These giants are my aforementioned mentors. All resolutions and conclusions within this thesis are just as much theirs as they are mine.

Where Am I Going? I will continue my career as an evolutionary ecologist, applying what I have learned and continue to learn from theory of ecology and evolutionary biology to conservation practice. I thank my mentors, in advance, for their unwavering support. The trusting relationships we have built will last our lifetimes—they are the single aspect of this thesis that I am most proud of.

Funding: I was fortunate to receive a variety of scholarships and grants from multiple agencies and organizations over the course of my graduate programme. Two Natural Sciences and Engineering Research Council (NSERC) Alexander Graham Bell Canada Graduate Scholarships (CGS Master's and CGS Doctoral) gave me the freedom to pursue a diversity of research projects over the course of my Ph.D. Funding from the University of Alberta, including the President's Doctoral Prize of Distinction and Walter H Johns Graduate Fellowship, and the Government of Alberta, including two Queen Elizabeth II Graduate Scholarships (Master's and Doctoral), was similarly supportive. A Senior Research Scholarship from the Canada-China Scholars' Exchange Program allowed me to spend half a year in China, developing multiple research projects that are still ongoing. Research grants from the Alberta Conservation Association (Grant in Biodiversity), the Northern Scientific Training Program (Northern

Research Grant), and UAlberta North (UofA Northern Research Award) helped to fund multiple projects included within this thesis and to kickstart an alpine butterfly genomics project (not included in this thesis). NSERC Discovery Grants to Scott Nielsen and Felix Sperling also supported much of the work included within this thesis.

Table of Contents

List of tables

whereas potential resident species subset was limited to species' co-occurrence on islands with their known larval food plant species.

Table 3-3. Island characteristics regressed on butterfly species richness (complete species assemblage and the potential resident species subset). The complete species assemblage included all occurrences, whereas potential resident species subset was limited to species' co-occurrence on islands with their known larval food plant species. Models were ranked for support using the corrected Akaike's Information Criterion (AIC_c) where smaller AIC_c values indicate better-supported models. Relative model weights based on AIC_c are given by w_i.

Table 3-4. Standardized coefficients (β) for parameters from the best-supported models shown in Table 3-3, where island characteristics were regressed on butterfly species richness for the complete species assemblage and the potential resident species subset. The complete species assemblage included all observed butterfly occurrences, whereas potential resident species subset was limited to butterfly species' co-occurrence on islands with their known larval food plant species.

Table 4-1. Summary of the nested set sampling design used to decouple degree of fragmentation from total island area across two distinct ranges of island sizes. Within island sets, total island area is maintained across sizes classes by halving the number of replicates per twofold increase in the individual areas of constituent islands. Degree of fragmentation decreased across increasing size classes. Aggregate vascular plant species richness and habitat richness are reported by size class.

Table 4-2. Standardized multigroup path coefficient estimates for the effects of habitat diversity (richness), island area (log-transformed), and island isolation (proportion of water within 500-m buffer) on vascular plant species richness. All unstandardized path coefficients were constrained

to single estimates for the small and large island set, without significantly reducing model fit, except for those measuring the direct effect of island area on species richness, which were estimated for the small and large island set independently. The indirect effect of island area on species richness (mediated by habitat diversity) was estimated as the product of the direct effect of island area on habitat diversity and the direct effect habitat diversity on species richness. The total effect of island area on species richness was then estimated as the sum of its direct and indirect effects. Multigroup path analysis model structure is given in Figure 4-4.

Table 5-1. Log likelihood values for linear mixed effects (abundance and occurrence) and linear (species richness) models. For species' abundances, species' occurrences, and species richness, separate models were built for a range of isolation buffers, measuring the proportion of open water within 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m of island shores. Model support significantly declined across increasing isolation buffer sizes in all instances. The most supported buffer size for each model set is highlighted in bold.

Table 5-2. Standardized regression coefficient estimates ("coef."), standard errors ("s.e."), and *P*-values ("*P*") from linear models fitting random placement residuals for species' abundances, species' occurrences, and species richness. Included in all models were island area, measured in m^2, and island isolation, measured as the proportion of water within the most supported buffer size (250 m for abundance and occurrence; 1500 m for species richness). Within abundance and occurrence models, the proportion suitable habitat ("habitat") was measured for each species as the area of suitable habitat on each island divided by the area of the island. Presence/absence of each species' preferred larval host plants ("plants") was included as a binary variable. Wingspan was included as a measure of species' body size and as a proxy of dispersal ability. Each species' total abundance and prevalence were used as an inverse measure of rarity in abundance and

occurrence models, respectively. Species identity was included as a random effect in abundance and occurrence models. Within the species richness model, habitat diversity ("habitat") was estimated as the total number of habitat types recorded on each island. Plant diversity ("plants") was measured as vascular plant species richness. Significant coefficients ($\alpha = 0.05$) are highlighted in bold.

Table 6-1. The number of Speyeria cybele and S. atlantis released at 30, 40, 50, and 60 m that were successful and unsuccessful in navigating to the target island. A subset of butterflies were exposed to a series of intense flashes immediately before release. This method induced flash blindness through bleaching of photoreceptive rhodopsins, without affecting olfaction.

Table 6-2. Summary of generalized linear model (GLM) and generalized linear mixed model (GLMM) results. Coefficient estimates and their corresponding *P*-values are given in parentheses. Coefficient estimates for continuous variables (release distance, wind speed, angular subtense, and distance flown) are standardized. Models 1 and 2 (GLMs) were fitted using a binomial distribution with a logit link function. Navigation success was measured as the success or failure of navigation to the target island. Models 3 and 4 (GLMs) were fit using a gamma distribution with a log link function. Flight speed was measured as total flight time to 10 m. Flight tortuosity was measured as the standard deviation of turn angles. Models 5 and 6 (GLMMs) were fitted using a Tweedie distribution with a log link function. Deviation from wind direction and deviation from island direction were measured as absolute deviations in flight orientations from the wind bearing and the bearing towards the target island's centre, respectively, after 2.5, 5, and 10 m of flight. Individual ID was treated as a random effect to account for lack of independence between successive flight orientations of individuals.

List of figures

smallest island to largest island (small-to-large curve, represented by closed circles connected by solid lines) and from the largest island to smallest island (large-to-small curve, represented by closed triangles connected by dashed lines). The saturation index was estimated as the area under the small-to-large curve relative to that of the large-to-small curve.

Figure 4-4. Multigroup path model structure accounting for species richness, habitat diversity (richness), island area (log-transformed), and island isolation (proportion of water within 500-m buffer). Habitat diversity, island area, and island isolation each directly affects species richness. Island area also directly affects habitat diversity, thereby having an additional indirect effect on species richness. All unstandardized path coefficients were constrained to single estimates for the small and large island set, without significantly reducing model fit, except for those measuring the direct effect of island area on species richness, which were estimated for the small and large island set independently. Residual variances ($1 - R^2$) for species richness and habitat diversity in the small and large island set are reported adjacent to arrows unconnected to other variables. Coefficients associated with the dashed double-headed arrow connecting island area and island isolation represent intercorrelation, which is not treated as a causal path. The direct, indirect, and total effects of habitat diversity, island area, and island isolation on species richness are reported in Table 4-2.

Figure 5-1. Map of the study area, located in Sabaskong Bay, Lake of the Woods, Canada. Butterfly abundance, occurrence, and species richness data were collected for 30 study islands, varying in area from 0.09 to 8.4 ha, using repeated full island surveys.

Figure 5-2. a) Standardized log likelihood values for linear mixed effects (abundance and occurrence) and linear (species richness) models where responding variables were random placement residuals. Explanatory variables for each model are listed in Table 5-2 and Figure 5-3.

Separate models were built using different island isolation buffer sizes, quantifying the proportion of open water within 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m of island shores. To permit comparisons of relationships between log likelihood values and buffer sizes among abundance, occurrence, and species richness model sets, log likelihood scores were standardized for each model set by subtracting the mean and dividing by standard deviation. Model support significantly declined across increasing isolation buffer sizes in all instances. b) Log-log island species-area relationship (ISAR) for butterflies occurring on 30 study islands. Random placement richness values were calculated using a random placement model (model 3; see Materials and Methods). Dashed green lines are 95% confidence intervals for random placement values, calculated using Coleman's (1981) formula for variance. Dashed purple lines represent 95% confidence intervals for the log-log ISAR linear regression (solid purple line), parameterized using observed richness values for all 30 islands.

Figure 5-3. Standardized regression coefficients and 95% confidence intervals from linear models relating random placement residuals to island characteristics and species' functional traits for a) species' abundances, b) species' occurrences, and c) species richness. Included in all models were island area, measured in m^2, and island isolation, measured as the proportion of water within the most supported buffer size (250 m for abundance and occurrence; 1500 m for species richness). Within abundance and occurrence models, the proportion suitable habitat ("suitable habitat") was measured for each species as the area of suitable habitat on each island divided by the area of the island. Presence/absence of each species' preferred larval host plants was included as a binary variable. Wingspan was included as a measure of species' body size and as a proxy of dispersal ability. Each species' total abundance and prevalence were used as an inverse measure of rarity in abundance and occurrence models, respectively. Species identity was

included as a random effect in abundance and occurrence models. Within the species richness model, habitat diversity was estimated as the total number of habitat types recorded on each island. Plant diversity was measured as vascular plant species richness. The shading of each variable's point estimate (coefficient) and confidence interval is proportional to its *P*-value, with darker shades indicating greater significance. Coefficients with 95% confidence intervals not overlapping zero were inferred to be significant at $\alpha = 0.05$.

Figure 6-1. Visual representation of experimental releases. The boat was secured at varying distances from the target island's shore (30, 40, 50, or 60 m), such that the bearing to the island's center was 90° to the wind direction. Butterflies were sexed, marked, and released one at a time. For each released individual, flight time and flight orientation at 2.5, 5, and 10 m of travel were recorded. Angular subtense of the target island, θ_s, was estimated as the angular difference between the left and right shore bearings. Deviations in flight orientations from wind direction (θw) and island direction (θi), given by θdw' and θdi', respectively, were estimated at 2.5, 5, and 10 m of travel using eq. 2.

Figure 6-2. The proportion (success rate) of *Speyeria cybele* and *Speyeria atlantis* that successfully navigated to the target island after experimental release at 30, 40, 50, and 60 m. Flashed butterflies were exposed to a series of intense flashes immediately before release. This method induced flash blindness through bleaching of photoreceptive rhodopsins, without affecting olfaction. Reduced success rates of flashed butterflies indicate that butterflies rely primarily on visual senses to detect and navigate to suitable habitat patches during interpatch and dispersal movements.

Figure 7-1. Visual representation of the methods used to generate geographic and environmental distances between the 161 sequenced *Papilio machaon dodi* included in this study. Various

spatial data layers (a – g) and 375 *P. m. dodi* records (161 sequenced individuals and 214 georeferenced *P. m. dodi* Global Biodiversity Information Facility (GBIF) records) were used to build a Maxent habitat suitability model, which was then used to predict habitat suitability across the study area, with higher values indicating greater suitability. High habitat suitability generally followed the eroding banks of major rivers in southern Alberta and Saskatchewan, Canada (Red Deer, Old Man, South Saskatchewan, and Milk Rivers). Euclidean distances (i) represent the minimum distances between sequenced individuals. A resistance surface was parameterized as the inverse of habitat suitability and used to estimate least-cost path and circuit distances (j and k). The background in inset j is a projected cost surface, representing the cumulative costs incurred by individuals moving across the landscape from each occurrence point. Environmental distances (e – g) were estimated by taking the absolute difference between values of environmental variables extracted from the occurrences of sequenced individuals. Inset pictures are the adult and larval stage of *P. m dodi*.

Figure 7-2. Population genetic structure within *Papilio machaon dodi* in Alberta and Saskatchewan, Canada, inferred using the model-based clustering program Structure. An optimal K value of 2 was best supported by the ΔK method and rate of change in the likelihood of K across $K = 1{:}10$ for the first set of Structure runs addressing all individuals. Hierarchical runs addressing the northern and southern clusters independently suggested no overt subclustering within the northern cluster and the existence of two subclusters within the southern cluster. We present admixture plots derived from the first set of Structure runs for both $K = 2$ (exhibiting the two primary clusters) and $K = 3$ (including the two southern subclusters) for simplicity. For $K = 2$, individuals collected near and north of Dorothy, Alberta, were generally assigned to a northern

cluster, while individuals collected within and south of Dinosaur Provincial Park, Alberta, were assigned to a southern cluster.

Figure 7-3. Pairwise heatmap visualizing reciprocal causal modelling (RCM) results. Values in each cell represent results of $R_{PM\text{-}A}$ - $R_{PM\text{-}B}$, with red and blue colours indicating positive and negative values, respectively. Rows and columns contain the focal and alternative variables, respectively, for model A within each reciprocal model. Therefore, the figure should be interpreted by rows and not columns; variables on the y-axis with more positive (red) values in their corresponding rows are better supported.

Figure 7-4. Associations between allele frequencies of loci (n = 1,382) and a) mean temperature of the warmest quarter of the year (temp.warm), mean temperature of the coolest quarter (temp.cool), and mean annual precipitation (precip.) in latent factor mixed models (LFMM) with K = 2. Black dots are loci with significant associations to the relevant environmental variable based on a q-value threshold of 0.05 ($-\log_{10}$ q-value ~1.3). Single loci often had multiple significant environmental associations due to spatial correlation of environmental variables. Open circles represent the strongest association (based on median $|z|$-scores) for each locus with a significant environmental association. Loci are arranged on the x-axis in order of position within scaffolds, which are in turn arranged by increasing size.

1 Chapter 1: Introduction

1.1 Consilience

"Consilience" is a principle, describing situations wherein evidence from multiple disciplines converge on similar inferences (*sensu* Wilson, 1998). The more distant the disciplines, the greater the probability that their shared inferences are accurate due to increased independence of their respective models. Discrepancies among such inferences are also informative, helping to resolve important biases that may be inherent to paradigms of one or both of the implicated disciplines. It follows that researchers, or at least of subset of them, should immerse themselves in multiple disciplines (or subdisciplines) that address overlapping questions that are of broad interest to their field. Of course, increases to research breadth come at a cost of reductions to depth—the optimal balance within the trade-off is unique to each researcher.

Throughout my Ph.D., I have tended to prioritize breadth of research, aiming to resolve inconsistencies among subdisciplines within ecology and evolutionary biology, including biogeography, landscape ecology, and population genomics. This dissertation is comprised of a series of studies that have played an integral part in my development as a practitioner of these subdisciplines. Throughout, I address a series of integral questions, including: i) what is biodiversity and how is it best measured? ii) how does variation in habitat configuration, habitat composition, and environmental conditions affect emergent patterns of species diversity? and iii) how do these same factors relate to genomic variation within single species?

1.2 A primer on biodiversity

A rainforest is more diverse than a monoculture. But what do we mean by this? Is a rainforest more diverse due to its greater number of species? What if we chose to compare two

rainforests that are equivalent in their total number of species, but, over a one hour stroll, an observer is likely to encounter a greater number of species in one rainforest compared to the other due to differences in their respective species abundance distributions? Clearly, there are multiple facets of biodiversity that are worthy of consideration, both from a theoretical and an applied conservation perspective.

Researchers within subdisciplines of ecology and evolutionary biology often differ in their working definitions of biodiversity. Indeed, we do not have a single measure at our disposal that adequately captures the multiple facets of the concept. However, most researchers, and particularly conservation biologists, tend to agree that, regardless of the particular measure, more biodiversity is generally better (*sensu* Soulé 1985). Within conservation biology, the terms "species diversity" and "biodiversity" are often used interchangeably, although genetic/allelic and landscape/ecosystem diversity are also included within the biodiversity concept (Magurran 2013) and their importance is increasingly recognized within conservation frameworks (e.g., see Chapter 7). Chapters 2 – 5 of this thesis consider diversity at the level of species. At this level of organization, two principal components of diversity must be considered: i) species richness, a count of the total number of species present; and ii) species evenness, a measure of how evenly the abundances of individuals are distributed among those species (Hurlbert 1971; Magurran 2013). Compound indices that simultaneously account for both species richness and evenness are often estimated in an attempt to distil variation in species diversity to a single number. The most common of these indices are information entropies, such as the Shannon-Wiener index, Simpson's index, and their various transformations (Hill 1973; Izsák & Papp 2000). Each of these indices give less weight to rare species, meaning index values will increase if either species richness or evenness increases. This has led some researchers to infer that compound indices

may represent panaceas for quantifying variation in biodiversity (e.g., see reviews in Rosenzweig 1995; Gotelli & Graves 1996; Magurran 2013). However, few studies have addressed the ability of compound indices to resolve variation in species diversity across space, and even fewer, through time (reviewed in Chapter 2).

Chapter 2 details a study (MacDonald et al. 2017) that explicitly evaluates the ability of different indices to resolve changes in the diversity of a species assemblage through time. To quantify diversity patterns in a single species assemblage, I aggregated 10 years of data from John Acorn's long-term butterfly monitoring programme in Edmonton, Canada, which I contributed to over the course of my undergraduate studies. I then used the resulting dataset to show that negative relationships between species richness and species evenness may compromise the efficacy of multiple compound indices. Year-to-year, as environmental conditions became more favorable, richness generally increased, while evenness decreased. Indices that conflate these two measures (e.g., information entropies) therefore failed to resolve important changes in the diversity of the assemblage through time. Species richness, although a very simple measure, adequately captured variation in the diversity of the species assemblage and has the added benefit of ease of interpretation. I therefore concluded that, when analyzing long-term diversity datasets, direct measures of species evenness may be used in conjunction with richness to deepen our understandings of trends in diversity. However, combining these two components into single measures (e.g., information entropies) may not produce measures that consistently align with our intuitive sense of biodiversity.

1.3 Habitat configuration and composition

Chapters 3 – 5 of this thesis focus on relationships between species diversity and the configuration and composition of habitat on fragmented landscapes. Habitat fragmentation is

generally defined as a process, wherein a large expanse of continuous suitable habitat is broken into a number of small fragments that are delineated by a matrix of unsuitable habitat (Fahrig 2003). The development of conservation biology is marked by an ongoing debate, addressing whether anthropogenic habitat fragmentation poses significant threats to species diversity. Although conservation biologists are most interested in habitat fragmentation from a process-oriented perspective, habitat fragmentation is most often addressed within the literature as a landscape pattern, quantifying the relative configuration (e.g., area and isolation) of individual habitat fragments (e.g., see reviews in Fahrig 2003; 2013; 2015; 2018; Hadley & Betts 2016; Fletcher et al. 2018).

Relationships between habitat configuration and species diversity have been of great interest since Levins's (1969) extrapolation of the theory of island biogeography to habitat fragments on terrestrial landscapes (MacArthur & Wilson 1963; Wilson & MacArthur 1967). If edges of terrestrial habitat fragments delimit populations similar to oceanic island shores, isolated populations occupying small fragments may be more prone to stochastic extinction and inbreeding depression than populations occupying larger fragments or unfragmented habitat (Diamond 1972; Diamond 1975; Wilson & Willis 1975; Connor & McCoy 1979; Saccheri et al. 1998; Gonzalez 2000). As per the theory of island biogeography, the relative configuration of habitat fragments is also predicted to affect species diversity because species immigration and rescue effects generally decrease as fragments become more isolated from neighboring habitat (Brown & Kodric-Brown 1977; Hanski 1998). Increasing proportions of habitat edges associated with fragmented habitat have also been linked to changes in species diversity, with reductions more common than promotions (Haddad et al. 2015; Hadley & Betts 2016).

By these and related mechanisms, the process of habitat fragmentation may result in a greater loss of species than what is associated with habitat loss alone (Haila 2002; Fletcher et al. 2018). Such applications of ecological theory are powerful heuristics, made apparent in the inferences of many ecologists: "Habitat fragmentation is considered by many biologists to be the single greatest threat to biological diversity" (Noss 1991 p. 27); "Habitat fragmentation is a major cause of biodiversity erosion" (Tabarelli et al. 1999 p. 119); "Habitat fragmentation is a leading cause of extinction" (Bruna & Oli 2005 p. 1816); and "… the pattern and process of habitat fragmentation been shown to have substantial and lasting effects on biodiversity" (Fletcher et al. 2018 p. 10). However, it is become increasingly recognized that many of these inferences are supported by observations or experimental designs that have not sufficiently decoupled the correlated effects of fragmentation *per se* from those of habitat loss (Fahrig 2003; 2013; 2015; 2018; Hadley & Betts 2016; Fletcher et al. 2018).

In a seminal review of the fragmentation literature, Fahrig (2003) concluded that, in contrast to widely held sentiments that fragmentation reduces species diversity, fragmentation effects are generally negligible after sufficiently controlling for deleterious effects of habitat loss. In the minority of reviewed studies where fragmentation effects were significant, 21 of 30 were positive, suggesting that fragmentation is actually more likely to increase, rather than reduce, species diversity. Fahrig (2013) later reviewed 14 studies addressing whether single large or several small habitat fragments, equivalent in total area, support a higher diversity of species (i.e., the Single Large Or Several Small "SLOSS" debate; Diamond 1975; Abele & Connor 1979; Ovaskainen 2002; Tjørve 2010). In all instances, several small habitat fragments were found to support at least as many species as single large fragments. Accordingly, Fahrig (2013) advanced the habitat amount hypothesis, which predicts that the diversity of species persisting on

fragmented landscapes is primarily a function of the amount of remaining habitat and not its degree of fragmentation. The primary mechanism underlying this prediction is the sample-area effect, describing that larger fragments contain more species simply because they sample more individuals from the abundance distribution of the regional species pool. Mechanistically, this hypothesis is indecipherable from the passive sampling hypothesis, developed in the context of oceanic islands more than 30 years prior (Connor & McCoy 1979). However, applied to conservation practice, the habitat amount hypothesis states that conservation biologists need not consider habitat fragmentation as a factor affecting species persistence and should instead focus on preserving the maximum amount of habitat possible, irrespective of its configuration.

More recently, Fahrig (2017) reviewed 381 studies addressing fragmentation effects on both species diversity and individual species. Studies were categorized as controlling for habitat loss either experimentally (n = 48), statistically (n = 273), or using SLOSS-based methods (n = 60). Overall, fragmentation effects were positive in 76% (290/381) of studies. For studies considering patterns or responses of species diversity (namely, species richness), 90% (114/127) of fragmentation effects were positive. Of particular interest, all 60 studies that used SLOSS-based methods to control for habitat loss reported positive fragmentation effects, with several small patches containing more species than single large patches of equal area (see Fahrig 2020 for further discussion on SLOSS-based inferences). In contrast, studies addressing patterns or responses of single-species were more variable in their results, with 68% (158/232) reporting positive fragmentation effects. Across all studies reviewed by Fahrig (2017), fragmentation effects were more often positive than negative or neutral for both species diversity and single species; however, disparity in the consistency of these findings is of interest (90% vs. 68%, respectively). Greater congruence may be expected if emergent patterns of species diversity are

indeed viable indicators of fragmentation effects on individual species—the level at which species diversity is ultimately structured.

Inferring fragmentation effects from emergent patterns of species diversity may be susceptible to the ecological fallacy (*sensu* Robinson 1950), which addresses biases that may arise when observed effects on aggregated variables (e.g., species richness) differ from causal relationships at finer levels of organization (e.g., single species). Indeed, responses to fragmentation have been shown to vary widely among species (Henle et al. 2004; Ewers & Didham 2006; Öckinger et al. 2009; Hanski 2015). Functional traits, including body size (Gehring & Swihart 2003; Henle et al. 2004; Prugh et al. 2008; Barbaro & Van Halder 2009; Warzecha et al. 2016), mobility/dispersal ability (Roland & Taylor 1997; Lens et al. 2002; Ewers & Didham 2006; Öckinger et al. 2009; MacDonald et al. 2018a), perceptual range (MacDonald et al. 2019), degree of ecological specialization (Tscharntke & Brandl 2004), rarity/conservation status (Ewers & Didham 2006), and trophic position (Tscharntke et al. 2002; Thies et al. 2005; Ewers & Didham 2006) are predicted to relate species' sensitivity to fragmentation. Still, there are very few empirical studies that have related interspecific variation in fragmentation effects to functional traits, or resolved how species-specific fragmentation effects scale to emergent patterns of species diversity (Melbourne et al. 2004; but see Barbaro & Van Halder 2009; Öckinger et al. 2009).

Even within single landscapes and taxa, species also vary widely in their habitat requirements. More than 50 years ago, Williams (1964) inferred that patterns of species diversity across ecological islands are likely driven by variation in their diversity of habitat types (i.e., the habitat diversity hypothesis), rather than area *per se*. Recent studies on both true islands and habitat fragments support this inference (Nilsson 1988; Rosenzweig 1995; Kadmon & Allouche

2007; Hortal et al. 2009; MacDonald et al. 2018b). This suggests that habitat composition is an important consideration in fragmentation research (MacDonald et al. 2018b) and that fragmentation effects may be most sensibly inferred on a species-by-species basis (*sensu* Hanski 2015).

In Chapters 3 and 4, I present two studies (MacDonald et al. 2018a; 2018b) that decouple the correlated effects of habitat fragmentation and habitat loss for butterfly and vascular plant species assemblages persisting on a naturally fragmented landscape of lake islands; Lake of the Woods, Canada. These studies utilize SLOSS-based analyses to show that highly fragmented sets of islands generally contain an equivalent or greater diversity of species than less fragmented sets of islands or a single island of an equal total area. However, I hypothesized that SLOSS-based analyses may obscure important fragmentation effects on individual species, and that interspecific variation in these effects may be related to measurable functional traits. These hypotheses were supported by a number of analyses. For example, when differentiating between potentially reproducing and transient butterfly species on each island based on the presence or absence of their preferred larval food plants, fragmentation effects on the diversity of potentially reproducing species became apparent (Chapter 3; MacDonald et al. 2018a). Interestingly, the probability of butterfly species occurring on islands without their preferred larval food plants was positively related to both wingspan and estimated mobility, suggesting that functional traits may be related to interspecific variation in species' responses to fragmentation. Together, these results suggest that inferring fragmentation effects from emergent patterns of species diversity may constitute an example of ecological fallacy. Additionally, using a structural equation modelling framework, I show in Chapter 4 (MacDonald et al. 2018b) that: i) mechanisms predicted by theory of island biogeography, the sample-area effect, and the habitat diversity

hypothesis may act in a complementarily manner to structure patterns of vascular plant diversity on fragmented landscapes; and ii) the relative importance of these mechanisms changes with island/fragment area.

In Chapter 5 (MacDonald et al. 2021), I present a novel modelling framework for disentangling mechanisms that underlie species-area relationships across true islands and habitat fragments. This is an important and timely contribution, as each of the described mechanisms informs a potentially different conservation directive regarding the relative importance of ameliorating habitat loss and habitat fragmentation. Specifically, the framework uses random placement models to control for variation in species' abundance, species' occupancy, and species richness associated with the sample-area effect, allowing deterministic effects of island/fragment area and isolation to be resolved using linear mixed effects models. Most interestingly, this framework facilitates the simultaneous evaluation of whether and how these effects of area and isolation vary among species in relation to their functional traits. Applying these models to our Lake of the Woods butterfly data showed that the island species-area relationship did not significantly deviate from random placement in relation to island area, isolation, or habitat diversity, supporting stochastic assembly mechanisms consistent with the sample-area effect. This result aligns with SLOSS-based analyses of Chapter 3 (MacDonald et al. 2018a) and conforms to predictions of habitat amount and passive sampling hypotheses (Connor & McCoy 1979; Fahrig 2013). However, further analyses, made possible by the novel modelling framework, show that species' abundances were significantly lower on smaller and more isolated islands than what is predicted by the sample-area effect. Integrating functional traits into this abundance model resolved that the effect of island area was significantly greater for smaller, less mobile, and rare butterfly species. Species' occurrences also significantly deviated from

predictions of the sample-area effect in relation to island isolation, but not area. This case study clearly demonstrates that richness-based analyses can result in incorrect inferences on the mechanisms underlying species-area relationships, obscuring important effects of island/fragment area and isolation on individual species.

1.4 Movement in the matrix

Movements of organisms among habitat fragments is a principal ecological process contributing to both metapopulation persistence and diversity patterns on fragmented landscapes (e.g., Hanski 1998; Wiens 2001; Stevens et al. 2012). Quantifying variation in species' ability to navigate fragmented landscapes is therefore integral to understanding how habitat fragmentation affects species diversity. Related investigations often involve estimating species' perceptual ranges—the maximum distance at which individuals are able to detect suitable habitat or critical resources using their sensory organs (Weins 1989). To infer perceptual range, a popular method is to release individuals at varying distances from habitat fragments and observe their movement through the matrix of unsuitable habitat. For butterflies in particular, it is often assumed that individuals rely on visual senses for habitat detection (see literature review in Chapter 6). However, this assumption has not been explicitly investigated.

In Chapter 6, I assess the extent and sensory determinants of perceptual range for two butterfly species that frequent lake islands of Lake of the Woods: the great spangled fritillary [*Speyeria cybele* (Fabricius, 1775)] and Atlantis fritillary [*Speyeria atlantis* (W.H. Edwards, 1862]. This was achieved by experimentally releasing individuals at various distances from a lake island ("release islands"), representing an isolated habitat fragment situated in a matrix of unsuitable habitat (open water). For each species, the probability of successfully navigating to the release island significantly decreased with increasing release distance, but at different rates.

No specific distance thresholds for perceptual range were observed for either species. Based on this result, I suggested that previous estimates of perceptual range thresholds for a variety of butterfly species are likely statistical artefacts. Instead, perceptual range may be best thought of as a continuum of probabilities, reflecting the likelihood of habitat detection across a range of distances, rather than a fixed distance measure. To infer whether butterflies rely on visual senses when navigating fragmented landscapes, I exposed a subset of individuals to a series of intense light flashes before release to induce flash blindness (compromising visual senses) without affecting olfactory senses. Flashing individuals completely inhibited individuals' ability to navigate to the release island, suggesting that visual senses are the principal determinant perceptual ranges of these species and are integral to detecting habitat fragments. In parallel work, not included in this thesis, I am investigating relationships between perceptual range and fragmentation effects (inferred from models detailed in Chapter 5) for 13 other butterfly species occurring on islands of Lake of the Woods.

1.5 Integrating landscape ecology and population genomics

An overarching conclusion of my research on lake islands was that an autecological approach, considering the responses of individual species to habitat configuration and composition, is most sensible from a theoretical perspective and most effective from a conservation perspective. For example, Chapter 5 shows that addressing variation in the abundances and occurrences of individual species can resolve important, species-specific fragmentation effects that are obscured by emergent patterns of species richness. However, it is difficult to infer from abundance and occurrence data whether individuals occupying isolated habitat fragments represent distinct populations, or how effective dispersal, resulting gene flow, is affected by habitat configuration and composition. Genetic data confer the greatest power for

addressing such questions (Saccheri et al. 1998; Sork et al. 1999; Keyghobadi et al. 2005; Cushman et al. 2006; Shirk et al. 2010; Richardson et al. 2016) and genomic data have the added benefit of resolving adaptive variation and adaptive potential within species (Schwartz et al. 2010; Rellstab et al. 2015; Balkenhol et al. 2017; Storfer et al. 2018). Resolving determinants of genomic variation within species requires equal parts landscape ecology, to map spatial variation habitat configuration, habitat composition, and environmental conditions, and population genomics, to quantify genomic variation at both the individual and population level. Landscape genomics represents the integration of these two disciplines (Sork et al. 1999; Richardson et al. 2016; Storfer et al. 2018).

Chapter 7 details a study (MacDonald et al. 2020) that uses a landscape genomics approach to quantify determinates of genetic divergence within a single butterfly species, Dod's Old World swallowtail (*Papilio machaon dodi* McDunnough, 1939), throughout its Canadian range (southern Alberta and Saskatchewan). A principal goal of this study was to evaluate three overarching hypotheses that explain genetic divergence: i) isolation by distance (Wright 1943), which predicts that geographic distance or physical barriers to dispersal moderate gene flow, thereby permitting drift and genetic divergence between spatially separated individuals or populations; ii) isolation by resistance (McRae 2006), which predicts that, in addition to spatial separation, variation in the relative resistance organisms experience when dispersing through heterogeneous landscapes moderates gene flow, drift, and genetic divergence; and iii) isolation by environment (Wang & Summers 2010), which predicts that differences in environmental conditions contribute to genetic divergence via the combination of: a) reduced fitness/negative selection on individuals that have dispersed across environmental gradients; b) reduced fitness/negative selection on dispersers' offspring in non-natal environments (i.e., outbreeding

depression); or c) reduced propensity of individuals to disperse across environmental gradients due to local adaptation to environmental conditions (Wang & Summers 2010; Wang et al. 2013). Together, predictions of isolation by distance, resistance, and environment may be evaluated in a multiple working hypothesis/strong inference framework (*sensu* Chamberlin 1890; Platt 1998) to infer the relative importance of geographic and environmental factors that affect species' demography, movement, and adaptation within heterogeneous landscapes and environments.

To quantify genomic variation within *P. m. dodi*, I used double digest restriction-site associated DNA sequencing (ddRADseq) and an assortment of bioinformatic pipelines to discover and map thousands of single nucleotide polymorphisms (SNPs) throughout the genomes of 192 sequenced individuals. Using a combination of these SNP data and ecological niche models, I identified two cryptic evolutionary lineages (north *vs* south) that are genetically and ecologically divergent, maintained by local adaptation to spatial variation in climatic conditions. Relating pairwise genetic divergence to a series of geographic and environmental distances resolved that mechanisms predicted by isolation by distance and isolation by environment are contributing to genomic variation within *P. m. dodi*. Interestingly, after controlling for climate-associated genetic variation, connectivity of suitable habitat (isolation by resistance) was not related to genetic connectivity. This unique result challenges a foundational method in landscape ecology: the use of habitat suitability models to infer connectivity among isolated populations when genetic data are unavailable. Following this inference, I suggested that it may be instructive to understand adult life stages of many terrestrial invertebrates as not only the life stage in which mating and reproduction occur, but also as "dispersal machines," exhibiting greater vagility and broader habitat tolerances than larval life stages. Such characteristics are likely to facilitate long-distance dispersal resulting in gene flow across heterogeneous

landscapes, varying in habitat configuration and composition. As a consequence of the dispersal machine concept, I concluded that considerable support for isolation by resistance in vertebrates cannot necessarily be extrapolated to organisms with disparate life histories; particularly, if the life cycles of focal taxa include a discrete dispersal life stage with substantially different habitat constraints than other life stages. The success of this study demonstrates clear utility for integrating theory and methods from landscape ecology and population genomics to address the cumulative effects of habitat fragmentation, habitat loss, and climate change on the ecology and evolution of species.

2 Chapter 2: Negative relationships between species richness and evenness render common diversity indices inadequate for assessing long-term trends in butterfly diversity

2.1 Abstract

Species richness and evenness, the two principal components of species diversity, are frequently used to describe variation in species assemblages in space and time. Compound indices, including variations of both the Shannon–Wiener index and Simpson's index, are assumed to intelligibly integrate species richness and evenness into all-encompassing measures. However, the efficacy of compound indices is disputed by the possibility of inverse relationships between species richness and evenness. Past studies have assessed relationships between various diversity measures across survey locations for a variety of taxa, often finding species richness and evenness to be inversely related. Butterflies are one of the most intensively monitored taxa worldwide, but have been largely neglected in such studies. Long-term butterfly monitoring programs provide a unique opportunity for analyzing how trends in species diversity relate to habitat and environmental conditions. However, analyzing trends in butterfly diversity first requires an assessment of the applicability of common diversity measures to butterfly assemblages. To accomplish this, we quantified relationships between butterfly diversity measures estimated from 10 years of butterfly population data collected in the North Saskatchewan River Valley in Edmonton, Alberta, Canada. Species richness and evenness were inversely related within the butterfly assemblage. We conclude that species evenness may be used in conjunction with richness to deepen our understandings of assemblage organization, but combining these two components within compound indices does not produce measures that consistently align with our intuitive sense of species diversity.

2.2 Introduction

Due to their charismatic nature and popularity among naturalists, butterflies are among the most intensively monitored taxonomic groups worldwide (Thomas 2005; Nowicki et al. 2014). Consequently, butterfly diversity has been invoked as an indicator of biological diversity and environmental health (Fleishman & Murphy 2009; Schmucki et al. 2015). Many ecological traits make butterflies promising ecological indicators. For example, because of their relatively short life cycles, butterfly populations respond rapidly to environmental stimuli, making them sensitive surrogates for trends in habitat and environmental conditions (van Swaay & Warren 1999; Nowicki et al. 2008). Additionally, given that many butterflies complete their life cycles within small patches of habitat, their movements and distributions can also be used to map habitat conditions at relatively fine scales (van Swaay et al. 2006). Finally, butterfly diversity generally correlates with the diversity of many other terrestrial, herbivorous insect groups (Thomas 2005). Together, such groups comprise a significant proportion of terrestrial biological diversity (Nowicki et al. 2008). These traits, coupled with high detectability, make butterflies excellent subjects for long-term diversity monitoring projects. Indeed, Thomas (2005, p. 340) suggests that butterflies "are often the most—or only—practical insect group to study across the world." The question therefore arises: how should the diversity of butterfly assemblages be assessed?

Ecologists have long struggled to find a simple index that is commensurate with the common notion of biological diversity, a.k.a. "biodiversity" (Humphries et al. 1995). The terms "species diversity" and "biodiversity" are often used interchangeably, although genetic/allelic and landscape/ecosystem diversity also fit within broad definitions of biodiversity (Magurran 2013). Species diversity is most often parsed into two principle components: i) species richness, a count of the total number of species; and ii) species evenness, a measure of how evenly

sampled individuals are distributed among species (Hurlbert 1971; Magurran 2013). If data on the relative abundances of species are available, compound indices accounting for both species richness and evenness, most commonly (but not always; see Jost 2006) giving less weight to rare species, may be calculated. The most common and widely used of these indices are variations of the Shannon-Wiener index and Simpson's index (Hill 1973; Izsák & Papp 2000). It has been suggested, however, that not all compound indices are true indices of diversity, but are entropies, reflecting distinct properties of species assemblages relating to diversity (Jost 2006).

Rooted in information theory, the Shannon-Wiener index is an entropy, measuring uncertainty in the outcome of a diversity sampling process. It has also been shown that most other nonparametric compound indices, including Simpson's index, are generalized entropies (Tóthmérész 1995; Ricotta 2003; Keylock 2005). Entropies measure properties of species assemblage data, but are not themselves true measures of species diversity (Jost 2006). Furthermore, these entropies are nonlinear, complicating their interpretations. Transformations of entropies to Hill numbers, indicating the "effective number of species," result in units akin to species richness, allowing for more intuitive, linear comparisons of species diversity (Table 2-1). The effective number of species represents the species richness of a theoretical assemblage, equivalent in diversity to a sampled assemblage (yielding the same value for the root entropy), but with a perfectly even species abundance distribution (Hill 1973; Jost 2006; Chao et al. 2014). The most common of these transformations is Simpson's reciprocal index, calculated as the inverse of Simpson's index. Jost (2006) recommends exponentiating the Shannon-Wiener index to give an effective number of species; however, comparisons of untransformed Shannon-Wiener index values are still common.

Table 2-1. Conversion of common indices to Hill numbers (true diversities), where p_i represents the proportion of individuals within an assemblage belonging to species *i*.

Index	Formula	Transformation to Hill numbers (true diversities)	Transformed index name
Species richness	$S = \sum_{i=1}^{S} p_i^0$	$H_0 = S$	Species richness (S)
Shannon-Wiener index	$H' = -\sum_{i=1}^{S} p_i \ln p_i$	$H_1 = \exp H'$	Exponential of Shannon-Wiener index ($\exp H'$)
Simpson's index	$\gamma = \sum_{i=1}^{S} p_i^2$	$H_2 = \frac{1}{\gamma} = D$	Simpson's reciprocal index (D)

Contrasting with their appeal as all-encompassing species diversity measures, both entropies and their transformations have been criticized for their potential to mask variation among the various components of diversity (Hurlbert 1971; Purvis & Hector 2000; Bock et al. 2007). For instance, species richness and evenness can counteract each other within compound indices if they are negatively correlated (Buzas & Hayek 1996), effectively hiding spatial or temporal gradients in species diversity. This potential for error within indices, coupled with conflicting behaviour between indices, has led some authors to conclude that compound indices are largely meaningless (Hurlbert 1971). Thus, species richness is frequently cited as the most reliable and straightforward measure of species diversity, and remains the central means for identifying biodiversity hotspots and monitoring trends in biodiversity worldwide (Andelman & Willig 2003; Wilsey et al. 2005; Magurran 2013). When using species richness as a single measure of species diversity, it is has been traditionally assumed that: i) species richness is positively correlated with evenness; and ii) species richness accounts for much of the spatial and temporal variation in diversity (Wilsey et al. 2005). These assumptions suggest that relationships between species richness and evenness are consistent between species assemblages, and that

these two components represent different interpretations of a coherent ecological property known as "diversity."

Empirical studies have tested the assumption that species richness and evenness are positively correlated, and whether inconsistencies in the relationships between the two components can compromise the efficacy of compound indices. In partitioning the Shannon-Wiener index into species richness and evenness, Buzas & Hayek (1996) found that species richness and evenness can affect the Shannon-Weiner index in a counteracting manner when negatively related. Stirling and Wilsey (2001) compared empirical relationships between species richness, species evenness, and the Shannon-Wiener index to null relationships generated from Caswell's neutral model (simulation of log-normal and log-series species abundance distributions; Caswell 1976) and found that relationships were generally weak and inconsistent across taxa. Through spatial comparisons of plant diversity, Ma (2005) found no consistent patterns between species richness and evenness, with the two components having different responses to edaphic factors. Wilsey et al. (2005) found that species richness and evenness were positively correlated within invertebrate communities, weakly positively correlated within vertebrate communities, and negatively correlated within plant communities. Bock et al. (2007) found neutral to moderately negative correlations between species richness and evenness within angiosperm, grasshopper, butterfly, bird, and rodent communities in a savanna landscape. In that study, butterfly diversity was assessed on four occasions in four different survey plots. No significant negative correlations between species richness and evenness were found for the butterfly assemblage, although the small sample size in this study brings into question their statistical power to resolve relationships. To our knowledge, all past studies addressing relationships between diversity measures analyzed data across multiple survey locations,

effectively addressing relationships between diversity indices across assemblages that may be influenced by different biophysical factors. In conservation biology, assessing temporal variability of diversity indices for single localities and species assemblages is just as important.

Long-term, standardized biodiversity monitoring programs are essential for assessing the conservation status of species and ecosystems (Schmucki et al. 2015), as well as for measuring the impacts of environmental change on biodiversity (van Swaay et al. 2011). Butterfly diversity monitoring programs, beginning with the establishment of the United Kingdom Butterfly Monitoring Scheme (UKBMS) in 1976, have appeared in a growing number of countries over the last two decades (Schmeller et al. 2009; Schmucki et al. 2015). Corresponding long-term population datasets are likewise expanding. Long-term population datasets are especially relevant for butterflies, among other invertebrate taxa, as they frequently exhibit inter-annual fluctuations in their populations and shifts in phenology that may indicate changes in habitat and environmental conditions (Roy et al. 2001; Saarinen et al. 2003; Westwood & Blair 2010; Schmucki et al. 2015). Butterfly monitoring programs generally place emphasis on the collection of density and abundance data, which can be used to generate species diversity indices (Stephens et al. 2015). Therefore, assessing the viability of species richness, species evenness, and compound indices as measures of butterfly diversity is paramount.

Expanding on the works of Bock et al. (2007), we implemented a higher sampling frequency and longer study duration to assess variability in butterfly species diversity indices through time. Specifically, we examined ten years of butterfly population data from surveys completed over a 16 year period within the largest protected urban green space in North America, the North Saskatchewan River Valley in Edmonton, Alberta, Canada. The objectives of this study were to: i) resolve how different indices of species diversity correlate through time;

and ii) measure the relative variability of diversity indices to identify which indices best capture temporal variation in butterfly assemblages. We hypothesized that butterfly assemblages have a propensity towards unevenness as species' abundances increase under more favourable environmental conditions. To test this hypothesis, butterfly abundance and measures of diversity were correlated with annual precipitation and mean growing season temperature—environmental factors have been shown to positively affect temperate butterfly abundances in past work (e.g., Pollard 1988; Roy et al. 2001).

2.3 Materials and Methods

2.3.1 Study area

Butterfly surveys were completed within the North Saskatchewan River Valley in Edmonton, Alberta, Canada, a municipally protected green space. The specific location of this ongoing study is a south-facing slope along the north bank of the North Saskatchewan River, from Government House Park to the mouths of Ramsay and McKinnon Ravines. The survey location has historically been maintained as a natural recreation area, and is set in a matrix of mixed forest comprised primarily of balsam poplar (*Populus balsamifera*), trembling aspen (*Populus tremuloides*), and white spruce (*Picea glauca*). Grassy areas line the transect route, as well as numerous willows (*Salix* spp.), alders (*Alnus* spp.), and caragana (*Caragana arborescens*). Approximately 60% of the grassy areas along the transect route are mowed regularly, with remaining sections left unkempt. Unkempt grassy areas border the mowed areas along the entire transect route. Nectar resources were generally concentrated within unkempt areas. However, mowed areas do not exceed more than 100 meters in breadth, meaning nectar resources were not substantially fragmented.

2.3.2 Survey methods

A modified form of fixed-route transects, or "Pollard walks," was employed to quantify butterfly diversity and abundances. This method was first conceptualized by Ernest Pollard in the early 1970s, and has since been implemented around the world in a wide array of butterfly monitoring schemes (Pollard 1977; Pellet et al. 2012). Butterfly surveys approximately 1 to 1.5 hours in duration and following a single transect route were completed in 10 different years over a 16-year period (1999, 2000, 2002, 2007, 2009, 2010, 2011, 2012, 2013, and 2014). Surveys with low butterfly activity relative to other surveys within a two week window were eliminated. From the remaining data, eight transect counts were randomly selected from four four-week blocks within each year (two surveys selected per block), from the period between May 1st and August 31st. Starting at the GPS coordinates 53°32.336'N, 113°32.432'W, c. 629 meters elevation, the transect route extends roughly 1,350 meters westward along the base of the south-facing hill slope, to the mouths of Ramsay and McKinnon Ravines, and about half way up a cleared grassy hillside to the stairs leading up to St. George's Crescent, before doubling back 1,050 meters eastward, running parallel to the North Saskatchewan River. Total length of the transect route was approximately 2,400 meters. Surveys started between 12:00 and 15:45, subject to the condition that butterfly activity was apparent. Surveys were only completed in sunny conditions (less than 40% cloud cover) if temperatures were between 13°C and 17°C. At temperatures over 17°C, surveys were completed regardless of cloud cover, but only if butterflies were obviously active. Transect counts were not completed in high wind conditions, or during any form of precipitation. Only butterflies occurring north of the transect route while traveling west were recorded. Similarly, only butterflies occurring to the south of the transect route while doubling back and travelling east were recorded. This method limits the possibility of "double

counts," where individuals are recorded twice in one survey. However, when observing the first individual of any species during a survey, there exists no possibility of a double count, thus individuals belonging to "new species" were recorded irrelevant of their orientation relative to the set transect route.

2.3.3 Diversity analyses

Eight variables related to species diversity were estimated for each survey year, including: butterfly abundance (*N*), species richness (*S*), Simpson's reciprocal index (*D*), the Shannon-Wiener index (H'), the exponential of the Shannon-Wiener index ($\exp H'$), Pielou's evenness (J'), species evenness (D/S), and proportion of rare species (*Rarity*). Butterfly abundance was calculated in each survey year as the sum of individuals observed across all eight surveys. Species richness was the total number of species observed in each survey year. Simpson's reciprocal index was calculated as $D = 1/\Sigma_{i=1}^{S} p_i^2$, where p_i represents the proportion of total butterfly abundance belonging to species *i*. The Shannon-Wiener index was calculated as $H' = -\Sigma_{i=1}^{S} p_i \ln p_i$. The exponential Shannon-Wiener index was calculated as $\exp H'$. Pielou's evenness was calculated as $J' = H'/\ln S$. Species evenness was calculated as D/S. Rarity was calculated in each year as the proportion of species that had relative abundances less than $1/S$ (Camargo 1992). Coefficients of variation (CV) and pairwise product-moment correlations between all eight variables over the 10 survey years were used to infer which measures best captured variation in species diversity and how the different measures covaried through time, respectively. Diversity variables were also correlated with annual precipitation and mean growing season temperature (April through August; calculated from daily highs) of both the same and the previous year, as these environmental factors have been shown to positively affect

butterfly abundances and diversity (Pollard 1988; Roy et al. 2001). Historical weather data was obtained through Environment Canada (http://climate.weather.gc.ca/).

2.4 Results

A total of 37 butterfly species were recorded over the 10 survey years. Table 2-2 summarizes estimates for eight variables relating to species diversity: butterfly abundance (*N*), species richness (*S*), Simpson's reciprocal index (*D*), the Shannon-Wiener index (*H* '), the exponential of the Shannon-Wiener index (exp *H* '), Pielou's evenness (*J* '), species evenness (*D* / *S*), and proportion of rare species (*Rarity*). Total annual butterfly abundance ranged from 285 to 1,270 individuals and species richness ranged from 16 to 30 species. Coefficients of variation (CV) show that, of the three measures of species diversity most commonly reported—species richness, Simpson's reciprocal index, and the Shannon-Weiner index (Hill 1973)—species richness was the most variable (sensitive) through time (CV = 21.6), with Simpson's reciprocal index only slightly less variable (CV = 20.1). The untransformed Shannon-Wiener index proved to be the least variable among common diversity indices (CV = 10.8); however, exponentiating the Shannon-Wiener index appeared to improve discriminating power (exponential of the Shannon-Wiener index; CV = 20.7). Species evenness, as measured by *D*/*S*, proved to be the most variable index of all (CV = 29.7), likely resulting from the negative correlation between the index's two constituent parts: Simpson's reciprocal index and species richness. Of all population variables, butterfly abundance was the most variable, with a coefficient of variation of 58.8.

Table 2-2. Butterfly abundance (*N*), species richness (*S*), Simpson's reciprocal index (*D*), the Shannon-Wiener index (*H* '), the exponential of the Shannon-Wiener index (exp *H* '), Pielou's evenness (*J* '), species evenness (*D / S*), and proportion of rare species (*Rarity*) calculated for 10

years of butterfly surveys completed within the North Saskatchewan River Valley in Edmonton, Alberta, Canada. Coefficient of variation (CV) is given for each metric.

Year	***N***	***S***	***D***	***H '***	**exp *H '***	***J '***	***D / S***	***Rarity***
1999	436	20	3.79	1.91	6.72	0.64	0.19	0.8
2000	419	18	3.6	1.68	5.34	0.58	0.2	0.72
2002	361	23	5.48	2.17	8.77	0.69	0.24	0.78
2007	434	19	4.04	1.79	5.98	0.61	0.21	0.78
2009	367	20	3.77	1.75	5.75	0.58	0.19	0.8
2010	318	17	3.7	1.76	5.81	0.62	0.22	0.82
2011	285	16	4.2	1.77	5.87	0.64	0.26	0.75
2012	970	28	3.43	1.61	4.99	0.48	0.12	0.85
2013	1 270	30	2.84	1.5	4.47	0.44	0.09	0.9
2014	840	23	2.79	1.58	4.85	0.5	0.12	0.82
CV	58.8	21.6	20.1	10.8	20.7	13.7	29.7	6.3

Table 2-3. Pairwise product-moment correlation coefficients for butterfly abundance (*N*) and seven indices related to species diversity— species richness (*S*), Simpson's reciprocal index (*D*), the Shannon-Wiener index (*H* '), the exponential of the Shannon-Wiener index (exp *H* '), Pielou's evenness (*J* '), species evenness (*D*/*S*), and proportion of rare species (*Rarity*)—derived from 10 years of butterfly population data.

	N	***S***	***D***	***H '***	**exp *H '***	***J '***	***D/S***
S	0.91***						
D	-0.67*	-0.35					
H '	-0.69*	-0.37	0.94***				
exp *H* '	-0.61	-0.28	0.93***	0.99***			
J '	-0.92***	-0.74*	0.85**	0.90***	0.85**		
D/*S*	-0.93***	-0.83**	0.79**	0.72*	0.67*	0.93***	
Rarity	0.82**	0.81**	-0.52	-0.47	-0.41	-0.71*	-0.79**

significance is denoted by * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Pairwise product-moment correlations show that species richness was positively correlated with butterfly abundance ($P < 0.001$) and of proportion rare species ($P < 0.01$), but was negatively correlated with Pielou's evenness ($P < 0.05$) and species evenness (D/S; $P < 0.01$; Figure 2-1; Table 2-3). Species richness correlated more strongly with untransformed butterfly abundances ($r = 0.91$) than natural log-transformed butterfly abundances ($r = 0.90$). No significant correlations were observed between species richness and Simpson's reciprocal index ($P > 0.05$), the Shannon-Weiner index ($P > 0.05$), or the exponential of the Shannon-Wiener index ($P > 0.05$), although correlation coefficients between species richness and compound indices were consistently negative. Butterfly abundance was positively correlated with proportion of rare species ($P < 0.01$), and was negatively correlated with all diversity indices accounting for species evenness (Simpson's reciprocal index [$P < 0.05$], the Shannon-Wiener index [$P < 0.05$], Pielou's evenness [$P < 0.001$], and species evenness [D/S; $P < 0.001$]), save the exponential of the Shannon-Wiener index, with which it was still negatively related ($P < 0.1$). Both measures of evenness were positively correlated with all three compound indices (Simpson's reciprocal index, the Shannon-Weiner index, and the exponential of the Shannon-Weiner index).

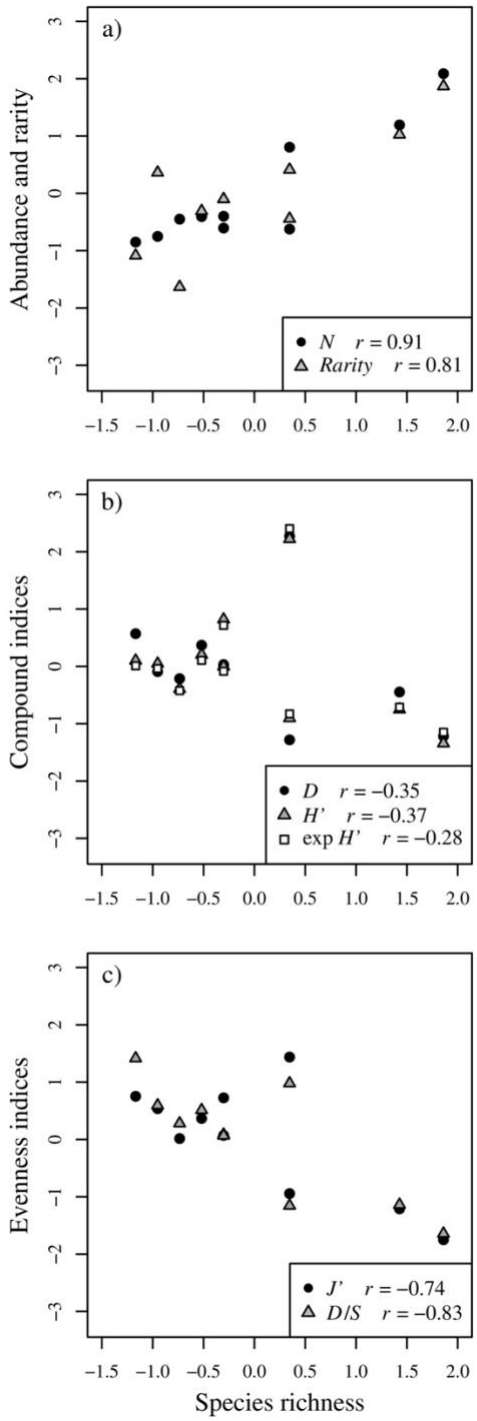

Figure 2-1. Scatter plots of a) Butterfly abundance (*N*) and proportion of rare species (*Rarity*) as they relate to species richness; b) Simpson's reciprocal index (*D*), the Shannon-Wiener index (*H* '), and the exponential of the Shannon-Wiener index (exp *H* ') as they relate to species richness;

and b) Pielou's evenness (J') and species evenness (D/S) as they relate to species richness. All eight variables are in units of standard deviation. r = corresponding pairwise product-moment correlation coefficient.

Relationships between diversity measures and environmental conditions (annual precipitation and mean growing season temperature) of the same year were consistently stronger than relationships between diversity measures and environmental conditions of the previous year. Butterfly abundance (N), species richness (S), and proportion of rare species (*Rarity*) were positively related to annual precipitation and mean growing season temperature of the same year (Table 2-4). Contrastingly, indices accounting for evenness (D, H', exp H') and measuring evenness explicitly (J', D/S) were negatively related to annual precipitation and mean growing season temperature of the same year.

Table 2-4. Pairwise product-moment correlation coefficients for eight variables related to butterfly diversity and annual precipitation (mm) and mean growing season temperature (°C; April – August; based on daily highs) of the same year for Edmonton, Alberta, Canada.

	N	***S***	***D***	***H'***	**exp *H'***	***J***	***D / S***	***Rarity***
Precipitation	0.52	0.25	-0.65*	-0.60	-0.58	-0.51	-0.45	0.34
Temperature	0.45	0.28	-0.48	-0.50	-0.52	-0.48	-0.43	0.42

significance is denoted by * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

2.5 Discussion

Species richness, butterfly abundance, and proportion of rare species were all positively correlated within our long-term butterfly population dataset. All three measures were also positively related to annual precipitation and to mean growing season temperature. These findings agree with other studies suggesting that butterfly populations respond quickly to favourable environmental conditions (i.e., higher precipitation and warmer temperatures; Pollard 1988; Roy et al. 2001). It seems likely that, in years of higher precipitation and warmer temperatures, increases in species richness are driven by increases in the abundances, and thus detection, of rare species. Notably, conditions for butterflies improved significantly in 2012, a year in which numerous irruptive species also made appearances, and with conditions also favourable the following year, 2013 produced the highest butterfly abundance and species richness values. By 2014, conditions were returning to less favourable measures.

In contrast, increases in butterfly abundance and species richness were associated with decreases in the two measures of species evenness (J' and D/S). These relationships can be explained in terms of either: i) increases in the abundances (and thus detection) of rare species, which decreases evenness; or ii) disproportional increases in abundances of common species (Pollard et al. 1995), which also decreases evenness. Our results support both possibilities, as the proportion of rare species was negatively correlated with both measures of evenness, and rapid population growth within common species was observed in years with favourable environmental conditions (e.g., *Thymelicus lineola*, *Glaucopsyche lygdamus*, and *Pieris rapae*). Our measure of rarity was not independent of the abundances of common species and thus separating the effects of these two factors on evenness was not possible.

Population increases in common species are expected to reduce species evenness when relative magnitudes of population increase are unequal among species within assemblages (*sensu*

Gosselin 2006). It is therefore predictable that butterfly assemblages may inherently become "less even" under favorable environmental conditions. Our results support this hypothesis. This relationship may indicate (but does not demonstrate) interspecific competition within assemblages, where rapid increases in the populations of common species moderate population increases in other species, effectively depressing species evenness (Stirling and Wilsey 2001; Mulder et al. 2004; Bock et al. 2007). Interspecific competition has been shown to decrease species evenness within plant assemblages (Mulder et al. 2004), but to our knowledge, this relationship remains to be empirically tested within butterfly assemblages. Considering that butterfly species vary in both host and nectar plant species (Hawkins and Porter 2003; Kitahara et al. 2008), niche overlap and interspecific competition are not expected to be the primary determinants of abundance distributions. More likely, negative relationships between butterfly abundance and measures of species evenness within our data may indicate differences in reproductive potential across species coupled with interspecific variation in environmental preferences. These hypotheses warrant further study.

Most interestingly, species richness in our study was not correlated with Simpson's reciprocal index, the Shannon-Weiner index, or the exponential of the Shannon-Wiener index. Both measures of species evenness, however, were positively correlated with the three compound indices. These results suggest that the compound indices weigh the evenness component of diversity more heavily than the richness component. This obscures the importance of species richness—the measure that accords best with our intuitive sense of biodiversity (Ma 2005; Magurran 2013). Indeed, coefficients of variation suggest that species richness captured more variability than any of the three compound indices, although Simpson's reciprocal index and the exponential of the Shannon-Wiener index proved to be almost as variable as species

richness. The untransformed Shannon-Weiner index had the lowest coefficient of variation among species richness, measures of species evenness, and the three compound indices, suggesting that it has the greatest potential to mask temporal variability in species diversity.

Besides the possibility of negative correlations between constituent parts, the Shannon-Wiener index has been criticized for its potential to "compress" data due to the log transformation of species proportion values, effectively weakening the discriminating power of the index (Magurran 2013). To compensate for this data compression, the Shannon-Wiener index may be exponentiated (Jost 2006), yielding a measure of the effective number of species. Within our butterfly population dataset, exponentiating the Shannon-Wiener index effectively improved the discriminating power while maintaining approximate relationships with other diversity measures. However, both untransformed and exponentiated Shannon-Wiener index values were weakly negatively related to species richness, suggesting the two indices are not consistent with our intuitive sense of diversity. The Shannon-Wiener index does, however, serve as an entropy, effectively measuring a distinct property of species assemblage data.

Using logarithms with a base of two in the Shannon-Wiener index results in the average minimum number of yes/no questions required to determine the species identity of a sampled individual (Jost 2006). A more even distribution of individuals among species will require more dichotomous questions, on average, to determine individuals' species identities. This indicates higher uncertainty within abundance distributions, and yields higher Shannon-Wiener index values. Additionally, as Simpson's index represents the probability that two species randomly selected from an assemblage will be of the same species, the inverse of this probability (Simpson's reciprocal index) is positively related to the uncertainty of individuals' species identities. These relationships effectively explain why compound indices were strongly related to

species evenness in our butterfly assemblage, with increases in evenness corresponding with increases in uncertainty in the identities of sampled individuals. Compound indices and species evenness therefore represent similar interpretations of abundance distributions. These results, coupled with weak, negative relationships to richness, suggest that compound indices do not integrate species richness and species evenness into all-encompassing measures that align well with the traditional biodiversity concept. While entropies and their respective transformations are mathematically related to species richness (Gosselin 2006; Jost 2006), their empirical relationships with richness have been shown to be largely inconsistent (e.g., Stirling and Wilsey 2001; Wilsey et al. 2005; Bock et al. 2007), and, as our study confers, often negative.

Distinctions between richness, evenness, and entropies (including their transformations to Hill numbers) are necessary because the three approaches to measuring diversity convey different information on distinct, but related, properties of assemblage organization. By analogy, Jost (2006, p 363) states that, "the radius of a sphere is an index of its volume but is not itself the volume, and using the radius in place of the volume in engineering equations will give dangerously misleading results. This is what biologists have done with diversity indices." However, this too may be misleading. Relationships between radii and volumes are consistent across spheres of all sizes, and a sphere's radius serves as a proxy for other properties relating to size (e.g., diameter, circumference, surface area, and volume). Although species evenness, species richness, and compound indices all measure different aspects of single species assemblages, the broad extension of the sphere metaphor to measures of assemblage diversity is unadvisable. Unlike the properties of a sphere, relationships between measures of species diversity are unpredictable, varying across space (Ma 2005), between taxa (Stirling and Wilsey 2001; Wilsey et al. 2005; Bock et al. 2007), and, as our study shows, through time. One measure

of species diversity cannot adequately proxy another; thus, single measures cannot be used to assess all properties of an assemblage's diversity in isolation. Despite this point, however, species richness is thought to be more central to the concept of diversity than other measures (Ma 2005; Jost 2006; Magurran 2013), and does, on its own, provide meaningful information on species assemblages.

In conclusion, when analyzing long-term butterfly population datasets, species evenness may be used in conjunction with richness to deepen our understandings of changes in butterfly diversity through time; however, combining these two components within compound indices does not produce measures that consistently align with our intuitive sense of biodiversity. Compound indices measure different, but not independent, properties (organization) of species assemblages. Entropies, including Simpson's index the Shannon-Wiener index, measure uncertainty, not diversity, which is a property of data, not ecological communities. Entropies may convey meaningful information about butterfly assemblage data if used in conjunction with more intelligible measures, such as richness, and the transformations of entropies into Hill numbers result in values that are easier to interpret. Species richness is the most viable measure of butterfly species diversity. However, other diversity indices accounting for abundance distributions can provide additional information, effectively deepening our understandings of the enigmatic ecological property that is diversity.

3 Chapter 3: Decoupling habitat fragmentation from habitat loss: butterfly species mobility obscures fragmentation effects in a naturally fragmented landscape of lake islands

3.1 Abstract

Since the publication of the theory of island biogeography, ecologists have postulated that fragmentation of continuous habitat presents a prominent threat to species diversity. However, negative fragmentation effects may be artifacts; the result of species diversity declining with habitat loss, and habitat loss correlating positively with degree of fragmentation. In this study, we used butterfly assemblages on islands of Lake of the Woods, Ontario, Canada to decouple habitat fragmentation from habitat loss and test two competing hypotheses: i) the island effect hypothesis, which predicts that decreasing fragment size and increasing fragment isolation reduces species diversity beyond the effects of habitat loss; and ii) the habitat amount hypothesis, which negates fragmentation effects and predicts that only total habitat area determines the diversity of species persisting on fragmented landscapes. Using eight independent size classes of islands (ranging from 0.1 to 8.0 ha) that varied in number of islands while holding total area constant, species diversity comparisons, species accumulation curves, and species–area relationship extrapolations demonstrated that smaller insular habitats contained at least as many butterfly species as continuous habitat. However, when highly mobile species occurring on islands without their larval food plants were excluded from analyses, island effects on potentially reproducing species became apparent. Similarly, generalized linear models suggested that effects of island isolation and vascular plant richness on insular butterfly richness were confounded by species of high mobility. We conclude that inter-fragment movements of highly mobile species may obscure important fragmentation effects on potentially reproducing populations, questioning support for the habitat amount hypothesis.

3.2 Introduction

Within continuous habitats, island archipelagoes, and fragmented landscapes, species richness increases with total area surveyed (MacArthur & Wilson 1963, Wilson & MacArthur 1967; Wilson & Willis 1975). Indeed, the positive species–area relationship (SAR) is widely cited as the closest thing ecology has to a law (Schoener 1976; Lomolino 2000). As a corollary of the SAR, loss of habitat results in loss of species (He & Hubbell 2011); however, the configuration of remaining habitat is widely thought to also have an effect on species diversity (Mendenhall et al. 2014; Haddad et al. 2017). Linked to the process of habitat loss, variation in the size and isolation of habitat fragments may be described as degree of habitat fragmentation (Fahrig 2003). While some have gone so far as to assert that "habitat fragmentation is considered by many biologists to be the single greatest threat to biological diversity" (e.g., Noss 1991 p. 27), in many studies, habitat fragmentation has not been distinguished from habitat loss. After decoupling habitat fragmentation from habitat loss, effects of fragmentation appear to vary widely within and among both landscapes and taxa (Quinn & Harrison 1988; Debinski & Holt 2000; Fahrig 2003; Mendenhall et al. 2014).

Over the past half-century, ecologists have related a variety of theories to the fragmentation problem; perhaps the most prevalent and influential being the theory of island biogeography (MacArthur & Wilson 1963, Wilson & MacArthur 1967). Drawing on the heuristic power of island biogeography, ecologists have frequently likened the ecologies of oceanic archipelagos to those of fragmented landscapes (Haila 1986 1990; Ovaskainen 2002; Fahrig 2013; Haddad et al. 2017). As with oceanic islands, species diversity within habitat fragments is predicted to reach equilibria between colonization and extinction rates, principally

determined by fragment size and isolation. We refer to this application of equilibrium theory as the "island effect hypothesis," which predicts that habitat fragmentation reduces species diversity below what is predicted based on habitat loss alone. A key assumption of this hypothesis is that fragment edges delimit species assemblages, such that population processes, including colonization and extinction, occur at the level of habitat fragments. If fragments are too small to support viable populations following their isolation, a gradual loss of species ("faunal relaxation") will reduce species diversity at a rate inversely related to fragment size (Diamond 1972; 1975; MacArthur & Wilson 1963; Wilson & MacArthur 1967; Connor & McCoy 1979; Gonzalez 2000). Consequentially, slopes of species–area curves (z values) across isolated habitat fragments are predicted to be steeper than those within continuous habitats (Gonzalez 2000; Haddad et al. 2017). In the context of habitat fragmentation, the ultimate "steady-state" legacy of an island effect will be several small habitat fragments supporting fewer species than continuous habitat of equivalent area (Gonzalez 2000; Fahrig 2013).

In contrast to the island effect hypothesis, the recently proposed "habitat amount hypothesis" suggests that the size and isolation of habitat fragments have little effect on species diversity (Fahrig 2013). The central premise of this hypothesis is that fragment edges do not delimit populations, such that only the aggregate amount of habitat determines the number of species persisting on fragmented landscapes. It follows that negative relationships between habitat fragmentation and species diversity are best interpreted as artifacts; the result of species diversity declining with habitat loss, and habitat loss correlating positively with degree of fragmentation (Harrison & Bruna 1999; Fahrig 2003; 2013; Yaacobi et al. 2007). Faunal relaxation (or the gradual loss of species) following habitat loss is therefore predicted as a landscape-level process, unrelated to the configuration of remaining habitat. In sum, predictions

of the habitat amount hypothesis are indistinguishable from the passive sampling hypothesis, developed in the context of oceanic islands (Connor & McCoy 1979). Both hypotheses predict that species richness increases with fragment/island size only because of the sample area effect: larger sample areas contain more individuals, which, for a given abundance distribution, belong to more species (Fahrig 2013). If the sample area effect best explains patterns of species diversity on fragmented landscapes, species–area curves across isolated habitats will approximate those of continuous habitats (Haddad et al. 2017). In consequence, fragmented and continuous habitat of equivalent total area should support equivalent numbers of species (Gonzalez 2000; Fahrig 2013).

While differences in species–area slopes between fragmented and continuous habitat have been interpreted as evidence against the habitat amount hypothesis (Haddad et al. 2017), Gotelli & Graves (1996 p. 227) states that, "…the most sensible view is that slopes of species–area curves are simply fitted constants, with little or no biological significance." Indeed, slopes of species–area curves have been found to vary unpredictably within and between both sampling locations and taxa (Connor & McCoy 1979), with *z* values clustering in certain ranges by chance and because of reporting biases in the literature (Gotelli & Graves 1996; Gonzalez 2000). In contrast with *z* value interpretations, comparing species diversity across sets of habitat fragments (while controlling for total habitat area) is a tractable method for assessing habitat fragmentation–species diversity relationships (Yaacobi et al. 2007; Gavish et al. 2012; Fahrig 2013). Such analyses bear on the long-standing SLOSS debate, addressing whether single large or several smaller fragments, equivalent in total area, contain (and therefore protect) more species (Diamond 1975; Abele & Connor 1979; Simberloff & Abele 1982; Ovaskainen 2002; Tjørve 2010). In a recent review of the fragmentation literature, Fahrig (2013) examined 14

studies addressing SLOSS directly. All studies reported equivalent or higher species richness within several smaller habitat fragments compared to fewer larger fragments or continuous habitat, suggesting that fragmentation does not reduce species diversity after habitat loss is controlled for. An additional study (Yaacobi et al. 2007) assessed patterns in species diversity on a fragmented agricultural landscape and reported similar findings.

Interestingly, studies supporting the habitat amount hypothesis (Yaacobi et al. 2007; Fahrig 2013) do not attempt to disentangle fragmentation effects on individual species, or differentiate between potentially reproducing species and transient species observed to occupy habitat fragments. While the sample area effect may best explain patterns of entire species assemblages on fragmented landscapes, assessments of fragmentation–species diversity relationships may be misleading if variation in fragmentation effects among species are not considered (Ewers & Didham 2006; Öckinger et al. 2009; Betzholtz & Franzén 2011; Franzén & Betzholtz 2012; Hanski 2015), and in particular, if potentially reproducing species are not distinguished from transient species temporarily occupying individual habitat fragments. Indeed, inter-fragment movements of highly mobile species from larger habitat fragments (supporting reproducing populations) to smaller fragments containing additional resources (e.g., Fretwell and Calver 1969; Dreisig 1995) have great potential to obscure fragmentation–species diversity relationships. Although several smaller and fewer (or single) large fragments may be observed to contain equivalent numbers of species, it should not be assumed that smaller fragments are capable of supporting viable populations in the absence of larger fragments present on the landscape.

In this study, we surveyed butterfly species diversity on lake islands in Sabaskong Bay, Lake of the Woods, Ontario, Canada. The tens of thousands of islands in Lake of the Woods

represent a suitable system for decoupling habitat fragmentation from habitat loss to assess how habitat configuration relates to patterns of species diversity. Islands of Sabaskong Bay are remnant fragments of continuous habitat that was flooded 3000–4000 years ago in the early Subboreal period (Yang and Teller 2005), permitting an assessment of fragmentation effects on a landscape where insular biotas have likely relaxed to equilibria following habitat loss and fragment isolation (MacArthur and Wilson 1963, Wilson and MacArthur 1967; Haila 2002). Furthermore, akin to oceanic islands, habitat boundaries in this system are strictly delimited by water, and butterflies cannot utilize surrounding aquatic habitats at any life stage. This effectively controls for "matrix effects," whereby the matrix of unsuitable habitat contributes to species diversity, or differentially facilitates the inter-fragment movements of individuals (Ricketts 2001; Haila 2002).

By differentiating between potential resident and transient butterfly species based on occurrences of larval food plant species, we were able to investigate fragmentation effects on both the complete species assemblage and a subset of potential resident (reproducing) species. The detection of negative fragmentation effects in our study system would indicate that several smaller insular habitats do not support butterfly species diversity to the same extent as continuous habitat of equivalent area. Such a result may be attributed either to an island effect, or to decreases in habitat suitability or habitat diversity within smaller islands (Gotelli & Graves 1996). In contrast, support for the habitat amount hypothesis would indicate that insular and continuous habitats contribute equally to butterfly species diversity, and that fragment edges do not delimit butterfly populations at the scales addressed.

3.3 Methods

3.3.1 Study design and focal taxon

Thirty islands within Sabaskong Bay were selected to represent a nested-set sampling design that effectively decoupled habitat fragmentation from habitat loss (Table 3-1; Figure 3-1). Specifically, islands were organized into two island sets that were used to assess fragmentation effects across two distinct ranges of island sizes. The first (small) island set contained 15 islands organized into four size classes, including eight 0.1-ha islands, four 0.2-ha islands, two 0.4-ha islands, and a single 0.8-ha island. The second (large) island set followed an identical pattern using islands ranging from 1.0 to 8.0 ha. By doubling the area of individual islands per twofold reduction in number of replicates, we were able to vary the degree of fragmentation across size classes while holding total habitat area constant. (See MacDonald et al. 2018a Electronic Supplementary Material [Appendix 1] for study island selection criteria.)

Table 3-1. Nested-set sampling design. Within island sets, aggregate area is maintained across sizes classes by doubling the size of constituent study islands when the number of replicates is reduced by half.

Island set	Size class (ha)	Number of replicates	Aggregate area (ha)
Small	0.1	8	0.8
Small	0.2	4	0.8
Small	0.4	2	0.8
Small	0.8	1	0.8
Large	1.0	8	8.0
Large	2.0	4	8.0
Large	4.0	2	8.0
Large	8.0	1	8.0

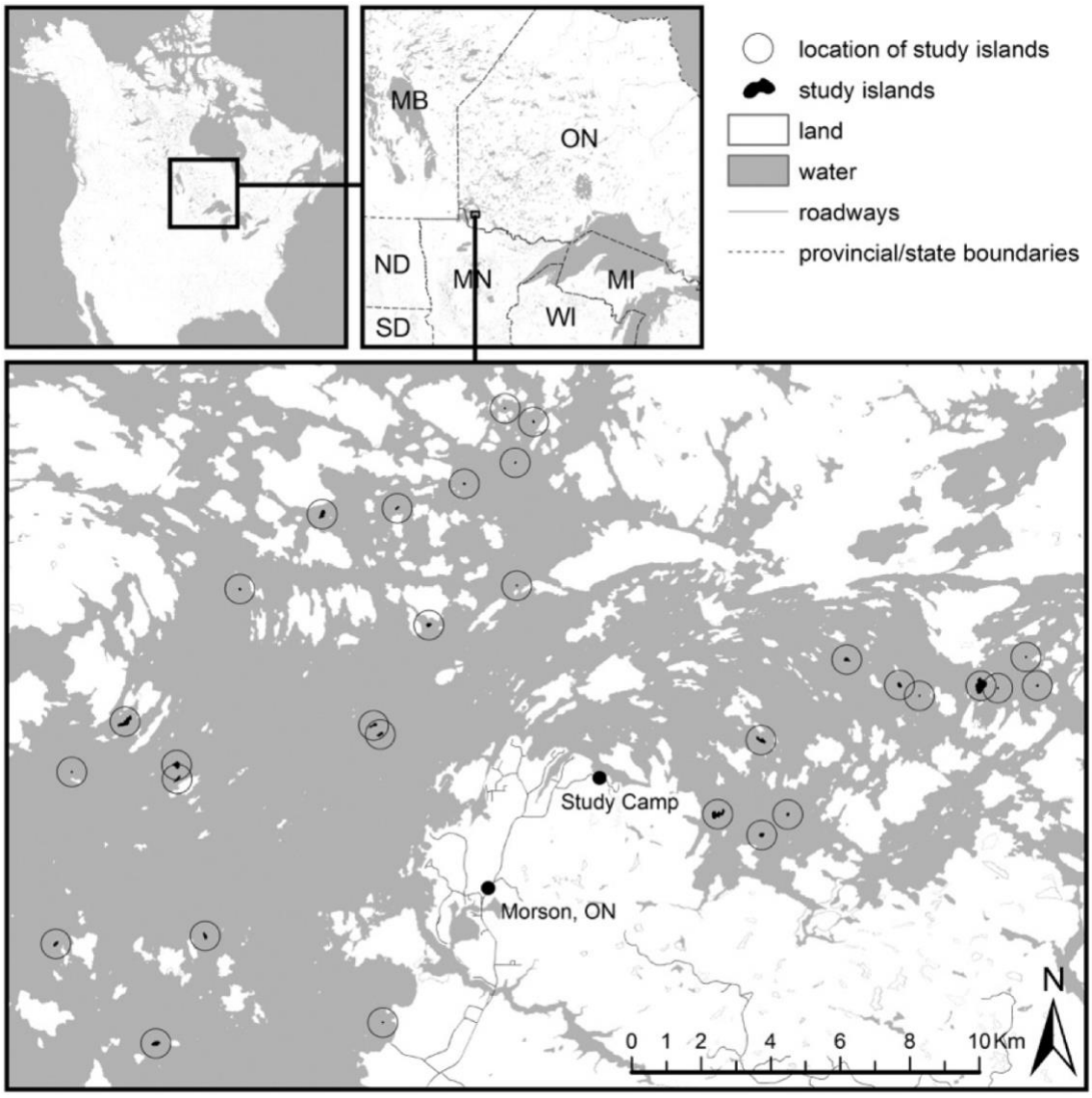

Figure 3-1. Map of the study area, located in Sabaskong Bay, Lake of the Woods, Ontario, Canada. All study islands were located within 20 km of the study camp, approximately 6 km northwest of Morson, Ontario.

Butterflies were chosen as the focal taxon for this study because: i) they are commonly used as model organisms in ecology and data on their biology are widely available (Baguette & Van Dyck 2007); ii) the majority of butterfly species complete their lifecycles within relatively

small patches of habitat (van Swaay et al. 2006; Nowicki et al. 2008), meaning butterfly diversity may serve as a proxy habitat suitability within islands; iii) butterflies cannot utilize aquatic habitats at any morphological stage, rendering the matrix of open water surrounding islands completely uninhabitable; and iv) butterfly occurrences representing potential resident (reproducing) and transient (non-reproducing) populations may be distinguished on individual islands by the presence or absence of their known larval food plants. Coupled with high detectability and established sampling methods (Pollard 1977), these traits make butterflies well-suited study organisms for assessing relationships between fragmentation and species diversity in our study system.

3.3.2 Survey methods

Butterfly species richness and abundance were estimated on each of the 30 islands through repeated full-island surveys, standardized to a survey time of 40 min per ha. This protocol ensured that sampling effort per unit area was consistent across islands of all sizes, eliminating the need for rarefaction, extrapolation, or other diversity corrections (Chao et al. 2014). Four rounds of butterfly surveys were completed by a single observer. Each island was visited at intervals between 10 and 14 days during peak flight season (from 01-June-2015 to 20-Aug-2015). Our survey protocol was similar to that outlined by Pollard (1977) to ensure that butterfly activity was optimal and consistent between surveys (see MacDonald et al. 2018a Electronic Supplementary Material [Appendix 1] for survey protocol details). Vascular plant species richness was surveyed by a second observer on all 30 islands using repeated vegetation surveys. Plant surveys were also standardized to a time of 40 min per ha, with four surveys completed per island (see MacDonald et al. 2018b [Thesis Chapter 4] for further details). Fourteen habitat classes, defined using vegetation and substrate characteristics, were used to

quantify habitat diversity within each island (see MacDonald et al. 2018a Electronic Supplementary Material [Appendix 1; Table A1] for habitat class descriptions).

3.3.3 Analyses

3.3.3.1 Comparisons of species diversity

For each of the eight island size classes, we estimated the effective number of species using species richness, the exponential of the Shannon–Wiener index, and Simpson's reciprocal index (Jost 2006; MacDonald et al. 2017; (see MacDonald et al. 2018a Electronic Supplementary Material [Appendix 1] for index equations). We hypothesized that if fragmentation decreased butterfly diversity, diversity measures would be lowest within the smallest sizes class (representing higher degrees of fragmentation) and increase with larger size classes (representing lower degrees of fragmentation). Such a result would support the island effect hypothesis. Other arrangements of species richness across size classes would suggest that fragmentation did not reduce species diversity, but would not necessarily support the habitat hypothesis. The habitat amount hypothesis specifically predicts that species diversity is unrelated to the number of habitat fragments when total habitat area is held constant. This prediction would equate to an even distribution of species diversity across size classes.

3.3.3.2 Species accumulation curves (SACs) and saturation index

For the two island sets separately and together, cumulative species richness was plotted against cumulative island area in two different ways: i) increasing order of island size (small to large); and ii) decreasing order of island size (large to small). Slight variation in island area within size classes allowed for the sorting of islands in a reasoned manner. Data points were connected with straight lines to generate SACs, which were constrained to pass through the

origin to allow for area-under-the-curve comparisons (Quinn & Harrison 1988; Gavish et al. 2012). Similar slopes between large-to-small and small-to-large SACs would suggest that species richness increased with cumulative area, irrespective of the number of fragments. Steeper slopes of large-to-small SACs would suggest that fewer/single larger islands contained more species, while steeper small-to-large SACs would indicate that several smaller islands contained more species.

Differences in slopes between large-to-small and small-to-large SACs were quantified using a saturation index (Quinn & Harrison 1988). This index is estimated as the area under the small-to-large SAC divided by that of the large-to-small SAC. To estimate area under the SACs, integrals were calculated using the trapezoidal rule. Saturation index estimates less than one would indicate negative fragmentation effects, which lend support to the island effect hypothesis. Index estimates equal to one would indicate that fragmentation did not affect species diversity, supporting the habitat amount hypothesis. Index estimates greater than one would indicate that fragmentation increased species diversity. This positive fragmentation effect is not predicted by either the island effect or habitat amount hypothesis, but may result from several smaller islands intersecting the distributions of more species than fewer or single larger island (Tjørve 2010; Fahrig 2013). Alternatively, several smaller islands may contain a higher diversity of habitat types than fewer or single larger island, which may support a higher diversity of species (Williams 1964; Nilsson et al. 1988).

3.3.3.3 SAR extrapolation and SLOSS index

For the two island sets separately and together, a log–log least-squares linear regression was applied to island area and species richness to attain a SAR (Yaacobi et al. 2007; Gavish et al. 2012). To allow the logarithmic transformation of a single 0.1-ha island with a species richness

of zero, a constant of one was added to all richness values. Each SAR was extrapolated to the aggregate area of all islands used to build the SAR (small island set = 3.21 ha; large island set = 32.13 ha; both island sets together = 35.34 ha). Substituting this aggregate area into the SAR regression (and subtracting a constant of one to account for the original transformation) yielded a species richness estimate for continuous habitat equivalent in area to all study islands. In SLOSS terms, this richness estimate (S_{sl}) represents the "single large" conservation strategy, while the aggregate observed richness of study islands (S_{ss}) represents the "several small" conservation strategy (Gavish et al. 2012).

To test whether estimated (S_{sl}) and observed (S_{ss}) richness values were significantly different, 95% confidence intervals were extrapolated for each of the SARs. If the SAR predicted a significantly higher number of species than observed ($S_{sl} > S_{ss}$), habitat fragmentation reduced species richness, supporting the island effect hypothesis. Conversely, if SAR regressions accurately predicted aggregate observed richness ($S_{sl} \approx S_{ss}$), habitat configuration was not related to species richness, supporting the habitat amount hypothesis. If the SAR predicted a significantly lower number of species than observed ($S_{sl} < S_{ss}$), habitat fragmentation increased species richness. A SLOSS index, estimated as $100\% \times (S_{ss} - S_{sl})/ S_{ss}$, indicates the proportion of species richness of several small fragments relative to that of a single large fragment (Boecklen 1997; Gavish et al. 2012). For example, a SLOSS index value of 20% (or − 20%) would indicate that the study islands representing the several small conservation strategy (S_{ss}) contained 20% more species (or 20% fewer species) from the species pool than continuous habitat representing the single large conservation strategy (S_{sl}).

3.3.3.4 Area-independent effects on butterfly species diversity (GLMs)

Generalized linear models (GLMs) and an information theoretic approach (Burnham & Anderson 2004) were used to test for effects of island isolation, habitat diversity, vascular plant species richness, and island shape (relative habitat edge) on butterfly species richness. Negative binomial regressions were used to account for overdispersion within island butterfly species richness data (Ver Hoef & Boveng 2007). Variance inflation factors (VIFs) were used to test for collinearity among explanatory variables, with a value of 10 used as a maximum cutoff (Craney & Surles 2002). Models were ranked for support using the small sample size corrected Akaike's Information Criterion (AIC_c), where smaller AIC_c values indicate higher relative model support (Burnham & Anderson 2004). Coefficients from our best-supported GLMs were standardized to permit comparisons of the relative importance of island attributes in structuring butterfly species richness.

Island area (log-transformed) was included as a covariate in most models to control for the expected positive SAR. A univariate log (area) model therefore represented an "ecological null" model for assessing the relative effects of other island characteristics on butterfly species richness. To test for the effects of island isolation on butterfly species richness, we quantified the proportion of open water (1 − proportion landmass) within various buffer sizes calculated from island edges (250, 500, 1000, 2500, and 5000 m). This measure of isolation is independent from the area of specific study islands. Proportion of uninhabitable matrix surrounding fragments has been shown to be a stronger predictor of dispersal and fragment immigration than distance-based metrics, justifying this measure (Moilanen & Nieminen 2002; Tischendorf et al. 2003). Beyond the size and isolation of islands, habitat diversity and plant diversity may contribute to patterns in butterfly species richness. Habitat diversity is expected to positively relate to butterfly richness

because butterfly species vary in their habitat requirements (i.e., the habitat diversity hypothesis, Williams 1964; Nilsson et al. 1988). Not unrelated, vascular plant species richness may serve as a proxy for larval food plant diversity, breadth and seasonal availability of nectar resources, and habitat diversity. Both habitat diversity and plant richness are expected to positively relate to island area (Nilsson et al. 1988; Gotelli & Graves 1996). However, habitat diversity and plant richness should make a statistical contribution to GLMs beyond the variation explained by island area if they contribute to patterns in butterfly richness (Gotelli & Graves 1996).

Within single islands, edges and interiors may differentially support butterfly populations. Edge effects therefore represent another causal mechanism that may affect patterns in butterfly species richness (Saunders et al. 1991; Murcia 1995; Stasek et al. 2008). To capture variability in the amount of island edge independent of island area, a relative edge index was estimated as the perimeter of a given study island made relative to the perimeter of a theoretical island identical in size but perfectly circular in shape. Values approaching one represented islands with minimal habitat edge, with higher values indicating increased habitat edge. Negative relationships between the relative edge index and butterfly richness would indicate that the increased habitat edge associated with fragmentation reduced butterfly diversity.

3.3.3.5 Potential resident and transient butterfly occurrences

While multiple studies addressing entire species assemblages support the habitat amount hypothesis (Fahrig 2013), responses to fragmentation may vary between species (Henle et al. 2004; Ewers and Didham 2006). Within single taxa, such variability in responses is often linked to species' mobility (Roland and Taylor 1997; Lens et al. 2002; Ewers and Didham 2006; Öckinger et al. 2009). In the case of lake islands, butterfly species of high mobility may utilize

resources on small islands that do not contain their larval food plants, thereby temporarily contributing to observed species diversity without constituting reproducing populations.

To test this hypothesis, we first used GLMs (logit link) to relate the probability of observing a butterfly species on at least one island where their larval food plant was not detected to: i) species' wingspans (mm; Burke et al. 2011; Hall et al. 2014); and ii) a species mobility index generated by Burke et al. (2011). Species' wingspans were log-transformed to improve model fit (Burke et al. 2011). The prevalence (number of occurrences) of both butterfly species and their larval food plants was controlled for in GLMs as covariates. We then classified butterfly species occurrences on all 30 study islands as either "potential resident" or "transient" based on the presence or absence of known larval food plants [larval food plant associations were compiled from records summarized by Hall et al. (2014) and Acorn and Sheldon (2017)]. We were not able to distinguish between potential resident and transient populations of one butterfly species, *Feniseca tarquinius*, based on food plant occurrences because larvae are known only to feed on woolly aphids (Eriosomatinae; Hall et al. 2014; Acorn & Sheldon 2017). *Feniseca tarquinius* was therefore excluded from the subset of potential resident species. *Danaus plexippus*, *Vanessa virginiensis*, *V. cardui*, and *V. atalanta* are migratory species and are not known to complete their life cycles within our study area (Hall et al. 2014; Acorn & Sheldon 2017), so were also excluded.

We proceeded to repeat our analyses ("comparisons of species diversity," "species accumulation curves [SACs] and saturation index," "SAR extrapolation and SLOSS index," and "area-independent effects on butterfly species diversity [GLMs]") using only the potential resident species subset of the complete species assemblage. Through this reanalysis, we were

able to investigate whether inter-island movements of highly mobile species obscured fragmentation effects on potentially reproducing populations.

3.3.3.6 Spatial patterns in species turnover

To further investigate ecological mechanisms structuring butterfly diversity, 28 pairwise comparisons of butterfly species turnover (using the complete species assemblage), habitat turnover, plant species turnover, and inter-island distance were made between islands within the 0.1- and 1.0-ha size classes. This isolated relationships between the complete butterfly assemblage, the vascular plant assemblage, habitat composition, and the spatial distribution of islands, while holding habitat area constant. Butterfly species turnover, plant species turnover, and habitat turnover were estimated using the Jaccard pairwise dissimilarity index (Baselga & Orme 2012). We tested for relationships between butterfly species turnover and inter-island distance because several small islands spread over a larger spatial extant may intersect the distributions of more species than fewer or single larger island (*sensu* Tjørve 2010; Fahrig 2013). This effect of several small islands "sampling" a larger species pool due to their broader spatial distribution may therefore obscure fragmentation effects using SLOSS-based analyses. Failure to detect positive relationships between species turnover and inter-island distance would indicate that the spatial distribution of small islands did not confound SLOSS-based analyses. Mantel tests (999 permutations) were used to assess whether relationships were significant. All statistical analyses were performed in the program R (R Core Team 2017).

3.4 Results

Complete species assemblage: A total of 82 butterflies belonging to 13 species and 786 butterflies belonging to 33 species were observed on islands within the small and large island set, respectively (Table 3-2). Butterfly diversity of the small island set was almost perfectly nested

within that of the large island set, with 12 of the 13 species observed in the small fragment also observed in the large island set. (See MacDonald et al. 2018a Electronic Supplementary Material for species occurrence/abundance data by island [Online Resource 1] and the documented range expansion of *Euphyes dion* [Appendix 1].)

Potential resident species subset: A total of 53 butterflies belonging to ten species and 684 butterflies belonging to 29 species were observed on islands within the small and large island sets, respectively. Potential resident butterfly diversity of the small island set was perfectly nested within that of the large.

Table 3-2. Butterfly abundance (*N*), species richness (*S*), the exponential of the Shannon-Wiener index (exp *H*'), and Simpson's reciprocal index (*D*) for the complete species assemblage and the potential resident species subset. The complete species assemblage included all occurrences, whereas potential resident species subset was limited to species' co-occurrence on islands with their known larval food plant species.

Complete species assemblage:					
Small island set	**Smallest eight islands (~0.1 ha)**	**Next four islands (~0.2 ha)**	**Next two islands (~0.4 ha)**	**Largest island (~0.8 ha)**	**Aggregate area (~3.2 ha)**
N	40	10	23	9	82
S	6	7	4	5	13
exp *H*'	3.02	5.74	2.68	4.33	5.64
D	2.35	4.55	2.37	3.86	3.58
Large island set	**Smallest eight islands (~1.0 ha)**	**Next four islands (~2.0 ha)**	**Next two islands (~4.0 ha)**	**Largest island (~8.0 ha)**	**Aggregate area (~32.0 ha)**
N	156	188	261	181	786
S	18	21	18	22	33
exp *H*'	6.99	12.46	9.34	9.94	13.73
D	4.19	10.02	6.47	5.74	10.2
Resident species subset:					
Small island set	**Smallest eight islands (~0.1 ha)**	**Next four islands (~0.2 ha)**	**Next two islands (~0.4 ha)**	**Largest island (~0.8 ha)**	**Aggregate area (~3.2 ha)**
N	16	6	23	8	53
S	3	5	4	4	10
exp *H*'	1.59	4.76	2.68	3.51	4.39
D	1.29	4.5	2.37	3.2	2.93
Large island set	**Smallest eight islands (~1.0 ha)**	**Next four islands (~2.0 ha)**	**Next two islands (~4.0 ha)**	**Largest island (~8.0 ha)**	**Aggregate area (~32.0 ha)**
N	130	161	217	176	684
S	16	18	17	21	29
exp *H*'	5.73	11.16	8.52	9.32	12.69
D	3.24	8.94	5.48	5.45	9.19

3.4.1 Comparisons of species diversity

Complete species assemblage: Each of the two smallest size classes contained more butterfly species than the two largest size classes in the small island set (Table 3-2). While the exponential of the Shannon–Wiener index and Simpson's reciprocal index did not exhibit this relationship, MacDonald et al. (2017) demonstrate that these indices may fail to capture variation in butterfly species diversity and caution their interpretation. Overall, effective numbers of species for the complete species assemblage showed no clear relationship to degree of fragmentation in the large island set.

Potential resident species subset: Effective numbers of species generally showed no clear relationship to island size class in either the small or large island set; the exception being the relatively high number of species in the largest size class (single 8-ha island; $S = 21$) compared with other size classes in the large island set.

3.4.2 Species accumulation curves (SACs) and saturation index

Complete species assemblage: When species accumulation was plotted against cumulative fragment area for the small island set, the small-to-large SAC lay considerably above the large-to-small SAC (saturation = 1.22; Figure 3-2). These results indicate that fragmented habitat generally contained more butterfly species than continuous habitat at this scale. Incongruently, slopes of SACs were quite similar for the large island set (saturation = 0.96) and for both island sets together (saturation = 0.98). This suggests that the positive fragmentation effect observed between the small island set's SACs was not preserved when both island sets were aggregated for SAC analysis.

Potential resident species subset: The small-to-large and large-to-small SACs were overlapping for the small island set (saturation = 1.03). This result indicates that fragmented and continuous habitat contained equivalent numbers of potential resident butterfly species at this scale. In contrast with this pattern, large-to-small SACs lay above small-to-large SACs for the large island set (saturation = 0.92) and both island sets together (saturation = 0.89). Such results indicate a negative fragmentation effect on potentially reproducing butterfly populations. Overall, saturation was lower for the resident species subset than for the complete species assemblage at all scales (small, large, and both island sets together).

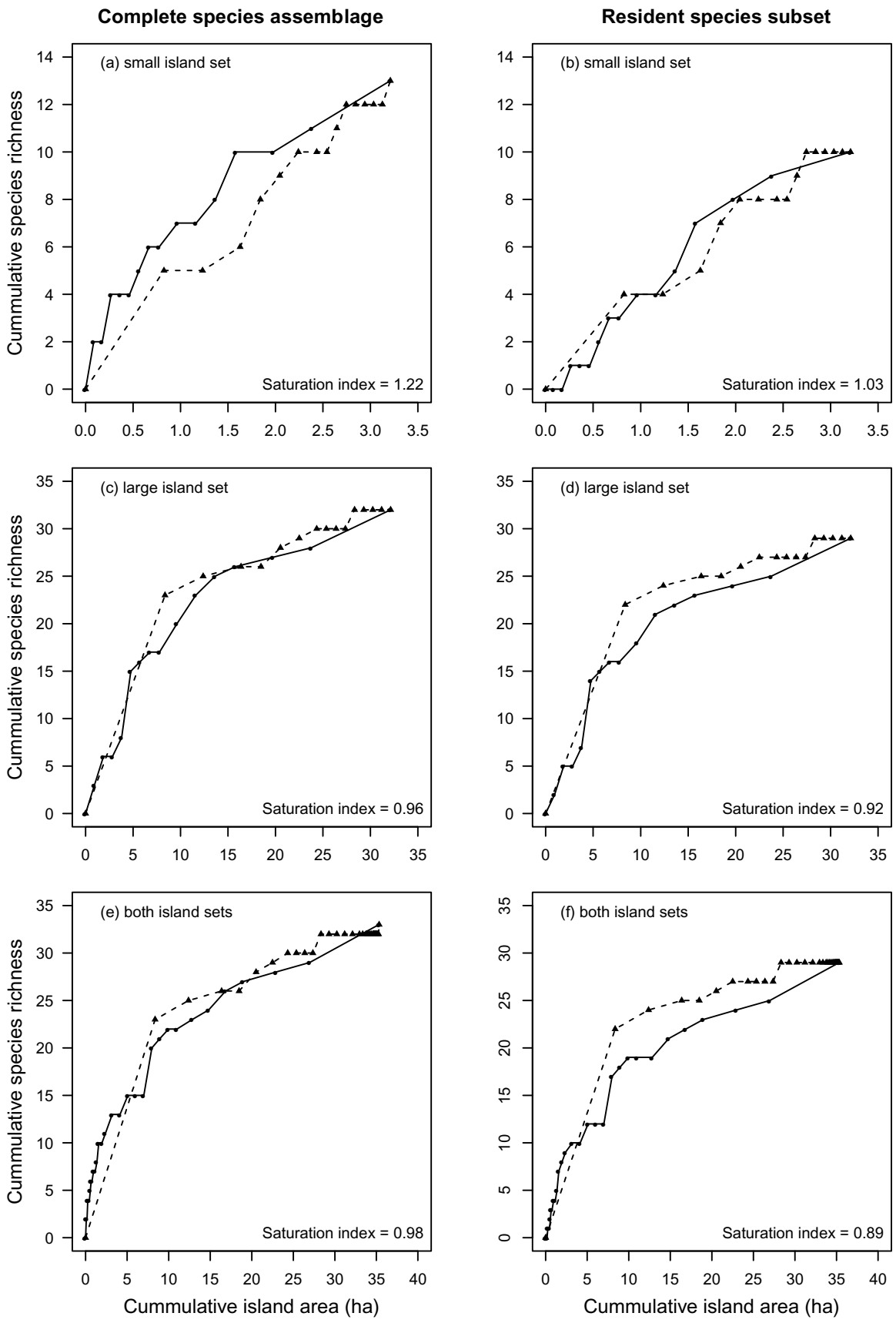
Complete species assemblage
Resident species subset
(a) small island set
(b) small island set
(c) large island set
(d) large island set
(e) both island sets
(f) both island sets
Saturation index = 1.22
Saturation index = 1.03
Saturation index = 0.96
Saturation index = 0.92
Saturation index = 0.98
Saturation index = 0.89
Cummulative species richness
Cummulative island area (ha)

Figure 3-2. Cumulative species richness relative to cumulative island area for the complete species assemblage (a, c, and e) and the potential resident species subset (b, d, and f). Accumulation of species richness occurs from the smallest to the largest island (small-to-large SAC; represented by closed circles connected by solid lines) and from the largest to smallest island (large-to-small; represented by closed triangles connected by dashed lines). Saturation index values are estimated as the area under the small-to-large SAC divided by that of the large-to-small SAC.

3.4.3 SAR extrapolation and SLOSS index

Complete species assemblage: Compared to SAR species richness estimates for continuous habitat, we observed higher aggregate richness across study islands in the small island set (S_{sl} = 6.14; S_{ss} = 13; SLOSS index = 53%) and lower aggregate richness across study islands in the large island set (S_{sl} = 64.42; S_{ss} = 33; SLOSS index = − 95%; Figure 3-3). When all 30 islands were combined for SAR extrapolation (both island sets), the SAR richness estimate for continuous habitat was very close to the aggregate richness observed across all study islands (S_{sl} = 33.33; S_{ss} = 34; SLOSS index = − 1%). For all three SAR extrapolations, aggregate richness observed across study islands fell within the extrapolated SAR's 95% confidence intervals.

Potential resident species subset: Compared to SAR species richness estimates for continuous habitat, we observed approximately equivalent aggregate richness across study islands in the small island set (S_{sl} = 9.55; S_{ss} = 10; SLOSS index = 5%), lower aggregate richness across study islands in the large island set (S_{sl} = 73.04; S_{ss} = 29; SLOSS index = − 152%), and lower aggregate richness when both island sets were considered together (S_{sl} = 40.20; S_{ss} = 29;

SLOSS index = − 40%). Again, however, aggregate richness observed across study islands fell within the extrapolated SAR's 95% confidence intervals for all thee SAR extrapolations. Of particular interest, SLOSS index estimates were lower for the potential resident species subset than for the complete assemblage at all scales (small, large, and both island sets together). Furthermore, slopes of SARs (*z* values) were greater for the potential resident species subset than for the complete species assemblage at all scales.

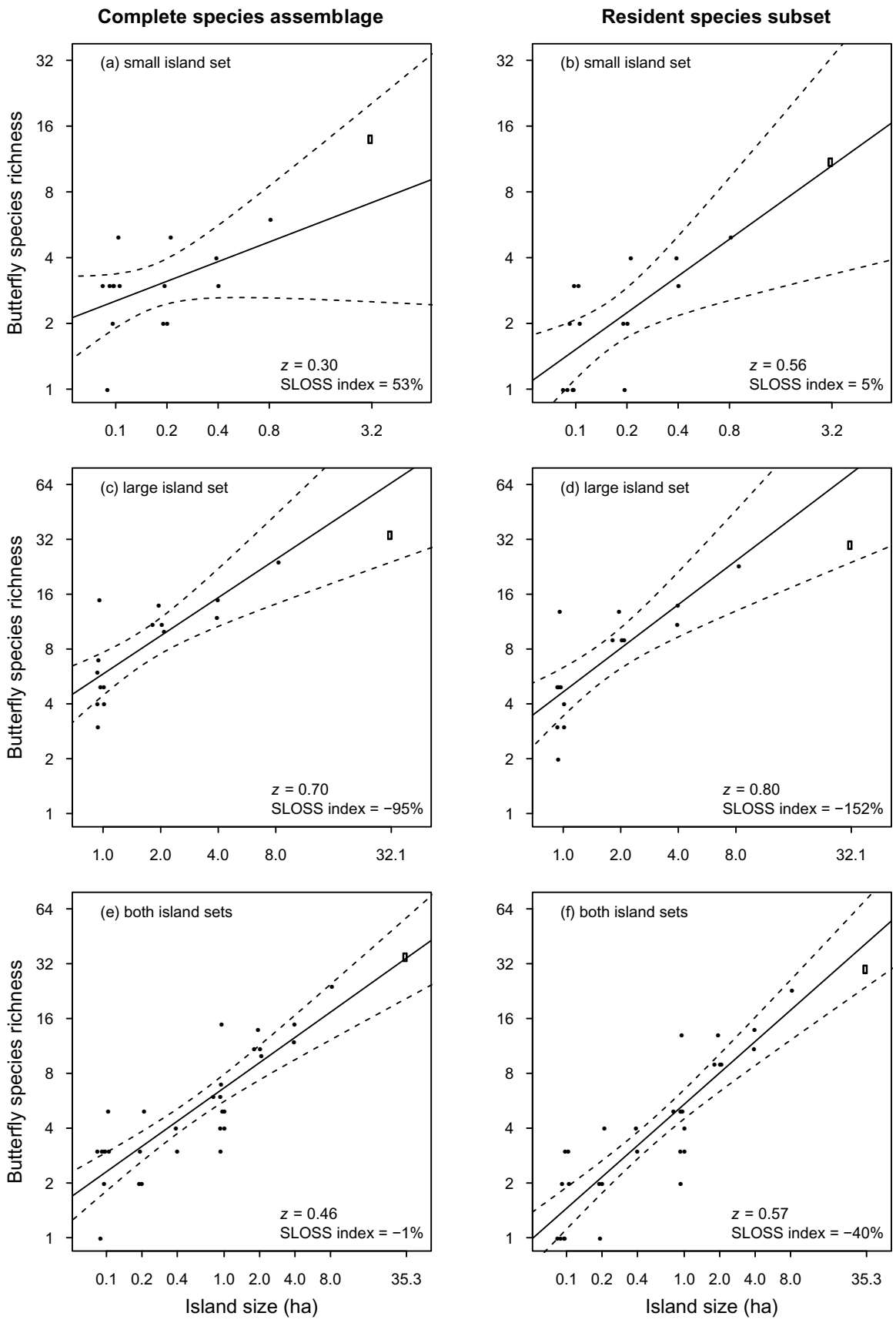
Complete species assemblage
Resident species subset
(a) small island set
(b) small island set
(c) large island set
(d) large island set
(e) both island sets
(f) both island sets
z = 0.30
SLOSS index = 53%
z = 0.56
SLOSS index = 5%
z = 0.70
SLOSS index = −95%
z = 0.80
SLOSS index = −152%
z = 0.46
SLOSS index = −1%
z = 0.57
SLOSS index = −40%
Butterfly species richness
Island size (ha)
0.1 0.2 0.4 0.8 3.2
1.0 2.0 4.0 8.0 32.1
0.1 0.2 0.4 1.0 2.0 4.0 8.0 35.3
1 2 4 8 16 32 64

Figure 3-3. SAR extrapolations for the complete species assemblage (a, c, and e) and the potential resident species subset (b, d, and f). Solid and dashed lines represent log-log least-squares linear SAR regressions and their 95% confidence intervals, respectively. Area coefficients of log-log SAR regressions are reported as *z*-values, approximating exponents of the species-area power model ($S = cA^z$).Closed and open circles represent species richness for individual islands and their aggregate richness, respectively (a constant of one was added to all richness values to allow for log-transformations). Axes were back-transformed from logarithmic to linear scales for straightforward interpretation of species richness and area values. SLOSS index values were estimated as $100\% \times (S_{ss} - S_{sl}) / S_{ss}$, where S_{ss} represents the aggregate observed richness of study islands and S_{sl} represents the SAR's richness estimate for continuous habitat of equivalent areal extent.

3.4.4 Area-independent effects on butterfly species diversity (GLMs)

Complete species assemblage: The best-supported model explaining patterns of butterfly species richness across study islands accounted for island area and vascular plant richness, with both variables relating positively to butterfly richness (Tables 3-3 & 3-4). Standardized coefficients indicate that island area had a greater effect than vascular plant richness on butterfly richness. The inclusion of a log(area):plant richness interaction term decreased model support, indicating the relationship was consistent across islands sizes. After controlling for the positive relationship between island area and butterfly species richness, the inclusion of island isolation measures (proportion of open water within 250-, 500-, 1000-, 2500-, and 5000-m buffers) decreased model support in all cases. However, three isolation measures (250-, 5000-, and 500-m buffers, in descending order of support) decreased model support by less than two AIC_c points,

indicating isolation effects were uncertain. As both the island effect and habitat amount hypothesis predict, the proportion of open water within the best-supported buffer size (250 m) was negatively related butterfly richness. Model support was reduced with inclusion of habitat richness and the relative island edge index, suggesting that habitat diversity and edge effects did not contribute to patterns in butterfly species richness beyond the variation explained by island area.

Potential resident species subset: When transient butterfly species were excluded from the complete assemblage, island area and island isolation (250-m buffer) were found to best explain variation in potential resident butterfly richness. In this model, island isolation was negatively related to butterfly richness and had a greater effect than island area. Models accounting for other isolation buffers decreased model support by more than two AIC_c points, indicating they were not well-supported. Vascular plant richness decreased model support wherever included, suggesting it was not related to insular patterns in richness of potential resident species. The standardized effects of both island area and isolation on butterfly richness were always greater for the resident species than for the complete species assemblage. VIFs were less than 10 for all models, suggesting collinearity was not problematic (Craney & Surles 2002).

Table 3-3. Island characteristics regressed on butterfly species richness (complete species assemblage and the potential resident species subset). The complete species assemblage included all occurrences, whereas potential resident species subset was limited to species' co-occurrence on islands with their known larval food plant species. Models were ranked for support using the corrected Akaike's Information Criterion (AIC_c) where smaller AIC_c values indicate better-supported models. Relative model weights based on AIC_c are given by w_i.

Complete species assemblage:	**AIC_c**	**ΔAIC_c**	**w_i**
log(area)[1] + plant richness[2]	129.37	0	0.16
log(area) (ECOLOGICAL NULL)	130.08	0.72	0.11
log(area) + isolation (250 m)[5]	130.28	0.92	0.1
log(area) + isolation (5000 m)[9]	130.46	1.09	0.09
log(area) + isolation (500 m)[6]	130.92	1.55	0.08
log(area) + plant richness + isolation (250 m)	131.02	1.65	0.07
plant richness	131.17	1.8	0.07
log(area) + plant richness + habitat richness[3]	131.33	1.96	0.06
log(area) + plant richness + log(area) × plant richness	131.43	2.07	0.06
log(area) + isolation (2500 m)[8]	132.19	2.83	0.04
log(area) + isolation (1000 m)[7]	132.46	3.09	0.03
log(area) + habitat richness	132.46	3.1	0.03
log(area) + relative edge[4]	132.53	3.17	0.03
log(area) + isolation (250 m) + log(area) isolation (250 m)	132.89	3.52	0.03
log(area) + habitat richness + log(area) × habitat richness	135.04	5.67	0.01
log(area) + relative edge + log(area) × relative edge	135.06	5.7	0.01
NULL	168.13	38.77	0
Resident species subset:			
log(area) + isolation (250 m)	114.34	0	0.35
log(area) + isolation (5000 m)	116.58	2.24	0.12
log(area) + isolation (2500 m)	116.58	2.24	0.12
log(area) + plant richness + isolation (250 m)	117.39	3.05	0.08
log(area) + isolation (250 m) + log(area) × isolation (250 m)	117.62	3.28	0.07
log(area) + isolation (500 m)	117.85	3.51	0.06
log(area) (ECOLOGICAL NULL)	118.35	4.01	0.05
log(area) + plant richness + log(area) × plant richness	118.46	4.12	0.05
log(area) + plant richness	118.88	4.54	0.04
log(area) + relative edge	120.64	6.3	0.02
log(area) + plant richness + habitat richness	120.75	6.41	0.01
log(area) + isolation (1000 m)	120.79	6.45	0.01
log(area) + habitat richness	120.83	6.49	0.01
log(area) + habitat richness + log(area) × habitat richness	121.68	7.34	0.01
log(area) + relative edge + log(area) × relative edge	122.1	7.76	0.01
plant richness	122.49	8.15	0.01
NULL	159.79	45.45	0

Table 3-4. Standardized coefficients (β) for parameters from the best-supported models shown in Table 3-3, where island characteristics were regressed on butterfly species richness for the complete species assemblage and the potential resident species subset. The complete species assemblage included all observed butterfly occurrences, whereas potential resident species subset was limited to butterfly species' co-occurrence on islands with their known larval food plant species.

Complete species assemblage:	**AIC_c**	**β log(area)**	**SE log(area)**	**β plant richness**	**SE plant richness**	**β isolation (250 m)**	**SE isolation (250 m)**
log(area)[1] + plant richness[2]	129.37	0.456	0.21	0.33	0.19		
log(area)	130.08	0.806	0.09				
log(area) + isolation (250 m)[3]	130.28	0.826	0.09			-2.25	1.44
Resident species subset:							
log(area) + isolation (250 m)	114.34	1.072	0.107			-3.982	1.608
log(area)	118.35	1.026	0.103				
log(area) + plant richness	118.88	0.692	0.256	0.299	0.215		

[1] natural log of island area (m^2), [2] island vascular plant species richness, [3] proportion of open water within a 250-m buffer (calculated from island edge)

3.4.5 Potential resident and transient butterfly occurrences

Logistic regressions indicate that the probability of observing a butterfly species on at least one island without their larval food plants was positively related to wingspan ($P = 0.011$) and species mobility ($P = 0.0075$; Figure 3-4). These relationships remained significantly positive after accounting for the prevalence (number of occurrences) of both butterfly species and their food plants.

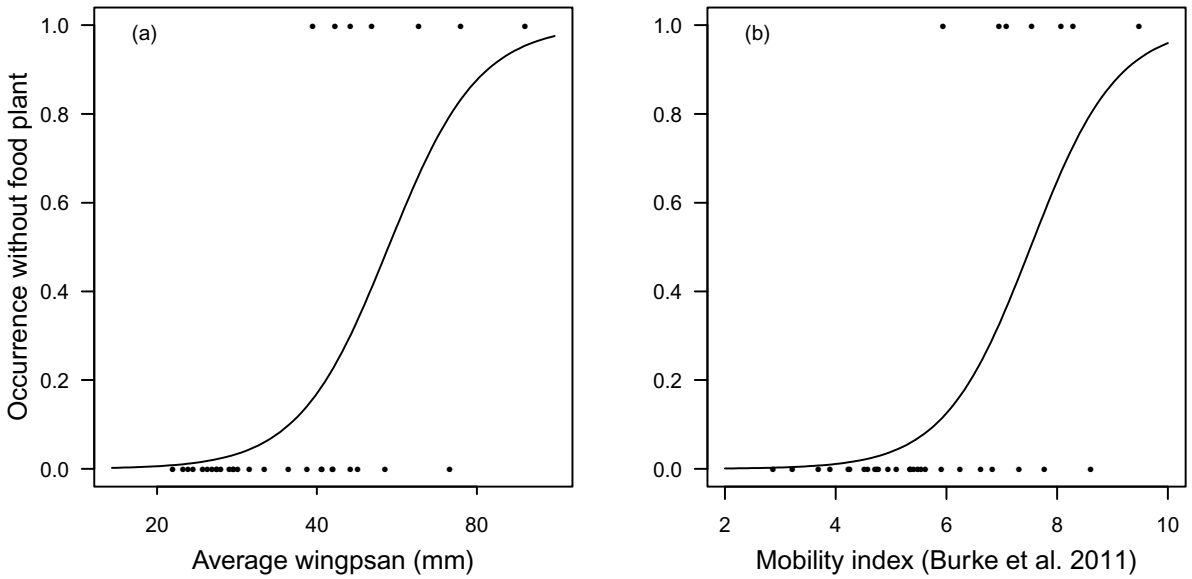

Figure 3-4. The probability of observing butterfly species on at least one island where their larval food plant was not detected relative to (a) average wingspan (mm) and (b) a species mobility index (Burke et al. 2011). Species' wingspans were log-transformed to improve model fit. Solid lines represent GLMs (logit link) used to assess relationships between variables. Relationships were significant for both average wingspan ($P = 0.011$) and species mobility ($P = 0.0075$).

3.4.6 Spatial patterns in species turnover

In accordance with GLMs addressing the complete species assemblage, Mantel tests indicated that butterfly species turnover was weakly positively related to plant species turnover in the 0.1- and 1.0-ha size classes ($r = 0.26$, $P = 0.18$ and $r = 0.32$, $P = 0.13$, respectively); although, these relationships were not significant. Contrasting with GLMs, butterfly species turnover and habitat turnover were significantly positively related in the 1.0-ha size class ($r = 0.41$, $P = 0.045$), but showed no strong relationship in the 0.1-ha size class ($r = 0.10$, $P = 0.33$).

Butterfly species turnover was unrelated to inter-island distance in both the 0.1- and 1.0-ha size classes ($r = 0.10$, $P = 0.35$ and $r = -0.019$, $P = 0.55$, respectively), indicating that the spatial distribution of several small islands did not contribute to their aggregate species richness.

3.5 Discussion

3.5.1 Fragmentation and species diversity

As predicted by the positive SAR, smaller islands were depauperate relative to larger islands at all scales addressed in this study. However, controlling for total habitat area demonstrated that this positive SAR is largely an artifact of the sample area effect: smaller sample areas contain fewer individuals, which for a given abundance distribution, belong to fewer species (Connor and McCoy 1979; Fahrig 2013). Direct comparisons of species richness across island size classes and SAC analyses of the complete species assemblage within the small island set suggest that several smaller islands actually contained more butterfly species than fewer larger islands, or a single large island of equivalent total area. Although this positive fragmentation effect was apparent in SAR extrapolation ($S_{sl} < S_{ss}$), the effect was not statistically significant, questioning the statistical power of the analysis. Most interestingly, this positive fragmentation effect was neutralized when transient butterflies were removed from analyses. Direct comparisons of species richness, SAC analysis, and SAR extrapolation all support the directionality of this pattern. This suggests that inter-island movements of highly mobile butterfly species, from larger island or mainland habitats with larval food plants to small islands without larval food plants, inflated the number of species small islands were observed to support. Positive relationships between both wingspan and mobility and the probability of observing butterfly species on islands without their larval food plants support this hypothesis.

To explain positive fragmentation effects in the context of the habitat amount hypothesis, Fahrig (2013) points out that several small fragments spread over a larger area are more likely to intersect the distributions of more species than single large fragments (e.g., Tjørve 2010). Given this relationship, the habitat amount hypothesis and underlying sample-area effect would predict positive relationships between degree of fragmentation and species diversity when SLOSS-based analyses are used. However, pairwise comparisons of butterfly species turnover across the smallest islands in both island sets (0.1- and 1.0-ha islands) showed no relationship to pairwise comparisons of distance. It is therefore reasonable to conclude that islands within our study area "sampled" a spatially consistent species pool and that the cumulative species richness of several small habitat fragments is generally unrelated to their spatial distribution at the scales addressed. A more likely explanation for positive fragmentation effects involves the inter-fragment movements of highly mobile species from larger habitat fragments, supporting reproducing populations, to smaller fragments, containing additional resources. Such movements are indeed predicted by ideal free distribution theory for nectar-feeding insects if smaller fragments contain higher densities of nectar resources (Dreisig 1995). Alternatively, several smaller islands may contain a higher diversity of habitat types compared with fewer or single larger islands, supporting a higher diversity of species (Williams 1964; Nilsson et al. 1988). However, GLMs accounting for both habitat diversity and habitat area do not support this latter hypothesis.

In contrast with the positive fragmentation effect observed in the small island set, the complete species assemblage shows no clear trend in species richness across island size classes in the large island set, supporting the habitat amount hypothesis. In partial conflict with this pattern, SAC analysis and SAR extrapolation both suggest a neutral to weakly negative fragmentation effect. Excluding transient species from the analyses revealed stronger negative

fragmentation effects on potentially reproducing butterfly populations for all analyses. The largest island in the large island set (8 ha of continuous habitat) contained a higher number of potential resident species than other size classes (8 ha of fragmented habitat). Similarly, SAC analyses show that fewer or single larger islands contained more potential resident species than several smaller islands summing to an equivalent area. Excluding transient species from SAR extrapolation in the large island set made negative fragmentation effects more apparent, although still not statistically significant.

Two principal conclusions may be drawn: i) negative fragmentation effects are more apparent in the large island set compared to the small, suggesting scale dependency; and ii) negative fragmentation effects are more apparent when excluding transient species that do not represent reproducing populations. This latter conclusion also holds true when both island sets were combined for SAC analysis and SAR extrapolation. We suspect that negative fragmentation effects observed across all 30 islands are likely a combined result of stochastic extinctions of populations isolated to smaller islands, as predicted by the island effect hypothesis, and decreased habitat suitability within these islands, such as the exclusion of potential larval food plants.

3.5.2 Implications of scale dependency

While SAC analysis may be the simplest method for archipelago- or landscape-wide comparisons of species diversity (Quinn & Harrison 1988), scale separation in our study shows that SACs have potential to obscure important aspects of fragmentation-species diversity relationships. We observed a considerable difference between small-to-large and large-to-small SACs for the complete species assemblage in the small island set, indicated by a saturation index estimate of 1.22. However, this pattern of species accumulation across small islands was not

preserved when both island sets were aggregated for analyses, where the saturation index was estimated at 0.98.

Although the small and large island sets contained equivalent numbers of islands, cumulative island area was tenfold greater in the large island set than in the small island set. Similarly, aggregate species richness was approximately twice as high in the large island set than in the small island set for the complete species assemblage. When all 30 islands were combined for SAC analysis, constrained integrals show that the small island set contributed only 3.09% to the area under the small-to-large SAC and 12.32% to the area under the large-to-small SAC. This demonstrates that patterns of species accumulation across larger fragments have potential to dominate those across smaller fragments, particularly when the range of fragment sizes is great and the abundance of small fragments is high—an arrangement common to many datasets. Within single archipelagos and landscapes, island/fragment areas may vary by several orders of magnitude, with smaller islands/fragments typically more abundant than larger ones (Quinn & Harrison 1988; Lomolino and Weiser 2001; Fahrig 2003; 2013). In the context of habitat fragmentation, it should not be assumed that ecological patterns and processes are consistent across these scales (Johnson 1980; Lomolino and Weiser 2001).

3.5.3 Scale separation and the SAR

SAR richness estimates for continuous habitat did not significantly differ from the aggregate observed richness of study islands for any of the six SAR extrapolations. Such a result may be attributed to either neutral fragmentation effects, or lack of statistical power to resolve relationships. While previous studies report *P*-values for SAR slope estimates (e.g., Yaacobi et al. 2007; Gavish et al. 2012), this method does not provide for the meaningful discrimination of insignificant results. Extrapolating SAR confidence intervals to infer the significance of

fragmentation effects is a novel approach that explicitly accounts for variation in statistical power between SAR extrapolations. As we observed in this study, confidence interval discrimination may increase the probability of a type II error (concluding that fragmentation was not related to diversity when in fact it was). However, this method reduces the possibility of type I error, because it accounts for uncertainty in SAR extrapolations stemming from small sample sizes or "noisy" species richness data. Reduced regression confidence and the subsequent broadening of confidence intervals likely explain why SAR extrapolation failed to detect both the positive fragmentation effect observed in the small island set for the complete species assemblage and the negative fragmentation effect observed in the large island set for the resident species subset. With other factors held constant, regression confidence increases with sample size—a relationship observed when all 30 islands were pooled for SAR extrapolation. However, when all 30 islands were included in the SAR, negative fragmentation effects were still not significant for the potential resident species subset. Contrasting with this result, SACs over all 30 islands demonstrate clearly that single or fewer larger islands contained more potential resident species than several small islands summing to an equivalent area. Accounting for regression confidence therefore brings into question the SAR extrapolation method's statistical power to resolve fragmentation effects.

Further questioning the viability of the SAR extrapolation method, the aggregation of fragment sizes required to raise SAR regression confidence to levels sufficient for resolving fragmentation effects has the inherent potential to obscure scale-dependent relationships. Scale separation should be considered necessary when the range of fragment sizes within datasets is great, such that fragmentation–species diversity relationships shift in relation to fragment size. However, many datasets do not contain a sufficient number of fragments to allow for scale

separation while maintaining an adequate sample size [e.g., Quinn & Harrison (1988) and Boecklen (1997), where datasets reviewed contained as few as six and five fragments, respectively]. Moreover, small sample sizes are often paired with broad ranges in fragment size [e.g., Rosin et al. (2011), where only 31 fragments ranged in size by over two orders of magnitude]. Such relationships also question whether *z* values should be used to infer fragmentation effects, as these analyses inherently assume uniformity in both SARs and fragmentation–species diversity relationships across broad ranges of fragment sizes.

Examples of shifts in SARs across island sizes are made clear by the small island effect (Lomolino & Weiser 2001), which states that insular species richness may not predictably increase with area below a threshold island or fragment size (Triantis et al. 2006). Below this threshold, species richness is largely determined by area-independent variables, such as intraspecific and interspecific interactions, stochastic events, island isolation, and habitat diversity (Nilsson et al. 1988; Lomolino 2000; Lomolino & Weiser 2001; Schoener et al. 2001; Triantis et al. 2006; Rosin et al. 2011). Area-independent variables influencing species richness add uncertainty to species–area regressions, further decreasing the probability of detecting fragmentation effects in SAR extrapolations or *z* value comparisons. Future studies should test for shifts in the SAR (e.g., Lomolino and Weiser 2001) before proceeding with such analyses. Theoretically, islands or fragments below the small island effect threshold should be excluded.

3.5.4 Area-independent relationships

The habitat amount hypothesis predicts that fragment isolation and species diversity will negatively correlate when mean isolation inversely relates to the amount of habitat on a landscape: if fragment edges do not delimit populations, negative relationships between fragment isolation and species richness may be an artifact of local species pools decreasing with habitat

amount (Fahrig 2013). However, the habitat amount hypothesis would not predict strong relationships between isolation and species richness if the spatial distribution of fragments is small, such that there is little spatial variation in the composition or abundances of species within a study area. This appears to be the case in our study system, as Mantel tests show no relationship between inter-island distances and differences in composition of the complete butterfly assemblage. Lack of strong relationships between island isolation and species richness in GLMs using the complete species assemblage is therefore best interpreted as support for the habitat amount hypothesis. Incongruently, when only considering the potential resident species subset, island isolation (250-m buffer) was found to be the most important factor structuring patterns of species richness in our best-supported model. This result suggests that when highly mobile, transient butterfly species occurring on islands without their food plants are excluded from fragmentation analyses, island effects on potentially reproducing populations become apparent. This result brings into question the neutral to positive fragmentation effects reported by multiple studies cited to support the habitat amount hypothesis (Fahrig 2013), as well as those concluded by Yaacobi et al. (2007).

Our best-supported GLM explaining variation in butterfly species richness within the complete species assemblage included both island area and vascular plant species richness. As previously documented (e.g., Erhardt 1985; Sparks & Parish 1995; Simonson et al. 2001; Croxton et al. 2005; Kitahara et al. 2008; but see Hawkins & Porter 2003), plant richness and butterfly richness were found to positively relate. Plant diversity may positively relate to butterfly diversity through the intermediate variables of food plant availability (Hawkins & Porter 2003), nectar resource availability (Kitahara et al. 2008), or habitat diversity (sensu Williams 1964; Nilsson et al. 1988). Interestingly, when excluding transient butterfly species

occurring on islands without their larval food plants, vascular plant richness was a poor predictor of butterfly richness. This difference suggests that the positive relationship between plant and butterfly richness within the complete species assemblage was driven by inter-island movements of highly mobile butterfly species to islands of particularly high plant diversity.

3.5.5 Conclusions

When considering the complete species assemblage, habitat fragmentation did not reduce butterfly species diversity in our study system. This result suggests that habitat configuration has little effect on the number of butterfly species persisting on fragmented landscapes, supporting the habitat amount hypothesis. However, butterfly species vary widely in mobility (Burke et al. 2011), and are therefore likely to vary widely in their responses to habitat fragmentation (Ewers & Didham 2006; Dover & Settele 2009). Our study shows that differentiating between potentially reproducing species and highly mobile, transient species observed within individual habitat fragments yields critical insight into the negative effects of habitat fragmentation on species diversity.

4 Chapter 4: The theory of island biogeography, the sample-area effect, and the habitat diversity hypothesis: complementarity in a naturally fragmented landscape of lake islands

4.1 Abstract

Aim: Investigate relationships between fragmentation and species diversity in the context of the theory of island biogeography, sample-area effect, and habitat diversity hypothesis.

Location: Lake of the Woods, Canada.

Taxon: Vascular plants

Methods: Vascular plant species diversity was inventoried on 30 islands, organized into two island sets. Each island set contained four size classes that varied in degree of fragmentation while controlling for the sample-area effect (small island set: 8 × 0.1-ha, 4 × 0.2-ha, 2 × 0.4-ha, and 1 × 0.8-ha islands; large island set: identical pattern utilizing 1.0-ha to 8.0-ha islands). Fragmentation effects were then examined using SLOSS-based analyses, addressing whether single large or several small islands contained more species/habitats: (a) direct comparisons of species and habitat richness across size classes; (b) extrapolations of species–area relationships; and (c) analyses of species and habitat accumulation curves. Multigroup path analysis was next used to quantify effects of habitat diversity, island area, and isolation on species richness for both island sets. Finally, pairwise and multiple-site dissimilarity was estimated for both species and habitats across 0.1-ha and 1.0-ha islands to investigate whether: (a) variation in species composition was related to habitat composition; and (b) species dissimilarity increased with inter-island distance.

Results: SLOSS-based analyses indicated that several small islands contained more species than single large islands in both island sets. This pattern was also observed for habitats, but only in

the small islands set. Path analysis suggested that island area had significant direct and indirect (mediated by habitat diversity) effects on species richness. Habitat diversity and island isolation had significant positive and negative effects on species richness, respectively, independent of island area. Species and habitat dissimilarities were significantly related across 0.1-ha but not 1.0-ha islands, and showed no relationship to inter-island distance.

Main conclusions: The overall positive relationship between fragmentation and species richness may be attributed to greater habitat diversity and increased species dissimilarity across smaller islands relative to larger islands. However, negative isolation effects indicate that landscape configuration is still an important conservation consideration. These results each align with different predictions of the theory of island biogeography, sample-area effect, and habitat diversity hypothesis, questioning the exclusivity of these theoretical frameworks.

4.2 Introduction

The diversity of species on islands has been a topic of considerable research in ecology for well over a century (e.g., Darwin 1859; Wilson & MacArthur 1967), resulting in a variety of theoretical explanations for variation in insular communities (Gotelli & Graves 1996; Rosenzweig 1995; Watson 2002). Ecologists have applied these explanations to diversity patterns on fragmented landscapes, interpreting isolated fragments as ecological islands situated in a sea of unsuitable habitat (Haila 2002). However, fragmentation effects on individual species and entire communities appear to be largely idiosyncratic, limiting the generality of fragmentation–species diversity relationships (Debinski & Holt 2000; MacDonald et al. 2018a). Still, a single recurrent pattern stands out; area is a good predictor of species richness at both the fragment and landscape level (Nilsson et al. 1988; Rosenzweig 1995; Fahrig 2013; but see Lomolino & Weiser 2001). Three principal frameworks addressing distinct ecological processes

have been proposed to account for this positive species–area relationship: i) the theory of island biogeography (MacArthur & Wilson 1963; Wilson & MacArthur 1967); ii) the sample-area effect (Connor & McCoy 1979; Fahrig 2013); and iii) the habitat diversity hypothesis (Williams 1964).

Developed in the context of oceanic islands, the theory of island biogeography interprets insular species richness as an equilibrium between extinction and immigration rates, arising from the effects of island area and isolation on demographic processes (MacArthur & Wilson 1963; Wilson & MacArthur 1967). Larger islands generally support larger populations relative to smaller islands, decreasing probabilities of inbreeding depression and stochastic extinction (Hanski 1999). Gilpin & Diamond (1976) add that larger islands also present larger dispersal targets, increasing probabilities of colonization (i.e., the target area effect). Of conservation interest, summing probabilities of colonization and persistence across species provides a mechanistic explanation for the species–area relationship across ecological islands, whether they are oceanic or terrestrial. Island configuration is also invoked as a predictor of species richness, as rates of species immigration (Simberloff & Wilson 1969) and rescue effects (Brown & Kodric-Brown 1977) generally decrease as islands become further isolated from sources of species immigration, such as the mainland or other islands.

While lack of empirical evidence and a plethora of competing models have led many ecologists to infer that the theory of island biogeography has been largely overturned (Gotelli & Graves 1996; Lomolino 2000), equilibristic interpretations of species richness on fragmented landscapes still appear to constitute a dominant scientific paradigm (*sensu* Kuhn 1967) in ecology (Haila 2002; Mendenhall et al. 2014). Demographic effects predicted by the theory of island biogeography suggest that decreasing fragment area and increasing fragment isolation

pose considerable threats to species diversity, thereby warranting continued investigation (Diamond 1975; May 1975; Wilson & Willis 1975; Rybicki & Hanski 2013; Haddad et al. 2017). If fragmentation indeed reduces species diversity via processes predicted by the theory of island biogeography, "island effects" should ultimately result in several smaller fragments containing fewer species than single larger fragments of equal area (Fahrig 2013). Independent of fragment area, species richness is also predicted to decrease as fragment isolation increases (Diamond 1975; Wilson & Willis 1975; Gotelli & Graves 1996). Such predictions are often framed in terms of the ongoing "SLOSS" debate, addressing whether conservation efforts should prioritize the protection of single large or several small conservation reserves (Diamond 1975; Simberloff & Abele 1982; Tjørve 2010). If fragmentation reduces species richness, finite conservation efforts may be best allocated to "single large" conservation strategies (Simberloff & Abele 1976; 1982) and maximizing connectivity within fragmented landscapes (Rybicki & Hanski 2013; Haddad et al. 2017).

Contrasting with the theory of island biogeography, the habitat amount hypothesis (Fahrig 2013) replaces fragment area and isolation with a single predictor of species richness, total habitat area. Not unlike the passive sampling hypothesis, developed in the context of oceanic islands (Connor & McCoy 1979), the habitat amount hypothesis uses the sample-area effect to explain positive species–area relationships across isolated fragments: larger sample areas generally contain more individuals, belonging to more species (Burns et al. 2010; Fahrig 2013). In SLOSS terms, the sample-area effect specifically predicts that single large and several small fragments will contain equivalent numbers of species when total area is held constant. The habitat amount hypothesis also interprets negative relationships between fragment isolation and species richness as sampling artefacts, based on two premises: i) total habitat area is the principal

determinant of local species pools because fragment edges do not typically delimit populations (i.e., extinction and colonization occur at the landscape level, and not within individual fragments); and ii) fragment isolation generally increases as total habitat area decreases. Species richness may therefore decrease with fragment isolation simply because of reductions in total habitat area at the landscape level, rather than increases in degree of fragmentation *per se* (Fahrig 2013).

Heuristics outlined by the habitat amount hypothesis offer a compelling gestalt switch (*sensu* Kuhn 1962) away from viewing fragments as natural spatial units for measuring and interpreting species richness. However, the hypothesis fails to account for variation in habitat composition both within and between fragments, as well as interspecific variation in habitat associations. Habitat associations differ considerably between species (Hortal et al. 2009), challenging applications of single habitat definitions to entire communities and begging the question as to whether relationships between habitat amount and species richness are even meaningful (Hanski 2015). Indeed, there is a strong theoretical and empirical basis to suggest that habitat diversity is a principal determinant of species richness within islands, fragments, and entire landscapes (Rosenzweig 1995; Kadmon & Allouche 2007; Hortal et al. 2009).

The habitat diversity hypothesis (Williams 1964) represents a third explanation of positive species–area relationships that is predicated on interspecific variation in habitat associations. Specifically, the habitat diversity hypothesis predicts that area *per se* has minor effects on demographic processes, and hence species richness, and instead serves as a surrogate variable for habitat diversity (Gotelli & Graves 1996). Larger sample areas generally contain more habitats, which support more species (Rosenzweig 1995; Williams 1964). A specific and testable prediction of the habitat diversity hypothesis is that species diversity and habitat

diversity will positively correlate independent of island or fragment area (e.g., MacArthur & MacArthur 1961). Nevertheless, it remains controversial whether habitat diversity or area *per se* is more important in structuring patterns of species richness on fragmented landscapes, and the habitat diversity hypothesis makes no specific predictions of fragmentation effects. While there is support for effects of habitat diversity on species richness independent of area (Kohn & Walsh 1994; Hortal et al. 2009; Burns et al. 2010), there is also support for direct effects of area *per se* on demographic processes affecting species richness (Buckley 1982; Nilsson et al. 1988). An important consideration in such investigations is scale, as the relative importance of habitat diversity and area *per se* has been shown to vary with island or fragment area (Rosenzweig 1995; Sfenthourakis & Triantis 2009). For instance, species–area relationships often become unpredictable below threshold island or fragment sizes (i.e., the small island effect; Lomolino & Weiser 2001; Triantis et al. 2006). Below these thresholds, habitat diversity and isolation frequently replace area as the strongest predictor of species richness (Sfenthourakis & Triantis 2009).

In this study, we estimated vascular plant species diversity and habitat diversity on 30 lake islands through repeated full-island surveys. We then used a series of SLOSS-based analyses, path analysis, and analyses of species and habitat dissimilarity (β-diversity) to investigate ecological processes underlying the species–area relationship and their implicated fragmentation effects. While species and habitat dissimilarity are widely understood as principal determinants of aggregate species richness (γ-diversity) in a variety of insular systems (Simberloff 1988; Gotelli & Graves 1996; Rosenzweig 1995), their importance is seldomly explicitly recognized in SLOSS-based investigations of fragmented landscapes (e.g., Yaacobi et al. 2007; Gavish et al. 2012; but see Wright & Reeves 1992; Tjørve 2010). Results of this study

indicate that patterns of species and habitat dissimilarity are important considerations that warrant continued investigation in the context of fragmentation–species diversity relationships.

4.3 Materials and Methods

4.3.1 Study area

Observations were made on islands within Sabaskong Bay at the southeastern corner of Lake of the Woods, Ontario, Canada (Figure 4-1). Sabaskong Bay is located in transitional zone between boreal forest to the north, Laurentian forest to the southeast, and, to a lesser extent, tallgrass prairie to the southwest. Local flora is therefore a mix of boreal tree species (e.g., *Pinus banksiana*, *Betula papyrifera*, and *Picea glauca*), Laurentian tree species (e.g., *Acer spicatum*, *Tilia americana*, and *Pinus strobus*), and few tree species from the Eastern prairies (e.g., *Quercus macrocarpa* and *Fraxinus pensylvanica*). All study islands are included within the Lake of the Woods Conservation Reserve, where residential and commercial developments are prohibited (Ontario Ministry of Natural Resources 2006). Island isolation is hypothesized to have occurred between 3000 and 4000 years ago, when differential rates of isostatic rebound and outlet restriction caused the progressive southward transgression of the remnants of Glacial Lake Agassiz, inundating Sabaskong Bay (Yang & Teller 2005). Islands within this system therefore represent “old high-contrast fragments,” appropriate for inferring long-term fragmentation effects on species richness (*sensu* Watson 2002).

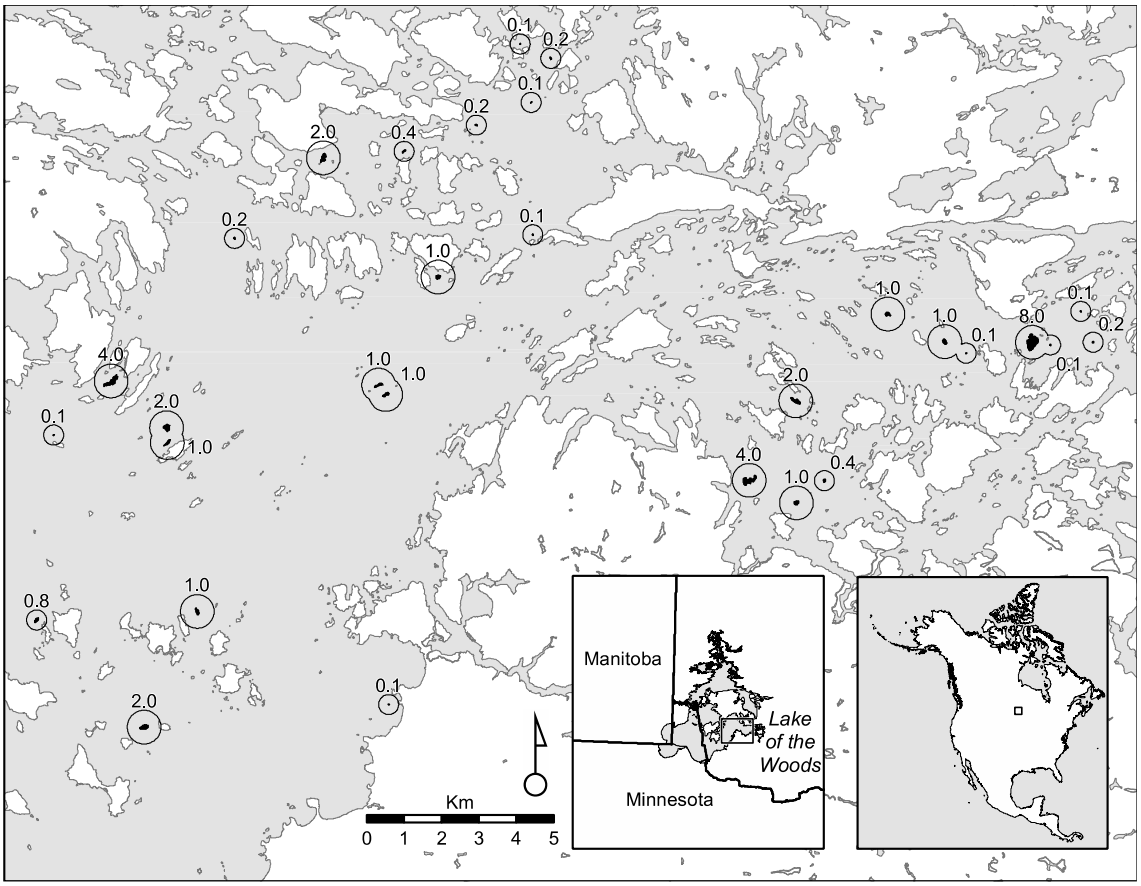

Figure 4-1. Map of the study area, located in Lake of the Woods, Ontario, Canada. Study islands (n = 30) in the small and large island sets are highlighted with small and large circles, respectively. Each island is labelled by size class (ha). Inset maps indicate the regional and continental location of the study area.

4.3.2 Sampling design

A nested set sampling design was used to decouple the effects of island configuration from those of island area (i.e., decouple the effects of fragmentation *per se* from those of habitat loss; sensu Fahrig 2003; 2013). Specifically, 30 study islands were randomly selected from a list of candidate islands and organized into two sets of non-overlapping size classes (Table 4-1; see

MacDonald et al. 2018b Supporting Information Appendix S1 for island selection criteria). Four size classes in the small island set consisted of eight 0.1-ha islands, four 0.2-ha islands, two 0.4-ha islands, and a single 0.8-ha island. The large island set followed an identical logarithmic pattern using islands ranging from 1.0 to 8.0 ha. Within each island set, degree of fragmentation decreased across increasing size classes, with the largest size class (a single island) representing the "single large" conservation strategy, and the smallest size class (a highly fragmented set of islands) representing the "several small" conservation strategy (Gavish et al. 2012).

Table 4-1. Summary of the nested set sampling design used to decouple degree of fragmentation from total island area across two distinct ranges of island sizes. Within island sets, total island area is maintained across sizes classes by halving the number of replicates per twofold increase in the individual areas of constituent islands. Degree of fragmentation decreased across increasing size classes. Aggregate vascular plant species richness and habitat richness are reported by size class.

Island Set	Size class (ha)	Number of islands	Total area (ha)	Species richness	Habitat richness
Small	0.1	8	0.8	114	9
Small	0.2	4	0.8	112	11
Small	0.4	2	0.8	95	8
Small	0.8	1	0.8	106	7
∑ Small		15	~3.2	179	12
Large	1.0	8	8.0	177	12
Large	2.0	4	8.0	195	13
Large	4.0	2	8.0	194	14
Large	8.0	1	8.0	167	12
∑ Large		15	~32.0	272	14
∑ Complete		30	~35.2	281	14

Vascular plant species diversity was estimated on each island through repeated full-island surveys, conducted between 1 June 2015 and 20 August 2015. Handheld GPS units were used to ensure adequate coverage of all areas and habitats on islands during each survey. To maintain consistency in survey effort across all islands, survey time was standardized to 40 min per ha per survey. Four repeated surveys were completed on each island, resulting in a seasonal total of two hours and 40 min per ha. This is consistent with recent sampling-effort recommendations for boreal plant communities (Zhang et al. 2014). Specimens that could not be identified in the field were collected and identified with a microscope and keys (e.g., Voss & Reznicek 2012; Chadde 2013) and voucher specimens were deposited in the University of Alberta Vascular Plant Herbarium. Twenty-three unidentified specimens were recorded as distinct morphospecies and included in richness totals.

Habitat diversity was estimated on each island using the number and relative area of 14 distinct habitat types, defined using structural properties of vegetation and geological features. (See MacDonald et al. 2018b Supporting Information Table S1.1 in Appendix S1 for habitat type descriptions.) While this habitat classification scheme is not entirely independent of plant diversity, no individual plant species were used in the delineation of habitat types. Only higher level taxonomic information was used (e.g., coniferous vs. deciduous forest), limiting the circularity of habitat diversity–species diversity relationships. The exclusion of all vegetation characteristics, in favour soil physical and chemical components, in a habitat classification scheme may be appropriate for testing environmental filtering and related hypotheses (e.g., Kraft et al. 2015), but not necessarily the habitat diversity hypothesis, which considers both biotic and abiotic factors affecting species richness (Williams 1964; Nilsson et al. 1988; Rosenzweig 1995).

4.3.3 Comparisons of species and habitat richness

To investigate the effects of fragmentation while controlling for the sample-area effect, aggregate species richness was compared across island size classes within the small and large island set. If fragmentation reduced species richness via island effects (*sensu* Diamond 1975; May 1975; Wilson & Willis 1975), aggregate species richness should be lowest within the smallest size classes in each island set (highest degree of fragmentation), and increase across larger size classes (lower degrees of fragmentation). Any other arrangement of species richness would suggest that fragmentation did not reduce species richness, but would not necessarily support the habitat amount hypothesis. The underlying sample-area effect specifically predicts species richness as unrelated to degree of fragmentation when total area is held constant, equating to an even distribution of species richness across size classes. A third possible result is species richness increasing with degree of fragmentation. This positive fragmentation effect would align with the habitat diversity hypothesis if several smaller islands contained a greater number of habitats than fewer or single larger islands. To test for this possibility, aggregate habitat richness was compared across island size classes within each island set. Pearson product-moment correlations were then used to assess relationships between species richness and habitat richness across islands in the 0.1-ha and 1.0-ha size classes, effectively controlling for island area.

4.3.4 Species–area relationship extrapolation

To further investigate the effects of fragmentation on species diversity, island species–area relationships (ISAR) were estimated using linear models for the small, large, and complete island sets (all 30 study islands together). As suggested by Rosenzweig (1995) for insular plant communities, semi-log ISARs were used. Each ISAR was extrapolated to generate a species

richness estimate for a single theoretical island, equivalent in area to all islands used to generate the ISAR (3.21, 32.13, and 35.34 ha for the small, large, and complete island set, respectively). This ISAR species richness estimate was then compared to the aggregate species richness of study islands used to generate the ISAR (e.g., Yaacobi et al. 2007; Gavish et al. 2012; Matthews et al. 2016). In SLOSS terms, the aggregate species richness of study islands is analogous to the "several small" conservation strategy (S_{ss}), while the ISAR species richness estimate for a single theoretical island is analogous to the "single large" conservation strategy (S_{sl}; Gavish et al. 2012; MacDonald et al. 2018). A SLOSS index, estimated as $100\% \times (S_{ss} - S_{sl})/ S_{ss}$, was used to compare the aggregate species richness of study islands to the ISAR species richness estimate in each island set (Boecklen 1997). Similar extrapolations were not used for island habitat–area relationships, as resulting habitat richness estimates for single theoretical islands exceeded the total number of defined habitat types in our a priori habitat classification scheme, indicating they were not meaningful.

If fragmentation reduced species richness at any given size scale (small, large, or complete island set), the aggregate species richness of study islands will be lower than the ISAR species richness estimate for the corresponding theoretical island ($S_{ss} < S_{sl}$). If islands passively sampled species, as predicted by the habitat amount hypothesis, the aggregate species richness of study islands will be approximately equivalent to the ISAR species richness estimate ($S_{ss} \approx S_{sl}$). If habitat fragmentation positively affected species richness, the aggregate species richness of study islands will be greater than the ISAR species richness estimate ($S_{ss} > S_{sl}$). Extrapolated ISAR 95% confidence intervals were used to determine the significance of fragmentation effects (MacDonald et al. 2018a).

4.3.5 Accumulation of species and habitats

To assess patterns of species and habitat accumulation across islands in the small, large, and complete island set, cumulative species richness and habitat richness were plotted against cumulative island area in two ways: i) increasing order of island area (small to large); and ii) decreasing order of island area (large to small; Quinn & Harrison 1988). The resulting accumulation curves were made to pass through the origin, permitting direct area-under-the-curve comparisons (Quinn & Harrison 1988; Gavish et al. 2012). A saturation index, estimated as the area under the small-to-large accumulation curve relative to that of the large-to-small accumulation curve, was used to quantitatively compare accumulation patterns. Integrals were calculated using the trapezoidal rule.

Steeper slopes of large-to-small accumulation curves relative to small-to-large accumulation curves (saturation index < 1) may be driven by two diversity patterns: i) a nested pattern of species or habitat richness with respect to island area (Matthews et al. 2016); or ii) fewer or single larger islands containing more species or habitats than several smaller islands (Quinn & Harrison 1988; Gavish et al. 2012). In either case, steeper large-to-small accumulation curves would suggest that fragmentation negatively affected the richness of species or habitats. Similarity between the slopes of small-to-large and large-to-small accumulation curves (saturation index ≈ 1) would indicate that numbers of species or habitats increased with cumulative area, irrespective of degree of fragmentation. This result would suggest that islands passively sampled species or habitats (Fahrig 2013). A third possibility, steeper small-to-large accumulation curves (saturation index > 1), would indicate that several smaller islands contained more species or habitats than fewer or single larger islands, suggesting a positive fragmentation effect on species or habitat richness.

4.3.6 Path analysis

If habitat diversity and island isolation contribute to patterns of insular species richness, they should make a statistical contribution to variation in species richness beyond that explained by area *per se* (Gotelli & Graves 1996). However, strong collinearity between habitat diversity and area questions the efficacy of multiple and residual regression techniques (Connor & Simberloff 1978; Freckleton 2002). Furthermore, area is hypothesized to be a principal determinant of habitat diversity, thereby having both direct and indirect effects on species richness (Williams 1964; Rosenzweig 1995; Triantis et al. 2006; Hortal et al. 2009). We therefore used path analysis with correlated causes (a structural equation model) to assess both direct and indirect effects of predictor variables according to an a priori model structure (Li 1975; Grace & Pugesek 1997; 1998). Path analysis is particularly useful for distinguishing the effects of multiple collinear variables (e.g., habitat diversity and area *per se*) on multiple response variables (for similar applications, see Kohn & Walsh 1994; Triantis et al. 2005; Triantis et al. 2006; Sfenthourakis & Triantis 2009; Burns et al. 2010). Multigroup path analysis, grouped by small and large island set, permitted comparisons of relationships between the two ranges of island sizes used in SLOSS-based analyses. This effectively determines the extent to which ecological processes within the small and large island set may be approximated by a single model.

Multigroup path analysis was completed by constructing an initial multigroup path model, wherein all path coefficients were estimated using maximum likelihood and permitted to vary between the small and large island set. Path coefficients were then iteratively constrained to a single estimate for the small and large island set together (complete island set), and chi-squared difference tests were used to assess whether model fit was significantly reduced ($\alpha = 0.05$)

relative to the unconstrained multigroup path model (Grace 2003). If model fit was not significantly reduced by a given constraint, we retained the single estimate, as this represents a more parsimonious model. This result would indicate that ecological processes moderating the relationship in question were consistent across the small and large island set. Alternatively, if model fit was significantly reduced by constraining coefficients for a given path to a single estimate, the respective estimates were permitted to vary between the small and large island set. This result would suggest that underlying ecological processes significantly differed between the small and large island set, suggesting size scale-dependency of the relationship.

Our *a priori* multigroup path model structure consisted of four variables: vascular plant species richness, habitat diversity, island area, and island isolation. Within this model structure, habitat diversity, island area, and island isolation each directly affects species richness. Island area also directly affects habitat diversity, thereby having an additional indirect effect on species richness. This indirect effect was estimated as the product of: i) the direct effect of island area on habitat diversity; and ii) the direct effect of habitat diversity on species richness. The total effect of island area on species richness was then estimated by summing direct and indirect effects (e.g., Kohn & Walsh 1994). Habitat diversity was included in competing models as either habitat richness or the exponential of Shannon's entropy, estimated using the relative area of habitat types on each island (Jost 2006). The best-supported measure of habitat diversity was determined using both R2SP RICH and R2HAB DIV, estimated as 1 – the standardized variance unexplained by the path model ("residual variance") for species richness and habitat diversity, respectively. This method effectively minimizes the proportion of variance in endogenous variables left unexplained by the path model. As with ISARs, island area was log-transformed to account for non-linear relationships (Rosenzweig 1995). Island isolation was estimated at

multiple scales as the proportion of water (1 – proportion of landmass) within 250-, 500-, 1000-, 2500-, and 5000-m buffers. Buffers were generated from island edges, meaning isolation estimates are independent from island area. Proportion-based measures have been shown to be better predictors of immigration rates and related ecological processes than distance-based measures to nearest neighbour or landmass (Fahrig 2013). The best-supported isolation buffer size was determined using Akaike's information criterion (AIC), where smaller AIC values indicate higher relative model support (Burnham & Anderson 2004). AIC comparison is possible in this instance because island isolation is exogenous within the multigroup path model. Finally, a likelihood ratio test was used to assess the overall fit of the multigroup path model. Here, a non-significant result ($\alpha = 0.05$) indicates that the covariance structure of the multigroup path model did not significantly differ from the observed covariance structure, equating to good model fit (Grace 2008; Grace et al. 2010). Multigroup path analysis and related statistical tests were completed using the R package 'lavaan' (Rosseel 2012).

4.3.7 Dissimilarity of species and habitats

The theory of island biogeography predicts that species composition may vary substantially across islands of comparable area and isolation, with little variation in species richness (MacArthur & Wilson 1963; Wilson & MacArthur 1967; Simberloff & Wilson 1969). The same may be true for habitats. Cryptic turnover of species and habitats may therefore obscure relationships predicted by the habitat diversity hypothesis (e.g., positive correlations between species richness and habitat richness across islands of equal area). To account for the identities of individual species and habitats, pairwise dissimilarity was estimated for both species and habitats across islands in the 0.1-ha and 1.0-ha size classes using the Jaccard pairwise dissimilarity index: $d_{\text{J-PAIR}} = [b + c/(a + b + c)]$, where a is the number of species or habitats

shared between two islands (*i* and *j*), *b* is the number of species or habitats occurring on *i* but not *j*, and *c* is the number of species or habitats occurring on *j* but not *i*. This index is a monotonic transformation of beta diversity, accounting for both turnover and nestedness, and reflects the proportion of unshared species or habitats observed on two islands (Anderson et al. 2011; Baselga 2012). Positive relationships between species dissimilarity and habitat dissimilarity would indicate that species diversity and habitat diversity were positively related independent of area, supporting the habitat diversity hypothesis. To investigate whether overall rates of species and habitat dissimilarity changed with island area, multiple-site dissimilarity was estimated for the 0.1-ha and 1.0-ha size classes using the Jaccard multiple-site dissimilarity index, referred to here as $d_{\text{J-MULT}}$, derived by Baselga (2012). Averages of pairwise dissimilarities are shown to produce misleading results, justifying this approach (Baselga 2012). All dissimilarity indices were estimated using the R package 'betapart' (Baselga & Orme 2012).

Pairwise species dissimilarity was also compared with inter-island (Euclidean) distance across islands in the 0.1- and 1.0-ha size classes. Several smaller fragments may be more likely to intersect the distributions of more species than fewer or single larger fragments, effectively sampling a higher diversity of species (Tjørve 2010; Fahrig 2013) Given this possibility, the sample-area effect may theoretically result in the spurious observation of positive fragmentation effects when using SLOSS-based analyses. Positive relationships between pairwise species dissimilarity and inter-island distance would indicate that the expanded spatial distribution of several smaller islands, relative to fewer or single larger islands, contributed to their aggregate species richness. Lack of such relationships would suggest that islands did not significantly differ in their pools of potential immigrants, and that their diversities were not significantly spatially autocorrelated. Simple Mantel tests (999 permutations) were used to assess whether relationships

between pairwise species dissimilarity, pairwise habitat dissimilarity, and inter-island distance were significant (Anderson et al. 2011). All statistical analyses were performed using the statistical software R version 3.4.3 (R Core Team 2017).

4.4 Results

4.4.1 Comparisons of species and habitat richness

A total of 179 and 272 vascular plant species were observed within the small (0.1 to 0.8 ha) and large (1.0 to 8.0 ha) island sets, respectively (Table 4-1). Aggregate species richness across all 30 study islands was 281, indicating that vascular plant diversity of the small island set was largely nested within that of the large island set. Although aggregate species richness did not consistently increase or decrease with degree of fragmentation in the small island set, each of the two smallest size classes (0.1 and 0.2 ha) contained more species than the two largest size classes (0.4 and 0.8 ha). No clear trend in aggregate species richness was observed across size classes in the large island set.

Similar relationships were observed for comparisons of habitat richness across size classes. In the small island set, the two smallest size classes contained more habitats than the two largest size classes, with habitat richness ranging from 7 to 12. Habitat richness was less variable in the large island set, ranging from 12 to 14 (all habitat types present). This suggests that a total area of 8.0 ha accumulates most of the defined habitat types to near saturation. There was no clear trend in habitat richness across size classes at this size scale. Species richness and habitat richness were positively correlated across 0.1-ha islands ($r_{\text{Pearson}} = 0.893$, $P = 0.003$), but not 1.0-ha islands ($r_{\text{Pearson}} = 0.056$, $P = 0.896$).

4.4.2 Species-area relationship extrapolation

The aggregate species richness of study islands was greater than the ISAR species richness estimate for a single theoretical island ($S_{ss} > S_{sl}$) for the small, large, and complete island set (Figure 4-2 a, b, and c, respectively). However, not all differences were significant. Aggregate species richness across islands in the small island set was observed to be 179; significantly higher than the ISAR species richness estimate of 126.93 for a theoretical 3.21-ha island (95% CI = [75.92, 177.94]). Here, the SLOSS index estimate indicated that a fragmented set of islands of this configuration is expected, on average, to contain 29.1% more species than a single large island of equal area. In the large island set, aggregate species richness was observed at 272, which did not significantly differ from the ISAR species richness estimate of 243.72 for a theoretical 32.13-ha island (95% CI = [210.73, 276.71]). Notwithstanding, the SLOSS index estimate indicated that a fragmented set of islands is expected to contain 10.4% more species than a single large island. In the complete island set (all 30 study islands), aggregate species richness was observed at 281; significantly greater than the ISAR species richness estimate of 188.78 for a theoretical 35.34-ha island (95% CI = [164.55, 213.01]). The SLOSS index estimate for the complete island set indicated that a fragmented set of islands is expected to contain 32.8% more species than a single large island.

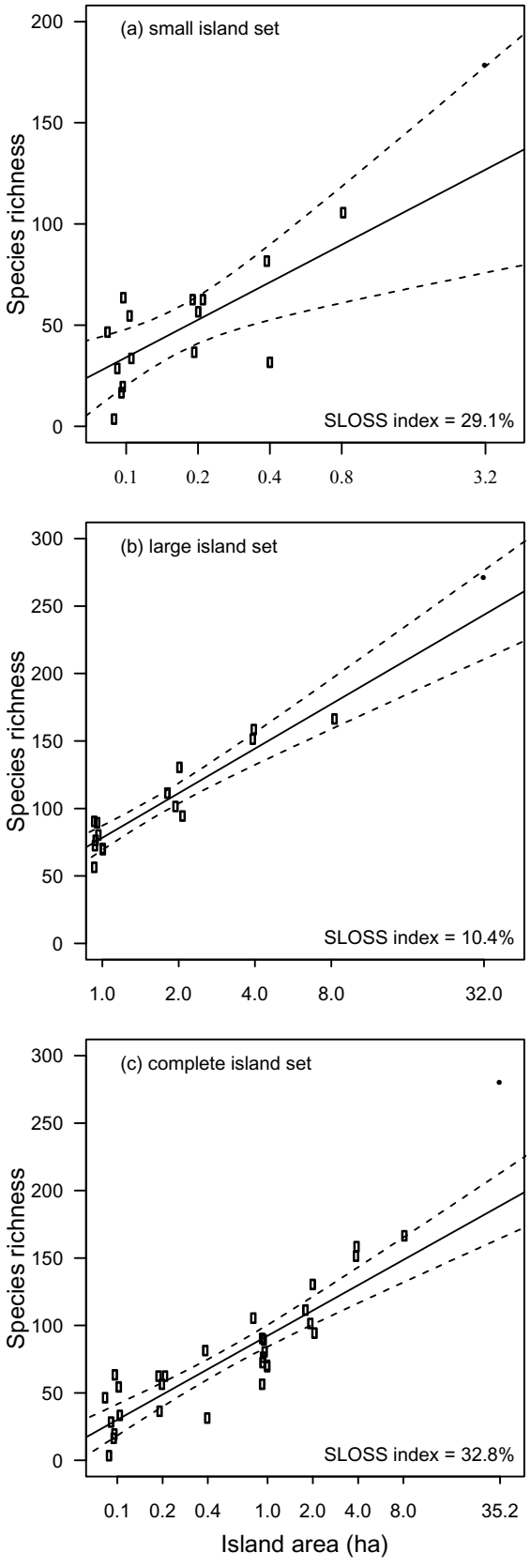
(a) small island set
Species richness
200
150
100
50
0
SLOSS index = 29.1%
0.1
0.2
0.4
0.8
3.2
(b) large island set
Species richness
300
250
200
150
100
50
0
SLOSS index = 10.4%
1.0
2.0
4.0
8.0
32.0
(c) complete island set
Species richness
300
250
200
150
100
50
0
SLOSS index = 32.8%
0.1
0.2
0.4
1.0
2.0
4.0
8.0
35.2
Island area (ha)

Figure 4-2. ISARs derived from the (a) small (n = 15), (b) large (n = 15), and (c) complete (n = 30) island sets. Open circles represent the observed vascular plant species richness for single islands, while filled circles represent the aggregate species richness of study islands used to generate the ISAR. Dashed lines represent 95% confidence intervals for ISAR regressions (estimated using least-squares). The SLOSS index was estimated as $100\% \times (S_{ss} - S_{sl}) / S_{ss}$, where S_{ss} is the aggregate species richness of study islands, and S_{sl} is the ISAR species richness estimate for a single theoretical island of equal area.

4.4.3 Accumulation of species and habitats

When cumulative species richness was plotted against cumulative island area, the small-to-large accumulation curve lay above the large-to-small accumulation curve in the small, large, and complete island sets (Figure 4-3 a, b, and c). This visual inspection of curves aligns with saturation index estimates of 1.071, 1.097, and 1.161, respectively. These results suggest two diversity patterns: i) species richness was not consistently nested in relation to island area; and ii) several smaller islands generally contained more species than fewer or single larger islands equivalent in areal extent. These observations equate to a positive effect of fragmentation on species richness (Gavish et al. 2012).

Visual inspection of habitat accumulation curves (Figure 4-3 d, e, and f) suggested that small-to-large and large-to-small curves only differed substantially in the small island set, where habitats accumulated with area more rapidly across small islands than large. In the large and complete island sets, habitats accumulated with area irrespective of island size, even though saturation index estimates were positive at all size scales. It is clear that saturation index estimates > 1 for the large and complete island sets stemmed from the constraint of passing

accumulation curves through the origin (Quinn & Harrison 1988, Gavish et al. 2012). Saturation index estimates under these circumstances (large variation in island area with few types of accumulating entities) are therefore unreliable. Considering only visual inspection of habitat accumulation curves, it is interesting that the positive fragmentation effect on habitat richness observed in the small island set was not persevered when all 30 islands were used to build habitat accumulation curves for the complete island set. To explain this result, MacDonald et al. (2018a) suggest that accumulation patterns across larger islands may dominate those across smaller islands; particularly, when the range of island sizes is great and the abundance of small islands is high.

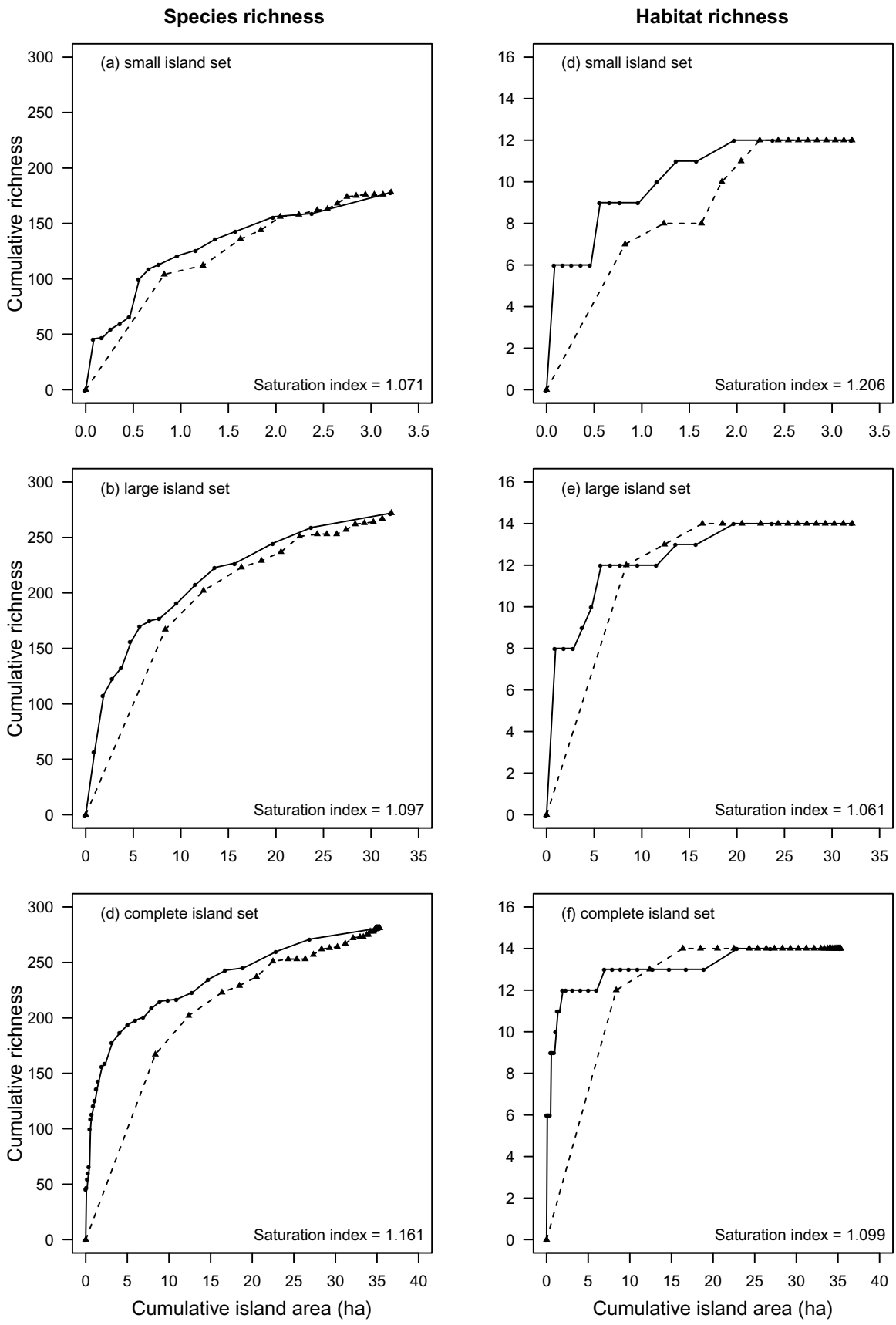
Species richness
Habitat richness
(a) small island set
Saturation index = 1.071
(d) small island set
Saturation index = 1.206
(b) large island set
Saturation index = 1.097
(e) large island set
Saturation index = 1.061
(d) complete island set
Saturation index = 1.161
(f) complete island set
Saturation index = 1.099
Cumulative richness
Cumulative island area (ha)

Figure 4-3. Cumulative number of vascular plant species (a, b, and c) and habitats (d, e, and f) relative to cumulative island area. Accumulation of species and habitats occurred from the smallest island to largest island (small-to-large curve, represented by closed circles connected by solid lines) and from the largest island to smallest island (large-to-small curve, represented by closed triangles connected by dashed lines). The saturation index was estimated as the area under the small-to-large curve relative to that of the large-to-small curve.

4.4.4 Path analysis

Habitat richness and the proportion of water within a 500-m buffer were the best-supported measures of habitat diversity and island isolation, respectively. The final multigroup path model accounting for these variables yielded a non-significant likelihood ratio test, indicating that the model's covariance structure provided an adequate description of the total observed covariance matrix. All coefficient estimates, including those measuring the indirect effect of island area on species richness, were significant at $\alpha = 0.05$ (Table 4-2, Figure 4-4). Unstandardized coefficients relating island area to habitat diversity, habitat diversity to species richness, and island isolation to species richness were constrained to single estimates for the small and large island set without significantly decreasing model fit (respective standardized coefficient estimates reported in Table 4-2 and Figure 4-4 vary between the small and large island set due to disparity in the variance of individual variables between the small and large island set). This result suggests that ecological processes underlying these relationships are approximately equivalent between the two ranges of island sizes. In contrast, constraining coefficients measuring the direct effect of island area on species richness to a single estimate for the small and large island set significantly reduced model fit. The direct effect of island area on

species richness was therefore estimated for the small and large island set independently. This effect of area *per se* was greater across islands in the large island set, suggesting size scale-dependency. Overall, the final multigroup path model explained 81.2% and 91.6% of the variation in species richness and 34.4% and 55.3% of variation in habitat diversity (richness) in the small and large island set, respectively.

Table 4-2. Standardized multigroup path coefficient estimates for the effects of habitat diversity (richness), island area (log-transformed), and island isolation (proportion of water within 500-m buffer) on vascular plant species richness. All unstandardized path coefficients were constrained to single estimates for the small and large island set, without significantly reducing model fit, except for those measuring the direct effect of island area on species richness, which were estimated for the small and large island set independently. The indirect effect of island area on species richness (mediated by habitat diversity) was estimated as the product of the direct effect of island area on habitat diversity and the direct effect habitat diversity on species richness. The total effect of island area on species richness was then estimated as the sum of its direct and indirect effects. Multigroup path analysis model structure is given in Figure 4-4.

Island set	Variable	Direct effect	Indirect effect	Total effect
Small	Habitat diversity	0.491***		
	Area	0.502***	0.288*	0.790***
	Isolation	-0.146*		
Large	Habitat diversity	0.323***		
	Area	0.675***	0.240*	0.915***
	Isolation	-0.151*		

Significance is denoted by * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

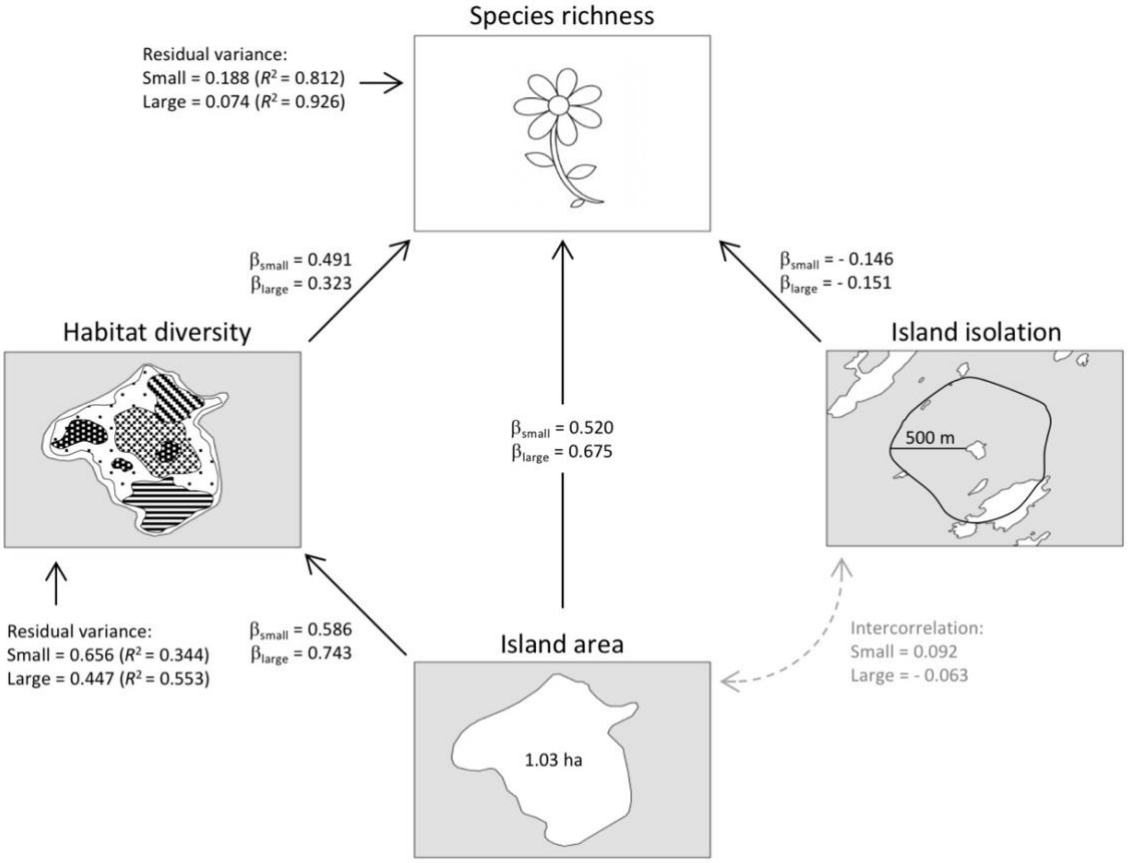

Figure 4-4. Multigroup path model structure accounting for species richness, habitat diversity (richness), island area (log-transformed), and island isolation (proportion of water within 500-m buffer). Habitat diversity, island area, and island isolation each directly affects species richness. Island area also directly affects habitat diversity, thereby having an additional indirect effect on species richness. All unstandardized path coefficients were constrained to single estimates for the small and large island set, without significantly reducing model fit, except for those measuring the direct effect of island area on species richness, which were estimated for the small and large island set independently. Residual variances ($1 - R^2$) for species richness and habitat diversity in the small and large island set are reported adjacent to arrows unconnected to other variables. Coefficients associated with the dashed double-headed arrow connecting island area and island

isolation represent intercorrelation, which is not treated as a causal path. The direct, indirect, and total effects of habitat diversity, island area, and island isolation on species richness are reported in Table 4-2.

4.4.5 Dissimilarity of species and habitats

Pairwise dissimilarities of species composition and habitat composition were significantly related across islands in the 0.1-ha size class (r_{Mantel} = 0.525, P = 0.021), but not in the 1.0-ha size class (r_{Mantel} = - 0.186, P = 0.756). This size scale-dependency may be driven by partial habitat saturation in areas approaching 1.0 ha, resulting in reduced habitat dissimilarity across larger islands. Greater multiple-site habitat dissimilarity across islands in the 0.1-ha size class than in the 1.0-ha size class ($d_{J\text{-}MULT}$ = 0.784 and 0.583, respectively) corroborate this hypothesis. Multiple-site species dissimilarity was also greater across islands in the 0.1-ha size class than in the 1.0-ha size class, although the difference was less pronounced ($d_{J\text{-}MULT}$ = 0.887 and 0.792, respectively). No significant relationship was observed between pairwise species dissimilarity and inter-island distance across islands in either the 0.1-ha or 1.0-ha size class (r_{Mantel} = - 0.221, P = 0.839 and r_{Mantel} = - 0.093, P = 0.659, respectively).

4.5 Discussion

Results of this study accord well with those of others, suggesting that fragmentation may not reduce species diversity (e.g., Yaacobi et al. 2007; Gavish et al. 2012; reviews in Fahrig; 2003; 2013; 2017). This expanding literature has led some ecologists to infer that the sample-area effect may adequately account for positive species-area relationships in the majority of fragmented landscapes (e.g., Fahrig 2013; 2017). Unexplained by the sample-area effect,

however, we observed a general trend of species richness actually increasing with degree of fragmentation after controlling for total area sampled with SLOSS-based analyses. This positive fragmentation effect was most pronounced in the small island set, where all SLOSS-based analyses indicated that species richness increased with increasing degrees of fragmentation. Fragmentation-species richness relationships were more ambiguous in the large island set, where no clear pattern in species richness was observed across size classes, and the aggregate species richness of study islands did not significantly differ from the ISAR species richness estimate. However, the SLOSS index estimate indicated that a fragmented set of islands, equivalent in configuration to the large island set, is still expected to contain 10.4% more species than a theoretical single large island. Others have interpreted similar differences between aggregate species richness and ISAR species richness estimates as meaningful, so long as the ISAR was significant (e.g., Yaacobi et al. 2007; Matthews et al. 2016). While we have cautioned interpretation of SLOSS index estimates under these circumstances (e.g., MacDonald et al. 2018a), species accumulation curves also indicated a weak positive fragmentation effect at this size scale, suggesting that the directionality of the SLOSS index estimate for the large island set was accurate. Considering all 30 islands together (complete island set), both ISAR extrapolation and species accumulation curves indicated that fragmented sets of smaller islands contained more species than fewer or single larger islands, suggesting that fragmentation at this scale may increase species richness.

To account for similar observations in the context of habitat amount hypothesis (i.e., sample-area effect), Fahrig (2013) points out that several smaller fragments may intersect the distributions of more species because of their expanded spatial distribution relative to fewer or single larger fragments. Several small fragments may thereby passively sample a higher diversity

of species, resulting in the spurious observation of positive fragmentation effects. Indeed, theoretical species-diversity models suggest that several small conservation reserves capture more species when species turnover increases with distance (Tjørve 2010). However, such processes are unlikely to operate within single fragmented landscapes; the scale at which fragmentation effects are most often inferred (review in Debinski & Holt 2000). In this study, pairwise species dissimilarity was not related to inter-island distance at either of the two island sizes addressed. We therefore find it reasonable to conclude that: i) the expanded spatial distribution of several small islands did not confound SLOSS-based analyses; and ii) the observation that several smaller islands contained a greater number of species than fewer or single larger islands is best interpreted as a positive effect of fragmentation on species richness. While the sample-area effect undoubtedly contributes to positive species-area relationships in a variety of systems, including this one, a more likely explanation of the positive fragmentation effect observed here involves a combination of habitat diversity and the small island effect.

In the small island set, SLOSS-based analyses of habitat richness suggested that several smaller islands contained more habitat types than fewer or single larger islands. This pattern of habitat richness aligned well with that of species richness, as would be predicted by the habitat diversity hypothesis. Further support for the habitat diversity hypothesis in the small island set is conferred by three additional relationships. First, species richness and habitat richness were positively correlated across 0.1-ha islands, demonstrating that species richness and habitat richness were positively related independent of island area at this size scale (MacArthur & MacArthur 1961; Williams 1964; Rosenzweig 1995). Second, pairwise species dissimilarity was significantly related to pairwise habitat dissimilarity across this same grouping of islands, indicating that the former correlation was not spurious. Third, path analysis demonstrated that

habitat richness had a significant positive effect on species richness, independent of area *per se*. Together, these relationships suggest that the positive fragmentation effect on species richness may be at least partially attributed to a positive fragmentation effect on habitat richness. Several smaller islands contained more habitats than fewer or single larger islands, and more habitats supported more species. These relationships were more prominent across islands within the small island set than those within the large, supporting that both fragmentation effects and ecological processes underlying species-area relationships are size scale-dependent (*sensu* Rosenzweig 1995; Lomolino & Weiser 2001, Triantis et al. 2006).

An additional, and perhaps complimentary, explanation of positive fragmentation effects is contributed by the small island effect. The small island effect specifically predicts the existence of threshold island sizes, below which, species richness does not consistently increase with area (e.g., Lomolino & Weiser 2001, Triantis et al. 2006; Sfenthourakis & Triantis 2009). A more generalized prediction here may be that the effect of area *per se* on species richness will be less prominent for smaller islands than larger islands. Multigroup path analysis supported this generalized prediction, indicating that direct and total effects of island area on species richness were smaller in the small island set than in the large island set. Working backwards through these area effects, it is clear that losses of species associated with reductions in area should be less prominent across smaller islands than across larger islands. Fragmented sets of smaller islands may therefore be expected to contain more species than fewer or single larger islands; particularly, if the areas of individual small islands are below a small island effect threshold.

Interestingly, this prediction that fragmented sets of smaller islands may contain more species than fewer or single larger islands may not be in opposition with the theory of island biogeography. In accordance with the small island effect, the theory of island biogeography

predicts that extinctions will become increasingly frequent and stochastic as island size decreases, to the eventual extent that extinction rates are decoupled from island area (Wilson & MacArthur 1967). These high, area-independent extinction rates may equilibrate with immigration rates at nonzero richness values, with particularly high rates of temporal species turnover. Assuming that demographic processes and assemblage dynamics are independent across islands (Wilson & MacArthur 1967; Hanski 1999; Leibold et al. 2004), high rates of temporal species turnover within small islands should translate to increased species dissimilarity across small islands, effectively increasing their aggregate species richness. Evidence for these relationships was observed in our study system, wherein multiple-site species dissimilarity was greater across smaller islands (0.1 ha) than larger islands (1.0 ha). Together with the positive fragmentation effect on habitat diversity, the small island effect and increased species dissimilarity across small islands may explain why the positive fragmentation effect on species richness was most pronounced in the small and complete island set, which both contained our smallest study islands.

If there is validity to these relationships, species richness within and among small fragments may not consistently decrease with reductions in fragment area. However, this conclusion should not be interpreted as conclusive evidence that fragmentation does not threaten species diversity. Species richness within small fragments may be largely comprised of early seral species of low conservation concern (Debinski & Holt 2000), with threatened species with higher extinction thresholds restricted to larger fragments or extirpated from fragmented landscapes entirely (Fukamachi et al. 1996; Rybicki & Hanski 2013). Furthermore, the prediction that high extinction rates may equilibrate with immigration rates at nonzero richness values for small islands is predicated on the existence of a mainland within the dispersal ranges

of species, serving as a continuous source of immigration (Wilson & MacArthur 1967). In the absence of mainland equivalents on many fragmented landscapes, similar source-sink dynamics are also predicted among fragments by metapopulation and metacommunity theory; particularly, when ranges of fragment areas within single landscapes are great (e.g., "mass effects"; Hanski 1999; Leibold et al. 2004). Ecologists should therefore proceed carefully when inferring the conservation significance of SLOSS, as large fragments may be necessary for the maintenance of species richness at the landscape level.

An additional shortfall of SLOSS-based analyses is that they do not permit direct investigation of the effects of isolation on species richness, which, in many respects, operate independent of the effects of area *per se* (Wilson & MacArthur 1967; Hanski 1999). Negative relationships between fragment isolation and species richness have been observed in other studies where SLOSS-based analyses suggested neutral to positive fragmentation effects overall (reviews in Fahrig 2013; 2017). To account for these relationships in the context of the habitat amount hypothesis, Fahrig (2013) suggests that species richness may decrease with fragment isolation simply because of reductions in: i) the total amount of habitat surrounding fragments; and ii) corresponding pools of potential immigrants (i.e., "source pools", *sensu* Gotelli & Graves 1996). However, we question whether it is reasonable to suppose that source pools vary within single landscapes, to the extent that individual fragments differ significantly in the number and composition of species they sample. If islands indeed sample species passively, dissimilarity in species composition between islands of similar areas should serve as an adequate proxy for dissimilarity in their respective source pools. In this study, pairwise species dissimilarity (accounting for both turnover and nestedness) was not related to inter-island distance. This suggests, but does not prove, a general homogeneity of source pools within our study area. The

negative isolation effect observed may therefore be best-interpreted as a negative fragmentation effect, as predicted by the theory of island biogeography, rather than a sampling artefact, as suggested by the habitat amount hypothesis.

In conclusion, while SLOSS-based observations indicated that fragmentation increased species richness overall, it should not be assumed that all aspects of landscape configuration associated with the fragmentation construct align in the directionality of their effects. Increased habitat diversity, the small island effect, and increased species dissimilarity may all positively affect aggregate species richness among several small fragments, while increasing isolation negatively affects species richness within individual fragments. The history of ecology is marked by an ongoing debate on the ecological processes moderating these relationships and how our understandings of such processes are best applied to conservation. Within this debate, the theory of island biogeography, the sample-area effect, and the habitat diversity hypothesis are most often framed as addressing mutually exclusive processes representing opposing schools of thought. Indeed, the epistemological comfort afforded by exclusive subscription to a scientific paradigm is attractive due to the relative ease of operating under a single framework. However, such perspectives are generally incompatible with the complexity of ecological systems; particularly, when considering emergent ecological properties such as species diversity. Results of this study, among others (e.g., Buckley 1982; Kadmon & Allouche 2007; Burns et al. 2010), suggest that multiple processes operate simultaneously to structure species diversity in insular systems, and ought to be viewed as mutually complementary, rather than exclusive.

5 Chapter 5: Distinguishing effects of area *per se* and isolation from the sample-area effect for true islands and habitat fragments

5.1 Abstract

The island species area relationship (ISAR) is an important tool for measuring variation in species diversity in variety of insular systems, from true-island archipelagoes to fragmented terrestrial landscapes. However, it suffers from several limitations. For example, due to the sample-area effect, positive relationships between species and area cannot be directly interpreted as evidence for deterministic effects of area *per se*. Additionally, richness-based analyses may obscure species-level responses to area and isolation that may better inform conservation practice. Here, we use random placement models to control for variation in abundance, occupancy, and richness associated with the sample-area effect, allowing deterministic effects of area and isolation, and how they vary with species' functional traits, to be resolved using linear mixed effects models. We demonstrate the utility of this approach using a butterfly assemblage persisting on a naturally fragmented landscape of lake islands. The ISAR did not significantly deviate from random placement in relation to island area, isolation, or habitat diversity, supporting stochastic assembly consistent with the sample-area effect. Such inferences support the habitat amount hypothesis, which prioritizes preserving the maximum amount of habitat irrespective of its fragmentation. However, species-level analyses demonstrated that species' abundances were significantly lower for both smaller and more isolated islands than what is predicted by the sample-area effect. Moreover, effects of area *per se* were significantly greater for smaller, less mobile, and rare species. Species' occurrences also significantly deviated from predictions of the sample-area effect in relation to island isolation. Thus, our approach illustrates that richness-based analyses not only result in incorrect inferences on mechanisms underlying

ISARs, but also obscure important effects of area *per se* and isolation on individual species that vary with functional traits. We therefore suggest that these effects should not be solely inferred from richness-based analyses, but rather evaluated on a species-by-species basis.

5.2 Introduction

From true islands to habitat fragments on terrestrial landscapes, positive relationships between species richness and the area of islands or fragments are among the oldest and most widely documented patterns in ecology (Arrhenius 1921; MacArthur & Wilson 1963; Rosenzweig 1995; Gotelli & Graves 1996; He & Legendre 2002; Hanski et al. 2013). These island species-area relationships (ISARs, *sensu* Triantis et al. 2003, or "Type IV" curves, *sensu* Scheiner 2003) have received considerable attention, in part due to their importance to conservation frameworks (e.g., see reviews in Shafer 1990 Lomolino 2000; Whittaker & Fernández-Palacios 2007). Although there are several documented divergences between the biogeographies of true islands and terrestrial habitat fragments (e.g., Laurance 2008; Mendenhall et al. 2014; Itescu 2019; Farneda et al. 2020), studies addressing true-island systems can still help resolve what mechanisms underlie ISARs and how fragmentation effects are best measured (Diamond 1975; Simberloff & Abele 1976; 1982; Haila 2002; Haddad et al. 2015; MacDonald et al. 2018a; 2018b). There are, however, at least two enduring problems with the use of ISARs in conservation that are generalizable to both true islands and habitat fragments: i) ISARs may emerge from a combination of different mechanisms, each of which potentially informs a different conservation directive (Connor & McCoy 1979; Kadmon & Allouche 2007; MacDonald et al. 2018b); and ii) ISARs are emergent patterns of diversity that can mask differential responses to habitat area among species that may require independent consideration for successful conservation planning (Ewers & Didham 2006 Öckinger et al. 2009; Franzén et al.

2012; Hanski 2015 MacDonald et al. 2018a). Here, we propose an approach to addressing each of these problems within a single modelling framework.

5.2.1 Mechanisms underlying ISARs

Three hypotheses, each with distinct underlying mechanisms and conservation implications, have been proposed to account for ISARs and related spatial patterns of species richness: i) the passive sampling hypothesis (Connor & McCoy 1979); ii) area *per se*, as outlined by the theory of island biogeography (MacArthur & Wilson 1963; Wilson & MacArthur 1967); and iii) the habitat diversity hypothesis (Williams 1964). The passive sampling hypothesis, originally developed within the context of oceanic islands, predicts that islands randomly sample individuals from the regional species pool in abundances proportional to their area (Connor & McCoy 1979). As larger islands sample more individuals, they sample more species according to the abundance distribution of the regional species pool (i.e., the "sample-area effect"). Thus, passive sampling serves as a useful null hypothesis, assuming random assembly of both individuals and species.

Area *per se* hypothesizes a disproportionate reduction in species richness as island area decreases, steepening the slope of ISARs within archipelagoes or fragmented terrestrial landscapes relative to species-area relationships within landscapes comprised of continuous habitat (Diamond 1972; Diamond 1975 Wilson & Willis 1975; Connor & McCoy 1979; Saccheri et al. 1998; Gonzalez 2000; Haila 2002; Haddad et al. 2015; MacDonald et al. 2018a; 2018b). From a mechanistic perspective, area *per se* essentially invokes the theory of island biogeography, where species richness arises as a dynamic equilibrium between rates of extinction and colonization, which in turn depend on island area and isolation (MacArthur & Wilson 1963; Wilson & MacArthur 1967). More isolated populations occupying small islands

are predicted to be more prone to stochastic extinction and small, isolated islands are less likely to be re-colonized from external source populations than larger, well-connected islands (Levins 1969; Hanski & Gyllenberg 1993; Orrock & Wattling; 2010). Thus, the theory of island biogeography addresses effects of both island area and isolation, predicting that immigration rates and rescue effects decrease as islands become increasingly isolated from neighboring habitat (i.e., the mainland or other islands), negatively affecting species' probabilities of occurrence and therefore species richness (MacArthur & Wilson 1963; Wilson & MacArthur 1967; Brown & Kodric-Brown 1977; Hanski 1994; 1998 1999).

As an alternative to dynamic balances between demographic rates predicted by the theory of island biogeography, Williams (1964) proposed that ISARs are driven by variation in habitat diversity among islands. The habitat diversity hypothesis predicts that island/fragment area correlates with species richness only insofar as area correlates with the intermediate variable of habitat diversity; larger sample areas generally contain more habitats, which support more species (Williams 1964; Nilsson 1988 Rosenzweig 1995; Gotelli & Graves 1996). It follows that the presence or proportion of suitable habitat within islands/fragments should affect abundances and occurrences of individual species (Buckley 1982; Haila & Järvinen 1983). However, few studies have investigated how habitat associations of single species relate to emergent patterns of species richness in this context (but see Haila et al. 1983). Still, multiple studies addressing species richness support the habitat diversity hypothesis (Nilsson 1988; Kadmon & Allouche 2007; Hortal et al. 2009). It has also been inferred that mechanisms predicted by the passive sampling hypothesis, theory of island biogeography, and habitat diversity hypothesis may simultaneously contribute to ISARs (Connor & McCoy 1979; Kadmon & Allouche 2007; MacDonald et al. 2018b; Chase et al. 2019).

5.2.2 Complications in extending ISARs to conservation

Due to its success on true islands, conservation biologists were quick to recognize the potential for the theory of island biogeography to be applied to fragmented habitat on terrestrial landscapes, initially in the design of nature reserves (e.g., Diamond 1972; Diamond 1975; Wilson & Willis 1975). However, the extension of ISAR-based inferences from true islands to habitat fragments is complicated by differences in their extents of insularity and interactions between habitat and non-habitat (i.e., matrix) areas (Itescu 2019). Whereas edges of true islands clearly delimit suitable habitat from a homogeneous matrix of unsuitable habitat (open water), species occurring on habitat fragments may utilize resources of heterogeneous terrestrial matrices and these matrices may differentially constrain or facilitate colonization rates of species (i.e., "matrix effects"; Dunning et al. 1992; Ricketts 2001). Thus, understanding habitat fragments as analogous to true islands may be problematic. More specifically, fragmentation effects observed within true-island systems may differ from those typical of fragmented terrestrial landscapes (Laurance 2008; Mendenhall et al. 2014; Farneda et al. 2020).

5.2.3 SLOSS-based inferences

Irrespective of the mechanisms underlying ISARs, observations that species richness generally decreases as island/fragment area decreases and isolation increases have contributed to long-standing inferences that habitat fragmentation poses a major threat to species diversity (Diamond 1972; 1975; Noss 1991; Haila 2002; Bruna & Oli 2005; Hanski 2015; Fletcher et al. 2018). However, many of these inferences are founded on observations or experimental designs that have not sufficiently decoupled the effects of area *per se* and isolation from the sample-area effect (i.e., decoupled habitat fragmentation from habitat loss; *sensu* Fahrig 2003; 2013; 2017; Hadley & Betts 2016). In the majority of studies successfully decoupling habitat fragmentation

from habitat loss, single large habitat fragments are generally found to contain an equivalent or lesser number of species than sets of several small habitat fragments summing to an equivalent total area (Quinn & Harrison 1988; Fahrig 2003; 2013; 2017; Yaacobi et al. 2007; MacDonald et al. 2018a; 2018b; Fahrig 2020; Deane et al. 2020). Such comparisons contribute to the ongoing Single-Large-Or-Several-Small ("SLOSS") debate, addressing how finite conservation efforts should prioritize the area and configuration of fragmented habitat and nature reserves (Diamond 1975; Abele & Connor 1979; Ovaskainen 2002; Tjørve 2010). In light of SLOSS-based observations that species richness is often equal or greater within sets of several small habitat fragments, Fahrig (2013) advanced the habitat amount hypothesis, predicting that the number of species persisting on fragmented landscapes is only a function of total habitat amount at the landscape scale irrespective of its spatial subdivision and configuration. The principal mechanism underlying the habitat amount hypothesis is the sample-area effect, as originally articulated by the passive sampling hypothesis (Connor & McCoy 1979). However, the habitat amount hypothesis extends implications of the sample-area effect to predict that there should also be no detectable effect of fragment isolation on species' abundances, species' occurrences, or species richness after the sample-area effect has been accounted for (Fahrig 2013).

5.2.4 The importance of understanding how species-level patterns affect ISARs

While the sample-area effect surely contributes to patterns of species richness within a variety of true-island systems and fragmented terrestrial landscapes, richness-based analyses may obscure important effects of both area *per se* and isolation on individual species (Ewers & Didham 2006; Öckinger et al. 2009; Franzén & Betzholtz 2012; Hanski 2015; MacDonald et al. 2018a). Indeed, area *per se* and isolation effects have been inferred to vary widely among species, even within single landscapes and taxa (Henle et al. 2004; Ewers & Didham 2006;

Nowicki et al. 2009; Öckinger et al. 2009 Hanski 2015; Hillebrand et al. 2018; MacDonald et al. 2018a). Functional traits, including body size (Gehring & Swihart 2003; Henle et al. 2004; Larsen et al. 2008; Prugh et al. 2008; Barbaro & Van Halder 2009; Warzecha et al. 2016), mobility/dispersal ability (Roland & Taylor 1997; Lens et al. 2002; Ewers & Didham 2006; Öckinger et al. 2009; MacDonald et al. 2018a; 2019), degree of ecological specialization (Tscharntke & Brandl 2004), rarity/conservation status (Ewers & Didham 2006), and trophic position (Tscharntke et al. 2002; Thies et al. 2005 Ewers & Didham 2006) are hypothesized to relate species' sensitivity to area *per se* and isolation. Still, relatively few empirical studies have investigated how functional traits relate to interspecific variation in responses to area *per se* and isolation or how this interspecific variation scales to emergent patterns of species richness, such as those reflected in ISARs (Melbourne et al. 2004; but see Barbaro & Van Halder 2009; Öckinger et al. 2009).

5.2.5 Distinguishing mechanisms underlying ISARs

Due to the sample-area effect, observations that species' abundances, species' probabilities of occurrence, or species richness positively correlate with island/fragment area cannot be directly interpreted as evidence of effects of area *per se* (Connor & McCoy 1979; Fahrig 2003; 2013; 2017; Fletcher et al. 2018). Three established methods may be used to control for the sample-area effect: i) comparing sets of islands/fragments that sum to equal areas but differ in degree of fragmentation, including the nested-set designs of Yaacobi et al. (2007) and MacDonald et al. (2018a; 2018b) and comparisons of species accumulation curves proposed by Quinn & Harrison (1988); ii) extrapolating a species-area regression to the total area of all islands/fragments used to build the regression and comparing predicted and observed species richness (β-diversity) (e.g., Rosenzweig 2004; Yaacobi et al. 2007; Santos et al. 2010; Gavish et

al. 2012; MacDonald et al. 2018a; 2018b); and iii) comparing equal-area sampling plots across islands/fragments (Westman 1983; Kelly 1989; Quinn et al. 1987 Stevens et al. 1986; Fahrig 2013; Watling et al. 2020). However, for methods 1 and 2 (SLOSS-based comparisons), substantial species turnover among several small islands/fragments can increase their aggregate richness relative to single large islands/fragments, such that important effects of area *per se* on individual species are obscured (*sensu* Simberloff 1976; MacDonald et al. 2018b; Deane et al. 2020). Assuming species are uniformly distributed within islands/fragments, inferences drawn from method 3 may be robust for sessile taxa (e.g., vascular plants; Westman 1983; Quinn et al. 1987; Kelly 1989), but remain tenuous for vagile species that move frequently within islands/fragments, as individual sampling plots may accumulate all vagile species within single islands/fragments if sampling effort is high. Additionally, if rare species are particularly sensitive to area *per se* or isolation, these effects are likely to go undetected when sampling at small spatial grain sizes dictated by method 3; rare species and important habitat types within islands/fragments may be missed entirely in comparisons of small sampling plots (Karger et al. 2014). Finally, each of the three established methods only contrast the sample-area effect with those of area *per se* (methods 1 & 2) or area *per se* and isolation (method 3); evaluating effects of isolation and habitat diversity requires additional analyses. A fourth method, recently proposed by Chase et al. (2019), employs parameters derived from individual-based rarefaction curves across various spatial scales within islands/fragments to distinguish between the sample-area effect and effects of area *per se* and habitat diversity. While this framework can effectively distinguish mechanisms underlying ISARs, it cannot simultaneously evaluate effects of isolation, assess whether area *per se* and isolation differentially affect species in relation to their functional

traits, or be applied to existing datasets lacking abundance data for subplots stratified within each island/fragment across diagnosable habitat heterogeneity.

In this paper, we present a novel application of random placement and linear mixed effects models that can simultaneously evaluate: i) how area *per se*, isolation, and habitat diversity affect species' abundances, species' occurrences, and species richness across true islands or terrestrial habitat fragments; and ii) whether interspecific variation in these responses relates to variation in species' functional traits. This modelling framework is applicable to any dataset for which abundance data were collected for sets of true islands or habitat fragments with sampling effort standardized per unit area. We assess the utility of the framework using a butterfly assemblage persisting on a naturally fragmented landscape of true islands; Lake of the Woods, Canada. Methodological developments and basic inferences presented here are equally applicable to both true islands and terrestrial habitat fragments, so long as the edges of habitat fragments can be consistently delimited.

5.3 Materials and Methods

5.3.1 Overview of the modelling framework

Starting with the assumption that all individuals of each species are randomly distributed across true islands or habitat fragments in abundances proportional to their areas, random placement models can be used to calculate expected species' abundances, expected probabilities of species' occurrences, and expected species richness for each island or fragment (Arrhenius 1921; Gotelli & Graves 1996 Coleman 1981; He & Legendre 2002). Resulting random placement values are equivalent to values of species' abundances, species' probabilities of occurrence, and species richness predicted by the sample-area effect. Predictions of the passive sampling/habitat amount hypotheses, theory of island biogeography, and habitat diversity

hypothesis may be then simultaneously evaluated by modelling relationships between random placement residuals (observed values minus random placement values) and the area, isolation, and habitat diversity of individual islands/fragments. Variables measuring species' functional traits may be introduced to abundance and occurrence models via interaction terms with island/fragment area and isolation to evaluate whether effects of area *per se* and isolation interspecifically vary contingent on the measured traits.

5.3.2 Random placement models

We present random placement models in ascending order of mathematical complexity, from species' abundances, to species' occurrences, to species richness. Within random placement models, a_j is the area of the j^{th} island/fragment, A_T is the total area of all islands/fragments, n_i is the abundance of species i summed across all islands/fragments, and S is the total number of species observed. Islands or fragments that were not surveyed are not included in random placement models. According to the sample-area effect, the expected abundance of species i on island/fragment j is simply proportional to j's area (model 1) and the occurrence probability follows the random placement model (model 2). The random placement model for expected richness on any island/fragment is simply the sum of the expected probabilities of occurrence over all species (model 3). This random placement richness model was first proposed a century ago by Arrhenius (1921) and later reinvented by Coleman (1981) with the inclusion of the variance.

(1) $$E(n_{ij}) = n_i(\frac{a_j}{A_T})$$

(2) $$E(O_{ij}) = 1 - (1 - \frac{a_j}{A_T})^{n_i}$$

(3) $$E(S_j) = \sum_{i=1}^{S} \left\{1 - (1 - \frac{a_j}{A_T})^{n_i}\right\}$$

5.3.3 Modelling of random placement residuals

Subtracting random placement abundance and occurrence probability values from observed abundance and occurrence values for each species on each island/fragment produces abundance and occurrence residuals. The direction and magnitude of these residuals measure how abundances and occurrences of each species deviate from predictions of the sample-area effect. In the modelling framework described below, all species' abundance residuals and all species' occurrence residuals are concatenated across species for use in single linear mixed effects models; one model addressing all species' abundances and one model addressing all species' occurrences. Abundance residuals require standardization (subtracting the mean and dividing by standard deviation) before concatenation across species, as the range of possible values is greater for common species than rare species. While this is not the case for occurrence residuals (values bound between -1 and 1), standardization is still recommended to generate values that are commensurate among species. Richness residuals are similarly calculated for each island/fragment by subtracting random placement richness values from observed richness values. Standardization is recommended to facilitate comparisons of effect sizes among abundance, occurrence, and richness analyses.

5.3.3.1 Species' abundances and occurrences

Linear mixed effects models are used to relate abundance and occurrence residuals to island/fragment variables while controlling for species identity as a random effect. If area *per se* affects species' abundances or occurrences beyond variation associated with the sample-area effect, residuals will be positively related to area, indicating a disproportionate concentration of

species' abundances or occurrences on larger islands/fragments. If isolation negatively affects species' abundances or occurrences, residuals will be negatively related to measures of isolation specific to individual islands/fragments, indicating a disproportionate concentration of species' abundances or occurrences on less isolated islands/fragments. Each of these results align with predictions of the theory of island biogeography. Alternatively, the absence of significant relationships between residuals and area or isolation would indicate that the sample-area effect sufficiently accounts for variation in species' abundances or occurrences across islands/fragments of varying area or isolation. The combination these results would confer support for the passive sampling/habitat amount hypotheses. The proportion of species-specific suitable habitat and presence of species-specific resources within each island/fragment may also be included in linear mixed effects models. Positive relationships between abundance or occurrence residuals and these variables would indicate that availability of specific habitats or resources within islands/fragments are important considerations that affect species' abundances or occurrences, as predicted by the habitat diversity hypothesis.

If data on species' functional traits are available, functional trait variables may be introduced to linear mixed effects models via interaction terms with island/fragment area and isolation. Here, a significant interaction between a functional trait variable and island/fragment area or isolation would indicate that area *per se* or isolation differentially affects species' abundances or occurrences contingent on the measured trait. Total abundance and number of occurrences (prevalence) for each species are also of interest, as rarity is cited as a predictor of species' sensitivity to fragmentation (Ewers & Didham 2006). A significant positive interaction between total abundance or prevalence and island/fragment area would indicate that rare species are disproportionately concentrated or likely to occur on larger islands/fragments. Similarly, a

significant negative interaction between total abundance or prevalence and island/fragment isolation would indicate that rare species are disproportionately concentrated or likely to occur on less isolated islands/fragments.

5.3.3.2 Species richness

A similar modelling process may be applied to species richness using linear models. Significant relationships between richness residuals and island/fragment area or isolation would indicate that area *per se* or isolation significantly affects richness after controlling for the sample-area effect. Each of these results align with predictions of the theory of island biogeography. Conversely, the absence of significant relationships between residuals and area and isolation would indicate that the sample-area effect sufficiently accounts for variation in species richness across islands/fragments of varying area or isolation. This combination of results would suggest that only habitat amount at the archipelago- or landscape-scale affects richness, as predicted by the passive sampling/habitat amount hypotheses. Predictions of the habitat diversity hypothesis may also be simultaneously evaluated by including measures of habitat diversity in linear models. Significant relationships between richness residuals and habitat diversity would indicate that, despite correlations between habitat diversity and island/fragment area, variation in habitat diversity among islands/fragments affects species richness beyond variation associated with both the sample-area effect and effects of area *per se*.

5.3.4 Application of the modelling framework

5.3.4.1 Study area

We assessed the utility of this modelling framework using a butterfly assemblage persisting on a ~1250 km^2 lake-island complex located in Sabaskong Bay, Lake of the Woods,

Canada (Figure 5-1). Differential isostatic rebound and outlet restriction resulted in the flooding of the study area and isolation of land-bridge islands approximately 3000 – 4000 YA (Yang & Teller 2005). Given this substantial time-since-isolation, we infer that species assemblages have relaxed to equilibria (>1000 generations; based on categories suggested by Fahrig et al. 2020). Butterflies represent a suitable taxon for this investigation, as most species complete their life cycles within relatively small patches of habitat, their detectability is generally high, and their diversity correlates with that of many other terrestrial taxa (Thomas 2005; Nowicki et al. 2008; MacDonald et al. 2017; 2018a). Butterflies do not utilize open water at any life stage, meaning the matrix separating islands in this system is entirely unsuitable. This effectively controls for matrix effects (Ricketts 2001; Dunning et al. 1992), but also limits the generalizability of our inferred fragmentation effects to terrestrial landscapes (Laurance 2008; Mendenhall et al. 2014; Farneda et al. 2020).

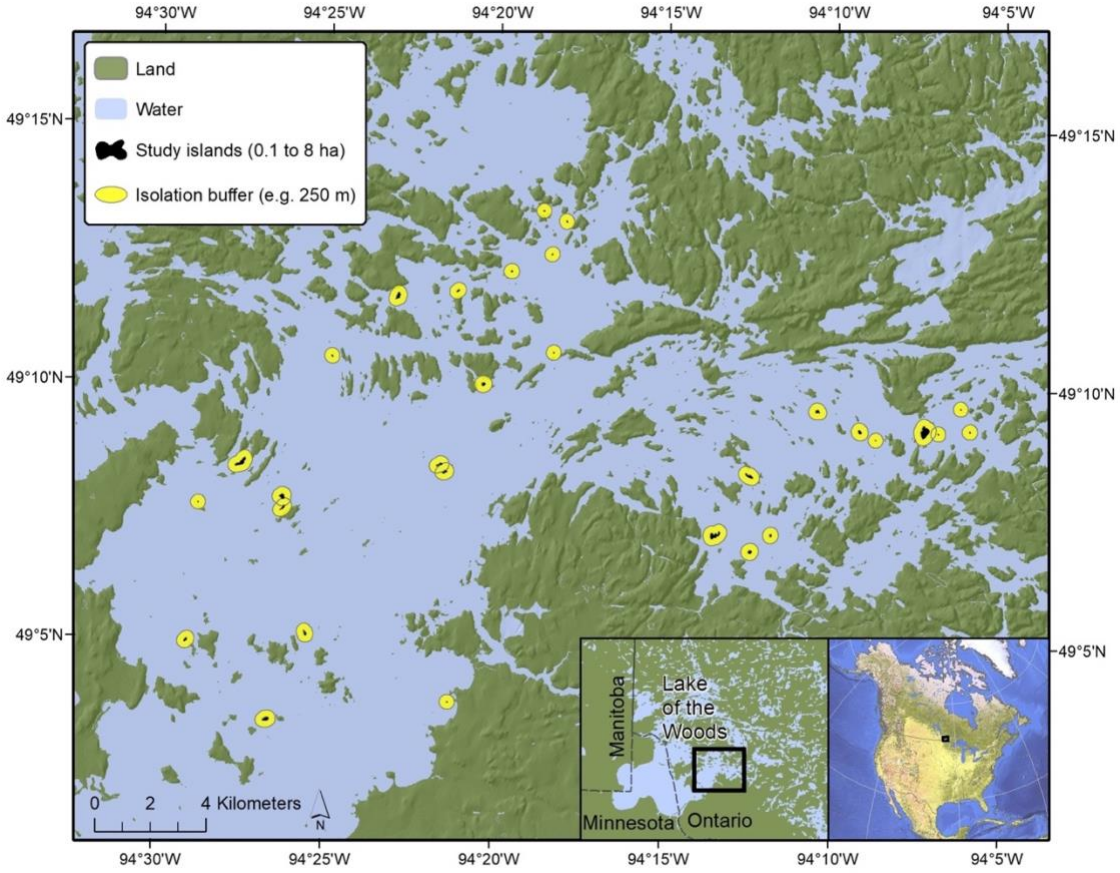

Figure 5-1. Map of the study area, located in Sabaskong Bay, Lake of the Woods, Canada. Butterfly abundance, occurrence, and species richness data were collected for 30 study islands, varying in area from 0.09 to 8.4 ha, using repeated full island surveys.

Thirty islands, ranging in area from 0.1 to 8.0 ha, were randomly selected from lists of candidate islands compiled according to the methods of MacDonald et al. (2018a). Islands were only considered as candidates if they were isolated from other landmasses by at least 100 m, beyond the inferred visual ranges of butterflies (Rutowski 2003; MacDonald et al. 2019). The relative isolation of each study island was quantified at multiple scales as the proportion of water within 250-, 500-, 1000-, 1500-, 2000-, 2500-, 3000-, 3500-, 4000-, 4500-, and 5000-m buffers.

Buffers were generated from island edges, meaning isolation measures are independent from and uncorrelated with island area. We opted for these proportion-based measures because they have been shown to predict immigration rates, rescue effects, and related ecological processes more accurately than distance-based measures (Moilanen & Nieminem 2002; Tischendorf et al. 2003; Prugh 2009). Habitat diversity was estimated on each island as the relative proportion of 14 habitat types, defined using structural properties of vegetation and geological features (see MacDonald et al. 2021 Supporting Information for habitat type descriptions).

5.3.4.2 Survey protocol

Butterfly abundance data were collected for each of the 30 islands using repeated full-island surveys. Each island was visited four times at intervals between 10 and 14 days during the peak flight season (01-June-2015 to 20-Aug-2015). Sampling effort was standardized to 40 min per ha per survey across all islands, eliminating the need for sampling effort and diversity corrections (e.g., rarefaction or extrapolation; e.g., Chao et al. 2014; Fahrig et al. 2020). Care was taken to visit all habitat types during each survey and handheld GPS units were used to ensure uniform coverage of islands. Recording observer tracks and capturing individuals whenever possible (kept as voucher specimens or released at the end of each survey) minimized the possibility of double counts, where individuals are recorded multiple times in single surveys. To ensure optimal and standardized butterfly activity, surveys were restricted to the hours of 10:45 to 15:45 and were not conducted in wind speeds over 15 km/hr or in temperatures below 13°C. If temperatures were below 17°C, surveys were only conducted in sunny conditions (cloud cover < 40%). Surveys were conducted in temperatures above 17°C, regardless of cloud cover (MacDonald et al. 2017).

The diversities of vascular plants and butterflies have been observed to positively correlate with one another in a variety of systems, primarily due to larval host plant associations (Erhardt 1985; Sparks & Parish 1995; Simonson et al. 2001; Croxton et al. 2005; Kitahara et al. 2008; Nowicki et al. 2009; MacDonald et al. 2018a; Riva et al. 2020). Accordingly, vascular plant species richness was quantified on each island using repeated full-island surveys (four surveys total), standardized to 40 min per ha per survey [see MacDonald et al. (2018b) for further details on vascular plant surveys]. A total survey time of two hr and 40 min per ha is consistent with recent recommendations for boreal plant communities (Zhang et al. 2014). Only presence-absence data were collected for vascular plants, precluding use of our modelling framework [but see Simberloff & Gotelli (1984) for random colonization models applied to presence-absence data].

5.3.4.3 Data analysis

We calculated values of species' abundances, species' probabilities of occurrence, and species richness predicted by random placement for each of the 30 study islands using random placement models (1, 2, and 3, respectively). Abundance, occurrence, and richness residuals were estimated as observed values minus random placement values. We next used linear mixed effects models to simultaneously: i) quantify relationships between either abundance residuals or occurrence residuals and island variables, including area, isolation, proportion of suitable habitat, and presence/absence of preferred larval host plants; ii) evaluate whether area *per se* or isolation differentially affects species' abundances or occurrences contingent on their functional traits. Separate models were built for each isolation buffer size (250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m). Each of these models had the same number of parameters (k) and only differed in the buffer size used to measure island isolation. We therefore directly

compared models with log likelihood (e.g., Lamb et al. 2018), where the model with the maximum log likelihood identified the optimal buffer size. Relationships between log likelihood and buffer sizes were quantified using Pearson's product moment correlation coefficients. The proportion of suitable habitat for each species was estimated as the total area of suitable habitat types divided by the total area of the island. Habitat types were classified as suitable for a species if we observed at least one individual within them during the repeated full-island surveys. The presence/absence of preferred larval host plants (compiled from Hall et al. 2014; Acorn & Sheldon 2017) for each species on each island was included as a binary variable. We included species' wingspan (mm; as reported in Burke et al. 2011) in models as a functional trait variable, serving as a measure of both body size and mobility/dispersal ability (Ewers & Didha 2006; Lens et al. 2002; Roland & Taylor 1997; Öckinger et al. 2009; MacDonald et al. 2018a; 2019). Other functional traits predicted to relate to species' sensitivity to fragmentation (e.g., degree of ecological specialization, trophic position) were either not measured or did not vary substantially among butterfly species and so were not investigated. As an inverse measure of species' rarity, we included species' total abundance and prevalence in abundance and occurrence models, respectively. To evaluate whether area *per se* or isolation differentially affected species' abundances or occurrences contingent on their functional traits, we included the following interaction terms: wingspan:area, wingspan:isolation, rarity:area, and rarity:isolation. All non-binary predictor variables were standardized (subtracting the mean and dividing by standard deviation), permitting comparisons of effect sizes. The structure for both abundance and occurrence linear mixed effects models was as follows, where "habitat" is the proportion of suitable habitat for each species on each island and "plants" is the presence/absence of each species' preferred larval host plants:

f(residuals) ~ β_1(area) + β_2(isolation) + β_3(habitat) + β_4(plants) + β_5(wingspan) + β_6(rarity) + β_7(wingspan:area) + β_8(wingspan:isolation) + β_9(rarity:area) + β_{10}(rarity:isolation) + (1|species id) + e

Linear models were fitted for species richness following the same basic protocol as abundance and occurrence linear mixed effects models. Predictor variables included island area, isolation, habitat diversity, and vascular plant species richness. The most supported isolation buffer size was again assessed using log likelihood comparisons among models differing only in buffer size. Habitat diversity was estimated as the number of habitat types on each island. All predictor variables were standardized. The structure for richness linear models was as follows, where "habitat" is the total number of habitat types recorded on each island and "plants" is vascular plant species richness:

f(residual) ~ β_1(area) + β_2(isolation) + β_3(habitat) + β_4(plants) + e

5.4 Results

A total of 869 butterflies belonging to 34 species were observed during repeated full island surveys. Butterfly abundance data are reported in MacDonald et al. (2021) Supporting Information. One species, *Feniseca tarquinius*, uniquely feeds on woolly aphids in its larval stage (Hall et al. 2014; Acorn & Sheldon 2017). Only one individual of this species was observed across all surveys. We excluded it from abundance and occurrence models, which included presence/absence of preferred larval host plants as a predictor variable. All individuals belonging to all species were included in richness models.

5.4.1 Species' abundances

Comparing log likelihood among linear mixed effects models differing only in isolation buffer size resolved that the proportion of water within 250 m (smallest buffer size) was most supported (Table 5-1; Figure 5-2 a). Model support significantly declined across increasing buffer sizes (r = -0.709; P = 0.015), indicating that the amount of habitat immediately surrounding individual islands better predicted species' abundances than the amount of habitat at broader spatial scales. Within the most supported model, abundance residuals were significantly related to both island area and isolation (Table 5-2; Figure 5-3). Area *per se* had a significant positive effect on species' abundances, while isolation had a significant negative effect, in accordance with mechanisms predicted by the theory of island biogeography. The absolute magnitude of the effects of area *per se* and isolation, inferred from standardized regression coefficients, were approximately equivalent. Together, these results indicate that the sample-area effect cannot account for variation in species' abundances across islands of varying area and that habitat configuration, and not just total area, has important effects on species' abundances in this system. Other island variables, including the proportion of suitable habitat and presence/absence of preferred larval host plants, were not related to species' abundances.

Coefficients of the wingspan:area and rarity:area interaction terms were significantly negative, indicating that effects of area *per se* systematically varied across species in respect to these functional traits. Causality behind the wingspan:area interaction is clear; effects of area *per se* on abundance were greater for smaller, less mobile butterfly species. However, for rarity:area, it cannot be resolved whether effects of area *per se* on abundance were greater for rare species, or whether these species were rare within the dataset because they experience greater effects of area *per se*. Comparisons of the relative abundances of species between the mainland

(continuous habitat) and islands (fragmented habitat) would help resolve the causal direction of this relationship; however, this was beyond the scope of this study. Relationships between abundance residuals and functional trait variables were not significant. This result is expected, as abundance residuals were standardized for each species individually before they were concatenated for use in linear mixed effects models.

Table 5-1. Log likelihood values for linear mixed effects (abundance and occurrence) and linear (species richness) models. For species' abundances, species' occurrences, and species richness, separate models were built for a range of isolation buffers, measuring the proportion of open water within 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m of island shores. Model support significantly declined across increasing isolation buffer sizes in all instances. The most supported buffer size for each model set is highlighted in bold.

Isolation buffer (m)	Abundance	Occurrence	Species richness
250	**-1329.83**	**-1340.37**	-36.79
500	-1330.22	-1340.71	-37.45
1000	-1330.99	-1342.89	-37.67
1500	-1332.11	-1341.77	**-36.61**
2000	-1332.00	-1340.52	-37.02
2500	-1332.05	-1340.72	-37.79
3000	-1332.17	-1341.71	-38.12
3500	-1332.13	-1342.61	-38.36
4000	-1332.01	-1343.28	-38.45
4500	-1331.86	-1343.60	-38.48
5000	-1331.87	-1343.53	-38.44

Table 5-2. Standardized regression coefficient estimates ("coef."), standard errors ("s.e."), and *P*-values ("*P*") from linear models fitting random placement residuals for species' abundances, species' occurrences, and species richness. Included in all models were island area, measured in

m^2, and island isolation, measured as the proportion of water within the most supported buffer size (250 m for abundance and occurrence; 1500 m for species richness). Within abundance and occurrence models, the proportion suitable habitat ("habitat") was measured for each species as the area of suitable habitat on each island divided by the area of the island. Presence/absence of each species' preferred larval host plants ("plants") was included as a binary variable. Wingspan was included as a measure of species' body size and as a proxy of dispersal ability. Each species' total abundance and prevalence were used as an inverse measure of rarity in abundance and occurrence models, respectively. Species identity was included as a random effect in abundance and occurrence models. Within the species richness model, habitat diversity ("habitat") was estimated as the total number of habitat types recorded on each island. Plant diversity ("plants") was measured as vascular plant species richness. Significant coefficients ($\alpha = 0.05$) are highlighted in bold.

	Abundance			**Occurrence**			**Species richness**		
Variable	**coeff.**	**s.e.**	***P***	**coeff.**	**s.e.**	***P***	**coeff.**	**s.e.**	***P***
area	**0.087**	0.032	0.007	0.005	0.033	0.884	0.294	0.292	0.323
isolation	**-0.069**	0.032	0.028	**-0.073**	0.032	0.022	-0.320	0.173	0.076
habitat	0.020	0.032	0.535	0.016	0.033	0.635	-0.763	0.395	0.075
plants	0.018	0.082	0.830	-0.070	0.083	0.400	0.110	0.484	0.822
wingspan	-0.001	0.035	0.973	-0.015	0.035	0.661			
rarity	-0.004	0.032	0.892	-0.001	0.033	0.988			
wingspan:area	**-0.117**	0.031	<0.001	-0.057	0.031	0.069			
wingspan:isolation	0.012	0.031	0.708	0.006	0.031	0.838			
rarity:area	**-0.081**	0.031	0.009	-0.009	0.032	0.783			
rarity:isolation	0.012	0.031	0.688	0.041	0.032	0.196			

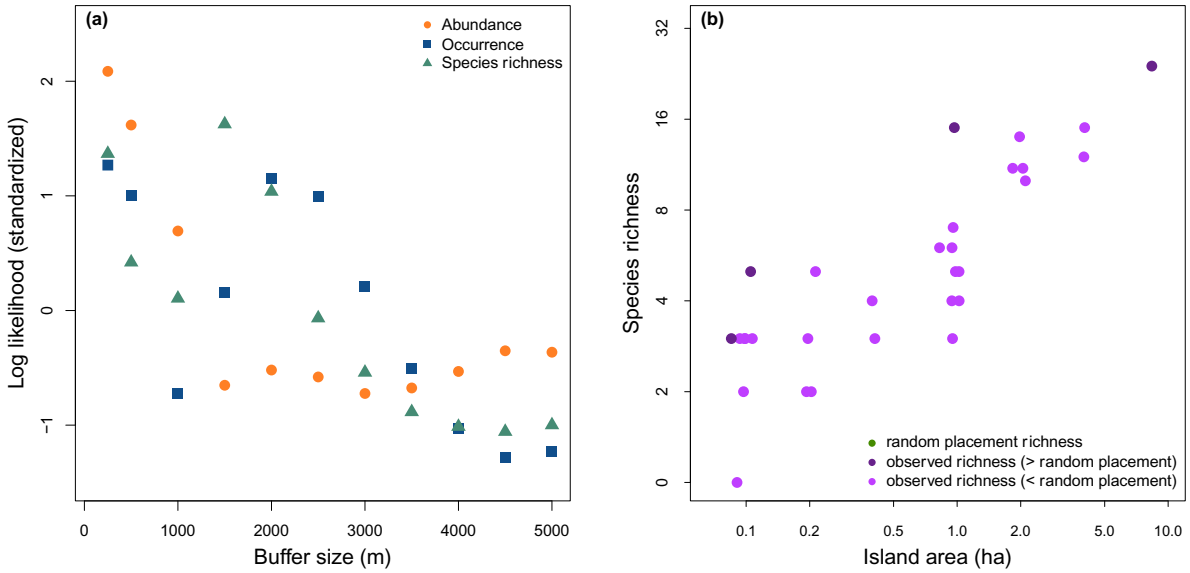

Figure 5-2. a) Standardized log likelihood values for linear mixed effects (abundance and occurrence) and linear (species richness) models where responding variables were random placement residuals. Explanatory variables for each model are listed in Table 5-2 and Figure 5-3. Separate models were built using different island isolation buffer sizes, quantifying the proportion of open water within 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m of island shores. To permit comparisons of relationships between log likelihood values and buffer sizes among abundance, occurrence, and species richness model sets, log likelihood scores were standardized for each model set by subtracting the mean and dividing by standard deviation. Model support significantly declined across increasing isolation buffer sizes in all instances. b) Log-log island species-area relationship (ISAR) for butterflies occurring on 30 study islands. Random placement richness values were calculated using a random placement model (model 3; see Materials and Methods). Dashed green lines are 95% confidence intervals for random placement values, calculated using Coleman's (1981) formula for variance. Dashed

purple lines represent 95% confidence intervals for the log-log ISAR linear regression (solid purple line), parameterized using observed richness values for all 30 islands.

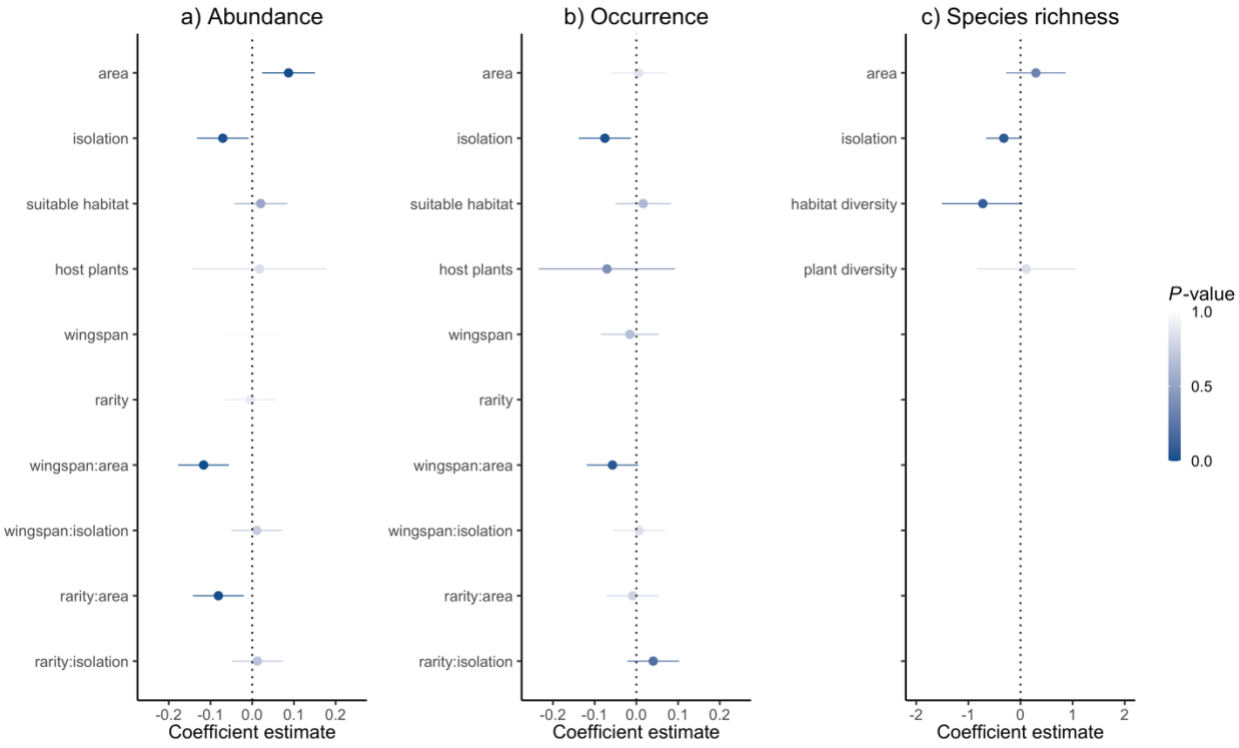

Figure 5-3. Standardized regression coefficients and 95% confidence intervals from linear models relating random placement residuals to island characteristics and species' functional traits for a) species' abundances, b) species' occurrences, and c) species richness. Included in all models were island area, measured in m^2, and island isolation, measured as the proportion of water within the most supported buffer size (250 m for abundance and occurrence; 1500 m for species richness). Within abundance and occurrence models, the proportion suitable habitat ("suitable habitat") was measured for each species as the area of suitable habitat on each island divided by the area of the island. Presence/absence of each species' preferred larval host plants

was included as a binary variable. Wingspan was included as a measure of species' body size and as a proxy of dispersal ability. Each species' total abundance and prevalence were used as an inverse measure of rarity in abundance and occurrence models, respectively. Species identity was included as a random effect in abundance and occurrence models. Within the species richness model, habitat diversity was estimated as the total number of habitat types recorded on each island. Plant diversity was measured as vascular plant species richness. The shading of each variable's point estimate (coefficient) and confidence interval is proportional to its *P*-value, with darker shades indicating greater significance. Coefficients with 95% confidence intervals not overlapping zero were inferred to be significant at $\alpha = 0.05$.

5.4.2 Species' occurrences

As with abundance linear mixed effects models, the most supported isolation buffer size for predicting occurrence residuals was 250 m (Table 5-1; Figure 5-2 a). Model support significantly declined across increasing buffer sizes ($r = -0.746$; $P = 0.008$). Within the most supported model, occurrence residuals showed no relationship to island area, suggesting that the sample-area effect sufficiently accounts for variation in species' occurrences across islands that vary in area (Table 5-2; Figure 5-3). This result confers support for the passive sampling/habitat amount hypotheses. However, occurrence residuals were significantly negatively related to island isolation, suggesting that island configuration has important effects on species' occurrences, as predicted by the theory of island biogeography. Other island variables, including the proportion of suitable habitat and presence/absence of preferred larval host plants, were not significantly related to occurrence residuals.

Effects of functional trait variables (wingspan and rarity) and their interaction with island variables were not significant at $\alpha = 0.05$. However, the coefficient of the wingspan:area interaction term was marginally significant at $P = 0.069$, suggesting that smaller, less mobile species were less likely than larger, more mobile species to occur on small islands. In other words, butterfly species richness on small islands may be disproportionately comprised of large butterfly species with high mobility.

5.4.3 Species richness

The most supported isolation buffer size for predicting species richness residuals was 1500 m, which was only marginally more supported than the smallest (250 m) buffer size (Table 5-1; Figure 5-2 a). Notwithstanding, model support still significantly declined across increasing buffer sizes ($r = -0.833$; $P = 0.001$). Within the most supported model, residuals were not significantly related to island area, isolation, habitat richness, or vascular plant species richness (Table 5-2; Figure 5-3). It should be noted, however, that the directionality of the island isolation coefficient aligned with predictions of the theory of island biogeography, and was significant at $\alpha = 0.10$. Failure to resolve a significant effect at $\alpha = 0.05$ suggests that richness-based analyses confer less analytical power than abundance- and occurrence-based analyses. If we adopt the most commonly used significance threshold of $\alpha = 0.05$, our linear model would suggest that the sample-area effect sufficiently accounts for variation in richness observed across our study islands, conferring support for the passive sampling/habitat amount hypotheses and random assembly of species.

5.5 Discussion

Distinguishing mechanisms that underlie variation in species' abundances, species' occurrences, and emergent patterns of species richness is not only of great interest within the context of fundamental ecology, but is also of paramount importance to applied ecology and understanding how habitat fragmentation affects species diversity (Diamond 1975; Simberloff & Abele 1976; 1982; Haila 2002; Haddad et al. 2015; Chase et al. 2019). Here, we detail a novel modelling framework to test hypotheses on which conservation directives are contingent (Fahrig 2003; 2013; 2017; Haddad et al. 2015; Hanski 2015; Fletcher et al. 2018). Applying this modelling framework to a butterfly assemblage persisting on a naturally fragmented landscape of true islands, we were able to resolve that: i) island area *per se* and isolation significantly affect species' abundances and occurrences contingent on their functional traits; and ii) important effects of area *per se* and isolation are not always apparent in aggregate diversity measures, such as those reflected in ISARs. Although there are several documented divergences between the biogeographies of true-island systems and fragmented habitat on terrestrial landscapes, findings from our study clearly demonstrate that fragmentation effects should not be inferred from richness-based analyses, but rather evaluated on a species-by-species basis.

5.5.1 Inferences from the ISAR

Our modelling framework resolved that spatial patterns in butterfly species richness (i.e., the ISAR) did not significantly deviate from random placement in relation to island area, isolation, habitat diversity, or vascular plant diversity. These ISAR-based inferences align with those of MacDonald et al. (2018a), who used a series of SLOSS-based analyses to infer that butterfly species richness in this naturally fragmented landscape approximately conform to predictions of the sample-area effect, with no significant effects of area *per se* or isolation. Therefore, all analyses addressing effects of area *per se* and isolation on butterfly species

richness in this system support the passive sampling/habitat amount hypotheses and Fahrig's (2003; 2013; 2017) general conclusion that effects of fragmentation (i.e., area *per se* and isolation) are generally negligible after controlling for deleterious effects of habitat loss; only habitat amount at the landscape-scale affects species richness via its influence on regional species pools.

While there seems to be considerable support for the passive sampling/habitat amount hypotheses in the literature (e.g., Fahrig 2003; 2013; Marti 2018; Watling et al. 2020), a recent meta-analysis by Fahrig (2020), including 157 SLOSS comparisons from 58 studies, found that several small islands/fragments contained more species than single large islands/fragments in 72% of comparisons, equivalent numbers of species in 22% of comparisons, and fewer species in 6% of comparisons. Removing studies with biased sampling effort in relation to island/fragment area shifted these figures to 58%, 37%, and 5%. Regardless, these results suggest that species richness varies with degree of fragmentation more often than not. Thus, the sample-area effect implicated in the passive sampling/habitat amount hypotheses cannot consistently account for SLOSS-based richness patterns. Furthermore, methods employed by Fahrig (2020)—specifically, comparisons of Quinn & Harrison (1988) species accumulation curves—suffer from an important limitation: substantial species turnover among several small islands/fragments can inflate their aggregate richness, such that important deviations from random placement (e.g., effects of area *per se*) are obscured (*sensu* Simberloff 1976 MacDonald et al. 2018b; Deane et al. 2020). This relationship may explain why several small islands/fragments are generally found to contain more species than single large fragments in the majority of SLOSS-based studies. While results of Fahrig's (2020) meta-analysis are of great interest to both fundamental and applied ecology, they cannot necessarily be used to distinguish effects of area *per se* from the sample-

area effect and cannot resolve additional effects of isolation or habitat diversity. Future investigations focusing on species richness would benefit from the inclusion of ISAR-based analyses that assess deviations from random placement on an island-by-island or fragment-by-fragment basis, such as those included within the modelling framework presented here.

Additional deviations from random placement may be resolved by visually examining the ISAR and random placement richness values (e.g., Figure 5-2 b). In this study, observed richness values were generally less than random placement richness values (all islands except four). This pattern cannot be attributed to effects of area *per se*, because the direction and magnitude of richness residuals was relatively consistent across islands of varying area (Figure 5-2 b), as indicated by the insignificant coefficient of island area within the species richness linear model (Table 5-2). Rather, this relationship is best explained by spatial species aggregation, wherein conspecific individuals are more likely to co-occur on islands than what is predicted by random placement, reducing the species richness of individual islands (He & Legendre 2002). Therefore, although spatial patterns of species richness did not significantly deviate from random placement in respect to either island area or isolation, they did deviate from random placement in respect to spatial species aggregation. It is therefore clear that failure to resolve significant effects of area *per se* and isolation on species richness cannot be taken as direct evidence for random assembly of individuals and species, the fundamental prediction of the habitat amount/passive sampling hypotheses (*c.f.* Fahrig 2003; 2013; 2017; 2020). Rigorous evaluation of random assembly requires comparison of observed richness to expected richness predicted by null models, such as the random placement models presented here.

5.5.2 Inferences from species' abundances and occurrences

Inferring whether fragmentation is "good" or "bad" (*sensu* Fahrig 2017; Fletcher et al. 2018) based on emergent patterns of species richness is potentially susceptible to "ecological fallacy" (*sensu* Robinson 1950), which describes biases that may arise when observed effects on aggregated variables (e.g., species richness) differ from causal relationships at more reductive and informative levels of organization (e.g., species' abundances and occurrences). Indeed, our abundance and occurrence models resolved important effects of area *per se* and isolation that were not apparent in either ISAR- or SLOSS-based analyses [this study and see MacDonald et al. (2018a), respectively]. This discrepancy among inferences suggests that conflating responses of all species into a single aggregate measure (e.g., species richness) reduces our power to detect important relationships on which conservation directives should be contingent. Two such relationships were resolved when considering the entire species assemblage in abundance and occurrence models: i) there was a disproportionate concentration of individuals on larger and less-isolated islands relative to what was predicted by the sample-area effect (passive sampling/habitat amount hypotheses); and ii) species were more likely to occur on less-isolated islands than what was predicted by the sample-area effect. It is therefore clear that the sample-area effect described by the passive sampling/habitat amount hypotheses cannot adequately account for spatial patterns in butterfly abundances and occurrences in this naturally fragmented landscape, which are better predicted by mechanisms outlined by the theory of island biogeography. It should be noted that the directionality of area *per se* and isolation effects were consistent among abundance, occurrence, and richness models, but only statistically significant ($\alpha = 0.05$) in the first two analyses, wherein species' responses were not aggregated into a single measure.

Most interestingly, our modelling framework simultaneously resolved that effects of area *per se* on species' abundances and occurrences varied significantly with species-specific functional traits, suggesting that mechanisms outlined by the theory of island biogeography are not neutral with respect to species identity. Whether species' sensitivity to fragmentation varies predictably with functional traits is a long-standing and pertinent question in conservation biology (Roland & Taylo 1997; Lens et al. 2002; Gehring & Swihart 2003; Henle et al. 2004; Tscharntke & Brandl 2004; Thies et al. 2005; Ewers & Didham 2006; Prugh et al. 2008; Barbaro & Van Halder 2009; Öckinger et al. 2009; Hanski 2015; Warzecha et al. 2016; MacDonald et al. 2018a; 2019). In this study, effects of area *per se* were significantly greater for butterfly species with smaller wingspans. For Canadian butterflies, wingspan is one of the strongest correlates of estimated mobility and dispersal ability (Burke et al. 2011); thus, we can infer that effects of area *per se* are greatest for small species of limited mobility (*c.f.* Larsen et al. 2008). This relationship may be explained by larger and more mobile butterfly species having the ability to move among multiple small islands to meet their resource requirements. These "transient" species may thereby exhibit patterns of abundance and occurrence that approximate random placement (MacArthur & Wilson 1967: Chapter 2; Rosenzweig 2004; MacDonald et al.. 2018a). Such patterns would be predicted by ideal free distribution theory (*sensu* Dreisig 1995) if two conditions are met: i) islands of varying area contain equivalent densities of resources; and ii) the mobility/dispersal ability of individuals is sufficient to render costs of inter-island movements negligible. By contrast, island edges may be perceived as impassible barriers for smaller and less mobile species, with energy expenditures and mortality risks associated with movement through the open-water matrix being too high for regular inter-island movements. Island edges may therefore delimit populations of smaller and less mobile species, which are generally restricted to

larger islands that contain all resources required for mate location, reproduction, resting, roosting, predator escape, and feeding (i.e., the functional resource-based habitat concept; Dennis et al. 2002). This hypothesis is supported by analyses of MacDonald et al. (2018a), who resolved that the probability of butterfly species occurring on at least one island without their preferred larval host plants was positively related to species' wingspan and estimated mobility. Considered together, these results suggest that functional traits may be used to predict species' sensitivity to fragmentation and that species identity should not be ignored when investigating mechanisms that underlie ISARs or in conservation planning. It is, however, important to recognize that the open-water matrix of this study landscape may be more unsuitable and less permeable than those of many fragmented terrestrial landscapes (Dunning et al. 1992; Ricketts 2001; Laurance 2008; Mendenhall et al. 2014; Itescu 2019; Farneda et al. 2020). It is therefore unclear the degree to which these relationships between species' functional traits and effects of area *per se* and isolation are generalizable to conservation efforts addressing terrestrial landscapes fragmented through anthropogenic activities.

The abundance and occurrence modelling framework proposed here may also be implemented on a species-by-species basis by regressing island/fragment variables (area, isolation, presence/amount of suitable habitat or specific resources, etc.) on abundance and occurrence residuals for each species in separate linear models. This method of analysis precludes the simultaneous integration of functional trait analyses within models, but has the added advantage of identifying single species that are particularly sensitive to habitat fragmentation (area *per se* and isolation) or other island/fragment variables of interest. This simple decomposition of our modelling framework may be used to resolve whether particular species require independent consideration within conservation frameworks.

5.5.3 Isolation vs. habitat amount

The relative isolation of islands addressed in this study was quantified across multiple scales as the proportion of water within 11 buffer sizes, ranging from 250 to 5000 m. These measures are equal to 1 minus the amount of landmass (habitat) within each buffer distance. Fahrig (2013) suggests that the habitat amount hypothesis would be supported by species' abundances, species' occurrences, or species richness of equal-area sampling plots (stratified across fragments of varying area and isolation) correlating with the amount of habitat on the surrounding landscape more strongly than with the area of the individual fragments on which the sampling plots are located. This is because landscapes containing less habitat should contain fewer individuals (belonging to fewer species) due to the sample-area effect, meaning fragments within such landscapes will have smaller species pools from which their own diversities are randomly sampled. However, the "appropriate distance" for quantifying the amount of habitat surrounding equal-area sampling plots is undefined and, most problematically, the area of individual fragments on which sampling plots are located becomes increasingly correlated with habitat amount as this distance is reduced. Thus, it may not be possible to decouple fragment area from habitat amount using Fahrig's (2013) proposed method; particularly, for taxa that respond to habitat amount and configuration at fine spatial scales, including butterflies (Thomas & Abery 1995; MacDonald et al. 2017; 2018a; 2019; Saura 2020).

Within our modelling framework, variation in species' abundances, species' occurrences, and species richness associated with full-island surveys and the sample-area effect is nullified in the calculation of random placement residuals. Abundance, occurrence, and richness residuals can therefore be correlated with island/fragment area and the amount of surrounding habitat in a fashion similar to Fahrig's (2013) proposed method of using equal-area sampling plots.

However, because the area of individual islands/fragments is not included in our measures of the amount of surrounding habitat (our buffers are generated from island/fragment edges), the problem of island/fragment area becoming increasingly correlated with habitat amount at fine spatial scales is avoided. After controlling for effects of area *per se*, abundance and occurrence residuals were significantly related to island isolation. Model support significantly declined across increasing isolation buffer sizes for species' abundances, species' occurrences, and species richness, suggesting that that the amount of habitat immediately surrounding islands, rather than the amount of habitat at broader landscape scales, has the greatest effect on butterflies in this system. Importantly, island area and isolation were effectively decoupled, as indicated by the absence of correlation between island area and isolation (e.g., 250 m buffer; $r = 0.082$; $P = 0.668$). Thus, in contrast with mechanisms outlined by the habitat amount hypothesis, isolation effects observed in this study are better attributed to the fact that individual butterflies moving through the open-water matrix are less likely to encounter more isolated islands (Andrén 1994), reducing species' abundances and probabilities of occurrence, as predicted by the theory of island biogeography. This inference is further corroborated by analyses of MacDonald et al. (2018a), which showed that butterfly species turnover among islands of equal areas—a proxy for variation in species pools if islands indeed randomly sample species—was unrelated to Euclidean distance between islands. This result suggests a uniform species pool throughout the study landscape. Again, the degree to which these findings apply to fragmented terrestrial landscapes, wherein the suitability and permeability of matrices may vary, is unclear. For studies addressing fragmented terrestrial landscapes, isolation measures should not only account for the proportion of suitable habitat within various buffer distances, but also include measures of matrix suitability and permeability, if they are available (e.g., MacDonald et al. 2020).

5.5.4 Habitat fragmentation and the narcissus effect

It is important to recognize a bias within this case study, and potentially other study designs addressing spatial patterns of species' abundances, species' occurrences, or species richness across true islands or terrestrial habitat fragments. Here, we investigated effects of area *per se* and isolation on a naturally fragmented landscape of true islands with substantial time-since-isolation (3000 – 4000 YA; Yang and Teller 2005). Therefore, species that are particularly sensitive to fragmentation are unlikely to occur on islands at all. Although our modelling framework resolved significant effects of area *per se* and isolation on butterfly species' abundances and occurrences, effects of area *per se* and isolation on the regional species pool were likely underestimated, as the species assemblage of adjacent continuous habitat was not quantified. This bias may be described as the "narcissus effect", which addresses situations wherein a null model or study design unintentionally accounts for or excludes effects that are of interest (*sensu* Colwell & Winkler 1984). It is possible that this bias contributed to results of Fahrig's (2017) review, where 68% (158/232) of studies addressing single species reported positive fragmentation effects; species that are particularly sensitive to fragmentation may be completely missed in many studies. We therefore suggest caution in interpreting results from study designs that are susceptible to the narcissus effect and encourage future studies to compare the identities and functional traits of species between islands/fragments and adjacent continuous habitat to assess potential biases resulting from the historic exclusion of fragmentation-sensitive species. This may be accomplished using our proposed modelling framework by surveying continuous habitat equal in area to the sum of all surveyed islands/fragments. Abundance data from continuous habitat may then be used in place of total abundances across all islands/fragments (n_i) to calculate abundance, occurrence, and richness random placement values

for each island/fragment using the random placement models described above. Subtracting these random placement values from observed abundance, occurrence, and richness values for each island/fragment will result in abundance, occurrence, and richness residuals that may be used in our linear mixed effects models (abundance and occurrence) and linear models (richness) to resolve whether there are additional effects of area *per se* and isolation on the regional species pool.

5.5.5 Conservation implications

Considerable uncertainty exists in the literature regarding the influence of area *per se* and isolation (i.e., habitat fragmentation) on populations and communities of wildlife (Fahrig 2003; 2013; 2017; Haddad et al. 2015; Hanski 2015; Fletcher et al. 2018). There is an immediate need to resolve this debate, as habitat loss and fragmentation are widespread and increasing (Hanski et al. 2013; Ibisch et al. 2016; Deane et al. 2020; Chase et al. 2020). We demonstrate here that ISAR- and SLOSS-based inferences, founded on emergent patterns of species richness, have the potential to obscure important interspecific variation in responses to area *per se* and isolation. To infer support for the passive sampling, habitat amount, and related hypotheses from emergent patterns of species richness that spuriously conform to predictions of the sample-area effect is to simplistically cut rather than carefully untie the Gordian knot of ecological complexity. We suggest that, in addition to emergent patterns of species richness, information at more reductive and informative levels of organization (e.g., species' abundances and occurrences) should be included in studies aiming to measure and understand effects of habitat fragmentation.

6 Chapter 6: Perceptual range, targeting ability, and visual habitat detection by greater fritillary butterflies *Speyeria cybele* (Lepidoptera: Nymphalidae) and *Speyeria atlantis*

6.1 Abstract

Butterflies are widely invoked as model organisms in studies of metapopulation and dispersal processes. Integral to such investigations are understandings of perceptual range; the maximum distance at which organisms are able to detect patches of suitable habitat. To infer perceptual range, researchers have released butterflies at varying distances from habitat patches and observed their subsequent flight behaviors. It is often assumed that butterflies rely on visual senses for habitat detection; however, this assumption has not been explicitly investigated. Here, we assess the extent and sensory determinants of perceptual range for the great spangled fritillary (*Speyeria cybele* (Fabricius, 1775)) and Atlantis fritillary (*Speyeria atlantis* (W.H. Edwards, 1862)). This was achieved by experimentally releasing butterflies over open water at various distances from a lake island, representing an isolated habitat patch in a dichotomous habitat-matrix landscape. To infer whether butterflies rely on vision for habitat detection, we exposed a subset of butterflies to a series of intense light flashes before release to induce flash blindness (bleaching of photoreceptive rhodopsins) without affecting olfaction. Flashed individuals were 30.1 times less likely to successfully navigate to the target island after release, suggesting butterflies rely primarily on visual senses to navigate fragmented landscapes. For unflashed butterflies, the likelihood of successful navigation decreased by a factor of 2.1 for every 10 m increase in release distance. However, no specific distance threshold for perceptual range was observed. We therefore suggest that perceptual range is best viewed as a continuum of probabilities (targeting ability), reflecting the likelihood of habitat detection across a range of distances.

6.2 Introduction

The movement of organisms between patches of suitable habitat is a principal ecological process contributing to both metapopulation persistence and diversity patterns on fragmented landscapes (Hanski 1998; Wiens 2001; Stevens et al. 2012). Butterflies (Lepidoptera: Papilionoidea), in particular, have proven to be important model organisms in related studies, as their adult movements are easily observable (e.g., Haddad 1999; Dover & Fry 2001; Riva et al. 2018), individuals may be marked for recapture (e.g., Ehrlich & Davidson 1960; Baguette 2003; Nowicki et al. 2014), and patch occupancy may be inferred due to their high detectability and well-documented host plant relationships (e.g., Hanski et al. 1996; Tiple et al. 2011; MacDonald et al. 2017; 2018a; Grant et al. 2018). However, despite a considerable history of study, information is generally lacking on how butterflies actually detect and navigate to patches of suitable habitat while moving through matrices of unsuitable habitat (Baguette & Van Dyck 2007; Schtickzelle et al. 2007). Related investigations are often predicated on estimating butterflies' perceptual range; the maximum distance at which individuals are able to detect patches of suitable habitat using their sensory organs (Wiens 1989).

To estimate perceptual range, butterflies may be released at varying distances from habitat edges, and their flight behaviors subsequently observed. Employing these or related methods, the perceptual ranges of multiple butterfly species have been estimated: the bay checkerspot (*Euphydryas editha bayensis* Sternitzky, 1937 (Lepidoptera: Nymphalidae)) at 50 m (Harrison 1989); the sleepy orange (*Eurema nicippe* (Cramer, 1779) (Lepidoptera: Pieridae)) and cloudless sulphur (*Phoebis sennae* (Linnaeus, 1758) (Lepidoptera: Pieridae)), both at 8 m (Haddad 1999); the Fender's blue (*Icaricia icarioides fenderi* (Macy, 1931) (Lepidoptera: Lycaenidae)), between 10 and 22 m (Schultz et al. 2001); the bog fritillary (*Boloria*

[*Proclossiana*] *eunomia* (Esper, 1800) (Lepidoptera: Nymphalidae)), between 15 and 30 m (Schtickzelle et al. 2007); and the speckled wood (*Pararge aegeria* (Linnaeus, 1758) (Lepidoptera: Nymphalidae)), at either 50 or 100 m, depending on whether individuals originated from fragmented or contiguous landscapes, respectively (Merckx & Van Dyck 2007). Contrasting with these studies, Fahrig & Paloheimo (1987) observed that female cabbage white butterflies (Pieris rapae (Linnaeus, 1758) (Lepidoptera: Pieridae)) did not orient towards patches of their host plants from distances greater than 1 m. While this was interpreted as evidence that visual senses of *P. rapae* are quite limited (Fahrig & Paloheimo 1987), it is unclear whether patches of host plants contrasted visually with the matrix in which butterflies were released, and whether the experiment facilitated use of olfactory senses.

Taken together, results of these studies demonstrate that perceptual ranges of butterflies are both variable and considerable, despite limitations of the compound eye (Rutowski 2003). Indeed, structural properties of butterfly ommatidia suggest that even large objects, several meters high, may not be resolvable at distances greater than 20–30 m (Rutowski 2003). Other senses, namely olfaction, may account for detection of suitable habitat and nectar resources beyond these distances (Cardé & Willis 2008). At finer spatial scales, visual and olfactory senses may work synergistically in larval host plant detection. For example, studies addressing the pipevine swallowtail (*Battus philenor* (Linnaeus, 1771) (Lepidoptera: Papilionidae)) in southeast Texas suggest that, while females identify suitable larval host plants by visual recognition of leaf shapes (Rausher 1978), individuals develop relevant search images by building associations between leaf shapes and appropriate chemical compositions (Papaj 1986). Other investigations of butterflies' senses and their relationships to habitat or resource detection have been largely limited to comparisons of genetic and morphological traits among species, populations, sexes, or

individuals that differ in movement, dispersal, or migratory behaviors (e.g., Hill et al. 1999; Berwaerts et al. 2006; Niitepõld et al. 2009; Altizer et al. 2010; Turlure et al. 2016; reviews in Silberglied 1984; Weiss 2001).

Despite this considerable body of literature, little has been done to experimentally decouple contributions of butterflies' multiple senses to detecting patches of suitable habitat while moving through matrices of unsuitable habitat. Offering some insight, Dover & Fry (2001) simulated suitable habitat corridors (hedgerows) in an agricultural landscape using windbreak materials, and passively observed flight behaviors of passing butterflies, including the scarce copper (*Heodes virgaureae* (Linnaeus, 1758) (Lepidoptera: Lycaenidae)), heath fritillary (*Mellicta athalia* (Rottemburg, 1775) (Lepidoptera: Nymphalidae)), high brown fritillary (*Argynnis* [*Fabriciana*] *adippe* (Dennis & Schiffermüller, 1775) (Lepidoptera: Nymphalidae)), and niobe fritillary (*Argynnis* [*Fabriciana*] *niobe* (Linnaeus, 1758) (Lepidoptera: Nymphalidae)). Their erected structures resembled habitat visually, but not chemically, and were still observed to influence flight patterns. This suggests that butterflies rely at least partially on visual senses to detect suitable habitat. However, individuals were only observed to fly along simulated hedgerows when they were encountered, and specific distances at which butterflies responded to or oriented towards hedgerows were not reported.

In this study, we investigate the extent and sensory determinants of perceptual range for two species of greater fritillary butterflies, the great spangled fritillary (*Speyeria cybele* (Fabricius, 1775) (Lepidoptera: Nymphalidae)) and Atlantis fritillary (*Speyeria atlantis* (W.H. Edwards, 1862) (Lepidoptera: Nymphalidae)), occurring in the Lake of the Woods region, Ontario, Canada. On islands of Lake of the Woods, both *S. cybele* and *S. atlantis* have been observed to consistently avoid open water during flight movements, indicating that they perceive

islands as discrete patches of suitable habitat situated in a matrix of unsuitable habitat (Z. G. MacDonald, unpublished data; and see MacDonald et al. 2018a). Preferred larval host plants of *S. cybele* and *S. atlantis* are *Viola* species. While *Viola* commonly occur on these islands (MacDonald et al. 2018b), we cannot be sure whether host plants exist in sufficient quantities within single islands to sustain isolated populations. Notwithstanding, we define islands as suitable habitat under the functional resource-based concept (*sensu* Dennis et al. 2003), as each contains resources sufficient for mate location, resting, roosting, feeding, and predator escape. Under this habitat concept, the open-water matrix is entirely unsuitable. The high-contrast nature of this relatively dichotomous habitat-matrix system thereby serves as a suitable natural arena for inferring perceptual range via experimental releases. Furthermore, the open-water matrix controls for unwanted matrix heterogeneity that might affect butterfly flight behavior (e.g., Nowicki et al. 2014).

To estimate perceptual range of both *S. cybele* and *S. atlantis*, we released individuals over open water at varying distances from a single island and observed their flight behaviors. To investigate the extent to which butterflies rely on visual senses to detect and navigate to patches of suitable habitat during dispersal movements, we developed a novel method of exposing individuals' photoreceptors to a series of intense light flashes before release. We hypothesized that this method would induce flash blindness through bleaching of photoreceptive pigments (rhodopsins; e.g., Bernard 1983a;b; Briscoe et al. 2003), reducing butterflies' ability to detect and navigate to the target island. Such a result would suggest visual senses are a primary means by which *S. cybele* and *S. atlantis* detect and navigate to patches of suitable habitat while moving through matrices of unsuitable habitat.

6.3 Methods

6.3.1 Study area and experimental design

Our study area was located at the southeast corner of Lake of the Woods, Ontario, Canada. We collected a total of 41 *S. cybele* and 54 *S. atlantis* at three mainland sites within 20 km of Morson, Ontario, between 1 July and 30 July 2016. All collected specimens were judged to be in good condition with minimal wear to wing margins. Collection of specimens was completed between 10:00 and 14:00 on days with ambient temperatures above 20°C, cloud cover less than 75%, and wind speeds below 25 km h−1. Collection sites were located at least 10 km from the lake shore, and were equivalent in habitat composition and structure, comprised of meadows situated within mixed stands of boreal and laurentian tree species (e.g., *Pinus strobus*, *P. banksiana*, *Betula papyrifera*, *Acer spicatum*, *Picea glauca*, and *Tilia americana*). After collection, butterflies were temporarily housed in small, polypropylene containers, kept within a cooler maintained between 20 and 25°C. Collected butterflies were then immediately transported via motorboat to a single island, located at 49.1139° N, −94.2071° E, for experimental release on the same day as collection. This 'target' island is approximately 1.0 ha in area, and was specifically selected for experimental releases because of its approximately circular shape, uniform habitat composition (mixed woodland and shoreline meadow), uniform habitat height (~25 m), and considerable isolation from other landmasses (>300 m). We secured the boat's position at varying distances from the target island's shore (30, 40, 50, or 60 m), using a laser range finder (RX-1200i TBR DNA; Leupold & Stevens, Inc., Beaverton) and a combination of anchors and a stern tie. For all releases, the boat was positioned relative to the target island such that the bearing to the island's center was 90° to the wind direction (Figure 6-1). This effectively controlled for biases that may arise from butterflies being blown towards or away from the island after their release, changing the effective release distance.

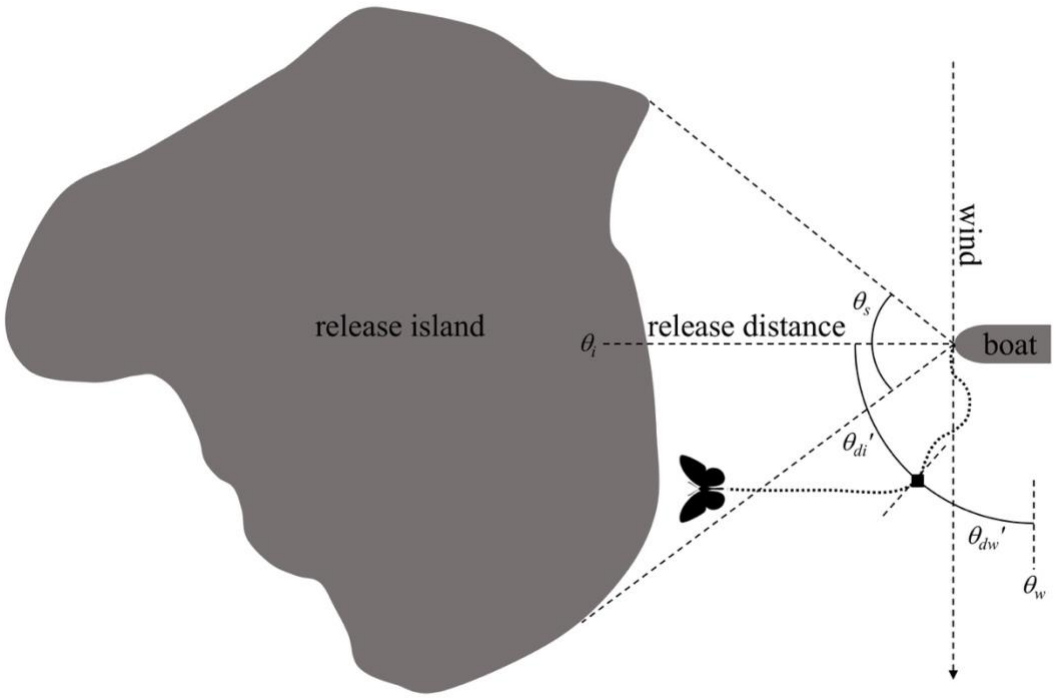

Figure 6-1. Visual representation of experimental releases. The boat was secured at varying distances from the target island's shore (30, 40, 50, or 60 m), such that the bearing to the island's center was 90° to the wind direction. Butterflies were sexed, marked, and released one at a time. For each released individual, flight time and flight orientation at 2.5, 5, and 10 m of travel were recorded. Angular subtense of the target island, θ_s, was estimated as the angular difference between the left and right shore bearings. Deviations in flight orientations from wind direction (θw) and island direction (θi), given by θdw' and θdi', respectively, were estimated at 2.5, 5, and 10 m of travel using eq. 2.

Once the boat was secured at specific distances from the target island, butterflies were sexed, marked, and released one at a time. All releases were completed between 14:00 and 18:00 on days with ambient temperatures above 20°C, cloud cover less than 75%, and wind speeds below 25 km h−1. Each butterfly was released from a 1.5-m aerial net, held away from the edge of the boat towards the target island's center. Keeping as still as possible, the extended net was left open for butterflies to leave at their own will. A second observer recorded the emerging orientation of butterflies, as well as their total flight time and flight orientation after 2.5, 5, and 10 m of travel. Each butterfly was then visually tracked, using a binocular, until it successfully navigated to the target island or flew out of sight (at distances greater than 100 m).

To investigate the role of vision in habitat detection, we sought to devise a method to inhibit butterflies' visual senses without affecting their other senses or flying ability. Painting over, or otherwise covering, butterflies' compound eyes would not achieve this, as this would introduce chemical compounds to sensory areas and add mass to the butterfly itself. To avoid these confounds, we attempted to induce flash blindness by exposing butterflies' photoreceptors to a series of intense light flashes before release. A powerful external photographic flash was used as a flash source (EM-140 DG Macro Flash, guide number = 14 at ISO 100, Sigma Corporation, Kawasaki, Japan). Triggering the flash for 1/6,000 s at a distance of 10 cm produced an estimated 100,352,000 lux (lumens m−2); approximately 1,024 times the intensity of ambient sunlight. Flashing butterflies at 10 cm from the left, right, dorsal, ventral, posterior, and anterior surfaces of their compound eyes, such that the majority of ommatidia were directly exposed to the flash. Butterflies were kept in their polypropylene containers within a dark cooler before flashing to maximize deleterious effects of flashing on rhodopsins (Bernard 1983a;

1983b; Briscoe et al. 2003). Approximately half of *S. cybele* and *S. atlantis* released at each distance were flashed immediately before release.

Flash blindness is caused by bleaching of rhodopsins involved in visual phototransduction within rhabdomeres (Przewłocki et al. 1983). Past work has shown that exposing butterflies' compound eyes to repeated flashes may bleach rhodopsins either temporarily or permanently, depending on the number and intensity of flashes (Bernard 1983a; 1983b; Briscoe et al. 2003). It is also reported that butterfly rhodopsins involved in the detection of ultraviolet, blue, green, and red light are all susceptible to bleaching (Bernard 1983a; 1983b; Briscoe et al. 2003). We did not complete work to assess whether this flash method induced temporary or permanent bleaching of rhodopsins (i.e., temporary or permanent flash blindness). However, if flashed butterflies consistently failed to detect and navigate to the target island from distances at which unflashed butterflies were generally successful, it would be reasonable to conclude that: i) the flash method effectively inhibited butterfly photoreceptor function through bleaching of rhodopsins, resulting in flash blindness; and ii) *Speyeria* spp. rely primarily on visual senses to detect and navigate to patches of suitable habitat while moving through matrices of unsuitable habitat.

6.3.2 Data analyses

A series of generalized linear models (GLMs) were used to assess what environmental and organism-specific variables affected butterflies' probability of successful navigation, flight speed and tortuosity, and flight orientation. All statistical analyses were performed using the statistical software R version 3.4.3 (R Core Team, 2017).

6.3.2.1 Probability of successful navigation

Model 1: Binomial GLMs (logit link) were used to measure the effects of release distance, wind speed, species identity, sex, and visual impairment (flashing) on the probability of successfully navigating to the target island (success/failure). Interactions between relevant experimental variables were assessed using first-degree interaction terms. Sunlight (direct vs diffuse), percentage cloud cover (estimated as 1 – the proportion of blue sky visible to observers), ambient temperature, butterfly collection location, day of release, and time in captivity were included as noise variables. All experimental variables were fitted regardless of their significance, with relevant interaction terms and noise variables included in the final model only if significant. Standardized coefficients were estimated for all continuous variables. To interpret the effects of experimental variables on the likelihood of navigation success, odds ratios were estimated using original units of experimental variables to permit straight-forward interpretation.

Model 2: Perceptual ranges of butterflies are most often reported as a single distance measures, irrespective of patch size or habitat characteristics (e.g., Harrison 1989; Schultz et al. 2001; Merckx & Van Dyck 2007; Schtickzelle et al. 2007). However, analyses of butterflies' ommatidial structures suggest that the sizes of objects may determine the maximum distances at which they are detectable (i.e., single object thresholds; Rutowski 2003). Therefore, the ability to detect habitat patches may decrease with increasing distance simply because of decreasing angular subtense, rather than increasing distance *per se*. To decouple these variables, we estimated the angular subtense of the target island for each experimental release as the angular difference between the left and right shore bearing (Figure 2-1). Although the target island was approximately circular in shape, angular subtense still varied independent of release distance,

depending on the location of the release boat (i.e., the direction from which the island was viewed). As would be expected, angular subtense and release distance were negatively correlated ($r = -0.775$). This strong correlation limited our ability to partition variance in the probability of successful navigation between angular subtense and release distance using residual or multiple regression techniques (Freckleton 2002). However, if sizes of habitat patches or islands determine the maximum distances at which they are detectable, a competing model accounting for angular subtense should explain more variation in navigation successes than a model accounting for release distance. We tested this hypothesis by substituting the target island's angular subtense into the previous binomial GLM built using release distance. The significance of the two variables, as well as McFadden's pseudo R2, was compared between the two competing models.

6.3.2.2 Flight speed and tortuosity

Models 3 and 4: GLMs were next used to measure the effects of release distance, wind speed, species identity, sex, and flashing on: i) flight speed, estimated as total flight time after 10 m of travel; and ii) flight tortuosity, estimated as the standard deviation of turn angles between first emergence, 2.5, 5, and 10 m of travel. Calculating the standard deviation of turn angles is non-trivial, since bearings wrap from 359° around to 0° (Batschelet 1981). Therefore, to estimate turn angles (θ_d'), we standardized flight orientations at each distance (2.5, 5, and 10 m; θ_d) relative to flight orientations at the previous distance (emergence, 2.5, and 5 m; θ_{d-1}), using the following conditional equation (eq.1):

$$\theta'_d = \begin{cases} \theta_d - \theta_{d-1}, & |\theta_d - \theta_{d-1}| \leq 180° \\ \theta_d - \theta_{d-1} + 360°, & \theta_d - \theta_{d-1} < -180° \\ \theta_d - \theta_{d-1} - 360°, & \theta_d - \theta_{d-1} > 180° \end{cases}$$

This equation produces reliable turn angle estimates, so long as absolute differences in sequential flight orientations are less than 180°, which they were in all instances. Within flight speed and flight tortuosity GLMs, experimental variables were fit regardless of their significance, with relevant interaction terms and noise variables included only if significant. Success/failure of navigation to the target island was fit as a binary covariate in both models, to assess whether flight speed and tortuosity varied between butterflies that were successful and unsuccessful in navigating to the target island. Total flight time after 10 m of travel and standard deviations of turn angles both took on positive continuous values that were best fit using a gamma distribution (log link).

6.3.2.3 Determinants of flight orientation

Models 5 and 6: Perceptual range is often inferred by determining the maximum distance at which the proportion of released butterflies orienting towards habitat significantly differs from what is expected under random flight orientations (e.g., Fahrig & Paloheimo 1987; Schtickzelle et al. 2007). Specifically, this random flight null assumption assumes that the proportion of butterflies failing to detect the a nearby habitat patch, but still flying towards it, will be proportional to the angular subtense of the patch divided by 360°. However, contrasting with this null assumption, wind direction appeared to determine initial flight orientations for the majority of released butterflies in our study, independent of whether butterflies detected and navigated to the target island or not. If wind direction indeed determined the flight orientations of the majority

released butterflies, the proportion of butterflies failing to detect a habitat patch, but still flying towards it, will be less than what is predicted by the random flight null assumption.

As a corollary of these relationships, we expect that, for butterflies that successfully detected and navigated to the target island (hereafter, "successful butterflies"), deviations in flight orientations from wind direction should increase with distance flown, while deviations in flight orientations from island direction should decrease with distance flown. Such relationships correspond to reorientation away from the wind direction, towards the target island. This reorientation is not predicted for butterflies that were unsuccessful in detecting and navigating to the target island ("unsuccessful butterflies"), and deviations in flight orientations from wind direction and island direction should not vary with distance flown. To build statistical models to test these predictions, we first estimated: i) absolute deviations in flight orientations at 2.5, 5, and 10 m (θ_d) from wind direction (θ_w), given by θ_{dw}'; and ii) absolute deviations in flight orientations at 2.5, 5, and 10 m (θ_d) from the bearing to the centre of the target island (θ_i), given by θ_{di}'. This was achieved using the following conditional equation (eq. 2):

$$\theta'_{dx} = \begin{cases} |\theta_d - \theta_x|, & |\theta_d - \theta_x| \leq 180° \\ |\theta_d - \theta_x + 360°|, & \theta_d - \theta_x < -180° \\ |\theta_d - \theta_x - 360°|, & \theta_d - \theta_x > 180° \end{cases}$$

where θ_w or θ_i takes the place of θ_x and θ_{dw}' or θ_{di}' takes the place of θ_{dx}' for estimating deviations in flight orientations from wind direction or island direction, respectively. Two generalized linear mixed models (GLMMs) were used to assess whether: i) deviations in flight orientations from wind direction increased with distance flown; and ii) deviations in flight

orientations from island direction decreased with distance flown. Dependent variables used in these GLMMs were: i) absolute deviations in flight orientations from wind direction; and ii) absolute deviations in flight orientations from island direction. Experimental variables in both GLMMs included distance flown (2.5, 5, and 10 m), wind speed, and success/failure of navigation to the target island. Release ID was treated as the random effect within GLMMs to control for lack of independence between successive flight orientations of individuals. An interaction term between distance flown and success/failure was used to assess whether relationships between deviations in flight orientations and distance flown differed between successful and unsuccessful butterflies. Flashed butterflies were not included within GLMMs to avoid introducing unwanted noise in flight orientations. Tweedie distributions (log link) were used to accommodate non-negative continuous response variables and right skew (Dunn & Smyth 2005).

Within the first GLMM, a significant positive interaction between distance flown and success would indicate that deviations in flight orientations from wind direction increased with distance flown for successful butterflies. Within the second GLMM, a significant negative interaction between distance flown and success would indicate deviations in flight orientations from island direction decreased with distance flown for successful butterflies. Nonsignificant main effects of flight distance in both models would suggest that these relationships were only observed for successful butterflies, that is, deviations in flight orientations from wind direction and island direction were unrelated to distance flown for unsuccessful butterflies. Together, these results would indicate that instances of successful navigation generally involved a reorientation away from wind direction and towards island direction after release, questioning the validity of the random flight null assumption.

6.4 Results

6.4.1 Probability of successful navigation

For both *S. cybele* and *S. atlantis*, the proportion of unflashed butterflies successfully navigating to the target island generally decreased with increasing release distance (Figure 6-2). At the maximum release distance of 60 m, 50.0% of unflashed *S. cybele* were successful (Table 6-1). At 50, 40, and 30 m, 54.5, 85.7, and 80.0% of unflashed *S. cybele* were successful. At 60 m, no unflashed *S. atlantis* were successful. This increased to 16.7% at 50 m, and to 50.0% at both 40 and 30 m. Flashing substantially reduced percentages of successful navigation for both species at all distances. Considering all releases at all distances, only 11.1% of flashed *S. cybele* and no flashed *S. atlantis* were successful in navigating to the target island. This contrasts with the 66.6% of unflashed *S. cybele* and 33.3% of unflashed S. atlantis that were successful in navigating to the target island overall.

Table 6-1. The number of *Speyeria cybele* and *S. atlantis* released at 30, 40, 50, and 60 m that were successful and unsuccessful in navigating to the target island. A subset of butterflies were exposed to a series of intense flashes immediately before release. This method induced flash blindness through bleaching of photoreceptive rhodopsins, without affecting olfaction.

		Speyeria cybele		*Speyeria atlantis*	
Release distance (m)	Navigation to island	Not flashed	Flashed	Not flashed	Flashed
30	successful:	4	1	4	0
	unsuccessful:	1	4	4	6
40	successful:	6	1	3	0
	unsuccessful:	1	8	3	6
50	successful:	6	1	1	0
	unsuccessful:	5	8	5	3
60	successful:	2	0	0	0
	unsuccessful:	2	4	4	2

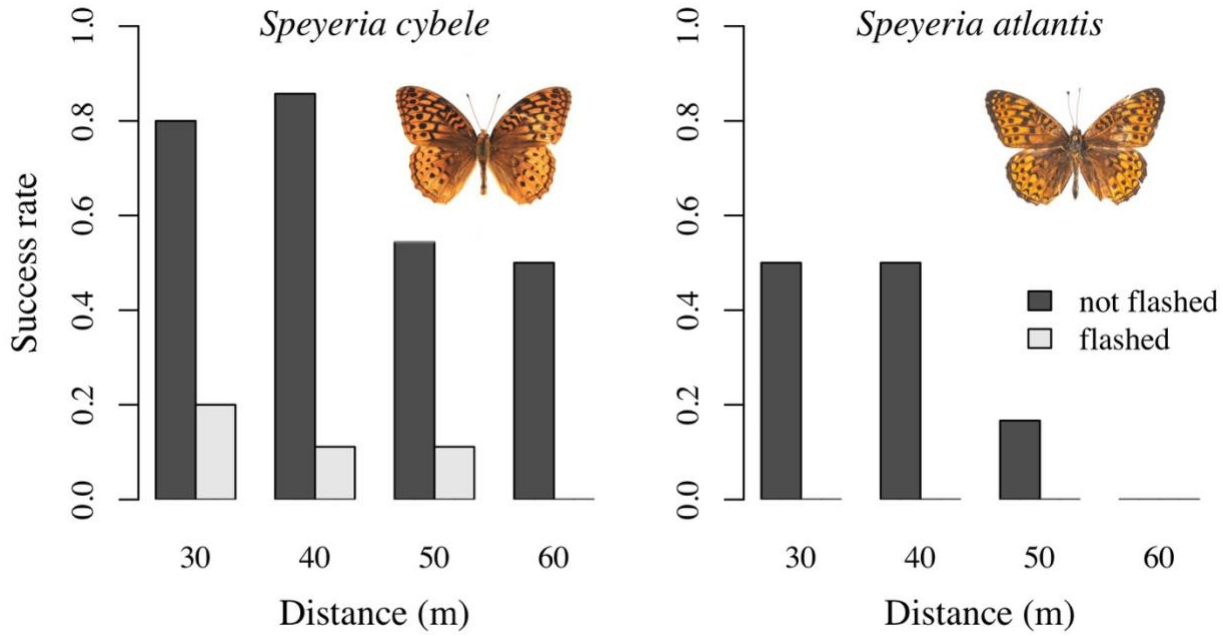

Figure 6-2. The proportion (success rate) of *Speyeria cybele* and *Speyeria atlantis* that successfully navigated to the target island after experimental release at 30, 40, 50, and 60 m. Flashed butterflies were exposed to a series of intense flashes immediately before release. This method induced flash blindness through bleaching of photoreceptive rhodopsins, without affecting olfaction. Reduced success rates of flashed butterflies indicate that butterflies rely primarily on visual senses to detect and navigate to suitable habitat patches during interpatch and dispersal movements.

The first binomial GLM accounting for probability of successful navigation to the target island corroborated these relationships (Model 1, Table 6-2). Release distance had a significant negative effect on the probability of successful navigation. An odds ratio of 0.93 (95% CI: 0.87, 0.98) indicates that the likelihood of successful navigation decreased by a factor of 1.08 for every 1 m increase in release distance, or by a factor of 2.15 for every 10 m increase in release distance. The effect of an interaction between species and distance was non-significant, suggesting that decreases in the likelihood of successful navigation associated with increases in release distance were approximately equivalent between species. However, *S. cybele* had a significantly higher probability of successful navigation than *S. atlantis* overall. An odds ratio of 2.48 (95% CI: 0.76, 9.13) indicates that released *S. cybele* were 2.48 times more likely to successfully navigate to the target island than *S. atlantis*. No significant difference between sexes was observed. Flashing butterflies had a significant negative effect on the probability of successful navigation. An odds ratio of 0.03 (95% CI: 0.01, 0.13) indicates that flashed butterflies were 30.13 times less likely to successfully navigate to the target island than unflashed butterflies. Substituting angular subtense of the target island into the binomial GLM

accounting for release distance reduced model fit (McFadden's pseudo R^2: Model 1 = 0.36; Model 2 = 0.31), and angular subtense was not significantly related to the probability of successful navigation (Model 2, Table 6-2).

Table 6-2. Summary of generalized linear model (GLM) and generalized linear mixed model (GLMM) results. Coefficient estimates and their corresponding *P*-values are given in parentheses. Coefficient estimates for continuous variables (release distance, wind speed, angular subtense, and distance flown) are standardized. Models 1 and 2 (GLMs) were fitted using a binomial distribution with a logit link function. Navigation success was measured as the success or failure of navigation to the target island. Models 3 and 4 (GLMs) were fit using a gamma distribution with a log link function. Flight speed was measured as total flight time to 10 m. Flight tortuosity was measured as the standard deviation of turn angles. Models 5 and 6 (GLMMs) were fitted using a Tweedie distribution with a log link function. Deviation from wind direction and deviation from island direction were measured as absolute deviations in flight orientations from the wind bearing and the bearing towards the target island's centre, respectively, after 2.5, 5, and 10 m of flight. Individual ID was treated as a random effect to account for lack of independence between successive flight orientations of individuals.

Dependent variable	Experimental variables					
	release distance[1]	wind speed[2]	species[3]	sex[4]	flashing[5]	
Mod. 1: navigation success	-0.777*	0.024	2.117**	0.909	-3.405***	
	angular subtense[6]	wind speed	species	sex	flashing	
Mod. 2: navigation success	0.378	-0.136	1.750**	0.78	-3.132***	
	release distance	wind speed	species	sex	flashing	success[7]
Mod. 3: flight speed	0.049'	-0.200***	-0.041	0.071	0.226***	0.111
Mod. 4: flight tortuosity	0.066	-0.522***	0.405'	-0.177	0.540*	0.524*
	distance flown	wind speed	species	sex	success	distance flown × success
Mod. 5: deviation from wind direction	-0.025	-0.523***	0.155	-0.115	0.537*	0.190'
Mod. 6: deviation from island direction	0.017	0.158**	0.054	-0.129	-0.429***	-0.176 ***

[1]continuous: distance at which butterflies were released from the island (m); [2]continuous: wind speed at time of release (kmph) ; [3]categorical: *Speyeria cybele* or *S. atlantis*, with *S. atlantis* treated as the reference category; [4]categorical: male or female, with females treated as the reference category; [5]categorical: exposure to a series of flashes to induce flash blindness (see 6.2 Methods), with flashing treated as the reference category; [6]continuous: difference between the left and right shore bearings of the target island estimated for each release; [7]categorical: eventual success/failure of navigation to the target island, with failure treated as the reference category; [8]continuous: distance travelled by butterflies from the release boat after release (2.5, 5, or 10 m). Significance is denoted by: ' $P < 0.10$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

6.4.2 Flight speed and tortuosity

Wind speed was observed to have a significant negative effect on flight time after 10 m, indicating that higher wind speeds were generally associated with faster movement (Model 3,

Table 6-2). Flashing butterflies had a significant positive effect on flight time to 10 m, indicating that flashed butterflies generally flew slower than unflashed butterflies. Flight time to 10 m was not observed to significantly relate to release distance, species, sex, or eventual success/failure of navigation to the target island.

Flight tortuosity, measured as the standard deviation of turn angles for each individual, significantly decreased with increasing wind speed (Model 4, Table 6-2). Flashing had a significant positive effect on flight tortuosity. These results, in combination with those of the flight speed model (Model 3), suggest that: i) higher wind speeds were generally associated with faster, less tortuous flights; and ii) flashed butterflies exhibited slower, more tortuous flights than unflashed butterflies. The positive relationship between success and flight tortuosity (Model 4) suggests that flight paths of successful butterflies were significantly more tortuous than those of unsuccessful butterflies. This is consistent with the hypothesis that instances of successful navigation generally involved reorientation from wind direction to island direction, increasing observed tortuosity. *S. cybele* may have exhibited more tortuous flights on average than *S. atlantis*, but this effect was only significant at $\alpha = 0.10$ ($P = 0.058$). Flight tortuosity did not significantly vary with release distance or sex. The effects of relevant interactions and noise variables in Models 3 and 4 were non-significant.

6.4.3 Determinants of flight orientation

GLMMs accounting for absolute deviations in flight orientations from wind direction (Model 5; Table 6-2) and island direction (Model 6; Table 6-2) each indicated that both wind speed and success were significantly related to flight orientations. Wind speed was observed to have a significant negative effect on absolute deviations in flight orientations from wind direction, and a significant positive effect on absolute deviations in flight orientations from

island direction. Flight orientations of successful butterflies were significantly nearer to island direction and further from wind direction than unsuccessful butterflies.

Within the GLMM accounting for absolute deviations in flight orientations from wind direction (Model 5), the positive interaction between distance flown and success suggests that deviations in flight orientations from wind direction increased with distance flown for successful butterflies. This interaction was, however, only significant at $\alpha = 0.10$ ($P = 0.068$). Within the GLMM accounting for deviations in flight orientations from island direction (Model 6), the negative effect of the interaction between distance flown and success indicated that deviations in flight orientations from island direction decreased with distance flown for successful butterflies. The main effects of distance flown on absolute deviations in flight orientations from both wind direction and island direction were near zero and non-significant in Models 5 and 6. Collectively, these results indicate that successive flight orientations of successful butterflies, but not unsuccessful butterflies, tended away from the wind direction and towards the island direction as butterflies travelled further from the release location. Therefore, instances of successful navigation generally involved a reorientation from wind direction to island direction after release. Flight orientations of unsuccessful butterflies were almost always aligned with the wind direction until they drifted out sight.

6.5 Discussion

6.5.1 Quantifying perceptual range

As expected, the probability of *S. cybele* and *S. atlantis* successfully detecting and navigating to the target island significantly decreased with increasing release distance. The effect of release distance on probability of success was substantial, with the likelihood of butterflies successfully navigating to the target island decreasing by a factor of 2.15 for every 10 m increase

in release distance. Similar observations have led many ecologists to infer the existence of maximum distance thresholds, beyond which, butterflies are unable to detect habitat patches or habitat features using their sensory organs. Based on such thresholds, Schtickzelle et al. (2007) and Dover & Settele (2009) suggest that a distinction between "apparent fragmentation" and "functional fragmentation" may be meaningful. While apparent fragmentation may describe any landscape with discrete habitat patches, functional fragmentation is reserved for landscapes wherein inter-patch distances exceed the perceptual ranges of focal taxa (Dover & Settele 2009).

In the context of this framework, single distance measures of perceptual range are appealing due to the relative simplicity of applying specific thresholds of patch isolation to infer whether landscapes are functionally fragmented—ecologists need only estimate single distance measures of perceptual range for focal taxa. For butterflies, in particular, a straight-forward method has been to determine the maximum distance at which the proportion of released individuals orienting towards habitat significantly differs from what is expected given random flight orientations (i.e., a random flight null assumption; e.g., Fahrig & Paloheimo 1987; Merckx & Van Dyck 2007; Schtickzelle et al. 2007). However, in this study, GLMMs (Models 5 and 6) suggested that instances of successful navigation generally involved a reorientation from wind direction (not random direction) to island direction after release. For unsuccessful butterflies, the mean of absolute deviations in flight orientations from wind direction after 2.5, 5, and 10 m of flight were 19.4°, 18.6°, and 18.0°, respectively. These deviations were less than 90° in all instances; far from 180°, the expected mean associated with random flight orientations. Therefore, if experimental releases are conducted with wind direction perpendicular to the direction of the target habitat patch (as would be recommended), the proportion of butterflies failing to detect habitat patches, but still orienting towards them, will be lower than what is

predicted by random flight orientations. These results suggest that the random flight null assumption is biased towards overestimating null proportions of butterflies successfully orienting towards adjacent habitat patches, and is therefore inappropriate for determining thresholds of perceptual range.

Qualitative observations of flight behaviour support this conclusion. In all instances of success navigation, we observed a surprisingly punctuated shift in flight behaviour, from a "fluttering" flight averaging with the wind direction, to a "directed" flight towards the target island. We interpret this change in flight behaviour as meaningful perception of, and reorientation towards, the target island based on several observations. Almost all individuals that did not successfully navigate to the target island vanished from sight following the wind direction. We did not observe a single instance wherein an individual exhibited a directed flight toward the target island, but failed to successfully navigate to it. Furthermore, there were very few instances (3 flashed *S. cybele*) wherein an individual maintained a fluttering flight in the direction of the target island, until it arrived at the island's shore, without adopting a directed flight. We therefore find it reasonable to conclude that: i) most, if not all, instances of failed to navigation to the target island represented failure to detect the target island; and ii) most instances of successful navigation to the target island represented meaningful detection of the target island. It is also reasonable to infer that unsuccessful and successful butterflies were similarly searching for suitable habitat, and that butterflies drifting with the wind direction were not simply exhibiting an escape response. While flight paths of successful butterflies were significantly more tortuous than unsuccessful butterflies (Model 4), this difference was caused by the consistent reorientation of successful butterflies from wind direction to island direction following detection of the target island (Models 5 and 6). Flight patterns of successful and

unsuccessful butterflies were generally indecipherable before successful butterflies detected and reoriented towards the target island.

While the probability of *S. cybele* and *S. atlantis* successfully detecting and navigating to the target island significantly decreased with increasing release distance, no obvious distance threshold was observed for either species. We see little reason to infer that single distance measures (i.e., thresholds) of perceptual range are ecological meaningful *a priori*. While ommatidial structure suggests the existence of single object thresholds (*sensu* Rutowski 2002), the probability of a dispersing butterfly detecting a nearby habitat patch is subject to a plethora of factors unique to landscapes, individuals, and environmental conditions. Indeed, the level of visual contrast between habitat patches and the matrix (*sensu* Rutowski 2002), the evolutionary history of individuals (e.g., Merckx & Van Dyck 2007), the perceived suitability of the matrix (e.g., Nowicki et al. 2014), and the wind speed and direction (this study) represent but a few factors that warrant continued investigation. Accounting for factors unique to landscapes, individuals, and environmental conditions (as in multiple logistic regression), perceptual range may be best viewed as a continuum of conditional probabilities, reflecting the likelihood that butterflies will detect habitat patches across a range of distances, rather than a single distance measure *per se*. This approach has the added benefit of permitting quantitative comparisons of probabilities of detecting habitat patches at distances below perceptual range thresholds, should they be found to exist in future research. To avoid confusion of terms in the literature, we suggest this concept may be referred to as "targeting ability."

6.5.2 Targeting ability and habitat fragmentation

In contrast with single distance measures of perceptual range, the concept of targeting ability does not evoke punctuated distinctions between apparent and functional fragmentation

(*sensu* Schtickzelle et al. 2007; Dover & Settele 2009). We find there is little evidence to suggest that this dichotomous distinction, predicated on thresholds of patch isolation and perceptual range, is ecologically meaningful. For example, the degree of asynchrony between subpopulation dynamics in metapopulations have been shown, both theoretically and empirically, to vary continuously with patch isolation (review in Hanski 1999). Furthermore, within fragmented landscapes, dispersal ranges of butterflies are commonly observed as 10 – 1000 fold greater than the greatest distance estimates of perceptual range (e.g., 37 km for the cranberry fritillary [*Boloria aquilonaris*, Baguette 2002]; see Introduction for review of perceptual range estimates). During dispersal events, the probability of a butterflies encountering habitat patches is well-approximated by a variety of functions, such as negative exponential or inverse power, where increasing patch isolation has continuous, rather than threshold, effects on the probability of patch colonization (Hanski et al. 2000; Baguette 2002; Nowicki et al. 2014). Even when inter-patch distances and movements do not exceed estimated perceptual range thresholds (i.e., short-range dispersal), organisms are still likely to experience increased mortality risk or deferred costs when moving between patches. Thus, punctuated distinctions between apparent and functional fragmentation may bear little resemblance to ecological patterns and processes on many fragmented landscapes.

Of greater ecological relevance, Baguette & Van Dyck (2007) advance a conceptual distinction between different perspectives of landscape connectivity, "structural" and "functional," without emphasizing specific thresholds of patch isolation in relation to perceptual range. Within this framework, structural connectivity addresses the spatial configuration of habitat patches and landscape elements, such as the vicinity and presence of barriers, while functional connectivity addresses how landscape structure affects behaviours of dispersing

individuals. In other words, functional connectivity contributes to the concept of structural connectivity by accounting for perceptual grain; the smallest spatial scale at which organisms perceive landscape heterogeneity (Weins 1989). Perceptual grain is most often inferred via estimates of single distance (threshold) measures of perceptual range (Baguette & Van Dyck; 2007). Including addition facets of habitat detection associated with the concept of targeting ability (e.g., factors unique to landscapes, individuals, and environmental conditions) may further the instructive power of the functional connectivity heuristic.

6.5.3 Targeting ability of *Speyeria cybele* and *S. atlantis*

Despite considerable overlap in their evolutionary and life histories (Hall et al. 2014; Acorn & Sheldon 2017), a significant difference in targeting ability was observed between *S. cybele* and *S. atlantis*. Overall, *S. cybele* were 2.48 times more likely to successfully navigate to the target island than *S. atlantis*. A noted difference in compound eye structure between the two species is colour, with compound eyes of live *S. cybele* and *S. atlantis* appearing brown and grey, respectively (Acorn & Sheldon 2017). This difference may be attributed to variation in the composition of screening pigments, which filter light passing both onto photoreceptive rhodopsins and between separate ommatidia (Stavenga 2002). Red screening pigments of the Japanese yellow swallowtail butterfly (*Papilio xuthus*) are inferred to act as short-wavelength absorbance filters, facilitating long-wave sensitivity of rhabdomeres (Arikawa et al. 1999). Via this mechanism, dark-orange screening pigments of the monarch butterfly (*Danaus plexippus*) have been shown to contribute to color discrimination in the long-wavelength range (Blackiston et al. 2011). Together, these studies suggest a link between screening pigment composition visual sensitivity under various light conditions. However, additional research will be required to

resolve whether interspecific variation in screening pigment composition among *Speyeria* species relates to variation in visual targeting ability.

Of greater interest to this study are relationships between interspecific variation in targeting ability and functional traits known to relate to dispersal, such as wingspan and estimates of mobility (Burke et al. 2011; Stevens et al. 2012). Wingspans of *S. cybele* and *S. atlantis* in Ontario have been measured at 70 – 100 and 55 – 70 mm, respectively (Acorn & Sheldon 2017). Burke et al. (2011) have estimated the mobility of *S. cybele* and *S. atlantis* at 7.10 and 7.00, respectively, using a qualitative index (based on expert opinion) ranging from 0 to 10. In accordance with positive interspecific relationships between wingspan, mobility, and dispersal ability of butterflies (e.g., Stevens et al. 2012), the larger and more mobile of the two species, *S. cybele*, had significantly greater targeting ability than did the smaller and less mobile of the two species, *S. atlantis*. In the context of the functional connectivity heuristic (*sensu* Baguette & Van Dyck 2007), inter-island movements and dispersal are likely to be less costly, both in terms of mortality risk and deferred costs, for *S. cybele* than *S. atlantis*. *Speyeria cybele* may therefore have greater a greater propensity and ability to navigate fragmented landscapes than *S. atlantis*. However, this inference is drawn from a single comparison of two congeneric species. More comprehensive studies, addressing disparity in targeting abilities across a greater number of species, is required to appropriately evaluate the hypothesis that targeting ability is a practical measure of the degree to which organisms perceive landscapes as fragmented. It is also worth noting that this study was completed in a landscape of extreme habitat-matrix contrast. Comparisons of related studies addressing terrestrial landscapes of greater complexity would be valuable for understanding how organisms perceive fragmented landscapes with lower habitat-matrix contrast (e.g., for a comparisons of diversity patterns on a true-island archipelago and an

anthropogenically fragmented landscape, see Mendenhall et al. 2014; or butterfly movement through different scales of linear forest fragmentation, see Riva et al. 2018).

6.5.4 Determinants of targeting ability

Given the prominence of morphological traits associated visual senses, butterflies have long been hypothesized, and even assumed, to rely primarily on vision to detect and navigate to habitat patches during dispersal movements (Silberglied 1984; Rutowski 2002; Turlure et al. 2016). However, to the best of our knowledge, this hypothesis has evaded explicit empirical investigation using experimental techniques. Thus, an interesting finding of this study is that repeated exposure to an intense flash significantly reduced the ability of both *S. cybele* and *S. atlantis* to detect and navigate to suitable habitat from a range of distances. This effect of flashing was substantial, with flashed individuals 30.13 times less likely to successfully navigate to the target island than unflashed individuals. The proportion of flashed butterflies successfully navigating to the target island was near zero or zero at all distances. Given these findings, we infer that visual senses of *S. cybele* and *S. atlantis* play a primary role in navigating fragmented landscapes when visual habitat-matrix contrast is high. However, we cannot rule out that olfaction may be used synergistically with vision, as demonstrated for long-range detection of nectar resources (Cardé & Willis 2008) and identification of larval host plants (Rausher 1981; Papaj 1986; Garlick 2007; Kinoshita et al. 2015). Notwithstanding, results of our study support the long-held assumption that visual senses are a primary means by which the butterflies detect and navigate to patches of suitable habitat while moving through matrices of unsuitable habitat.

In light of this conclusion, it is both unexpected and interesting that angular subtense of the target island did not explain more variation in navigation success than distance *per se*. If visual senses do indeed account for long-range habitat detection in butterflies, the apparent sizes

of habitat patches (angular subtense) should relate to their probability detection (i.e., single object thresholds; *sensu* Rutowski 2002). However, GLMs indicated that probability of successful navigation was not significantly related to angular subtense, despite the fact that angular subtense was strongly correlated with release distance. Taken at face value, this finding suggests that: i) there are intrinsic effects of distance *per se* on butterflies' ability to detect habitat patches; and ii) perceptual range and targeting ability may not vary with patch size. However, this latter conclusion contrasts with the common and reasonable assumption that patches or islands of larger areas present larger dispersal targets (*sensu* Wilson & MacArthur 1967; Hanski 1999). It is worth noting here that the target island was approximately circular in shape, meaning angular subtense did not vary substantially independent of release distance in this study. We therefore question whether the conclusion, that perceptual range and targeting ability may not vary with patch size, is meaningful. Relationships between patch size and patch detectability, and their relevance to the dispersal process, requires further investigation.

A superior assessment of relationships between patch size and patch detectability would empirically determine, across a range of release distances, variation in the probability of butterflies detecting habitat patches that vary substantially in area and thus angular subtense. Including habitat patches that also vary in habitat height would permit two-dimensional estimations of angular subtense, deepening inferences that may be drawn. Measures should be taken to quantify relative levels of visual contrast between habitat patches and their immediate surroundings if patch or matrix compositions are heterogeneous. As butterflies are inferred to have colour vision (Silberglied; 1984; Kinoshita; 1999; Arikawa; 2003; Blackiston et al. 2011), quantifying visual contrast across a variety of wavelengths may permit decoupling of specific visual ques used by butterflies to detect habitat or resource patches in heterogeneous landscapes.

7 Chapter 7: Gene flow and climate-associated genetic variation in a vagile habitat specialist

7.1 Abstract

Previous work in landscape genetics suggests that geographic isolation is of greater importance to genetic divergence than variation in environmental conditions. This is intuitive when configurations of suitable habitat are a dominant factor limiting dispersal and gene flow, but has not been thoroughly examined for habitat specialists with strong dispersal capability. Here, we evaluate the effects of geographic and environmental isolation on genetic divergence for a vagile invertebrate with high habitat specificity and a discrete dispersal life stage: Dod's Old World swallowtail butterfly, *Papilio machaon dodi*. In Canada, *P. m. dodi* are generally restricted to eroding habitat along major river valleys where their larval host plant occurs. A series of causal and linear mixed effects models indicate that divergence of genome-wide single nucleotide polymorphisms is best explained by a combination of environmental isolation (variation in summer temperatures) and geographic isolation (Euclidean distance). Interestingly, least-cost path and circuit distances through a resistance surface parameterized as the inverse of habitat suitability were not supported. This suggests that, although habitat associations of many butterflies are specific due to reproductive requirements, habitat suitability and landscape permeability are not equivalent concepts due to considerable adult vagility. We infer that divergent selection related to variation in summer temperatures has produced two genetic clusters within *P. m. dodi*, differing in voltinism and diapause propensity. Within the next century, temperatures are predicted to rise by amounts greater than the present-day difference between regions of the genetic clusters, potentially affecting the persistence of the northern cluster under continued climate change.

7.2 Introduction

A principal aim in ecology and evolutionary biology is to resolve factors and understand processes that influence genetic divergence at both the individual and population level (Mayr 1963; Coyne & Orr 2004; Nosil 2012). When genetic divergence has a strong spatial component, causes are generally attributed to spatial variation in evolutionary processes, such as gene flow, genetic drift, and selection (Slatkin 1987; Rousset 1997; Bohonak 1999; Schwartz et al. 2010). However, inferring the relative contributions of these processes is challenging. Landscape genetics addresses this by quantitatively relating patterns of genetic divergence to geographic and environmental landscape factors (Sork et al. 1999; Cushman et al. 2006; Shirk et al. 2010; Richardson et al. 2016).

Multiple heuristics have been invoked to conceptualize relationships between genetic divergence and landscape factors, with each implicating specific evolutionary processes. The first, isolation-by-distance (IBD; Wright 1943), predicts that geographic distance or physical barriers to dispersal reduce gene flow and permit drift between spatially separated individuals or populations. Because dispersal is limited in most species (Greenwood 1980), Euclidean and genetic distances are often positively correlated, supporting IBD (Rousset 1997; Vekemans & Hardy 2004; but see Meirmans, 2012). A second heuristic, isolation-by-resistance (IBR; McRae 2006), may be seen as a modification of IBD and predicts that patterns of genetic divergence will be best explained by geographic distances accounting for variation in the relative resistance organisms experience when dispersing through heterogeneous landscapes. To assess IBR, resistance surfaces are often parameterized as the inverse of habitat suitability, generally modelled using occurrences of focal taxa and geographic and environmental/ecological predictor

variables (McRae & Beier 2007; Wang et al. 2008; Storfer et al. 2010; Wang et al. 2012a; but see Peterman et al. 2014). Optimal pathways (e.g., least-cost path distances) or a multitude of possible pathways of varying probability derived from circuit theory (circuit distances) may then be estimated across resistance surfaces and related to patterns of genetic divergence (McRae & Beier 2007). When resistance surfaces are parametrized in this way, positive correlations between genetic distance and resistance-based distances suggest that organisms are more likely to disperse within suitable habitat and experience high movement resistance/cost within unsuitable habitat. IBR thereby equates the concept of habitat suitability to that of landscape permeability (the ease with which organisms move through a landscape), or in circuit theory terms, conductance. A third heuristic, isolation-by-environment (IBE; Wang & Summers 2010), predicts that spatial variation in environmental/ecological conditions contributes to genetic divergence via the combination of a) reduced fitness and negative selection on individuals that have dispersed across environmental gradients, b) reduced fitness and negative selection on dispersers' offspring in non-natal habitats (outbreeding depression), or c) reduced tendency of individuals to disperse due to local adaptation to environmental conditions (Dobzhansky 1937; Nosil 2004; 2012; Nosil et al. 2005; Crispo et al. 2006; Lee & Mitchell-Olds 2011; Wang & Bradburd 2014). After controlling for geographic distance, positive correlations between genetic distance and differences in environmental/ecological conditions confer support for IBE.

Although often framed as competing hypotheses, complementarity of IBD, IBR, and IBE has been documented in multiple studies (Coyne & Orr 2004; Thorpe et al. 2008; Crispo et al. 2006; Wang et al. 2012a; Sánchez-Ramírez et al. 2018; Van Buskirk & van Rensburg 2020). Within such investigations, it can be instructive to invoke Euclidean, least-cost path, and circuit distances as measures of geographic isolation (IBD + IBR) for contrast with measures of

environmental/ecological isolation (IBE) estimated as differences in biotic or abiotic conditions. For example, Wang et al. (2012a) compared the effects of geographic isolation (estimated as least-cost path and circuit distances) and ecological isolation (estimated as differences in values of 24 environmental variables) on genetic divergence for 17 *Anolis* species. Genetic divergence was significantly related to geographic and ecological isolation for 15 and 13 species, respectively, with inferred effects of geographic isolation being, on average, more than twice as strong as those of ecological isolation. Similar results have been reported for a variety of other vertebrate taxa, including the Trinidad Guppy (*Poecilia reticulata*; Crispo et al. 2006), Agassiz's Desert Tortoise (*Gopherus agassizii*; Sánchez-Ramírez et al. 2018), and the Common Frog (*Rana temporaria*; Van Buskirk & van Rensburg 2020), suggesting that geographic distance, spatial features, and arrangements of suitable habitat are of greater importance to genetic divergence than variation in environmental/ecological conditions (but see support for IBE in epigenetic data, Wogan et al. 2019).

Across these studies, greater support for IBD + IBR than IBE contrasts with the hypothesis that factors contributing to genetic divergence are predominantly environmental/ecological (Foll & Gaggiotti 2006; Thorpe et al. 2008; Nosil 2012). However, past comparisons of geographic and environmental isolation have tended to address vertebrate taxa with relatively specific habitat associations that are maintained through their life cycles. It is therefore intuitive that patterns of genetic divergence are generally best explained by resistance-based geographic distances, as configurations of suitable habitat typically moderate spatial variation in movement and dispersal in such taxa (Coulon et al. 2004; Vignieri 2005; Broquet et al. 2006; Crispo et al. 2006; McRae 2006; Epps et al. 2007; McRae & Beier 2007; Wang et al. 2008; 2009; Sánchez-Ramírez et al. 2018). But other taxa, including many terrestrial

invertebrates, have discrete dispersal life stages (generally the adult) with broader habitat tolerances than larval stages, which may have important consequences for processes affecting genetic divergence (Phillipsen et al. 2015). For example, Keller & Holderegger (2013) found that short-distance movements of the southern damselfly (*Coenagrion mercuriale*) generally followed corridors of reproductive and larval habitat (streams), while long-distance dispersal, inferred from patterns of genetic divergence, was best explained by Euclidean distance across unsuitable habitat (agricultural land). Although IBE was not evaluated for *C. mercuriale*, genetic divergence within this taxon and similar taxa with discrete dispersal life stages may be expected to show stronger relationships to environmental isolation than geographic isolation, as evolutionary processes predicted by IBD and IBR become subsidiary to those predicted by IBE.

The aim of our study was to evaluate IBD, IBR, and IBE for a taxon with high habitat specificity, a discrete dispersal stage, and distribution across a variable environment. The Old World swallowtail butterfly (*Papilio machaon* L.) species group has been the subject of considerable study in North America (e.g., Sperling 1987; 1990; Dupuis & Sperling 2015; 2016; Dupuis et al. 2016; 2019). One subspecies in particular, *P. m. dodi* McDunnough 1939, is well suited for this investigation. Adult *P. m. dodi* search for mates by hilltopping along prominent edges of river valleys, leading to clustering of occurrence records along the Red Deer, South Saskatchewan, Old Man, and Milk Rivers in southern Alberta and Saskatchewan, Canada (Sperling 1987; Bird et al. 1995; Dupuis et al. 2019; occurrence records are visualized in Figure 7-1, inset h). After mating, females travel downslope from hilltops to oviposit on their larval host plant, *Artemisia dracunculus* L., which is generally restricted to south-facing eroding slopes of river valleys. We therefore hypothesize that these river valleys constitute suitable habitat under the functional resource-based concept (*sensu* Dennis et al. 2002), providing resources sufficient

for mate location, reproduction, resting, roosting, and feeding. This unique habitat configuration may be described as a dendritic ecological network of suitable habitat corridors situated in a matrix of unsuitable agricultural and prairie habitat. Such configurations have proven practical for decoupling Euclidean and resistance-based distances (e.g., Keller & Holderegger 2013). Additionally, occurrences of *P. m. dodi* in Canada span an area of approximately 53,000 km^2 that is sufficiently variable in environmental conditions to assess IBE (environmental gradients are visualized in Figure 7-1, insets e – g).

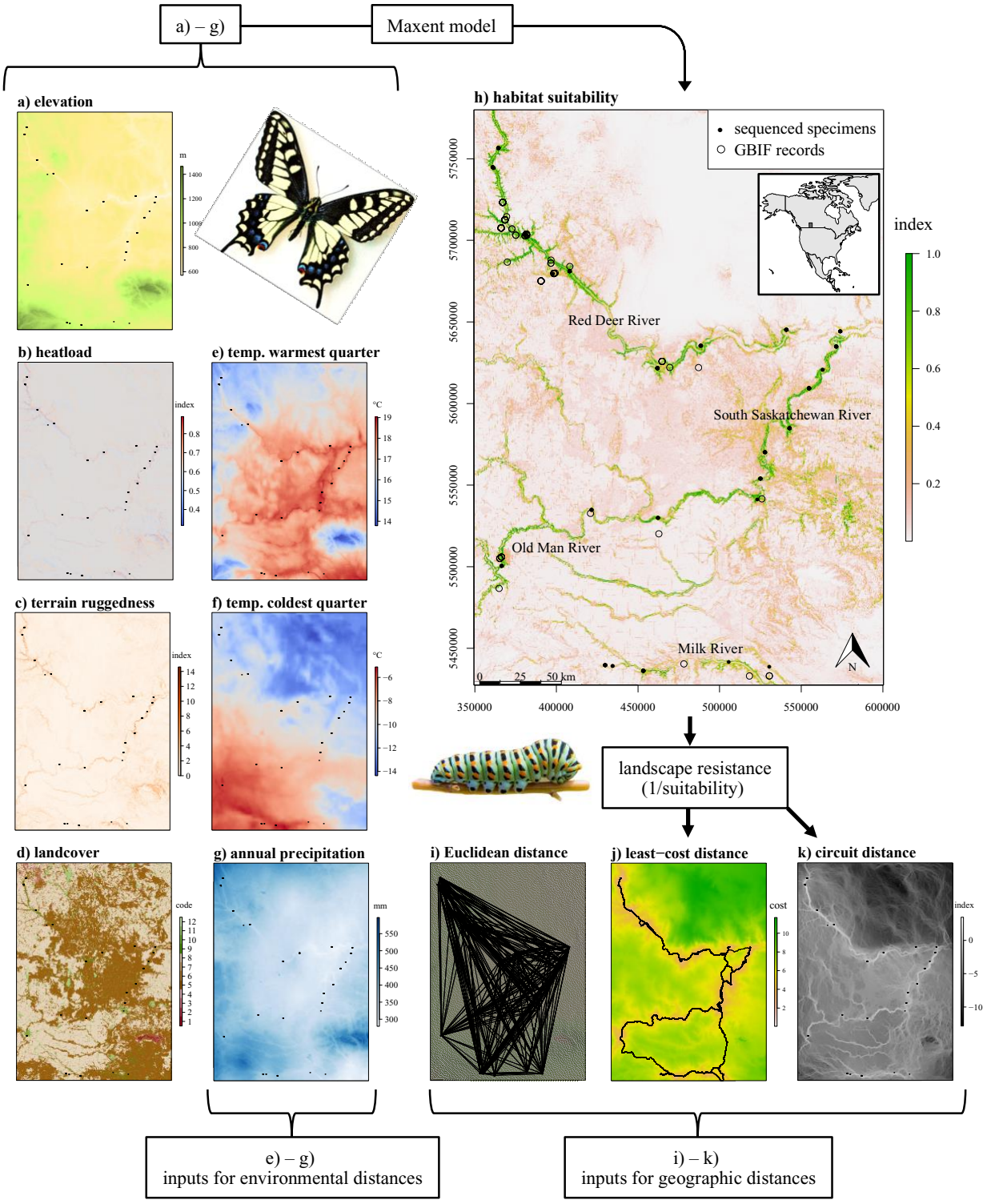

Figure 7-1. Visual representation of the methods used to generate geographic and environmental distances between the 161 sequenced *Papilio machaon dodi* included in this study. Various

spatial data layers (a – g) and 375 *P. m. dodi* records (161 sequenced individuals and 214 georeferenced *P. m. dodi* Global Biodiversity Information Facility (GBIF) records) were used to build a Maxent habitat suitability model, which was then used to predict habitat suitability across the study area, with higher values indicating greater suitability. High habitat suitability generally followed the eroding banks of major rivers in southern Alberta and Saskatchewan, Canada (Red Deer, Old Man, South Saskatchewan, and Milk Rivers). Euclidean distances (i) represent the minimum distances between sequenced individuals. A resistance surface was parameterized as the inverse of habitat suitability and used to estimate least-cost path and circuit distances (j and k). The background in inset j is a projected cost surface, representing the cumulative costs incurred by individuals moving across the landscape from each occurrence point. Environmental distances (e – g) were estimated by taking the absolute difference between values of environmental variables extracted from the occurrences of sequenced individuals. Inset pictures are the adult and larval stage of *P. m dodi*.

7.3 Materials and methods

We used a series of causal models to assess and contrast the effects of geographic isolation (IBD + IBR) and environmental isolation (IBE) on genetic divergence within *P. m. dodi*. Resistance-based distances were estimated using a resistance surface parameterized as the inverse of predicted habitat suitability to assess whether adult butterflies are more likely to disperse within suitable habitat. Such a result would suggest that: i) configurations of suitable habitat are important considerations for predicting gene flow on heterogeneous landscapes; and ii) habitat suitability may be used as a proxy of landscape permeability. However, although habitat associations of *P. m. dodi* appear to be specific and geographically restricted, the

dispersal ability of *P. machaon* is estimated to be among the greatest of all Canadian butterflies (Burke et al. 2011). We therefore hypothesize that, considering only measures of geographic isolation, relationships between genetic divergence and Euclidean distance (IBD) will be stronger than those between genetic divergence and least-cost path or circuit distances (IBR), despite a resistance surface predicting greatest landscape permeability (lowest resistance) along the dendritic ecological network of suitable habitat.

Environmental isolation may also play a significant role in structuring genetic divergence within *P. m. dodi*. Sperling (1987) noted distinct differences in the butterfly's phenology, diapause propensity, and voltinism across its Canadian range, possibly implicating divergent selection related to variation in summer temperatures as a driver of spatial genetic divergence. If IBE is detected, mechanisms by which environmental isolation structures genetic divergence may be inferred from patterns of population clustering, and we can make several subsequent predictions. First, if IBE is primarily driven by reduced fitness and negative selection on individuals that have dispersed across environmental gradients, genetic clustering should indicate some spatial discordance of individual cluster assignments (i.e. migrants found in non-natal populations) without admixture between clusters indicative of successful hybridization. Second, if IBE operates via reduced fitness and negative selection on genetically intermediate individuals, some admixture indicative of F1 hybridization between migrant and natal individuals may be evident, but substantial admixture among clusters should be absent due to selection against these admixed genotypes. Third, if individuals exhibit a reduced tendency to disperse across environmental gradients due to local adaptation, there should be little or no spatial discordance of individual cluster assignments or admixture among clusters. Finally, we also identified loci under putative divergent selection between genetic clusters (F_{ST} outliers) and

assessed environmental associations of allele frequencies for individual loci. If population structure is driven by local adaptation to environmental conditions, we hypothesize that F_{ST} outlier and environmental association analyses will identify similar sets of candidate loci under putative selection.

7.3.1 Sample collection

We collected adult and larva *P. m. dodi* with aerial net surveys and host plant searches, respectively, between May 15 and August 31, 2017. We visited most known *P. m. dodi* occurrence locations in Canada and aimed to collect 5-10 individuals every 25-50 km along the Red Deer, South Saskatchewan, Old Man, and Milk Rivers, where suitable habitat is present (collection locations are visualized in Figure 7-1, inset h). As with previous studies (Sperling 1987; 1990; Dupuis & Sperling 2015; 2016; Dupuis et al. 2016; 2019), we attempted to locate and collect *P. m. dodi* between major river valleys (e.g., smaller eroding slopes with *A. dracunculus* and the largest hilltopping features within sight). However, consistent with our past work and historical records (e.g., Bird et al. 1995), we did not observe any individuals outside of known suitable habitat along major river valleys. Adults were generally collected on prominences along major river valleys used as hilltopping features. Adult females descend from hilltops immediately after mating while males remain in search of additional mates, which means that males are more frequently encountered during sampling (Dupuis et al. 2019). Larvae were collected from patches of *A. dracunculus* on eroding slopes, typically 100–500 m^2 in area, below hilltopping features. Individuals collected within 500 m of each other were given the same collection location, recorded as the centroid of the hilltopping feature or *A. dracunculus* patch. After collection, adult individuals were frozen live and stored at -20°C. Larvae were raised to 4^{th} or 5^{th} instar on clippings of *A. dracunculus*, preserved in 95% ethanol, and stored at -20°C.

7.3.2 DNA extraction and library preparation

We extracted genomic DNA from thoracic tissue of adults (n = 148) and larvae (n = 32) using DNeasy Kits (Qiagen, Hilden, Germany). Extractions followed the manufacturer's protocol, with the addition of bovine pancreatic ribonuclease A treatment (RNaseA, 4 ul at 100mg/ml; Sigma-Aldrich Canada Co., Canada). Genomic DNA was then ethanol precipitated and stored in 50 ul Millipore water at -20°C. Double digest restriction-site associated DNA sequencing (ddRADseq) libraries were prepared from 200 ng input DNA and *MspI* and *PstI.* We followed a modified version of Poland et al. (2012) for wet lab procedures and used a standard dual index Illumina adapter system following Peterson et al. (2012). Details of our library preparation protocol and adapters are provided in MacDonald et al. (2020) Data S1 and S2, respectively. A final, pooled library of 180 individuals was sequenced with single-end, 75-bp sequencing on a single high output flowcell of an Illumina NextSeq 500.

7.3.3 Bioinformatic processing

Following Illumina sequencing, we used "process_radtags" in the program Stacks 2.0 (Rochette et al. 2019) to demultiplex FASTQ reads and filter those with quality scores below 20 within a sliding window 15% of the read length. All reads were truncated to 67 bp after removing the 8-bp Illumina index sequences, identified with one mismatch permitted. We then searched for and removed remnant Illumina adaptor sequences and removed the first 5 bp from the 5' end of each read (corresponding to the PstI restriction site) using the program Cutadapt 1.9.1 (Martin 2011). Filtered and trimmed reads were aligned to a *P. machaon* reference genome comprised of 63,187 scaffolds (NCBI Accession GCA_001298355.1) using Burrows-Wheeler Aligner 0.7.17 (BWA-MEM) (Li and Durbin 2009; Li 2013). We then converted files from SAM to BAM format using SAMtools 1.9 (Li et al. 2009) and used "gstacks" and "populations" within

Stacks 2.0 to call SNPs and generate output files, stipulating a single population containing all individuals. Genotype calls were exported in variant call format (VCF), and individuals with more than 50% missing data were removed from the dataset (three adults and three larvae). Finally, we used VCFtools 0.1.14 (Danecek et al. 2011) to filter genotypes with read depths less than five and filter loci with minor allele frequencies less than 0.05, percentages of missing data greater than 5%, and those within <10 kb of each other to reduce the probability of retaining loci that are in physical linkage. This thinning interval was based on linkage decay documented in other butterfly species; e.g., linkage decays to baseline within 1 kb – 10 kb in *Heliconius* spp. (Martin et al. 2013) and within 100 bp in *Danaus plexippus* (Zhan et al. 2014).

Multiple larval samples were often collected from single or adjacent *A. dracunculus* plants. Full-sibling relationships between individuals are therefore probable due to females ovipositing multiple eggs on single plants, which may bias estimates of genetic relatedness and inferences of population structure (O'Connell et al. 2019). To identify full sibs, we used the package "SNPRelate" (Zheng et al. 2009) implemented in the R environment (v3.5.1; R core team) to estimate kinship coefficients for all pairs of sequenced individuals. For diploid organisms, the expected kinship coefficient between full sibs is 0.25. Only pairs of larvae collected from single locations had kinship coefficients greater than 0.22, where a natural break in values occurred. For each of these larval pairs, we removed the individual with the greater percentage of missing data (13 individuals total). We then reverted to the original BAM files, recalled SNPs with Stacks, and filtered VCF files as above using the reduced dataset of 161 individuals comprised of 136 adult males, 8 adult females, and 17 unsexed larvae.

7.3.4 Population structure

Two independent methods were used to quantify population structure. We first used the model-based clustering program Structure 2.3.4 (Pritchard et al. 2000) in a hierarchical fashion (Vähä & Primmer 2006) to assess *K*-values ranging from 1 to 10. For the first set of runs in the hierarchical analysis, 10 independent runs were completed for each value of *K* using the admixture model and correlated allele frequencies. The burn-in period and number of Markov chain Monte Carlo (MCMC) repetitions were set to 100,000 and 1,000,000, respectively. Location priors (n = 27 collection sites) were used to inform the MCMC algorithm. The alpha prior (relative admixture levels between populations) was set to 0.5, based on the inverse of the expected value of $K = 2$ informed by overt clustering in preliminary principal component analysis (PCA; see Data S1, Fig. S1) (Wang 2017). Two approaches were then used to determine the optimal *K* from Structure outputs; the ΔK method (Evanno et al. 2005), implemented in the program Structure Harvester (Earl & vonHoldt 2011), and the rate of change in the likelihood of *K* across $K = 1{:}10$ (Pritchard et al. 2000). For the second set of runs in the hierarchical analysis, we completed independent Structure analyses on each cluster identified in the first analysis using settings identical to the first set of runs. Q-values > 0.8 from the first set of runs denoted assignment of individuals to specific clusters, effectively excluding individuals with substantial admixture.

In addition to Structure analyses, we also assessed population structure using discriminant analysis of principal components (DAPC), which conducts discriminant analysis (DA) on principal components (PCs) generated in PCA (Jombart et al. 2010). Assignments of individuals to *a priori* population clusters for DAPC were inferred using the "find.clusters" function (adegenet package) with default parameters, retaining all PCs, to find the optimal *K*

value based on Bayesian Information Criterion (BIC) scores in successive *K*-means clustering analysis across $K = 1{:}20$. DAPC was then completed using the R package "adegenet" v2.1.1 (Jombart 2008). We used the "xvalDapc" function (adegenet package), stipulating 100 replicates, to determine the optimal number of PCs to retain in DAPC using stratified cross-validation of DAPC across increasing numbers of principal components (PCs). Missing genotypes were imputed as the mean of the available data per locus for this cross-validation.

7.3.5 Habitat suitability

To map habitat suitability for *P. m. dodi* within our study landscape, we created a habitat suitability model using Maxent software (Phillips et al. 2006), implemented through the R package "dismo" (Hijmans et al. 2011). Briefly, Maxent uses machine-learning maximum entropy modelling to predict habitat suitability across a landscape using georeferenced occurrence localities and a set of geographic information systems (GIS) predictor variables (spatial data layers). For georeferenced occurrence localities, we used both the collection locations of the 161 sequenced individuals from this study and georeferenced *P. m. dodi* occurrences downloaded from the Global Biodiversity Information Facility (GBIF; accessed from https://doi.org/10.15468/dl.axez0s, 5th December 2018). Of the 259 occurrences available from GBIF, 214 were within the study landscape. Geographic and environmental GIS data layers included elevation, a terrain ruggedness index, a heat load index (based on terrain slope and aspect), land cover (12 categories), and three Worldclim 2 (Fick and Hijmans 2017) bioclimatic variables: mean temperature of warmest quarter (temp.warm), mean temperature of coldest quarter (temp.cool), and mean annual precipitation (precip.). These environmental variables were selected based both on biological relevance and to minimize collinearity (see MacDonald et al. 2020 Data S1 for further details). Each GIS data layer was reprojected to an equal-area

projection (NAD83[CSRS98]/UTM zone 12N) at 30-m resolution using the R package "raster" (Hijmans & van Etten 2012). Further details and sources for GIS data layers may be found in MacDonald et al. 2020 Data S1. Correlation coefficients were less than 0.7 for all pairs of GIS data layers, and so all seven were included in the habitat suitability model.

To evaluate the predicative power of the habitat suitability model, 20% of occurrence localities were withheld for cross-validation and receiver operating characteristic (ROC) analysis (Phillips et al. 2006). Following evaluation, we used the model to predict habitat suitability across the study landscape using the "predict" function (raster package), with each grid cell receiving habitat suitability values ranging from 0–1, where higher values indicate greater habitat suitability. Information on the validity of the Maxent process and its application to *P. m. dodi* is available in MacDonald et al. (2020) Data S1.

7.3.6 Geographic distance

We estimated geographic isolation between sequenced individuals using three different pairwise distance metrics; Euclidean distance, least-cost path distance, and circuit distance. Euclidean distances represent minimal distances required to travel between locations and do not account for landscape characteristics. In contrast, least-cost path distances are estimated by searching for single, optimal routes that minimize cumulative costs associated with travelling through heterogeneous landscapes (Wang et al. 2009). Least-cost path analysis thereby assumes that organisms have complete knowledge of such landscapes and are able to consistently navigate optimal routes. Circuit-based analyses relax this assumption, with distances estimated by summarising costs associated with all possible paths through heterogeneous landscapes (McRae & Beier 2007).

Pairwise Euclidean distances between the collection locations of the 161 sequenced individuals were estimated using the "spDists" function in the R package "sp" (Pebesma and Bivand 2005). To estimate least-cost path and circuit distances, we first parameterized a resistance surface as the inverse of habitat suitability scores predicted by the habitat suitability model (McRae & Beier 2007; Wang et al. 2008; Storfer et al. 2010; Wang et al. 2015). We then estimated pairwise least-cost path distances using the "costDistance" function in the R package "gdistance" (van Etten 2018) and pairwise circuit distances using the program Circuitscape 5.0 (McRae 2006), both with an eight-neighbour connection scheme. To increase computational efficiency, we aggregated the resolution of the resistance surface to 300 m, as connectivity inferences are shown to be generally robust to such aggregations (McRae & Beier 2007). Collectively, these analyses produced three pairwise matrices of geographic distances.

7.3.6 Environmental distance

Environmental isolation between sequenced individuals were estimated using the same Worldclim 2 bioclimatic variables included in the habitat suitability model (temp.warm, temp.cool, and precip.). We extracted values for each of the three bioclimatic variables for collection locations of sequenced individuals using the "extract" function in the R package "raster". Following Wang et al. (2012a), environmental distances were estimated by taking absolute differences of these values. Resulting differences were organized into three pairwise matrices of environmental distances.

7.3.7 Determinants of genetic divergence

To evaluate how variation in geographic and environmental isolation relate to that of genetic divergence within *P. m. dodi*, we used two forms of causal modelling and an individual-based (c.f. population-based) approach. To quantify genetic divergence, we first used the "dist"

function within the R package "adegenet" to estimate pairwise genetic distance (sum of squared Euclidean distances between i^{th} and the j^{th} genotype) between sequenced individuals, commensurate with the geographic and environmental distance matrices generated above. This simple individual-based measure of genetic distance has been shown to effectively quantify genetic divergence in a variety of simulations (e.g., Shirk et al. 2017) and in field studies (e.g., Sánchez-Ramírez et al. 2018).

7.3.7.1 Reciprocal causal modelling (RCM)

Our first set of causal models addressed relationships between genetic distance and geographic and environmental distances using reciprocal causal modelling (RCM) with partial Mantel tests (Cushman et al. 2006; 2013). We used the "mantel.partial" function within the R package "vegan" (Oksanen et al. 2007) to perform partial Mantel tests (999 permutations) for genetic distance and each combination of the 6 geographic and environmental distances, totalling to 30 tests organized into 15 reciprocal causal models. For each comparison of relationships between genetic distance and two geographic/environmental distances (one reciprocal model), we first estimated the partial Mantel's *R* coefficient (R_{PM}) between genetic distance and one geographic/environmental distance (focal variable) conditioned on the other geographic/environmental distance (alternative variable), comprising partial Mantel test A. We then estimated the reciprocal R_{PM}, comprising partial Mantel test B. If $R_{PM\text{-}A} > R_{PM\text{-}B}$, the focal variable from partial Mantel test A is better supported, and *vice versa*. Results of $R_{PM\text{-}A} - R_{PM\text{-}B}$ were summarized in a heatmap similar to Ruiz-Gonzalez et al. (2105). Notwithstanding this RCM framework, if both $R_{PM\text{-}A}$ and $R_{PM\text{-}B}$ were significant, we inferred partial support for relationships between genetic distance and each of the two geographic or environmental distances used in the reciprocal model.

7.3.7.2 Structural equation modelling (SEM)

Our second set of causal models employed structural equation modelling (SEM), a method originally developed by Wright (1921), to quantify the relative strength of effects of geographic distance and environmental distance on genetic distance according to an *a priori* causal path network. SEM analyses have proven particularly useful for distinguishing effects of multiple collinear variables (e.g., geographic and environmental distances; Grace 2006; Wang et al. 2012a). Our causal path network included two regression pathways; one from geographic distance to genetic distance and one from environmental distance to genetic distance. Geographic and environmental distance were linked by a covariance pathway. Results of the RCM analysis were used to infer which single measures of geographic and environmental distance were most strongly related to genetic distance. To test whether geographic and environmental distance contributed meaningfully to observed variation in genetic distance, we compared Akaike's information criterion (AIC) scores for the full model, a model excluding geographic distance, and a model excluding environmental distance (*sensu* Wang et al. 2012a). Lower AIC scores indicated superior model fit. Models with AIC scores exceeding the best supported model by 10 or more points were unsupported (Burnham & Anderson 1998). SEM analyses were completed using maximum-likelihood estimation in the R package "lavaan" (Rosseel 2012). To account for nonindependence among pairwise data, we randomly permuted rows and columns of distance matrices to generate null distributions for path coefficients assuming no relationships between variables exist (Fourtune et al. 2018). Unbiased standard errors and *P*-values for path coefficients were then calculated by comparing observed coefficients to their null distributions. This was completed using a modification of the "permutation.based.pathanalysis" R code provided by Fourtune et al. (2018).

7.3.7.3 Validation of the causal model framework

To investigate whether inferences from casual models were contingent on method of analysis, we also employed linear mixed effects models with maximum likelihood population effects (MLPE; Clarke et al. 2002). MLPE linear mixed effects models have been shown to be one of the highest performing methods for quantifying relationships between distance matrices while controlling for nonindependence among pairwise data (Shirk et al. 2017). Relationships between genetic distance and each of the six measures of geographic and environmental isolation were quantified in independent models. The identities of sequenced individuals involved in pairwise distance values were included in models as mixed effects to control for nonindependence within distance matrices (Clarke et al. 2002). MLPE linear mixed effects models were fit using the "MLPE.lmm" function within the R package "ResistanceGA" (Peterman 2018). We set REML=FALSE to allow for the estimation of valid AIC scores that were used to evaluate relative model support (Row et al. 2017; Shirk et al. 2017; Peterman 2018).

7.3.8 Population divergence of candidate loci

To identify candidate loci under putative divergent selection, we used Bayescan 2.1 (Foll & Gaggiotti 2008) to estimate allele frequencies and F_{ST} values for 1,382 loci. When population assignment of individuals is sensible, Bayescan is generally recognized as the most effective method for identification of F_{ST} outlier loci (Narum & Hess 2011; De Mita et al. 2013; Lotterhos & Whitlock 2014). Q-values > 0.8 from the Structure analysis ($K = 2$) denoted assignment of individuals to either the northern or southern cluster, effectively excluding individuals with substantial admixture from this analysis. We used default parameters to run Bayescan (prior odds set to 10, thinning interval to 10, number of pilot runs to 20, length of pilot runs to 5000, and

burn-in length to 50,000), except for the number of outputted iterations, set to 10,000. To reduce the likelihood of false positives associated with multiple tests, we assessed significance of F_{ST} outliers using q-values generated by Bayescan according to the False Discover Rate (FDR) criterion (Benjamini & Hochberg 1995). F_{ST} outlier loci were identified using a q-value-threshold of 0.05 ($-\log_{10}$ q-value ~ 1.3). Fifteen Bayescan runs were completed using this protocol, and a union of the resulting lists of F_{ST} outlier loci was taken to generate a final list of loci under putative divergent selection. We also used the "snpzip" function (adegenet package) with default settings to identify which loci contributed most significantly to between-population structure in DAPC, with population assignment based on *K*-means clustering analysis. This analysis uses the relative contribution of each SNP to DAPC to perform hierarchical clustering and classify loci as either "structural" or "non-structural".

7.3.9 Environmental associations of individual loci

While Bayescan is effective for identifying loci under putative divergent selection among discrete populations, an individual-based approach may be more effective for identifying candidate loci potentially under selection across environmental gradients (Frichot et al. 2013). To accomplish this, we used latent factor mixed modelling (LFMM), implemented in LFMM 1.3 (Frichot et al. 2013) via the R package "LEA" (Frichot & François 2015). LFMM tests for correlations between allele frequencies of individual loci and environmental variables (each included in an independent model as a fixed effect), while controlling for background population structure using latent factors equal in number to the optimal value of *K*. This reduces the likelihood of false positives arising from spurious relationships between allele frequencies and environmental variables due to autocorrelation of space, demography, and the environment (Frichot et al. 2013; Lotterhos & Whitlock 2014), which can be problematic for other analysis

methods, such as those employed in BayEnv2 (Günther & Coop 2013; Lotterhos & Whitlock 2014).

Environmental variables included in LFMM analyses were temp.warm, temp.cool, and precip.. We completed five independent LFMM runs with 10,000 iterations and a burn-in of 5,000, stipulating two latent factors ($K = 2$ inferred from both Structure runs and DAPC). Results were then combined by calculating the median $|z|$-scores across the five LFMM runs, which represent the strength of the genetic-environment association for each locus. To validate the number of latent factors used in LFMM, we visually inspected adjusted *P*-values histograms for each environmental variable, estimated using the genomic inflation factor (λ) procedure (Devlin and Roeder 1999). Distributions that are relatively flat with a peak near zero indicate the selected number of latent factors adequately controlled for potentially confounding effects of spatial genetic structure (Frichot & François 2015). Finally, to control for false positives associated with multiple tests, we again used the FDR criterion (Benjamini & Hochberg 1995), producing q-values for each association. Loci with q-values < 0.05 were inferred as having significant environmental associations.

7.3.10 Genomic contexts of candidate loci

To map the location of candidate loci within the *P. m. dodi* genome, we used BEDTools v2.27.1 (Quinlan & Hall 2010) to extract 5 kb of flanking sequence on the 5' and 3' ends of each candidate locus identified by either Bayescan or LFMM analyses. This length of flanking sequence was selected in reference to previous thinning of loci within 10 kb of each other. We then used the BLAST function within Lepbase (Challis et al. 2016) to match resulting sequences to annotated genes within Lepbase's butterfly and moth CDS databases. As this search queried multiple species' genomes, we evaluated possible interspecific matches based on percent match

of the query, phylogenetic distance to the matched species, and the number of distinct genomes (multiple species) in which each gene was reported. Putative gene functions were compiled from the UniProt Consortium (2018) using gene accession codes included within the Lepbase output.

7.4 Results

ddRAD sequencing resulted in a total of 293,036,249 single-end, 75-bp reads across the original set of 180 sequenced individuals. After running "process_radtags" and associated filters, 273,664,014 reads remained, of which, 192,902,810 were aligned to the *P. machaon* reference genome. After removing individuals with >50% missing data and putative full-sibs, 108,049,831 reads were used to call 104,038 SNPs for the final set of 161 sequenced individuals. Filtering of loci resulted in a total of 1,382 SNPs with a mean read depth of 71.99 (min = 11.76, max = 1950.7), comprising the dataset used in all subsequent analyses.

7.4.1 Population structure

Our first set of Structure runs predicted an optimal *K*-value of $K = 2$ using both the ΔK method (Evanno et al. 2005) and the rate of change in the likelihood of *K* across $K = 1{:}10$ (Pritchard et al. 2000; see MacDonald et al. 2020 Data S1, Fig. S2 a). Individuals collected near and north of Dorothy, Alberta, were generally assigned to a northern cluster, while individuals collected within and south of Dinosaur Provincial Park, Alberta, were assigned to a southern cluster (Figure 7-2). Spatial discordance of two individuals' cluster assignments (i.e. migrants found in non-natal populations) suggests that dispersal between the regions occurs. Nine individuals had an approximate 50/50 split of Q-values, suggesting that hybridization between migrant and natal individuals occurs. However, little admixture was observed beyond these putative F1 hybrids. In the second set of Structure runs, no subclustering was evident in the northern cluster, while the existence of two subclusters was best supported within the southern

cluster (MacDonald et al. 2020 Data S1, Fig. S2 b, c). Q-values for individuals of these two subclusters (K = 2 in the southern cluster only analysis) were very similar to those for K = 3 in the first set of runs (including all individuals). For simplicity, we therefore present in Figure 7-2 admixture plots from the first set of Structure runs (all individuals) for both K = 2 and K = 3.

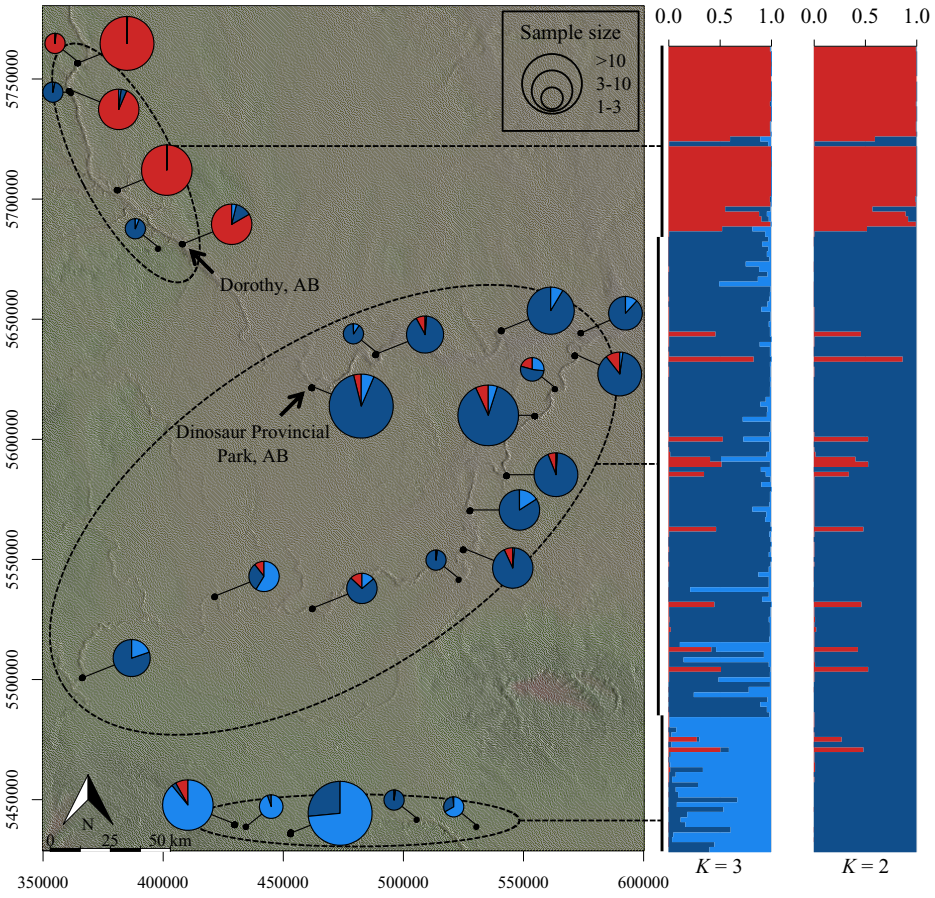

Figure 7-2. Population genetic structure within *Papilio machaon dodi* in Alberta and Saskatchewan, Canada, inferred using the model-based clustering program Structure. An optimal K value of 2 was best supported by the ΔK method and rate of change in the likelihood of K across K = 1:10 for the first set of Structure runs addressing all individuals. Hierarchical runs addressing the northern and southern clusters independently suggested no overt subclustering

within the northern cluster and the existence of two subclusters within the southern cluster. We present admixture plots derived from the first set of Structure runs for both $K = 2$ (exhibiting the two primary clusters) and $K = 3$ (including the two southern subclusters) for simplicity. For $K = 2$, individuals collected near and north of Dorothy, Alberta, were generally assigned to a northern cluster, while individuals collected within and south of Dinosaur Provincial Park, Alberta, were assigned to a southern cluster.

Similar to the first set of Structure runs, DAPC based on *K*-means clustering analysis and BIC across $K = 1{:}20$ suggested $K = 2$ was best supported. An optimal number of 20 PCs was retained for DAPC. Assignments of individuals to northern and southern clusters were nearly identical between the first set of Structure runs and DAPC, save four admixed individuals with Q-values around 0.5 that were assigned to the southern cluster by Structure (based on Q-values > 0.5) and the northern cluster by DAPC.

7.4.2 Habitat suitability

The Maxent habitat suitability model adequately predicted habitat suitability across our study landscape, indicated by an out-of-sample AUC score of 0.948. As hypothesized, high suitability generally followed the eroding banks of major rivers, taking on the form of a dendritic ecological network (Figure 7-1, inset h). Contributions of each variable to the habitat suitability model were estimated by measuring the drop in AUC after values of each variable were randomly permuted (permutational importance): terrain ruggedness = 47.7, temp.cool = 34.2, elevation = 4.8, precip. = 4.4, temp.warm = 0.2, and heat load = 0.1. A resistance surface parameterized as the inverse of habitat suitability scores permitted estimation of least-cost path and circuit distances (Figure 7-1, insets j and k).

7.4.3 Determinants of genetic divergence

7.4.3.1 Reciprocal causal modelling (RCM)

Results of RCM are summarized in a heatmap (Figure 7-3), with red and blue colours indicating positive and negative values for $R_{PM\text{-}A}$ - $R_{PM\text{-}B}$, respectively (*sensu* Ruiz-Gonzalez et al. 2105). Focal and alternative variables used in partial Mantel test A for each reciprocal model are on the y- and x-axes, respectively. For ease of interpretation, variables on the y-axis with more positive (red) values in their corresponding rows are better supported. Overall, the strongest correlates of genetic distance after partialling out relationships with alternative variables were Euclidean distance and temp.warm distance. Euclidean distance was significantly correlated with genetic distance after partialling out temp.warm distance (R_{PM} = 0.23; P = 0.001) and the reciprocal partial Mantel test was also significant (R_{PM} = 0.19; P = 0.001). All other partial Mantel tests using either the Euclidean or temp.warm distances as alternative variables were not significant, indicating other measures of geographic and environmental distances were unsupported.

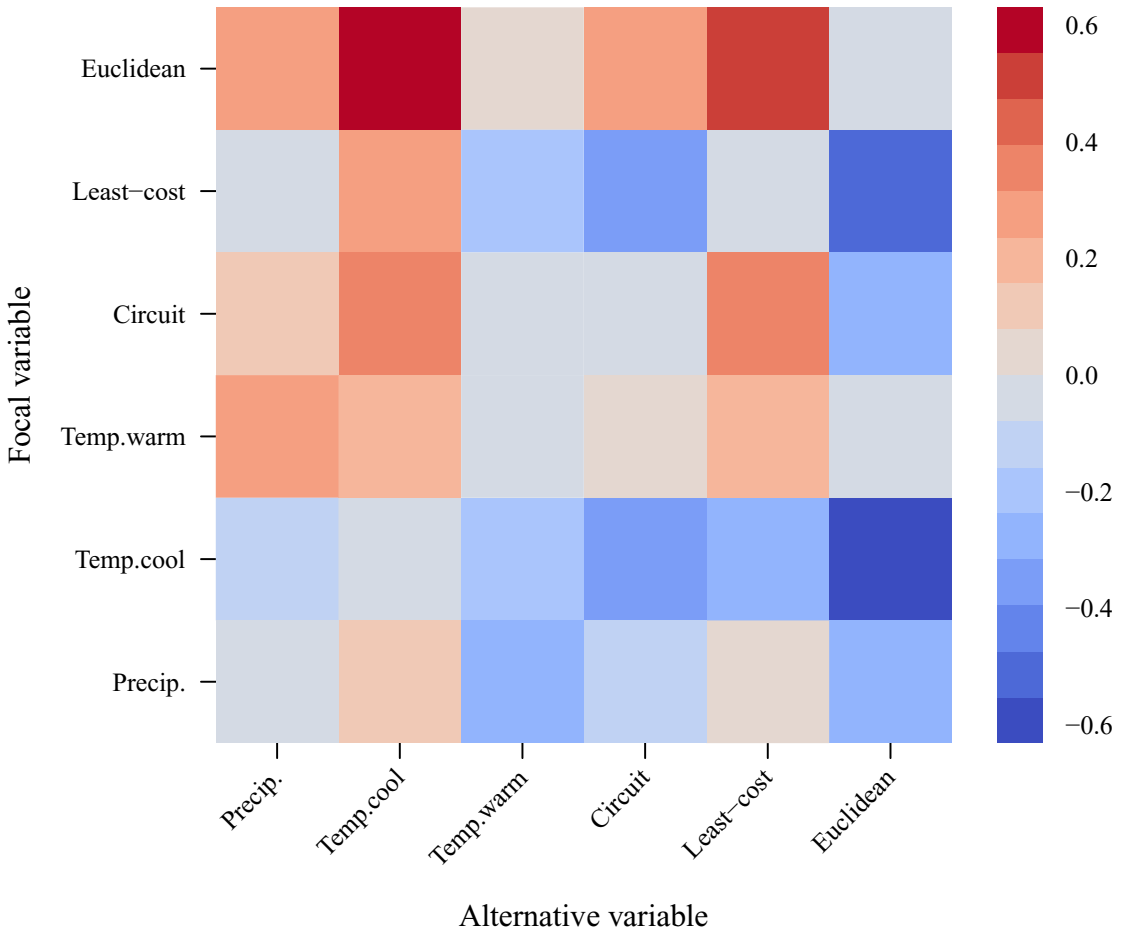

Figure 7-3. Pairwise heatmap visualizing reciprocal causal modelling (RCM) results. Values in each cell represent results of $R_{PM\text{-}A}$ - $R_{PM\text{-}B}$, with red and blue colours indicating positive and negative values, respectively. Rows and columns contain the focal and alternative variables, respectively, for model A within each reciprocal model. Therefore, the figure should be interpreted by rows and not columns; variables on the y-axis with more positive (red) values in their corresponding rows are better supported.

7.4.3.2 Structural equation modelling (SEM)

Based on results of RCM analysis, Euclidean distance and temp.warm distance were used as single measures of geographic and environmental distance, respectively, in our causal path network. The full model, including direct paths from both geographic distance (Euclidean distance) and environmental distance (temp.warm distance) to genetic distance, was better supported than alternative models excluding either geographic or environmental distance (Table 7-1). Path coefficients for Euclidean distance and temp.warm distance were 0.120 ($P < 0.001$) and 0.331 ($P < 0.001$), respectively, suggesting each is positively related to genetic distance. Covariance of the two predictor variables in the model was 0.592; thus, relative effect sizes should be interpreted with caution. However, model selection based on AIC indicated that variation in Euclidean distance and temp.warm distance each have important relationships to variation in genetic distance beyond that associated with their covariance structure.

Table 7-1. Relative support for path analysis structural equation models (SEMs) inferred using Akaike's information criterion (AIC). The causal path network in the full model included two regression pathways; one from geographic distance to genetic distance and one from environmental distance to genetic distance. Reciprocal causal models (RCM) were used to infer which single measures of geographic and environmental distance were most strongly related to genetic distance and used in place of geographic and environmental distance (Euclidean distance and temp.warm distance, respectively). Akaike's information criterion (AIC) scores are reported for the full model, a model excluding environmental distance (geography only), and a model excluding geographic distance (environment only).

SEM structure	AIC	Δ AIC
Full model	34147.52	0
Environment only	34288.77	141.25
Geography only	35203.77	1056.25

7.4.3.3 Validation of the causal model framework

Results of MLPE linear mixed effects models aligned with those of our causal models. Overall, the best supported variable affecting genetic distance was temp.warm distance followed by Euclidean distance (Table 7-2). This result supported the use of Euclidean and temp.warm distances as single distance measures for geographic and environmental distances, respectively, in SEM. AIC scores for MLPE linear mixed effects models with Euclidean and temp.warm distances were > 2 points lower than all other models, suggesting that alternative measures of geographic and environmental isolation were unsupported (Burnham & Anderson 1998).

Table 7-2. Relative support for effects of geographic and environmental distances on genetic distance inferred using linear mixed effects models with maximum likelihood population effects (MLPE). Akaike's information criterion (AIC) scores are reported for each model.

Distance measure	AIC	Δ AIC
Temp.warm	32955.55	0
Euclidean	33147.24	191.69
Circuit	33152.41	196.86
Least-cost	34623.89	1668.34
Precip.	35055.58	2100.03
Temp.cool	35590.9	2635.35

7.4.4 Population divergence of candidate loci

All 15 independent Bayescan runs identified the same list of 33 outlier loci based on elevated F_{ST} values (2.39% of 1,382 SNPs). Positive alpha values for each of these 33 loci indicated they were under putative divergent selection, rather than purifying or balancing selection. To visualize the data, -$\log_{10}$ q-values were plotted against F_{ST} values estimated in the first Bayescan run for all 1,382 SNPs (see MacDonald et al. 2020 Data S1, Fig. S3). The mean of F_{ST} values for putative outlier loci was 0.169 (SD = 0.092), ranging from 0.030 to 0.272. The "snpzip" function (adegenet package) identified 17 structural loci that significantly contributed to between-population structure in DAPC. Each of these 17 structural loci were contained within the list of 33 candidate loci identified by Bayescan.

7.4.5 Environmental associations of individual loci

Histograms of adjusted *P*-values were uniformly distributed for all three environmental variables, suggesting $K = 2$ adequately controlled for confounding effects of spatial genetic structure (see MacDonald et al. 2020 Data S1, Fig. S4). LFMM identified a total of 78 loci with

significant environmental associations (q-values < 0.05; Figure 7-4). Single loci often had multiple significant environmental associations, likely due to spatial correlation of environmental variables. We therefore only considered the strongest association for each locus based on median $|z|$-scores (De Kort et al. 2015; Martins et al. 2018). In total, 52 of 57 loci significantly associated with temp.warm were more strongly associated with temp.warm than any other environmental variable, 12 of 27 for temp.cool, and 14 of 43 for precip.. A total of 25 loci were identified in both LFMM and Bayescan analyses; 23 of which were most strongly associated with temp.warm, 0 with temp.cool, and 2 with precip.. LFMM identified 56 loci with significant environmental associations that were not identified as candidate loci by Bayescan, while Bayescan identified 8 candidate loci that were not identified by LFMM as having significant environmental associations.

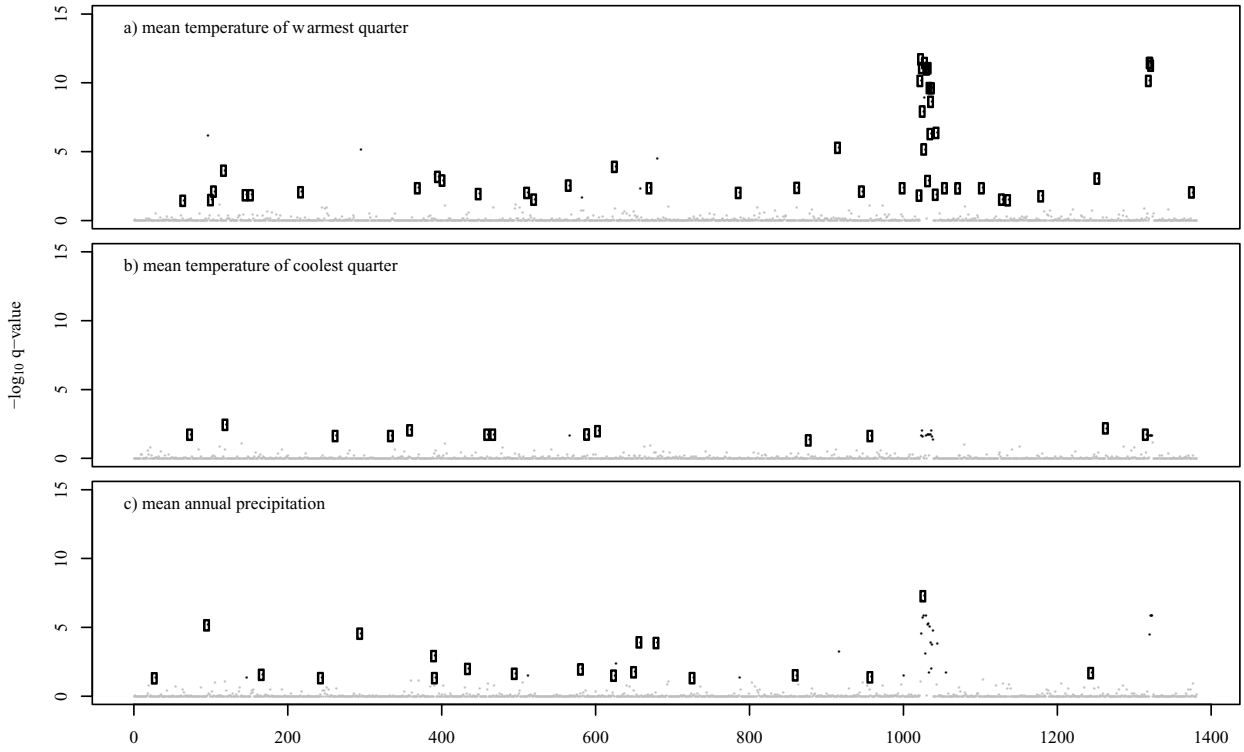

Figure 7-4. Associations between allele frequencies of loci (n = 1,382) and a) mean temperature of the warmest quarter of the year (temp.warm), mean temperature of the coolest quarter (temp.cool), and mean annual precipitation (precip.) in latent factor mixed models (LFMM) with K = 2. Black dots are loci with significant associations to the relevant environmental variable based on a q-value threshold of 0.05 ($-\log_{10}$ q-value ~1.3). Single loci often had multiple significant environmental associations due to spatial correlation of environmental variables. Open circles represent the strongest association (based on median $|z|$-scores) for each locus with a significant environmental association. Loci are arranged on the x-axis in order of position within scaffolds, which are in turn arranged by increasing size.

7.4.6 Genomic contexts of candidate loci

Using both Bayescan and LFMM, a total of 86 candidate loci were identified as being under putative divergent selection or having significant environmental associations. Results of our Lepbase BLAST search are summarized in MacDonald et al. (2020) Data S2, including putative genes matched to each of the 86 sequences and their functional annotation. Although several biological and cellular processes were evident from these candidates, no striking patterns were apparent to justify specific narratives of local adaptation.

7.5 Discussion

7.5.1 Geographic isolation and dispersal machines

Overt clustering of *P. m. dodi* occurrences along major river valleys in southern Canada is generally understood to be a function of both the ecologically restricted occurrences of its larval host plant, *A. dracunculus*, and the presence of hilltopping features along river valley

edges (Sperling 1987; Bird et al. 1995; Dupuis et al. 2019). Accordingly, results of our habitat suitability model support the inference that suitable habitat takes on the form of a dendritic ecological network with the intervening landscape comprised primarily of unsuitable habitat. Examination of a corresponding cost surface (Figure 7-1, inset j) shows that individuals dispersing between river valleys are predicted to experience high resistance/costs. IBR based on habitat suitability therefore predicts that dispersal and resultant gene flow should follow the dendritic ecological network (Figure 7-1, insets j and k). However, both RCM and MLPE linear mixed effects models suggested that variation in genetic distance was better explained by Euclidean distance than by resistance-based distances accounting for arrangements of suitable habitat. It is possible that other landscape features, beyond configurations of suitable habitat, influence dispersal of *P. m. dodi* and thereby support IBR. For example, Peterman et al. (2014) outlines a methodological approach for optimizing resistance surfaces using a nonlinear optimization algorithm and any combination of spatial variables. However, we did not have *a priori* mechanistic hypotheses predicting what variables might influence *P. m. dodi* dispersal beyond those associated with habitat suitability. Lack of support for IBR based on habitat suitability in this system was of principal interest and demonstrates a clear decoupling of landscape permeability from habitat suitability (*cf*. Coyne & Orr 2004; Crispo et al. 2006; McRae 2006; McRae & Beier 2007; Thorpe et al. 2008; Wang et al. 2009; Sánchez-Ramírez et al. 2018).

In addition to mating and reproduction associated with adult life stages of invertebrates, it may be instructive to think of adult *P. m. dodi* as dispersal specialists within the taxon's greater life cycle, with greater vagility and more general habitat tolerances than larval stages. Such traits enable long-distance dispersal of adults across considerable stretches of unsuitable habitat,

implicated in the spatial discordance of the cluster assignments of three adult individuals (i.e. migrants found in non-natal populations; see Figure 7-2 and MacDonald et al. 2020 Data S1 Fig. S1). Indeed, for many vagile butterfly species occupying fragmented landscapes, instances of long-distance dispersal between patches or corridors of suitable habitat are rare, but are known to have important consequences for gene flow, metapopulation dynamics, and emergent diversity patterns (Hanski 1998; Wiens 2001; Nowicki et al. 2014; MacDonald et al. 2018). We therefore propose that habitat suitability and landscape permeability should be evaluated as distinct concepts for taxa with discrete dispersal life stages, even if they are habitat specialists. As a mental shortcut, we suggest a "dispersal machine" concept, similar to Dawkins's (1978) selfish gene perspective on evolution by natural selection, in which genes metaphorically build "survival machines" (i.e., bodies of organisms) to facilitate their own stability and replication. In landscape genetics, it may be instructive to understand adult life stages of many terrestrial invertebrates as not only the life stage in which mating and reproduction occur, but also as "dispersal machines," exhibiting greater vagility and broader habitat tolerances than larval life stages. Such characteristics are likely to facilitate long-distance dispersal resulting in gene flow across heterogeneous landscapes. As a consequence of the dispersal machine concept, considerable support for IBR in past research cannot necessarily be extrapolated to organisms with disparate life histories; particularly, if the life cycles of focal taxa include a discrete dispersal life stage with substantially different habitat constraints.

7.5.2 Environmental isolation

Beyond the effects of geographic isolation on genetic divergence, this study provides multiple lines of evidence for strong associations between genetic variation and environmental conditions. First, analyses of population structure within *P. m. dodi* indicated the presence of two

prominent genetic clusters, with their spatial separation corresponding to a cline in mean temperature of the warmest quarter (temp.warm). Although any correlative relationship between spatial genetic structure and an environmental variable may be a product of spatial autocorrelation with a third unconsidered but causative variable, our results accord our *a priori* hypotheses relating summer temperatures to variation in phenology, diapause propensity, and voltinism within *P. m. dodi*, suggesting this relationship is both meaningful and worthy of further consideration. Second, RCM, SEM, and MLPE linear mixed effects models indicated that genetic distance among individuals was strongly associate with environmental distance (specifically, temp.warm distance) beyond its covariance structure shared with geographic distance (Euclidean distance). Finally, Bayescan identified 33 F_{ST} outlier loci inferred to be under putative spatially divergent selection between the two primary genetic clusters. Of these 33 F_{ST} outlier loci, 23 were also identified by LFMM analysis as having allele frequencies that were significantly associated with temp.warm values, meaning most loci identified as being under putative spatially divergent selection were also significantly associated with variation in summer temperatures.

Referring back to our original hypotheses, mechanisms by which environmental isolation structures genetic divergence may be inferred from patterns of population structure. Spatial discordance of a few individuals' cluster assignments suggests that some dispersal of adult individuals between natal and non-natal regions of the study area occurs (and see Dupuis and Sperling 2016). Additionally, some admixture indicative of F1 hybridization between migrant and natal individuals is evident (individuals with ~50/50 split of Q-values in Structure plots). However, further admixture among the two primary genetic clusters was not prevalent, possibly due to reduced fitness and negative selection on admixed genotypes. We therefore infer that

mechanisms by which environmental isolation contributes to genetic divergence in *P. m. dodi* may be a combination of reduced fitness and negative selection on both individuals that have dispersed across environmental gradients and genetically intermediate individuals resulting from hybridization.

7.5.3 Adaptation to local environmental conditions

Our combination of analyses suggests that variation in environment conditions has significant effects on genetic structure within *P. m. dodi*, possibly attributed to divergent selection across environmental gradients resulting in local adaptation. BLAST searches of candidate loci (both F_{ST} outliers and loci with significant associations) and their flanking sequences did not resolve annotated genes with biological functions sufficient to justify specific narratives of local adaptation. However, a total of 52 loci were significantly and most strongly associated with mean temperature of the warmest quarter (temp.warm)—the period in which development, reproduction, and diapause initiation occur. Nineteen of these 52 loci were evenly distributed along a single scaffold (NW_014538813.1, length 6.9 Mb) within the *P. machaon* reference genome, corresponding to the spike in $-\log_{10}$ q-values observed in LFMM analysis (Figure 7-1, inset a). Seventeen of these 19 loci were identified by Bayescan as being under putative divergent selection based on elevated F_{ST} values. Considered together, these results suggest the existence of an island of genomic differentiation (*sensu* Turner et al. 2005; Harr 2006) between northern and southern *P. m. dodi* populations, possibly corresponding to local adaptation to environmental conditions. However, the relatively low quality of the *P. machaon* genome assembly (>60 thousand scaffolds) precludes more in-depth comparative genomic analyses of this region.

Local adaptation to environmental conditions has been documented in a number of butterfly species, with variation in voltinism and diapause propensity across environmental gradients being a focal point of much past work (e.g., Friberg et al. 2008; Aalberg Haugen & Gotthard 2015; Pruisscher et al. 2018; Ryan et al. 2018). *Papilio machaon dodi* is known to exhibit variation in voltinism and diapause propensity across its Canadian range, with northern and southern populations exhibiting one and two generations per year, respectively (Sperling 1987; Bird et al. 1995). Additionally, we have noted that a proportion (~25%) of pupae reared from northern populations (near Drumheller, AB) require two distinct cooling cycles before emergence occurs, while pupae reared from southern populations (near Taber, Alberta) consistently emerge after a single cooling cycle (unpublished data, and see Sperling 1987; Dupuis et al. 2016). We hypothesize that this facultative second diapause in northern populations represents an ecological "hedging of bets" (*sensu* Seger & Brockmann 1987), distributing risks of high mortality and low fecundity due to poor environmental conditions across multiple years (Hanski 1988; Tuljapurkar 1990; Dupuis et al. 2016). Specific mechanisms by which variation in voltinism and diapause propensity might contribute to divergent selection between northern and southern *P. m. dodi* populations are not entirely clear, but there are several possibilities. Following hybridization between natal and migrant individuals in the northern extent of the study area, genetically intermediate offspring may experience high mortality if over-winter diapause is not induced in the first generation, as individuals of a second generation will lack sufficient day-degree accumulation and resources to complete their life cycle and enter over-winter diapause before temperatures drop to lethal levels. Conversely, genetically intermediate individuals emerging in the southern extent of the study area may experience reduced fecundity and fitness relative to natal individuals if only one generation of genetically intermediate

offspring emerge annually and/or a proportion of genetically intermediate individuals undergo a second diapause, resulting in a partially semivoltine lifecycle. By these mechanisms, divergent selection on voltinism and diapause propensity may maintain the integrity of the two genetic clusters. However, further work is required to evaluate the validity of these hypotheses.

7.5.4 Environmental determinants of genetic diversity vs. species occurrence

The environmental variables that were most strongly associated with genetic variation in *P. m. dodi* (temp.warm, followed by precip., followed by temp.cool) differed from those identified as the best for predicting occurrences in our habitat suitability model (temp.cool, followed by precip., followed by temp.warm). While summer temperatures may influence population structure in *P. m. dodi* via spatially divergent selection related to phenology, diapause propensity, and voltinism, winter temperatures may influence habitat suitability and limit the range of *P. m. dodi* due to limited cold tolerance of pupae. Indeed, winter temperatures have been inferred to limit ranges of other *Papilio* species (e.g., Kukal et al. 1991; Yoshio & Ishii 2001; Scriber et al. 2012). Considered together, our inferences suggest that the environmental/ecological conditions that influence divergent selection and possibly facilitate ecological speciation (Foll & Gaggiotti 2006; Thorpe et al. 2008; Nosil 2012) may differ from those that limit species' ranges and structure emergent patterns of species diversity (e.g., due to environmental filtering; *sensu* Kraft et al. 2015). Landscape genetic analyses, comparing the environmental/ecological conditions that influence divergent selection to those that limit species' ranges, are required for multiple species to assess the generality of this finding.

7.5.5 Anticipated changes to genetic structure in *Papilio machaon dodi*

Currently, no members of the *P. machaon* species group are listed as being of conservation concern in Canada. However, recognition of cryptic evolutionary significant units

(*sensu* Ryder; 1986), such as the northern and southern genetic clusters identified in this study, may affect future conservation in light of continued climate change. Our study resolved that a northern genetic cluster, which includes the type specimen for *P. m. dodi* (Kondla 1981), is geographically restricted and occupies a climatic niche that is distinct from more southerly populations in Alberta and Saskatchewan. If genetic divergence between the northern and southern genetic clusters is indeed driven by local adaptation to environmental conditions, continued climate change and rising summer temperatures may lead to the displacement of the northern genetic cluster as genotypes and associated traits of southern populations become more favorable across the northern extent of the range of *P. m. dodi*.

Quantitative data support these predictions. Mean temperature of the warmest quarter differed by an average of 1.60°C between collection locations of individuals belonging to the northern and southern genetic clusters (excluding migrants). Based on the ClimateWNA model (Wang et al. 2012b), which provides climate data from 24 general circulation models, mean annual temperature for Alberta is predicted to rise by 2.8 – 4.2 °C by the end of the century, contingent on emission scenarios (Schneider 2013). Accordingly, growing degree-days, based on a break point of 5°C, are estimated to increase 33-56% (Schneider 2013). Within Alberta and Saskatchewan, changes in mean annual temperature, growing degree-days, and vegetation composition are expected to be most pronounced in central and southern regions, including the current range of *P. m. dodi* (Schneider 2013; Zhang et al. 2015; Barber et al. 2016). There is some evidence that vagile North American butterfly species may track their climatic niches poleward as temperatures warm; however, more often than not, these range expansions are not sufficient to offset contractions toward the equator (Lewthwaite et al. 2019). Indeed, there exist few opportunities for *P. m. dodi* to track its climatic niche northward as temperatures rise. Steep

south-facing riverbanks that might provide adequate *A. dracunculus* habitat are sparse north of the current range of *P. m. dodi* and successional changes to the composition of riverbank vegetation that could provide suitable habitat are unlikely to match the pace of the southern genetic cluster's northward expansion. We therefore hypothesize that genotypes unique to the northern genetic cluster may be displaced by the end of the century.

8 Chapter 8: Conclusions and synthesis

What is biodiversity and how is it best measured? How does variation in habitat configuration, habitat composition, and environmental conditions affect emergent patterns of species diversity? How do these same factors relate to genomic variation within single species? These three questions were the focal points of my Ph.D. thesis. Addressing them required a combination of biogeography, landscape ecology, and population genomics, with sprinklings of philosophy here and there. The overarching aim of this work was to contribute to consilience of these subdisciplines within ecology and evolutionary biology, helping to resolve consistencies and inconsistencies among their inferences.

In Chapter 2, my analyses of temporal patterns of butterfly diversity showed that negative relationships between species richness and evenness can compromise the efficacy of many diversity indices; particularly, those based on information entropies. This was, to my knowledge, the first study to explicitly evaluate the ability of common diversity indices to resolve changes in the diversity of species assemblages through time. Overall, species richness was one of the best performing indices, exhibiting substantial interannual variation. Furthermore, it is widely thought to best align with most peoples' intuitive sense of species diversity. I therefore recommended that, for citizen-science and related long-term butterfly monitoring programmes, species richness may serve as a single, viable indicator of diversity trends. Additionally, if abundance data are available, direct measures of species evenness should be used in conjunction with richness to deepen our understandings of variation in species assemblages. However, conflating species richness and evenness in compound indices (specifically, information entropies) produces measures that can fail to capture variation in species diversity through time. Furthermore, these measures generally do not align with peoples' intuitive sense of species diversity. To understand

why, it is instructive to revisit the heuristic question: “Why is a rainforest more diverse than a monoculture?” Is it because, on average, the number of yes/no questions required to determine the species identity of a randomly selected individual will be greater for a rainforest than a monoculture? Is it because the probability that two randomly selected individuals will belong to the same species is lower for a rainforest than a monoculture? Clearly not, but these are verbal translations of measures given by information entropies (Shannon-Winer and Simpson’s index, respectively). While such measures quantify interesting properties of abundance distributions that may be used to characterize species assemblages, they do not conform to most peoples’ intuitive sense of species diversity. In citizen-science and related long-term butterfly monitoring programmes, deferring to simpler measures, namely species richness, not only helps to erode an overwhelming statistical machismo that pervades ecology, but also adequately captures variation in the diversity of species butterfly assemblages through time.

Chapters 3 and 5 refute the basic inference of Chapter 2, showing that patterns of butterfly species richness have great potential to obscure important response of individual species to habitat amount, fragmentation, and configuration. These studies compared abundances and diversities of species across 32 lake islands varying in both area and isolation. After controlling for total habitat amount, species richness was generally unrelated to degree of fragmentation, supporting stochastic species assembly mechanisms. However, further analyses, using a novel modelling framework detailed in Chapter 5, resolved that habitat fragmentation has important effects on smaller, less-mobile butterfly species. These effects were not apparent in richness-based analyses. This basic finding questions previous, richness-based support for the recently proposed and widely debated habitat amount hypothesis (e.g., Fahrig 2003; 2013; 2017; 2020), which posits that conservation efforts should focus solely on preserving the maximum

amount of habitat irrespective of its degree of fragmentation. In Chapter 6, I carried out a series of experimental releases of butterflies in the lake-island matrix. Tracking movements of released individuals suggested there is significant disparity in species' ability to navigate fragmented landscapes and that visual senses play a primary role in habitat detection. Considered together, our work addressing butterfly assemblages on lake islands suggests that species' identities and functional traits are necessary considerations in conservation frameworks and that efforts to minimize habitat fragmentation should continue as foundations for mitigating biodiversity loss. Chapter 4, which addressed patterns of vascular plant diversity across the same set of lake islands, resolved that scale of fragmentation and habitat heterogeneity within individual fragments may also be important considerations in this field of research.

The last section of my thesis, Chapter 7, addressed gene flow and climate-associated genomic variation within a single butterfly species: Dod's Old World swallowtail butterfly, *Papilio machaon dodi*. Genomic analyses and habitat suitability models were used to identify two distinct evolutionary lineages (north *vs* south) that were previously unrecognized. A series of landscape genetic and genomic analyses resolved that the integrity of the evolutionary lineages is likely maintained by spatially divergent selection resulting in local adaptation to climatic conditions. Interestingly, after controlling for climate-associated genetic variation, configurations of suitable habitat were unrelated to genetic connectivity within *P. m. dodi*. This challenges a foundational method in ecology: the use of habitat suitability models to infer patterns of connectivity between isolated populations when genetic data are unavailable. Although landscape composition and the scale of analysis differ considerably between Chapters 3 – 6 and Chapter 7, they share a common inference that habitat configuration may not be an important consideration for the conservation of highly vagile butterfly species. Landscape genetic/genomic

analyses of less vagile species are a necessary next step to inferring whether effects of habitat configuration on genomic patterns vary with species' functional traits, as suggested in Chapter 5.

Overall, the combination of my thesis projects demonstrates clear utility for integrating biogeography, landscape ecology, and population genomics to address cumulative effects of changes to habitat configuration, habitat composition, and environmental conditions. I believe that continued consilience among these subdisciplines is required for the successful conservation of butterflies, among other taxa. Much of the work presented in this thesis has suggested that an autecological approach may be most sensible from a theoretical perspective and most effective from a conservation perspective. However, continued work and consilience among subdisciplines in ecology and evolutionary biology will resolve whether generalizations across taxa are viable and applicable to conservation practice.

References

Aalberg Haugen, I. M., & Gotthard, K. (2015) Diapause induction and relaxed selection on alternative developmental pathways in a butterfly. Journal of Animal Ecology, 84(2), 464-472.

Abele, L. G., & Connor, E. F. (1979) Application of island biogeography theory to refuge design: making the right decision for the wrong reasons. In: Proceedings of the first conference on scientific research in the national parks (Vol. 1, pp. 89-94) US Department of the Interior Washington, D.C.

Abele, L., & Connor, E. F. (1979) Application of island biogeography theory to refuge design: Making the right decision for the wrong reasons. In: Proceedings of the first conference of scientific research in the National Parks I, Washington, pp 89-94

Acorn, J., & Sheldon, I. (2017) Butterflies of Ontario and Eastern Canada. Lone Pine Publishing, Vancouver, BC, Canada.

Altizer, S., & Davis, A. K. (2010) Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. Evolution: International Journal of Organic Evolution, 64(4), 1018-1028.

Andelman, S. J., & Willig, M. R. (2003) Present patterns and future prospects for biodiversity in the Western Hemisphere. Ecology Letters, 6(9), 818-824.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., … & Harrison, S. P. (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. Ecology Letters, 14, 19-28.

Andren, H. (1994) Effects of habitat fragmentation on birds & mammals in landscapes with different proportions of suitable habitat: a review. Oikos, 71(3), 355-366.

Arikawa, K. (2003) Spectral organization of the eye of a butterfly, Papilio. Journal of Comparative Physiology A, 189(11), 791-800.

Arikawa, K., Scholten, D. G., Kinoshita, M., & Stavenga, D. G. (1999) Tuning of photoreceptor spectral sensitivities by red and yellow pigments in the butterfly Papilio xuthus. Zoological science, 16(1), 17-24.

Arrhenius, O. (1921) Species & area. Journal of Ecology, 95-99.

Baguette, M. (2003) Long distance dispersal and landscape occupancy in a metapopulation of the cranberry fritillary butterfly. Ecography, 26(2), 153-160.

Baguette, M., & Van Dyck, H. (2007) Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. Landscape ecology, 22(8), 1117-1129.

Balkenhol, N., Dudaniec, R. Y., Krutovsky, K. V., Johnson, J. S., Cairns, D. M., Segelbacher, G., ..., & Joost, S. (2017) Landscape genomics: understanding relationships between environmental heterogeneity and genomic characteristics of populations. In Population genomics (pp. 261-322) Springer, Cham.

Barbaro, L., & Van Halder, I. (2009) Linking bird, carabid beetle & butterfly life-history traits to habitat fragmentation in mosaic landscapes. Ecography, 32(2), 321-333.

Barber, Q. E., Nielsen, S. E., & Hamann, A. (2016) Assessing the vulnerability of rare plants using climate change velocity, habitat connectivity, and dispersal ability: A case study in Alberta. Canada. Regional Environmental Change, 16(5), 1433-1441.

Baselga, A. (2012) The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Global Ecology and Biogeography, 21, 1223-1232.

Baselga, A., & Orme, C. D. L. (2012) betapart: an R package for the study of beta diversity. Methods in Ecology and Evolution, 3(5), 808-812.

Batschelet, E. (1981) Circular statistics in biology. Academic Press, New York.

Benjamini, Y., & 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(1), 2891300.

Bernard, G. D. (1983) Bleaching of rhabdoms in eyes of intact butterflies. Science, 219(4580), 69-71.

Bernard, G. D. (1983) Dark-processes following photoconversion of butterfly rhodopsins. Biophysics of Structure and Mechanism, 9(4), 277-286.

Berwaerts, K., Aerts, P., & Van Dyck, H. (2006) On the sex-specific mechanisms of butterfly flight: flight performance relative to flight morphology, wing kinematics, and sex in Pararge aegeria. Biological Journal of the Linnean Society, 89(4), 675-687.

Betzholtz, P. E., & Franzen, M. (2011) Mobility is related to species traits in noctuid moths. Ecological Entomology, 36(3), 369-376.

Bird, C. D., Hilchie, G. J., Kondla, N. G., Pike, E. M., & Sperling, F. A. H. (1995) Alberta butterflies. Edmonton, Alberta: The Provincial Museum of Alberta.

Blackiston, D., Briscoe, A. D., & Weiss, M. R. (2011) Color vision and learning in the monarch butterfly, Danaus plexippus (Nymphalidae). Journal of Experimental Biology, 214(3), 509-520.

Bock, C. E., Jones, Z. F., & Bock, J. H. (2007) Relationships between species richness, evenness, and abundance in a southwestern savanna. Ecology, 88(5), 1322-1327.

Boecklen, W. J. (1997) Nestedness, biogeographic theory, and the design of nature reserves. Oecologia, 112, 123-142.

Bohonak, A. J. (1999) Dispersal, gene flow, and population structure. The Quarterly Review of Biology, 74(1), 21-45.

Briscoe, A. D., Bernard, G. D., Szeto, A. S., Nagy, L. M., & White, R. H. (2003) Not all butterfly eyes are created equal: Rhodopsin absorption spectra, molecular identification, and localization of ultraviolet-, blue-, and green-sensitive rhodopsin-encoding mRNAs in the retina of Vanessa cardui. Journal of Comparative Neurology, 458(4), 334-349.

Broquet, T., Ray, N., Petit, E., Fryxell, J. M., & Burel, F. (2006) Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*) Landscape Ecology, 21(6), 877-889.

Brown, J. H., & Kodric-Brown, A. (1977) Turnover rates in insular biogeography: effect of immigration on extinction. Ecology, 58, 445-449.

Buckley, R. (1982) The habitat-unit model of island biogeography. Journal of Biogeography, 9(4), 339-344.

Burke, R. J., Fitzsimmons, J. M., & Kerr, J. T. (2011) A mobility index for Canadian butterfly species based on naturalists' knowledge. Biodiversity and Conservation, 20(10), 2273-2295.

Burnham, K. P., & Anderson, D. R. (1998) Model selection and inference: A practical information-theoretic approach. New York: Springer-Verlag.

Burnham, K. P., & Anderson, D. R. (2004) Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods & Research, 33, 261-304.

Burns, K., Berg, J., Bialynicka-Birula, A., Kratchmer, S., & Shortt, K. (2010) Tree diversity on islands: Assembly rules, passive sampling and the theory of island biogeography. Journal of Biogeography, 37, 1876-1883.

Buzas, M. A., & Hayek, L. C. (1996) Biodiversity resolution: an integrated approach. Biodiversity Letters, 4, 40-43.

Camargo, J. A. (1992) Can dominance influence stability in competitive interactions?. Oikos, 64, 605-609.

Cardé, R. T., & Willis, M. A. (2008) Navigational strategies used by insects to find distant, wind-borne sources of odor. Journal of chemical ecology, 34(7), 854-866.

Caswell, H. (1976) Community structure: a neutral model analysis. Ecological Monographs, 46(3), 327-354.

Chadde, S. W. (2013) Minnesota flora: An illustrated guide to the vascular plants of Minnesota. Scotts Valley, CA: CreateSpace Publishing.

Challis, R. J., Kumar, S., Dasmahapatra, K. K. K., Jiggins, C. D., & Blaxter, M. (2016) Lepbase: The Lepidopteran genome database. BioRxiv, 56994.

Chamberlin, T. C. (1890) The method of multiple working hypotheses. Science, 15(366), 92-96.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs, 84(1), 45-67.

Chase, J. M., Blowes, S. A., Knight, T. M., Gerstner, K., & May, F. (2020) Ecosystem decay exacerbates biodiversity loss with habitat loss. Nature, 584(7820), 238-243.

Chase, J. M., Gooriah, L., & May, F. (2019) A framework for disentangling ecological mechanisms underlying the island species-area. Frontiers of Biogeography, 11(1), e40844.

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002) Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. Journal of Agricultural, Biological, and Environmental Statistics, 7(3), 361.

Coleman, B. D. (1981) On random placement and species-area relations. Mathematical Biosciences 54, 191-215.

Colwell, R. K., & Winkler, D. W. (1984) A null model for null models in biogeography. In: Strong Jr, D. R., Simberloff, D., Abele, L. G., & Thistle, A. B. (Eds.) *Ecological communities: conceptual issues and the evidence*. Princeton University Press, Princeton. (pp. 344-359).

Connor, E. F., & McCoy, E. D. (1979) The statistics and biology of the species-area relationship. The American Naturalist, 113, 791-833.

Connor, E. F., & Simberloff, D. (1978) Species number and compositional similarity of the Galapagos flora and avifauna. Ecological Monographs, 48, 219-248.

Coulon, A., Cosson, J. F., Angibault, J. M., Cargnelutti, B., Galan, M., Morellet, N., … & Hewison, A. J. M. (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: An individual-based approach. Molecular Ecology, 13(9), 2841-2850.

Coyne, J. A., & Orr, H. A. (2004) Speciation. Sinauer Associates, Sunderland.

Craney, T. A., & Surles, J. G. (2002) Model-dependent variance inflation factor cutoff values. Quality Engineering, 14(3), 391-403.

Crispo, E., Bentzen, P., Reznick, D. N., Kinnison, M. T., & Hendry, A. P. (2006) The relative influence of natural selection and geography on gene flow in guppies. Molecular Ecology, 15, 49-62.

Croxton, P., Hann, J., Greatorex-Davies, J., & Sparks, T. (2005) Linear hotspots? The floral and butterfly diversity of green lanes. Biological Conservation, 121, 579-584.

Cushman, S. A., McKelvey, K. S., Hayden, J., & Schwartz, M. K. (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. The American Naturalist, 168(4), 486-499.

Cushman, S., Wasserman, T., Landguth, E., & Shirk, A. (2013) Re-evaluating causal modeling with Mantel tests in landscape genetics. Diversity, 5(1), 51-72.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011) The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.

Darwin, C. (1859) On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life. London, UK: John Murray.

Dawkins, R. (1976) The selfish gene. Oxford University Press, Oxford.

De Kort, H., Vandepitte, K., Mergeay, J., Mijnsbrugge, K. V., & Honnay, O. (2015) The population genomic signature of environmental selection in the widespread insect-

pollinated tree species Frangula alnus at different geographical scales. Heredity, 115(5), 414-425.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013) Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. Molecular Ecology, 22(5), 1383-1399.

Deane, D. C., Nozohourmehrabad, P., Boyce, S. S., & He, F. (2020) Quantifying factors for understanding why several small patches host more species than a single large patch. Biological Conservation, 249, 108711.

Debinski, D. M., & Holt, R. D. (2000) A survey and overview of habitat fragmentation experiments. Conservation Biology, 14, 342-355.

Dennis, R. L., Shreeve, T. G., & Van Dyck, H. (2003) Towards a functional resource-based concept for habitat: A butterfly biology viewpoint. Oikos, 417-426.

Devlin, B., & Roeder, K. (1999) Genomic control for association studies. Biometrics, 55(4), 997-1004.

Diamond, J. M. (1972). Biogeographic kinetics: estimation of relaxation times for avifaunas of southwest Pacific islands. Proceedings of the National Academy of Sciences, 69(11), 3199-3203.

Diamond, J. M. (1975) The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. Biological Conservation, 7, 129-146.

Dobzhansky, T. G. (1937) Genetics and the origin of species. New York: Columbia University Press.

Dover, J. W., & Fry, G. L. A. (2001) Experimental simulation of some visual and physical components of a hedge and the effects on butterfly behaviour in an agricultural landscape. Entomologia Experimentalis et Applicata, 100(2), 221-233.

Dover, J., & Settele, J. (2009). The influences of landscape structure on butterfly distribution and movement: a review. Journal of Insect Conservation, 13(1), 3-27.

Dreisig, H. (1995) Ideal free distributions of nectar foraging bumblebees. Oikos, 72, 161-172

Dunn, P. K., & Smyth, G. K. (2005) Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing, 15(4), 267-280.

Dunning, J. B., Danielson, B. J., & Pulliam, H. R. (1992) Ecological processes that affect populations in complex landscapes. Oikos, 65(1), 169-175.

Dupuis, J. R., & Sperling, F. A. H. (2015) Repeated reticulate evolution in North American *Papilio machaon* group swallowtail butterflies. PLoS One, 10, e0141882.

Dupuis, J. R., & Sperling, F. A. H. (2016) Hybrid dynamics in a species group of swallowtail butterflies. Journal of Evolutionary Biology, 29, 1932-1951.

Dupuis, J. R., Cullingham, C. I., Nielsen, S. E., & Sperling, F. A. H. (2019) Environmental effects on gene flow in a species complex of vagile, hilltopping butterflies. Biological Journal of the Linnean Society, 127(2), 417-428.

Dupuis, J. R., Mori, B. A., & Sperling, F. A. H. (2016) *Trogus* parasitoids of *Papilio* butterflies undergo extended diapause in western Canada (Hymenoptera, Ichneumonidae) Journal of Hymenoptera Research, 50, 179.

Earl, D. A., & vonHoldt, B. M. (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4(2), 359-361.

Ehrlich, P. R., & Davidson, S. E. (1960) Techniques for capture-recapture studies of Lepidoptera populations. Journal of the Lepidopterists' Society, 14(4), 227-229.

Epps, C. W., Wehausen, J. D., Bleich, V. C., Torres, S. G., & Brashares, J. S. (2007) Optimizing dispersal and corridor models using landscape genetics. Journal of Applied Ecology, 44(4), 714-724.

Erhardt, A. (1985) Diurnal Lepidoptera: sensitive indicators of cultivated and abandoned grassland. Journal of Applied Ecology, 22(3), 849-861.

Evanno, G., Regnaut, S., & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 14(8), 2611-2620.

Ewers, R. M., & Didham, R. K. (2006) Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews, 81, 117-142.

Fahrig, L. (2003) Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics, 34(1), 487-515.

Fahrig, L. (2013) Rethinking patch size and isolation effects: The habitat amount hypothesis. Journal of Biogeography, 40, 1649-1663.

Fahrig, L. (2017) Ecological responses to habitat fragmentation *per se*. Annual Review of Ecology, Evolution, and Systematics, 48, 1-23.

Fahrig, L. (2020) Why do several small patches hold more species than few large patches? Global Ecology and Biogeography, 29(4), 615-628.

Fahrig, L., & Paloheimo, J. E. (1987) Interpatch dispersal of the cabbage butterfly. Canadian Journal of Zoology, 65(3), 616-622.

Farneda, F. Z., Grelle, C. E., Rocha, R., Ferreira, D. F., López-Baucells, A., & Meyer, C. F. (2020) Predicting biodiversity loss in island and countryside ecosystems through the lens of taxonomic and functional biogeography. Ecography, 43(1), 97-106.

Fick, S. E., & Hijmans, R. J. (2017) WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37(12), 4302-4315.

Fleishman, E., & Murphy. D. D (2009) A realistic assessment of the indicator potential of butterflies and other charismatic taxonomic groups. Conservation Biology, 23(5), 1109-1116.

Fletcher, R. J., Didham, R. K., Banks-Leite, C., Barlow, J., Ewers, R. M., Rosindell, J., ..., & Melo, F. P. (2018) Is habitat fragmentation good for biodiversity? Biological Conservation, 226, 9-15.

Foll, M., & Gaggiotti, O. (2006) Identifying the environmental factors that determine the genetic structure of populations. Genetics, 174, 875-891.

Foll, M., & 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.

Fourtune, L., Prunier, J. G., Paz-Vinas, I., Loot, G., Veyssière, C., & Blanchet, S. (2018) Inferring causalities in landscape genetics: An extension of Wright's causal modeling to distance matrices. The American Naturalist, 191(4), 491-508.

Franzén, M., & Betzholtz, P. E. (2012). Species traits predict island occupancy in noctuid moths. Journal of Insect Conservation, 16(2), 155-163.

Franzén, M., Schweiger, O., & Betzholtz, P. E. (2012) Species-area relationships are controlled by species traits. PloS one, 7(5), e37359.

Freckleton, R. P. (2002) On the misuse of residuals in ecology: regression of residuals vs. multiple regression. Journal of Animal Ecology, 71(3), 542-545.

Fretwell, S. D. (1969). On territorial behavior and other factors influencing habitat distribution in birds. Acta biotheoretica, 19(1), 45-52.

Friberg, M., Bergman, M., Kullberg, J., Wahlberg, N., & Wiklund, C. (2008) Niche separation in space and time between two sympatric sister species: A case of ecological pleiotropy. Evolutionary Ecology, 22, 1-18.

Frichot, E., & François, O. (2015) LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution, 6(8), 925-929.

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. Molecular Biology & Evolution, 30(7), 1687-1699.

Fukamachi, K., Iida, S., & Nakashizuka, T. (1996) Landscape patterns and plant species diversity of forest reserves in the Kanto region, Japan. Vegetatio, 124, 107-114.

Garlick, K. M. (2007) Visual and olfactory sensory systems employed by monarch butterflies (*Danaus plexippus*) to locate their milkweed host plants. M.Sc Thesis, Queen's University, Kingston, ON, Canada.

Gavish, Y., Ziv, Y., & Rosenzweig, M.L. (2012) Decoupling fragmentation from habitat loss for spiders in patchy agricultural landscapes. Conservation Biology, 26(1), 150-159.

Gehring, T. M., & Swihart, R. K. (2003) Body size, niche breadth, and ecologically scaled responses to habitat fragmentation: mammalian predators in an agricultural landscape. Biological Conservation, 109(2), 283-295.

Gilpin, M. E., & Diamond, J. M. (1976) Calculation of immigration and extinction curves from the species-area-distance relation. Proceedings of the National Academy of Sciences, 73, 4130

Gonzalez, A. (2000) Community relaxation in fragmented landscapes: the relation between species richness, area and age. Ecology Letters, 3, 441-448.

Gosselin, F. (2006) An assessment of the dependence of evenness indices on species richness. Journal of Theoretical Biology, 242(3), 591-597.

Gotelli, N. J. G., & Graves, G. R. (1996) Null Models in Ecology. Smithsonian Institution Press.

Grace, J. B. (2003) Comparing groups using structural equations. In Pugesek, B. H., Tomer, A. & Eye, A. (Eds.) *Structural equation modelling*. Cambridge, UK: Cambridge University Press. (pp. 291-296).

Grace, J. B. (2006) Structural equation modeling and natural systems. Cambridge University Press, Cambridge.

Grace, J. B. (2008) Structural equation modeling for observational studies. Journal of Wildlife Management, 72, 14-22.

Grace, J. B., & Pugesek, B. H. (1997) A structural equation model of plant species richness and its application to a coastal wetland. The American Naturalist, 149, 436-460.

Grace, J. B., & Pugesek, B. H. (1998) On the use of path analysis and related procedures for the investigation of ecological problems. The American Naturalist, 152, 151-159.

Grace, J. B., Anderson, T. M., Olff, H., & Scheiner, S. M. (2010) On the specification of structural equation models for ecological systems. Ecological Monographs, 80, 67-87.

Grant, T. J., Parry, H. R., Zalucki, M. P., & Bradbury, S. P. (2018) Predicting monarch butterfly (Danaus plexippus) movement and egg-laying with a spatially-explicit agent-based model: the role of monarch perceptual range and spatial memory. Ecological Modelling, 374, 37-50.

Greenwood, P. J. (1980) Mating systems, philopatry and dispersal in birds and mammals. Animal Behaviour, 28, 1140-1162.

Günther, T., & Coop, G. (2013) Robust identification of local adaptation from allele frequencies. Genetics, 195(1), 205-220.

Haddad, N. M. (1999) Corridor use predicted from behaviors at habitat boundaries. The American Naturalist, 153(2), 215-227.

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., ..., & Cook, W. M. (2015) Habitat fragmentation and its lasting impact on Earth's ecosystems. Science Advances, 1(2), e1500052.

Haddad, N. M., Gonzalez, A., Brudvig, L. A., Burt, M. A., Levey, D. J., & Damschen, E. I. (2017) Experimental evidence does not support the habitat amount hypothesis. Ecography, 40, 48-55.

Hadley, A. S., & Betts, M. G. (2016) Refocusing habitat fragmentation research using lessons from the last decade. Current Landscape Ecology Reports, 1(2), 55-66.

Haila, Y. (1986). On the semiotic dimension of ecological theory: the case of island biogeography. Biology and Philosophy, 1(4), 377-387.Haila Y (1990) Toward an ecological definition of an island: a northwest European perspective. J Biogeogr 17:561-568

Haila, Y. (2002) A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. Ecological Applications, 12, 321-334.

Haila, Y., & Järvinen, O. (1983) Land bird communities on a Finnish island: species impoverishment and abundance patterns. Oikos, 41(2), 255-273.

Haila, Y., Jarvinen, O., & Kuusela, S. (1983) Colonization of islands by land birds: prevalence functions in a Finnish archipelago. Journal of Biogeography, 10(6), 499-531.

Hall, P. W., Jones, C. D.. Guidotti, A., & Hubley, B. (2014) The ROM field guide to butterflies of Ontario. Royal Ontario Museum Press, Toronto.

Hanski, I. (1988) Four kinds of extra long diapause in insects: A review of theory and observations. Annales Zoologici Fennici, 25, 37-53.

Hanski, I. (1994) A practical model of metapopulation dynamics. Journal of Animal Ecology, 63, 151-162.

Hanski, I. (1998) Metapopulation dynamics. Nature, 396, 41-49.

Hanski, I. (1999) Metapopulation ecology. Oxford University Press, Oxford.

Hanski, I. (2015) Habitat fragmentation and species richness. Journal of Biogeography, 42, 989-993.

Hanski, I., & Gyllenberg, M. (1993) Two general metapopulation models and the core-satellite species hypothesis. American Naturalist, 142, 17-41.

Hanski, I., Moilanen, A., Pakkala, T., & Kuussaari, M. (1996) The quantitative incidence function model and persistence of an endangered butterfly metapopulation. Conservation Biology, 10(2), 578-590.

Hanski, I., Alho, J., & Moilanen, A. (2000) Estimating the parameters of survival and migration of individuals in metapopulations. Ecology, 81(1), 239-251.

Hanski, I., Zurita, G. A., Bellocq, M. I., & Rybicki, J. (2013) Species-fragmented area relationship. Proceedings of the National Academy of Sciences, 110(31), 12715-12720.

Harr, B. (2006) Genomic islands of differentiation between house mouse subspecies. Genome Research, 16, 730-737.

Harrison, S., & Bruna, E. (1999). Habitat fragmentation and large-scale conservation: what do we know for sure?. Ecography, 22(3), 225-232.Harrison, S. (1989) Long-distance dispersal and colonization in the bay checkerspot butterfly, Euphydryas editha bayensis. Ecology, 70(5), 1236-1243.

Hawkins, B. A., & Porter, E. E. (2003) Does herbivore diversity depend on plant diversity? The case of California butterflies. The American Naturalist, 161(1), 40-49.

He, F., & Legendre, P. (2002) Species diversity patterns derived from species-area models. Ecology, 83, 1185-1198.

He, F., & Hubbell, S. P. (2011). Species–area relationships always overestimate extinction rates from habitat loss. Nature, 473(7347), 368-371.

Henle, K., Davies, K. F., Kleyer, M., Margules, C., & Settele, J. (2004) Predictors of species sensitivity to fragmentation. Biodiversity and Conservation, 13, 207-251.

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2011) Package 'dismo'. Available online at: http://cran.r-project.org/web/packages/dismo/index.html

Hijmans, R. J., & van Etten, J. (2012) raster: Geographic analysis and modeling with raster data. R package version 2.0-12. http://CRAN.R-project.org/package=raster

Hill, J. K., Thomas, C. D., & Blakeley, D. S. (1999) Evolution of flight morphology in a butterfly that has recently expanded its geographic range. Oecologia, 121(2), 165-170.

Hill, M. O. (1973) Diversity and evenness: a unifying notation and its consequences. Ecology, 54(2), 427-432.

Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., ..., & Lewandowska, A. M. (2018) Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. Journal of Applied Ecology, 55(1), 169-184.

Hortal, J., Triantis, K.A., Meiri, S., Thébault, E., & Sfenthourakis, S. (2009) Island species richness increases with habitat diversity. The American Naturalist, 174, E205-E217.

Humphries, C. J., Williams, P. H., & Vane-Wright, R. I. (1995) Measuring biodiversity value for conservation. Annual Review of Ecology and Systematics, 26(1), 93-111.

Hurulbert, S. (1971) The nonconcept of species diversity: a critique and alternatives parameters. Ecology, 52, 577-586.

Ibisch, P. L., Hoffmann, M. T., Kreft, S., Pe'er, G., Kati, V., Biber-Freudenberger, L., ..., & Selva, N. (2016) A global map of roadless areas and their conservation status. Science, 354(6318), 1423-1427.

Itescu, Y. (2019) Are island-like systems biologically similar to islands? A review of the evidence. Ecography, 42(7), 1298-1314.

Izsák, J., & Papp, L. (2000) A link between ecological diversity indices and measures of biodiversity. Ecological Modelling, 130(1), 151-156.

Johnson, D. H. (1980). The comparison of usage and availability measurements for evaluating resource preference. Ecology, 61(1), 65-71.

Jombart, T. (2008) adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403-1405.

Jombart, T., Devillard, S., & Balloux, F. (2010) Discriminate analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics, 11, 94. pmid:20950446

Jost, L. (2006) Entropy and diversity. Oikos, 113(2), 363-375.

Kadmon, R., & Allouche, O. (2007) Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: A unification of island biogeography and niche theory. The American Naturalist, 170, 443-454.

Karger, D. N., Weigelt, P., Amoroso, V. B., Darnaedi, D., Hidayat, A., Kreft, H., & Kessler, M. (2014) Island biogeography from regional to local scales: evidence for a spatially scaled echo pattern of fern diversity in the Southeast Asian archipelago. Journal of Biogeography, 41(2), 250-260.

Keller, D., & Holderegger, R. (2013) Damselflies use different movement strategies for short- and long-distance dispersal. Insect Conservation and Diversity, 6(5), 590-597.

Kelly, B. J., Wilson, J. B., & Mark, A. F. (1989) Causes of the species-area relation: a study of islands in Lake Manapouri, New Zealand. The Journal of Ecology, 77(4), 1021-1028.

Keylock, C. (2005) Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy. Oikos, 109(1), 203-207.

Kinoshita, M., Shimada, N. A. O. K. O., & Arikawa, K. (1999) Colour vision of the foraging swallowtail butterfly *Papilio xuthus*. Journal of Experimental Biology, 202(2), 95-102.

Kinoshita, M., Shimohigasshi, M., Tominaga, Y., Arikawa, K., & Homberg, U. (2015) Topographically distinct visual and olfactory inputs to the mushroom body in the Swallowtail butterfly, Papilio xuthus. Journal of Comparative Neurology, 523(1), 162-182.

Kitahara, M., Yumoto, M., & Kobayashi, T. (2008) Relationship of butterfly diversity with nectar plant species richness in and around the Aokigahara primary woodland of Mount Fuji, central Japan. Biodiversity and Conservation, 17(11), 2713-2734.

Kohn, D. D., & Walsh, D. M. (1994) Plant species richness - the effect of island size and habitat diversity. The Journal of Ecology, 82, 367-377.

Kondla, N. G. (1981) Type localities of the Badlands Old World Swallowtail in Alberta. Blue Jay, 39, 144.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015) Community assembly, coexistence and the environmental filtering metaphor. Functional Ecology, 29, 592-599.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015) Community assembly, coexistence and the environmental filtering metaphor. Functional Ecology, 29, 592-599.

Kukal, O., Ayres, M. P., & Scriber, J. M. (1991) Cold tolerance of the pupae in relation to the distribution of swallowtail butterflies. Canadian Journal of Zoology, 69(12), 3028-3037.

Lamb, C. T., Mowat, G., Reid, A., Smit, L., Proctor, M., McLellan, B. N., ..., & Boutin, S. (2018) Effects of habitat quality and access management on the density of a recovering grizzly bear population. Journal of Applied Ecology, 55(3), 1406-1417.

Larsen, T.H., Lopera, A., & Forsyth, A. (2008) Understanding trait-dependent community disassembly: dung beetles, density functions, and forest fragmentation. Conservation Biology, 22(5), 1288-1298.

Laurance, W. F. (2008) Theory meets reality: how habitat fragmentation research has transcended island biogeographic theory. Biological Conservation, 141(7), 1731-1744.

Lee, C. R., & Mitchell-Olds, T. (2011) Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. Molecular Ecology, 20, 4631-5464.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., … Loreau, M. (2004) The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters, 7, 601-613.

Lens, L., Van Dongen, S., Norris, K., Githiru, M., & Matthysen, E. (2002) Avian persistence in fragmented rainforest. Science, 298, 1236-1238.

Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America, 15, 237-240.

Lewthwaite, J. M. M., Angert, A. L., Kembel, S. W., Goring, S. J., Davies, T. J., Mooers, A. Ø., … & Kerr, J. T. (2018) Canadian butterfly climate debt is significant and correlated with range size. Ecography, 41(12), 2005-2015.

Li, C. C. (1975) Path analysis: A primer. Boxwood Press, Pacific Grove.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … & Durbin, R. (2009) The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), 2078-2079.

Lomolino, M. V. (2000). A call for a new paradigm of island biogeography. Global Ecology and Biogeography, 9(1), 1-6.

Lomolino, M. V. (2000) Ecology's most general, yet protean pattern: the species-area relationship. Journal of Biogeography, 27(1), 17-26.

Lomolino, M., & Weiser, M. (2001) Towards a more general species-area relationship: Diversity on all islands, great and small. Journal of Biogeography, 28, 431-445.

Lotterhos, K. E., & Whitlock, M. C. (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Molecular Ecology, 23(9), 2178-2192.

Ma, M. (2005) Species richness vs evenness: independent relationship and different responses to edaphic factors. Oikos, 111(1), 192-198.

MacArthur, R. H., & MacArthur, J. W. (1961) On bird species diversity. Ecology, 42, 594-598.

MacArthur, R. H., & Wilson, E. O. (1963) An equilibrium theory of insular zoogeography. Evolution, 17, 373-387.

MacDonald, Z. G., Nielsen, S. E., & Acorn, J. H. (2017) Negative relationships between species richness and evenness render common diversity indices inadequate for assessing long-term trends in butterfly diversity. Biodiversity and Conservation, 26(3), 617-629.

MacDonald, Z. G., Anderson, I. D., Acorn, J. H., & Nielsen, S. E. (2018a) Decoupling habitat fragmentation from habitat loss: Butterfly species mobility obscures fragmentation effects in a naturally fragmented landscape of lake islands. Oecologia, 186, 11-27.

MacDonald, Z. G., Anderson, I. D., Acorn, J. H., & Nielsen, S. E. (2018b) The theory of island biogeography, the sample-area effect, and the habitat diversity hypothesis: complementarity in a naturally fragmented landscape of lake islands. Journal of Biogeography, 45(12), 2730-2743.

MacDonald, Z. G., Acorn, J. H., Zhang, J., & Nielsen, S. E. (2019) Perceptual Range, Targeting Ability, and Visual Habitat Detection by Greater Fritillary Butterflies Speyeria cybele (Lepidoptera: Nymphalidae) and Speyeria atlantis. Journal of Insect Science, 19(4), 1-10.

MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, J. H., Nielsen, S. E., & Sperling, F. A. (2020) Gene flow and climate-associated genetic variation in a vagile habitat specialist. Molecular Ecology, 29(20), 3889-3906.

MacDonald, Z. G., Deane, D. C., Lamb, C. T., He, F., Acorn, J. H., Nielsen, S. E. (2021) Distinguishing effects of area *per se* and isolation from the sample-area effect for true islands and habitat fragments. Ecography, In Press.

Magurran, A. E. (2013) Measuring biological diversity. Wiley, New Jersey.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003) Landscape genetics: Combining landscape ecology and population genetics. Trends in Ecology & Evolution, 18(4), 189-197.

Martin, C. A. (2018) An early synthesis of the habitat amount hypothesis. Landscape Ecology 33(11), 1831-1835.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet Journal, 17(1), 10-12.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., … & Jiggins, C. D. (2013) Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Research, 23(11), 1817-1828.

Martins, K., Gugger, P. F., Llanderal-Mendoza, J., González-Rodríguez, A., Fitz-Gibbon, S. T., Zhao, J. L., … & Sork, V. L. (2018) Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of *Quercus rugosa*. Evolutionary Applications, 11(10), 1842-1858.

Matthews, T. J., Triantis, K. A., Rigal, F., Borregaard, M. K., Guilhaumon, F., & Whittaker, R. J. (2016) Island species-area relationships and species accumulation curves are not equivalent: An analysis of habitat island datasets. Global Ecology and Biogeography, 25, 607-618.

May, R. M. (1975) Island biogeography and the design of wildlife preserves. Nature, 254, 177-178.

Mayr, E. (1963) Animal species and evolution. Cambridge, MA: Harvard University Press.

McRae, B. H. (2006) Isolation by resistance. Evolution, 60, 1551-1561.

McRae, B. H., & Beier, P. (2007) Circuit theory predicts gene flow in plant and animal populations. Proceedings of the National Academy of Sciences USA, 104, 19885-19890.

Meirmans, P. G. (2012) The trouble with isolation by distance. Molecular Ecology, 21(12), 2839-2846.

Melbourne, B.A., Davies, K.F., Margules, C.R., Lindenmayer, D.B., Saunders, D.A., Wissel, C., & Henle, K. (2004) Species survival in fragmented landscapes: where to from here? Biodiversity and Conservation, 13, 275-284.

Mendenhall, C. D., Karp, D. S., Meyer, C. F., Hadly, E. A., & Daily, G. C. (2014) Predicting biodiversity change and averting collapse in agricultural landscapes. Nature, 509, 213-217.

Merckx, T., & Van Dyck, H. (2007). Habitat fragmentation affects habitat-finding ability of the speckled wood butterfly, Pararge aegeria L. Animal Behaviour, 74(4), 1029-1037.

Moilanen, A., & Nieminen, M. (2002) Simple connectivity measures in spatial ecology. Ecology, 83(4), 1131-1145.

Mulder, C. P. H., Bazeley-White, E., Dimitrakopoulos, P. G., Hector, A., Scherer-Lorenzen, M., & Schmid, B. (2004) Species evenness and productivity in experimental plant communities. Oikos, 107(1), 50-63.

Murcia, C. (1995). Edge effects in fragmented forests: implications for conservation. Trends in ecology & evolution, 10(2), 58-62.

Narum, S. R., & Hess, J. E. (2011) Comparison of FST outlier tests for SNP loci under selection. Molecular Ecology Resources, 11, 184-194.

Niitepõld, K., Smith, A. D., Osborne, J. L., Reynolds, D. R., Carreck, N. L., Martin, A. P., ... & Hanski, I. (2009) Flight metabolic rate and Pgi genotype influence butterfly dispersal rate in the field. Ecology, 9.

Nilsson, S. G., Bengtsson, J., & As, S. (1988) Habitat diversity or area *per se*? Species richness of woody plants, carabid beetles and land snails on islands. The Journal of Animal Ecology, 57, 685-704.

Nosil, P. (2004) Reproductive isolation caused by visual predation on migrants between divergent environments. Proceedings of the Royal Society of London. Series B: Biological Sciences, 271(1547), 1521-1528.

Nosil, P. (2012) Ecological speciation.Oxford University Press, Oxford.

Nosil, P., Vines, T. H., & Funk, D. J. (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution, 59, 705-719.

Noss, R.F. (1991) Landscape connectivity: different functions at different scales. In: Soulé, M. E., Hay, K., Csuti, B., Harris, L., Scott, J. M., Brown, B., ..., & Chadwick, D. (Eds.) *Landscape linkages and biodiversity*. Island Press.

Nowicki, P., Pepkowska, A., Kudlek, J., Skórka, P., Witek, M., Settele, J., & Woyciechowski, M. (2007) From metapopulation theory to conservation recommendations: lessons from spatial occurrence and abundance patterns of Maculinea butterflies. Biological Conservation, 140, 119-129.

Nowicki, P., Settele, J., Henry, P. Y., & Woyciechowski, M. (2008) Butterfly monitoring methods: the ideal and the real world. Israel Journal of Ecology and Evolution, 54(1), 69-88.

Nowicki, P., Vrabec, V., Binzenhöfer, B., Feil, J., Zakšek, B., Hovestadt, T., & Settele, J. (2014) Butterfly dispersal in inhospitable matrix: rare, risky, but long-distance. Landscape Ecology, 29(3), 401-412.

O'Connell, K. A., Mulder, K. P., Maldonado, J., Currie, K. L., & Ferraro, D. M. (2019) Sampling related individuals within ponds biases estimates of population structure in a pond-breeding amphibian. Ecology and Evolution, 9(6), 3620-3636.

Öckinger, E., Franzén, M., Rundlöf, M., & Smith, H.G. (2009) Mobility-dependent effects on species richness in fragmented landscapes. Basic and Applied Ecology, 10, 573-578.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. A. S. S. (2007) The vegan package. Community Ecology Package, 10, 631-637.

Ontario Ministry of Natural Resources. (2006) Lake of the woods conservation reserve (C2366) resource management plan. Retrieved from http://www.ontla.on.ca/library/repository/mon/15000/267564.pdf

Orrock, J. L., & Watling, J. I. (2010) Local community size mediates ecological drift and competition in metacommunities. Proceedings of the Royal Society B: Biological Sciences, 277, 2185-2191.

Ovaskainen, O. (2002) Long-term persistence of species and the SLOSS problem. Journal of Theoretical Biology, 218(4), 419-433.

Papaj, D. R. (1986) Conditioning of leaf-shape discrimination by chemical cues in the butterfly, *Battus philenor*. Animal Behaviour, 34(5), 1281-1288.

Pebesma, E. J., & Bivand, R. S. (2005) Classes and methods for spatial data in R. R News, 5(2), 9-13. https://cran.r-project.org/doc/Rnews/.

Pellet, J., Bried, J. T., Parietti, D., Gander, A., Heer, P. O., Cherix, D., & Arlettaz, R. (2012) Monitoring butterfly abundance: beyond Pollard walks. PloS one, 7(7), e41396.

Peterman, W. E. (2018) ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. Methods in Ecology and Evolution, 9(6), 1638-1647.

Peterman, W. E., Connette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014) Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. Molecular Ecology, 23(10), 2402-2413.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One, 7(5), e37135.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190, 231-259

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015) Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. Molecular Ecology, 24(1), 54-69.

Platt, J. R. (1998) Strong inference. In Maritime Archaeology (pp. 155-166) Springer, Boston, MA.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One, 7(2), e32253.

Pollard, E. (1977) A method for assessing changes in the abundance of butterflies. Biological Conservation, 12(2), 115-134.

Pollard, E. (1988) Temperature, rainfall and butterfly numbers. Journal of Applied Ecology, 819-828.

Pollard, E., Moss, D., & Yates, T. J. (1995) Population trends of common British butterflies at monitored sites. Journal of Applied Ecology, 9-16.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. Genetics, 155(2), 945-959.

Prugh, L. R. (2009) An evaluation of patch connectivity measures. Ecological Applications, 19(5), 1300-1310.

Prugh, L. R., Hodges, K. E., Sinclair, A. R., & Brashares, J. S. (2008) Effect of habitat area and isolation on fragmented animal populations. Proceedings of the National Academy of Sciences 105(52), 20770-20775.

Pruisscher, P., Nylin, S., Gotthard, K., & Wheat, C. W. (2018) Genetic variation underlying local adaptation of diapause induction along a cline in a butterfly. Molecular Ecology, 27(18), 3613-3626.

Przewlocki, R., Lason, W., Konecka, A. M., Gramsch, C., Herz, A., & Reid, L. D. (1983) The opioid peptide dynorphin, circadian rhythms, and starvation. Science, 219(4580), 71-73.

Purvis, A., & Hector, A. (2000) Getting the measure of biodiversity. Nature, 405(6783), 212-219.

Quinlan, A. R., & Hall, I. M. (2010) BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics, 26(6), 841-842.

Quinn, J. F., Harrison, S. P. (1988) Effects of habitat fragmentation and isolation on species richness: evidence from biogeographic patterns. Oecologia 75,132-140.

Quinn, S. L., Wilson, J. B., & Mark, A. F. (1987) The island biogeography of Lake Manapouri, New Zealand. Journal of Biogeography, 14(6), 569-581.

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rausher, M. D. (1978) Search image for leaf shape in a butterfly. Science, 200(4345), 1071-1073.

Rausher, M. D. (1981) Host plant selection by *Battus philenor* butterflies: the roles of predation, nutrition, and plant chemistry. Ecological monographs, 51(1), 1-20.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015) A practical guide to environmental association analysis in landscape genomics. Molecular Ecology, 24(17), 4348-4370.

Richardson, J. L., Brady, S. P., Wang, I. J., & Spear, S. F. (2016) Navigating the pitfalls and promise of landscape genetics. Molecular Ecology, 25(4), 849-863.

Ricketts, T. H. (2001) The matrix matters: effective isolation in fragmented landscapes. The American Naturalist, 158, 87-99.

Ricotta, C. (2003) Parametric scaling from species relative abundances to absolute abundances in the computation of biological diversity: a first proposal using Shannon's entropy. Acta Biotheoretica, 51(3), 181-188.

Riva, F., Acorn, J. H., & Nielsen, S. E. (2018) Narrow anthropogenic corridors direct the movement of a generalist boreal butterfly. Biology Letters, 14(2), 20170770.

Riva, F., Pinzon, J., Acorn, J. H., & Nielsen, S. E. (2020) Composite effects of cutlines and wildfire result in fire refuges for plants and butterflies in boreal treed peatlands. Ecosystems, 23, 1-13.

Robinson, W. S. (1950) Ecological Correlations and the Behavior of Individuals. American Sociological Review, 15(3), 351-357.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019) Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular Ecology, 28(21), 4737-4754.

Roland, J., & Taylor, P. D. (1997) Insect parasitoid species respond to forest structure at different spatial scales. Nature, 386, 710.

Rosenzweig, M. L. (1995) Species diversity in space and time. Cambridge University Press, Cambridge.

Rosenzweig, M. L. (2004) Applying species-area relationships to the conservation of species diversity. In Lomolino, M. V., & Heaney, M. V. (Eds.) *Frontiers in biogeography, new directions in the geography of nature*. Sinauer Associates. (pp. 325-344).

Rosin, Z. M., Skórka, P., Lenda, M., Moroń, D., Sparks, T. H., & Tryjanowski, P. (2011). Increasing patch area, proximity of human settlement and larval food plants positively affect the occurrence and local population size of the habitat specialist butterfly *Polyommatus coridon* (Lepidoptera: Lycaenidae) in fragmented calcareous grasslands. European Journal of Entomology, 108(1).

Rosseel, Y. (2012) Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA) Journal of Statistical Software, 48(2), 1-36.

Rousset, F. (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics, 145, 1219-1228.

Row, J. R., Knick, S. T., Oyler-McCance, S. J., Lougheed, S. C., & Fedy, B. C. (2017) Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. Ecology and Evolution, 7(11), 3751-3761.

Roy, D. B., Rothery, P., Moss, D., Pollard, E., & Thomas, J. A. (2001) Butterfly numbers and weather: predicting historical trends in abundance and the future effects of climate change. Journal of Animal Ecology, 70(2), 201-217.

Ruiz-Gonzalez, A., Cushman, S. A., Madeira, M. J., Randi, E., & Gómez-Moliner, B. J. (2015) Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape. Molecular Ecology, 24, 5110-5129.

Rutowski, R. L. (2003) Visual ecology of adult butterflies. Butterflies: ecology and evolution taking flight. In Boggs, C. L., Watt, W. B., Ehrlich, P. R., & Ehrlich, P. R. (Eds.) University of Chicago Press, Chicago. (pp. 9-25).

Ryan, S. F., Valella, P., Thivierge, G., Aardema, M. L., & Scriber, J. M. (2018) The role of latitudinal, genetic and temperature variation in the induction of diapause of *Papilio glaucus* (Lepidoptera: Papilionidae) Insect Science, 25(2), 328-336.

Rybicki, J., & Hanski, I. (2013) Species-area relationships and extinctions caused by habitat loss and fragmentation. Ecology Letters, 16, 27-38.

Ryder, O. A. (1986) Species conservation and systematics: The dilemma of subspecies. Trends in Ecology and Evolution, 1, 9-10.

Saarinen, K., Lahti, T., & Marttila, O. (2003) Population trends of Finnish butterflies (Lepidoptera: Hesperioidea, Papilionoidea) in 1991-2000. Biodiversity and Conservation, 12(10), 2147-2159.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. (1998) Inbreeding and extinction in a butterfly metapopulation. Nature, 392, 491-494.

Sánchez-Ramírez, S., Rico, Y., Berry, K. H., Edwards, T., Karl, A. E., Henen, B. T., & Murphy, R. W. (2018) Landscape limits gene flow and drives population structure in Agassiz's desert tortoise (Gopherus agassizii) Scientific Reports, 8(1), 11231.

Santos, A. M., Whittaker, R. J., Triantis, K. A., Borges, P. A., Jones, O. R., Quicke, D. L., & Hortal, J. (2010) Are species-area relationships from entire archipelagos congruent with those of their constituent islands? Global Ecology and Biogeography, 19(4), 527-540.

Saunders, D. A., Hobbs, R. J., & Margules, C. R. (1991). Biological consequences of ecosystem fragmentation: a review. Conservation biology, 5(1), 18-32.

Saura, S. (2021) The Habitat Amount Hypothesis implies negative effects of habitat fragmentation on species richness. Journal of Biogeography, 48(1), 11-22.

Scheiner, S. M. (2003) Six types of species-area curves. Global Ecology and Biogeography, 12(6), 441-447.

Schmeller, D. S., Henry, P. Y., Julliard, R., Gruber, B., Clobert, J., Dziock, F., ..., & Henle, K. (2009) Advantages of volunteer-based biodiversity monitoring in Europe. Conservation Biology, 23(2), 307-316.

Schmucki, R., Pe'Er, G., Roy, D. B., Stefanescu, C., Van Swaay, C. A., Oliver, T. H., ..., & Julliard, R. (2016) A regionally informed abundance index for supporting integrative

analyses across butterfly monitoring schemes. Journal of Applied Ecology, 53(2), 501-510.

Schneider, R. R. (2013) Alberta's natural subregions under a changing climate: Past, present, and future. Biodiversity Management and Climate Change Adaptation Project.

Schoener, T. W. (1976). The species-area relation within archipelagos: models and evidence from island land birds. In 16th International Ornithological Congress, Canberra, Australia, 12 to 17 August 1974 (pp. 629-642). Australian Academy of Sciences.

Schoener, T. W., Spiller, D. A., & Losos, J. B. (2001). Natural restoration of the species-area relation for a lizard after a hurricane. Science, 294(5546), 1525-1528.

Schtickzelle, N., Joiris, A., Van Dyck, H., & Baguette, M. (2007) Quantitative analysis of changes in movement behaviour within and outside habitat in a specialist butterfly. BMC Evolutionary Biology, 7(1), 1-15.

Schultz, C. B., & E. E. Crone. (2001) Edge-mediated dispersal behavior in a prairie butterfly. Ecology. 82, 1879-1892.

Schwartz, M. K., McKelvey, K. S., Cushman, S. A., & Luikart, G. (2010) Landscape genomics: A brief perspective. In Cushman, S. A., & Huettmann, F. (Eds.) *Spatial complexity, informatics, and wildlife conservation*. Springer Publishing, New York. (pp. 165-174).

Scriber, J. M., Maher, E., & Aardema, M. L. (2012) Differential effects of short term winter thermal stress on diapausing tiger swallowtail butterflies (*Papilio* spp.) Insect Science, 19(3), 277-285.

Seger, J., & Brockmann, H. J. (1987) What is bet-hedging? In Harvey, P. H., & Partridge, L. (Eds.), *Oxford surveys in evolutionary biology*. Oxford University Press, Oxford. (pp. 182-211).

Sfenthourakis, S., & Triantis, K. A. (2009) Habitat diversity, ecological requirements of species and the Small Island Effect. Diversity and Distributions, 15, 131-140.

Shafer, C. L. (1990) Nature reserves: island theory and conservation practice. Smithsonian Institution Press.

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2017) A comparison of individual-based genetic distance metrics for landscape genetics. Molecular Ecology Resources, 17(6), 1308-1317.

Shirk, A. J., Wallin, D. O., Cushman, S. A., Rice, C. G., & Warheit, K. I. (2010) Inferring landscape effects on gene flow: A new model selection framework. Molecular Ecology, 19, 3603-3619.

Silberglied, R. E. (1984) Visual communication and sexual selection among butterflies, pp. 207-223. In Vane-Wright, R. I., & Ackery, P. R. (Eds.) *The biology of butterflies*. Princeton University Press, New Jersey.

Simberloff, D. (1976) Species turnover and equilibrium island biogeography. Science 194(4265), 572-578.

Simberloff, D. S. (1988) The contribution of population and community biology to conservation science. Annual Review of Ecology, Evolution, and Systematics, 19, 473-511.

Simberloff, D. S., & Wilson, E. O. (1969) Experimental zoogeography of islands: The colonization of empty islands. Ecology, 50, 278-296.

Simberloff, D. S., & Abele, L. G. (1976) Island biogeography theory and conservation practice. Science, 191, 285-286.

Simberloff, D. S., & Abele, L. G. (1982) Refuge design and island biogeographic theory: Effects of fragmentation. American Naturalist, 120, 41-50.

Simberloff, D., & Gotelli, N. (1984) Effects of insularisation on plant species richness in the prairie-forest ecotone. Biological Conservation, 29(1), 27-46.

Simonson, S. E., Opler, P. A., Stohlgren, T. J., & Chong, G. W. (2001) Rapid assessment of butterfly diversity in a montane landscape. Biodiversity and Conservation, 10, 1369-1386.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. Science, 236, 787-792.

Sork, V. L., Nason, J., Campbell, D. R., & Fernandez, J. F. (1999) Landscape approaches to historical and contemporary gene flow in plants. Trends in Ecology & Evolution, 14(6), 219-224.

Soulé, M. E. (1985) What is conservation biology? BioScience, 35(11), 727-734.

Sparks, T., & Parish, T. (1995) Factors affecting the abundance of butterflies in field boundaries in Swavesey fens, Cambridgeshire, UK. Biological Conservation, 73, 221-227.

Sperling, F. A. H. (1990) Natural hybrids of *Papilio* (Insecta: Lepidoptera): Poor taxonomy or interesting evolutionary problem? Canadian Journal of Zoology, 68(8), 1790-1799.

Sperling, F. A. H. (1987) Evolution of the *Papilio machaon* species group in western Canada (Lepidoptera: Papilionidae) Quaestiones Entomologica, 23, 198-315.

Stasek, D. J., Bean, C., & Crist, T. O. (2008). Butterfly abundance and movements among prairie patches: the roles of habitat quality, edge, and forest matrix permeability. Environmental entomology, 37(4), 897-906.

Stavenga, D. G. (2002) Colour in the eyes of insects. Journal of Comparative Physiology, 188, 337-348.

Stephens, P. A., Pettorelli, N., Barlow, J., Whittingham, M. J., & Cadotte, M. W. (2015) Management by proxy? The use of indices in applied ecology. Journal of Applied Ecology, 52(1), 1-6.

Stevens, G. C. (1986) Dissection of the species-area relationship among wood-boring insects and their host plants. The American Naturalist, 128(1), 35-46.

Stevens, V. M., Trochet, A., Van Dyck, H., Clobert, J., & Baguette, M. (2012) How is dispersal integrated in life histories: a quantitative analysis using butterflies. Ecology Letters, 15(1), 74-86.

Stirling, G., & Wilsey, B. (2001) Empirical relationships between species richness, evenness, and proportional diversity. The American Naturalist, 158(3), 286-299.

Storfer, A., Murphy, M. A., Spear, S. F., Holderegger, R., & Waits, L. P. (2010) Landscape genetics: Where are we now? Molecular Ecology, 19(17), 3496-3514.

Storfer, A., Patton, A., & Fraik, A. K. (2018) Navigating the interface between landscape genetics and landscape genomics. Frontiers in Genetics, 9, 68.

Thies, C., Roschewitz, I., & Tscharntke, T. (2005) The landscape context of cereal aphid-parasitoid interactions. Proceedings of the Royal Society B: Biological Sciences, 272, 203-210.

Thomas, C.D., & Abery, J.C.G. (1995) Estimating rates of butterfly decline from distribution maps: the effect of scale. Biological Conservation, 73(1), 59-65.

Thomas, J. A. (2005) Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1454), 339-357.

Thorpe, R. S., Surget-Groba, Y., & Johansson, H. (2008) The relative importance of ecology and geographic isolation for speciation in anoles. Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 3071-3081.

Tiple, A. D., Khurad, A. M., & Dennis, R. L. (2011) Butterfly larval host plant use in a tropical urban context: Life history associations, herbivory, and landscape factors. Journal of Insect Science, 11(1), 65.

Tischendorf, L., Bender, D.J., & Fahrig, L. (2003) Evaluation of patch isolation metrics in mosaic landscapes for specialist vs. generalist dispersers. Landscape Ecology, 18(1), 41-50.

Tjørve, E. (2010) How to resolve the SLOSS debate: Lessons from species-diversity models. Journal of Theoretical Biology, 264, 604-612.

Tóthmérész, B. (1995) Comparison of different methods for diversity ordering. Journal of Vegetation Science, 6(2), 283-290.

Triantis, K. A., Mylonas, M., Weiser, M. D., Lika, K., & Vardinoyannis, K. (2005) Species richness, environmental heterogeneity and area: A case study based on land snails in Skyros archipelago (Aegean Sea, Greece) Journal of Biogeography, 32, 1727-1735.

Triantis, K. A., Vardinoyannis, K., Tsolaki, E. P., Botsaris, I., Lika, K., & Mylonas, M. (2006) Re-approaching the small island effect. Journal of Biogeography, 33, 914-923.

Triantis, K.A., Guilhaumon, F., & Whittaker, R.J. (2012) The island species-area relationship: biology and statistics. Journal of Biogeography, 39, 215-231.

Tscharntke, T., Steffan-Dewenter, I., Kruess, A., & Thies, C. (2002) Characteristics of insect populations on habitat fragments: a mini review. Ecological Research, 17, 229-239.

Tscharntke, T., & Brandl, R. (2004) Plant-insect interactions in fragmented landscapes. Annual Reviews in Entomology, 49, 405-430.

Tuljapurkar, S. (1990) Delayed reproduction and fitness in variable environments. Proceedings of the National Academy of Sciences, 87(3), 1139-1143.

Turlure, C., Schtickzelle, N., Van Dyck, H., Seymoure, B., & Rutowski, R. (2016) Flight morphology, compound eye structure and dispersal in the bog and the cranberry fritillary butterflies: an inter-and intraspecific comparison. PLoS One, 11(6), e0158073.

Turner, T. L., Hahn, M. W., & Nuzhdin, S. V. (2005) Genomic islands of speciation in Anopheles gambiae. PLOS Biology, 3, e285.

UniProt Consortium. (2018) UniProt: A worldwide hub of protein knowledge. Nucleic Acids Research, 47(D1), D506-D515.

Vähä, J. P., & Primmer, C. R. (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology, 15(1), 63-72.

Van Buskirk, J., & van Rensburg, A. J. (2020) Relative importance of isolation-by-environment and other determinants of gene flow in an alpine amphibian. Evolution, 74, 962-978.

van Etten, J. (2018) gdistance: Distances and routes on geographical grids. R package version (p. 1.2-2.)

van Swaay, C., & Warren, M. (1999) Red data book of European butterflies (Rhopalocera) Council of Europe, Strasbourg.

van Swaay, C., Warren, M., & Loïs, G. (2006). Biotope use and trends of European butterflies. Journal of Insect Conservation, 10(2), 189-209.

van Swaay, C., Maes, D., Collins, S., Munguira, M. L., Šašić, M., Settele, J., ..., & Cuttelod, A. (2011) Applying IUCN criteria to invertebrates: How red is the Red List of European butterflies? Biological Conservation, 144(1), 470-478.

Vekemans, X., & Hardy, O. J. (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology, 13, 921-935.

Ver Hoef, J. M., & Boveng, P. L. (2007). Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology, 88(11), 2766-2772.

Vignieri, S. N. (2005) Streams over mountains: Influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*) Molecular Ecology, 14(7), 1925-1937.

Voss, E. G., & Reznicek, A. A. (2012) Field manual of Michigan flora. Ann Arbor, MI: University of Michigan Press.

Wang, I. J., Savage, W. K., & Bradley Shaffer, H. (2009) Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*) Molecular Ecology, 18(7), 1365-1374.

Wang, I. J., & Summers, K. (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. Molecular Ecology, 19, 447-458.

Wang, I. J., Glor, R. E., & Losos, J. B. (2012) Quantifying the roles of ecology and geography in spatial genetic divergence. Ecology Letters, 16(2), 175-182.

Wang, I. J., & Bradburd, G. S. (2014) Isolation by environment. Molecular Ecology, 23(23), 5649-5662.

Wang, J. (2017) The computer program STRUCTURE for assigning individuals to populations: Easy to use but easier to misuse. Molecular Ecology Resources, 17(5), 981-990.

Wang, T., Hamann, A., Spittlehouse, D. L., & Murdock, T. Q. (2012) ClimateWNA—high-resolution spatial climate data for western North America. Journal of Applied Meteorology and Climatology, 51(1), 16-29.

Wang, Y. H., Yang, K. C., Bridgman, C. L., & Lin, L. K. (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. Landscape Ecology, 23(8), 989-1000.

Warzecha, D., Diekötter, T., Wolters, V., & Jauker, F. (2016) Intraspecific body size increases with habitat fragmentation in wild bee pollinators. Landscape Ecology, 31(7), 1449-1455.

Watling, J. I., Arroyo-Rodríguez, V., Pfeifer, M., Baeten, L., Banks-Leite, C., Cisneros, L. M., ..., & Lens, L. (2020) Support for the habitat amount hypothesis from a global synthesis of species density studies. Ecology Letters, 23(4), 674-681.

Watson, D. M. (2002) A conceptual framework for studying species composition in fragments, islands and other patchy ecosystems. Journal of Biogeography, 29, 823-834

Watson, J. E., Whittaker, R. J., & Dawson, T. P. (2004) Avifaunal responses to habitat fragmentation in the threatened littoral forests of south-eastern Madagascar. Journal of Biogeography, 31(11), 1791-1807.

Weiss, M. R. (2001) Vision and learning in some neglected pollinators: beetles, flies, moths and butterflies. In Chittka, L. & Thomson, J. D. (Eds.), *Cognitive ecology of pollination: animal behavior and floral evolution*. Cambridge University Press, Cambridge. (pp. 106-126).

Westman, W. E. (1983) Island biogeography: studies on the xeric shrublands of the inner Channel Islands, California. Journal of Biogeography, 10, 97-118.

Westwood, A. R., & Blair, D. (2010) Effect of regional climate warming on the phenology of butterflies in boreal forests in Manitoba, Canada. Environmental Entomology, 39(4), 1122-1133.

Whittaker, R. J., & Fernández-Palacios, J. M. (2007) Island biogeography: ecology, evolution, and conservation. 2nd edn. Oxford University Press, Oxford.

Wiens, J. A. (1989). Spatial scaling in ecology. Functional ecology, 3(4), 385-397.

Wiens, J. A. (2001) The landscape context of dispersal. In Clobert, J., Danchin, E., Dhondt, A. A., & Nichols, J. D. (Eds.) *Dispersal, individual, population, and community*. Oxford University Press, Oxford. (pp. 96-109).

Wilcove, D. S., McLellan, C. H., & Dobson, A. P. (1986) Habitat fragmentation in the temperate zone. Conservation Biology, 6, 237-256.

Williams, C. B. (1964) Patterns in the balance of nature. Academic Press, London.

Wilsey, B. J., Chalcraft, D. R., Bowles, C. M., & Willig, M. R. (2005) Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. Ecology, 86(5), 1178-1184.

Wilson, E. O., & MacArthur, R. H. (1967) The theory of island biogeography. Princeton University Press, Princeton.

Wilson, E. O., & Willis, E. O. (1975) Applied biogeography. In Cody, M. L., & Diamond, J. M. (Eds.) *Ecology and evolution of communities*. Belknap Press, Cambridge. (pp. 522-534).

Wogan, G. O., Yuan, M. L., Mahler, D. L., & Wang, I. J. (2020) Genome-wide epigenetic isolation by environment in a widespread Anolis lizard. Molecular Ecology, 29(1), 40-55.

Wright, D. H., & Reeves, J. H. (1992) On the meaning and measurement of nestedness of species assemblages. Oecologia, 92, 416-428.

Wright, S. (1921) Correlation and causation. Journal of Agricultural Research, 20, 557-585.

Wright, S. (1943) Isolation by distance. Genetics, 28(2), 114.

Yaacobi, G., Ziv, Y., & Rosenzweig, M. L. (2007) Habitat fragmentation may not matter to species diversity. Proceedings of the Royal Society of London B: Biological Sciences, 274, 2409-2412.

Yang, Z., & Teller, J. T. (2005) Modeling the history of Lake of the Woods since 11,000 cal yr B.P. using GIS. Journal of Paleolimnology, 33, 483-497.

Yoshio, M., & Ishii, M. (2001) Relationship between cold hardiness and northward invasion in the great mormon butterfly, Papilio memnon L. (Lepidoptera: Papilionidae) in Japan. Applied Entomology and Zoology, 36(3), 329-335.

Zhan, S., Zhang, W., Niitepold, K., Hsu, J., Haeger, J. F., Zalucki, M. P., … & Kronforst, M. R. (2014) The genetics of monarch butterfly migration and warning colouration. Nature, 514(7522), 317-321.

Zhang, J., Nielsen, S. E., Grainger, T. N., Kohler, M., Chipchar, T., & Farr, D. R. (2014) Sampling plant diversity and rarity at landscape scales: Importance of sampling time in species detectability. PLoS ONE, 9, e95334.

Zhang, J., Nielsen, S. E., Stolar, J., Chen, Y., & Thuiller, W. (2015) Gains and losses of plant species and phylogenetic diversity for a northern high-latitude region. Diversity and Distributions, 21(12), 1441-1454.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics, 28(24), 3326-3328.