Minimum sample size, population genomics and morphological variation in the terrestrial gastropod *Webbhelix multilineata* (Mollusca, Polygyridae)

by

Jermaine Mahguib

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Ecology and Evolutionary Biology

Program of Study Committee:
Kevin Roe, Major Professor
John Nason
Matthew Hufford
Tracy Heath
Dennis Lavrov

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2021

Copyright © Jermaine Mahguib, 2021. All rights reserved.

DEDICATION

To Hilary and Jake, without whom I would not be writing this dedication now. For months leading up to my final oral exam, I struggled with the most crippling anxiety and worst depression of my entire life as a result of trying to write this dissertation, at times to the point where I did not want to live anymore. It was so bad that three weeks out from my defense I still had not written most of this document, and I was going to quit on 7 years of work at that point because I could not manage anymore on my own. And then Hilary reached out to me to see how I was doing, and I vomited everything that I was going through onto her. In response, she and Jake insisted that I pack a bag and drive a couple of hours to their house in the countryside to stay with them for a long weekend, where they hugged me and talked me down before setting me up in a spare office with plenty of sunlight. They told me not to worry about anything else, and that they would come in to check on my progress every couple of hours. They cooked for me, engaged in meaningful conversation with me, provided me with cats, gave me feedback on my writing as I worked to catch up, and arranged a spare bedroom for me to sleep in. When that first weekend was done, they hugged me again before I made my way back home to continue working, and the following weekend they invited me back to do it all over again. With their incredible support, I managed to write the bulk of this dissertation before heading into my defense. It is not hyperbole to say that I would not have gotten to this point without them.

I would also like to dedicate this accomplishment to Team Colon, the amazing people from my graduate cohort who became life-long friends. To the aforementioned Hilary, Queen of Prairies; to my main broseph Andrew, who is also a prairie expert and is now upset that he was not designated the ruler of prairies; to The Lioness, Monica; to Tori the Monarch; and to The Little Princess, Clare.

To Alexstrasza, who is also known to me as Aqua Princess, The Dancing Bear Queen, The Mother of Silly Gooses, and The Patron Saint of the Theatah. But also as just Alex, who for several years personally bore witness to my deepest struggles and who shared her own with me, and so much more.

To my dear friends Stacey, Jessica and Jen, who flew from California to Iowa to personally attend my dissertation defense, and all of whom I adore for reasons beyond counting.

To my former Masters advisor, who now feels kinda bad for driving me into this PhD program; don't worry Ángel, I finally made it through and I did it in part for you.

To the guys, who since high school have bore the incredible weight of my near constant existential crises, a weight that over the past 7 years ramped up to a critical mass and ultimately drove one of them away. I dedicate this insane accomplishment to them, for themselves being insane enough to be my friends. This graduate program charged me for air, and the guys made sure I kept my bill paid.

To my Dad and my Brother, who have never really understood my academic pursuits but who nevertheless always wanted me to accomplish great things while patiently waiting for me to come back home.

And to my Mom. When this program became too much for me, she told me that she loved me and that no degree or any other accomplishment could ever mean as much to her as my being happy and healthy. Even when I was 95% done and couldn't stand to write another word, she looked me in the eyes and told me that if I was done then I was done. Which of course worked like a charm and made me want to finish it, for her. Ma, you have always been my biggest supporter and my best friend. I love you.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

NOMENCLATURE

NGS	Next Generation Sequencing
GBS	Genotyping by Sequencing
RADseq	Restriction Site Associated DNA Sequencing
PCR	Polymerase Chain Reaction
SNP	Single Nucleotide Polymorphism
ANOVA	Analysis of Variance
LGM	Last Glacial Maximum

ACKNOWLEDGMENTS

I would like to acknowledge the person who ultimately was most responsible for me becoming capable of completing the exhaustive work painstakingly detailed in the dissertation chapters below: Dr. Kyung Seok Kim. Kyung Seok was hired as a post doc in our lab sometime around my third year in the program and was tasked with the superhuman undertaking of teaching me how to do genomics, from lab work to bioinformatic processing to computational analyses. Heck, he basically taught me the general life skill of how to troubleshoot problems that invariably arise when trying to do almost anything. He was incredibly patient, particularly while trying to teach me how to do bioinformatics, which was excruciatingly difficult for me to wrap my head around. Kyung Seok was an excellent mentor who on more than one occasion when I thanked him for all his help, told me that my success was his success. I cannot express enough appreciation for all that he did for me.

I would also like to acknowledge Dr. David Hufnagel, without whom it would not have been possible for me to accomplish the minimum sample size chapter of this dissertation with any kind of efficiency. While in the midst of crunching to finish his own dissertation work just a few short months before his defense, Dave volunteered half a dozen hours of his time to help me create the custom Python script that was used for resampling subsets of the genomic data for that study. And when I say that he helped me to create it, I mean that I described to him what I needed and he told me step by step exactly what to write in order to code what I needed the script to do. Dave's generous contribution was so critical to the timely and efficient success of that project that I wanted to add him as an author in the eventual manuscript that would be submitted for publication in a scientific journal, but unfortunately I was overruled on that decision. The

most that I could do was to acknowledge him here as fully as possible, and to thank him a thousand times over.

Lastly, I would like to acknowledge my major professor, Dr. Kevin Roe. For 7 years, through all of our ups and downs, he always made sure that I was funded throughout each academic year, and even secured funding for me through most of the summers that I spent in the graduate program. Kevin and the EEOB department at ISU made it possible for me to accomplish this degree while also being able to financially support myself and my Mom back home in California. He and my program of study committee were also very generous in their decision to grant me an extra semester to finish writing this document after my final oral exam, and I appreciate their understanding and patience. Without it, the following dissertation would not have been completed.

ABSTRACT

Patterns in evolutionary and ecological biology help shape our view of the physical world both in the present and its past. Biogeographical studies in particular afford understanding of how biological systems have been shaped by processes over geological time, both genetically and morphologically. Some systems, like terrestrial gastropod mollusks, are well suited for these types of studies owing to their near ubiquitous occupation of terrestrial habitats, while also individually being very locally restricted and having poor reproductive dispersal. A notable species is the striped white-lip snail *Webbhelix multilineata*, a morphologically charming and biogeographically interesting study system because of its somewhat unusually broad geographic range spanning large regions of North America previously glaciated and unglaciated during the Pleistocene. It is also quite restricted to moist, shaded riparian woodlands, a historically profuse habitat that has been largely lost and become highly fragmented due to human development. These features and others make *Webbhelix* an enticing system for biogeographical and population study, and the chosen focus of the studies presented here. While there are many different molecular and morphological approaches that can be taken to pursue such studies, this dissertation focused primarily on the generation and use of genomic data for a non-model system, and secondarily on generating and utilizing morphological data, to accomplish three major aims: (1) test for a minimum sample size to accurately estimate basic population genomic parameters in *Webbhelix multilineata*; (2) generate the first genomic dataset for *Webbhelix* and utilize it for population genomic assessment, as well as coalescent simulations to assess models of post-glacial refugial expansion; and (3) test variation in *Webbhelix* shell size and brightness within the context of established ecogeographic patterns. Addressing the first aim using SNPs generated through a ddRADseq approach showed that as few as 6 individuals per sampled

locality was sufficient to accurately estimate observed and expected heterozygosity, inbreeding coefficient, and pairwise differentiation in *Webbhelix*. Pursuit of the second aim netted an 8,480 SNP dataset composed of 18 sample localities from across the species range, and its assessment for population genomic structure revealed 4 population clusters: two Midwestern clusters, a southern Mississippi River cluster, and an eastern cluster on the shores of Lake Erie in Ontario. Estimates of genetic diversity found it was highest in the southern-most sampled localities and generally decreased further north and away from the Mississippi River, indicative of serial founder effects resulting from range expansion. Analysis of coalescent simulations suggested that post-glacial range expansion in *Webbhelix* occurred along major river corridors rather than openly in all directions, and that expansion either began as early as some 20,000 years ago around the LGM or a more northerly refugia existed for this species than initially considered. Lastly, work toward the third aim resulted in a moderately strong negative correlation being found between shell size and latitude in *Webbhelix*, indicating a pattern described by a reverse Bergmann cline. Additionally, a weak positive correlation was found between shell brightness and latitude, suggesting (with heavy caveats) that *Webbhelix* may exhibit the pattern described by Gloger's ecogeographic hypothesis.

CHAPTER 1. GENERAL INTRODUCTION

Studies in Evolutionary and Ecological Biology

Evolutionary and ecological studies often seek to identify patterns in nature that help us construct realistic views of the physical world through time. Such patterns aid our understanding of biological systems, from studies on relationships between intelligence, brain size and environment (Gonda et al, 2013), to studies uncovering connections between behavioral suites and intraspecific population constraints (Sih et al, 2004), association studies linking single nucleotide polymorphisms (SNPs) to genome regions responsible for morphological color variation (Andrews et al, 2016), correlative studies on physiology and geography (Blackburn et al, 1999), and everything in-between. Relating biology and geography in particular can offer interesting views into the history and present condition of biological systems, such as how species ranges have expanded and contracted, or how plastic morphological traits can be across species distributions. One group of organisms often well suited for investigating biogeographic patterns is terrestrial gastropod mollusks, for reasons that include their restricted mobility and reproductive dispersal, and the fact that they leave behind calcium carbonate shells that generally remain well preserved in sediments over geologic time. A notable terrestrial gastropod with a variety of interesting features, and the subject of the present dissertation, is the striped white-lip snail *Webbhelix multilineata*.

Webbhelix has a broad geographical range stretching across much of the Midwestern United States and down into regions of the American South, with historical populations in parts of the eastern U.S. and southern Ontario, Canada (La Rocque, 1966). This species' broad range is interesting for a couple of reasons: one, *W. multilineata* is highly restricted to moist, shaded riparian woodlands, a habitat which is contemporarily scarce and fragmented across much of the

species range due to extensive agricultural development; two, the range encompasses a vast region that is more or less split in terms of land area that has been previously glaciated (in the northern part of the range) and non-glaciated land (to the south) through the most recent ice age during the Pleistocene. Additionally, the contemporary range is so broad that it encompasses regions with varying seasonal weather extremes. These features and others, combined with a limited vagility, make this molluscan system an interesting one for biogeographic study, as the proceeding chapters of this dissertation will show.

There are many approaches that can be taken to study biogeographical patterns in systems like *Webbhelix*, both morphological and molecular. While the work presented here does touch on approaches of the former, it primarily focuses on the latter. Molecular approaches to biogeographical and population studies have traditionally relied on genetic markers such as amplified fragment length polymorphisms or microsatellites to look for patterns of interest; although these techniques have served as incredible tools for biological research and some continue to be utilized, the past decade has seen a proliferation of studies using genomic markers generated with ever-advancing and increasingly affordable sequencing technologies.

The Genomics Era: A Quick Overview

Genomics has revolutionized biology and studies on model organisms through advances in next generation sequencing (NGS) which have facilitated practical and efficient generation of massive amounts of genetic data (Hudson, 2008; Allendorf et al, 2010; Avise, 2010; Funk et al, 2012; Narum et al, 2013; Ekblom & Wolf, 2014; Larson et al, 2014; McMahon et al, 2014; Russello et al, 2015; Ali et al, 2016; Andrews et al, 2016; Jones & Good, 2016). And these advancements in NGS technologies have, over time, done the same for many fields in evolutionary and ecological biology of non-model organisms. The cost of NGS has steadily decreased year by year such that more labs of modest budget have been able to pursue genomic

study (to one degree or another) on a broader range of systems (Hudson, 2008; Stein et al, 2014; Andrews et al, 2016; van Nimwegen et al, 2016).

Genomics can be pursued in a couple of different ways: either through whole genome sequencing or reduced genome sequencing (Ali et al, 2016). Next generation sequencing allows for *de novo* assembly, discovery and typing of genetic variation across entire genomes, but to date is still prohibitively expensive to do for most labs studying non-model organisms (and particularly so if sampling of many individuals is desired). Fortunately, reduced genome representation methods exist which allow for subsets of genomes to be sequenced for study. Broadly, there are two such approaches: methods that utilize previously established genome data, such as sequence capture or targeted capture, and restriction enzyme-based methods.

In capture-like methods, tailored oligonucleotide primers are used to target specific genomic regions of interest for sequencing, allowing for cost effective sequencing of many more individuals and at increased sequence depth for targeted loci (Ali et al, 2016). Restriction enzyme-based approaches, on the other hand, make use of regular enzyme cut sites found throughout genomes to facilitate predictable genome complexity reduction without the need for any prior genomic data. Although this comes with the tradeoff of less control over the number and location of loci obtained compared to capture methods, library preparation for restriction methods is less expensive when samples are barcoded and multiplexed early in the process and can allow for more individuals to be sequenced per reaction (further reducing cost).

Genomics for Non-Model Taxa in Ecology and Evolutionary Biology

Since ecological and evolutionary research on non-model organisms is typically bereft of previously established genome resources (particularly so for invertebrates, such as mollusks; Lydeard et al, 2004; Gomes-dos-Santos et al, 2020), many researchers have leaned into restriction enzyme-based methods to transition their study systems into the realm of genomics

(Andrews et al, 2016). Among the various types of restriction approaches (Miller et al, 2007; Baird et al, 2008; Van Tassell et al, 2008; Elshire et al, 2011), restriction site-associated DNA sequencing (RADseq) has emerged as the most widely used approach to generate large amounts of SNP data for non-model studies (Andrews et al, 2016; Jones & Good, 2016). Variations of RADseq typically consist of a relatively straightforward protocol to process numerous samples for library preparation in a cost-effective set of experiments. The greater depth of coverage per locus and increased feasibility for inclusion of higher numbers of samples (over whole genome sequencing), the lack of requirement of previously established genome data, and the simplicity and low cost of RADseq have made it the go-to tool in ecology and evolutionary biology for bringing non-model taxa into the genomics era.

Genomics in Molluscan Studies

Genomics studies over the past decade have been dominated by vertebrate systems, with the majority of those being focused on mammalian species, despite these taxa representing less than 5% of all metazoan diversity (Hudson, 2008; Gayral et al, 2013; GIGA, 2014; Lopez et al, 2019; Gomes-dos-Santos, 2020). Invertebrates in general have received comparably far less research attention than vertebrate groups despite comprising most of all animal diversity (Lydeard et al, 2004; Gomes-dos-Santos et al, 2020). Among invertebrates, arthropods comprise the majority of the species diversity and constitute the bulk of available genomics studies; in contrast, mollusks are the second most diverse group (by far) but constitute a fraction of the studies and available genomic resources despite their global ecological importance and economic value, comprehensively discussed in Gomes-dos-Santos et al (2020). Within the Mollusca, gastropods are the largest and most diverse group, comprising ~80% of all mollusks (Ponder & Lindberg, 2008). Representatives of this group are found in nearly all terrestrial and aquatic habitats, from some of the deepest sea floors to some of the highest mountains on the planet; yet

in their overview of molluscan genomics, Gomes-dos-Santos et al (2020) showed that gastropods trail behind bivalves in terms of application of different genetic and genomic approaches.

Expanding the application of genomic techniques to a broader range of taxa is an important collective step for advancing our understanding of biological systems (Gayral et al, 2013; Rossetto et al, 2021). Generating and disseminating genomic data for taxa in underrepresented groups such as Gastropoda is a good way to do this, as availability of genomic resources may attract additional research interest to these groups.

Population Genomics for a Terrestrial Gastropod

The first major aim of the present dissertation was to utilize appropriate genomic tools to generate a dataset of thousands of SNPs for population genomic assessment of the terrestrial gastropod *Webbhelix multilineata* (a monotypic taxon in the family Polygyridae). *Webbhelix's* restrictive and fragmented habitat, spread out across such a broad range, and combined with a low individual and reproductive dispersal ability, has led us to question what kind of genetic structure is present in this species. A second aim of this dissertation then was to use SNP data to assess population genomic structure across as much of the species range as could be managed.

The range of *Webbhelix* relative to previously glaciated regions, as well as fossil evidence (La Rocque, 1966), tells us that the species has experienced range contractions and expansions as the North American ice sheet grew southward and receded back north during the Pleistocene. Such contraction and subsequent expansion is likely to have resulted in a series of founder effects leading to a gradation in population genetic diversity across the contemporary range (Slatkin & Excoffier, 2012). And so a third aim for this dissertation was to estimate population genomic diversity to look for any apparent gradation suggesting a specific directionality of population expansion from Pleistocene refugia.

On discovering a suggestive genetic diversity gradient, a fourth aim became to model different refugial expansion scenarios and test them against each other using coalescent simulations. Knowing something about the mode of expansion could provide insight into some of the geographic conditions present during that process, based on *Webbhelix's* habitat requirements.

Minimum Sample Sizes for Population Genomics

Before the previous aims could be pursued, however, the question of appropriate population sampling size begged consideration. At the time of this consideration, only one empirical study (Nazareno et al, 2017) had been published investigating minimum sample sizes required for estimating population genomic parameters for a non-model system. Although Nazareno et al (2017) found that as few as 8 samples could be used to accurately estimate all parameters they tested (and as few as 2 samples for estimating F_{ST}), it was decided for the present study that pursuing a similar investigation of minimum sample sizes for *Webbhelix multilineata* was warranted for a couple of reasons. The first was that the aforementioned study was performed on a non-model Amazonian plant species (*Amphirrox longifolia*), so pursuing a similar investigation for *W. multilineata* presented itself as an opportunity to test minimum sampling in a non-model animal system. The second reason was a basic desire to replicate and corroborate these findings to (potentially) reinforce confidence in the recommendations. As a result, testing for minimum population genomic sampling sizes became a major aim of this dissertation.

Testing Ecological Biogeographic Rules on *Webbhelix multilineata*

The final aim of the present dissertation came about serendipitously through field observations while sampling *Webbhelix* for population genomics. It was observed that some populations had noticeably larger individuals than others, and that some populations seemed to

have more heavily banded (striped) individuals, making them appear darker overall and more difficult to notice against a backdrop of mud or leaf litter. These observations, combined with an overall geographic sampling distribution from about N35° to N45°, presented an opportunity to test established ecological biogeographic rules using latitude as a geographic variable.

Bergmann's Rule and Gloger's Rule are ecogeographic hypotheses that state temperature and humidity are relevant factors determining variation in body size and pigmentation, respectively, and are often tested with respect to latitude. Bergmann's Rule states that larger body morphs should be found in colder climates, while Gloger's Rule states that more heavily pigmented morphotypes should be found in warmer, more humid climates. Hence, testing variation in *Webbhelix* shell size and overall shell brightness in the context of these ecogeographical hypotheses became an additional major aim of this study.

Dissertation Overview

The proceeding chapters focus on each of the three major aims outlined above. Chapter 2 focuses on identification of minimum sample sizes for estimating population genomic parameters, since doing so directly informed the volume of genomic data (in terms of population sampling) that was to be generated for population genomic assessment. Chapter 3 then focuses on population genomics of *Webbhelix multilineata*, addressing all the aims outlined in that subsection above. Finally, Chapter 4 is devoted to testing ecogeographical hypotheses related to shell size and brightness variation in *Webbhelix*.

References

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping. *Genetics*, *202*(2), 389-400.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, *11*(10), 697-709.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81.

Avise, J. C. (2010). Perspective: conservation genetics enters the genomics era. *Conservation genetics*, *11*(2), 665-669.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PloS one*, *3*(10), e3376.

Blackburn, T. M., Gaston, K. J., & Loder, N. (1999). Geographic gradients in body size: a clarification of Bergmann's rule. *Diversity and distributions*, *5*(4), 165-174.

Ekblom, R., & Wolf, J. B. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary applications*, *7*(9), 1026-1042.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PloS one*, *6*(5), e19379.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in ecology & evolution*, *27*(9), 489-496.

Gayral, P., Melo-Ferreira, J., Glemin, S., Bierne, N., Carneiro, M., Nabholz, B., Lourenco, J. M., Alves, P. C., Ballenghien, M., Faivre, N., Belkhir, K., Cahais, V., Loire, E., Bernard, A., & Galtier, N. (2013). Reference-Free Population Genomics from Next-Generation Transcriptome Data and the Vertebrate–Invertebrate Gap. *PLoS Genet*, *9*(4), e1003457.

GIGA Community of Scientists. (2014). The Global Invertebrate Genomics Alliance (GIGA): Developing Community Resources to Study Diverse Invertebrate Genomes. *Journal of Heredity*, *105*(1), 1-18.

Gomes-dos-Santos, A., Lopes-Lima, M., Castro, L. F. C., & Froufe, E. (2020). Molluscan genomics: the road so far and the way forward. *Hydrobiologia*, *847*(7), 1705-1726.

Gonda, A., Herczeg, G., & Merilä, J. (2013). Evolutionary ecology of intraspecific brain size variation: a review. *Ecology and Evolution*, *3*(8), 2751-2764.

Hudson, M. E. (2008). Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular ecology resources*, *8*(1), 3-17.

Jones, M. R., & Good, J. M. (2016). Targeted capture in evolutionary and ecological genomics. *Molecular ecology*, *25*(1), 185-202.

La Rocque, A. (1966). *Pleistocene Mollusca of Ohio*. State of Ohio Department of Natural Resources Division of Geological Survey, Bulletin 62 Part 1 + 2

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., & Seeb, J. E. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (Oncorhynchus tshawytscha). *Evolutionary Applications*, *7*(3), 355-369.

Lopez, J. V., Kamel, B., Medina, M., Collins, T., & Baums, I. B. (2019). Multiple Facets of Marine Invertebrate Conservation Genomics. *Annual review of animal biosciences*, *7*, 473-497.

Lydeard, C., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Clark, S. A., Cummings, K. S., Frest, T. J., Gargominy, O., Herbert, D. G., Hershler, R., Perez, K. E., Roth, B., Seddon, M., Strong, E. E., & Thompson, F. G. (2004). The Global Decline of Nonmarine Mollusks. *BioScience*, *54*(4), 321-330.

McMahon, B. J., Teeling, E. C., & Höglund, J. (2014). How and why should we implement genomics into conservation?. *Evolutionary applications*, *7*(9), 999-1007.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research*, *17*(2), 240-248.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular ecology*, *22*(11), 2841.

Nazareno, A., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*.

Ponder, W. F., & Lindberg, D. R. (2008). *Phylogeny and Evolution of the Mollusca*. University of California Press.

Rossetto, M., Yap, J. Y. S., Lemmon, J., Bain, D., Bragg, J., Hogbin, P., Gallagher, R., Rutherford, S., Summerell, B., & Wilson, T. C. (2021). A conservation genomics workflow to guide practical management actions. *Global Ecology and Conservation*, *26*, e01492.

Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, *3*, e1106.

Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: an ecological and evolutionary overview. *Trends in ecology & evolution*, *19*(7), 372-378.

Slatkin, M., & Excoffier, L. (2012). Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift. *Genetics*, *191*(1), 171-181.

Stein, E. D., Martinez, M. C., Stiles, S., Miller, P. E., & Zakharov, E. V. (2014). Is DNA Barcoding Actually Cheaper and Faster than Traditional Morphological Methods: Results from a Survey of Freshwater Bioassessment Efforts in the United States?. *PloS one*, *9*(4), e95525.

van Nimwegen, K. J., van Soest, R. A., Veltman, J. A., Nelen, M. R., van der Wilt, G. J., Vissers, L. E., & Grutters, J. P. (2016). Is the $1000 Genome as Near as We Think? A Cost Analysis of Next-Generation Sequencing. *Clinical chemistry*, *62*(11), 1458-1464.

Van Tassell, C. P., Smith, T. P., Matukumalli, L. K., Taylor, J. F., Schnabel, R. D., Lawley, C. T., Haudenschild, C. D., Moore, S. S., Warren, W. C., & Sonstegard, T. S. (2008). SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature methods*, *5*(3), 247-252.

CHAPTER 2. MINIMUM SAMPLE SIZE FOR POPULATION GENOMICS IN THE TERRESTRIAL GASTROPOD *WEBBHELIX MULTILINEATA* (MOLLUSCA, POLYGYRIDAE)

Jermaine Mahguib

Iowa State University, Ames, Iowa 50011

Modified from a manuscript to be submitted to *Molecular Ecology*

Abstract

Knowledge of minimum sample sizes for accurately estimating population parameters is a valuable component of the planning process for population studies. It enables researchers to fine-tune the balance between study design and budgetary constraints, since sampling scheme can impact both the types of questions that can be addressed and how financially feasible a study can be. Traditional population genetic markers like microsatellites feature a deep literature covering their use, including sampling recommendations, but only a handful of studies to date have explicitly tested effective minimum sampling for estimating population parameters using genomic markers. Large datasets of genome-wide markers like SNPs hold the promise of dramatically expanding the types of ecological questions that can be pursued, but can still be prohibitively expensive for many labs to generate for many samples and has been slow to translate to conservation biology due to the need for destructive sampling in order to obtain large amounts of high quality DNA. Luckily, genomic minimum sample size studies have begun to show that between 4 to 8 individuals per population can accurately estimate common population genetic parameters, making genomic approaches even more broadly appealing. The present study adds to this growing body of evidence by testing the effect of sample size on population parameter estimates in the terrestrial gastropod *Webbhelix multilineata*, the striped white-lip snail. We found that as few as 6 individuals was enough to accurately estimate observed and

expected heterozygosity, inbreeding coefficient and pairwise differentiation in this molluscan species.

Introduction

Sample Size Recommendations for Population Genomics

Population genetic studies have traditionally relied on markers such as amplified fragment length polymorphisms or microsatellites to estimate population parameters, the latter being the most commonly used for this purpose (Tsykun et al, 2017). The type of genetic markers used impacts population sampling requirements because the number of loci that are typically discovered and the amount of allelic variation per locus varies between different marker types. Microsatellites can be highly variable, having as many as 10 or more alleles per locus, and it has been shown that a per population sample size of between 25 and 30 individuals is enough to accurately estimate population parameters like allele frequency and expected heterozygosity (Hale et al, 2012). In contrast to microsatellites, single nucleotide polymorphism (SNP) markers have very little allelic variation (typically 2 alleles) but thousands to hundreds of thousands of loci can be discovered for a study system depending on the method for generating the sequence data. Using a double digest restriction site-associated DNA sequencing (ddRADseq) approach, Nazareno et al (2017) found that as few as 8 individuals per population could be used to accurately estimate genetic diversity measures and as few as 2 individuals for estimating F_{ST} in a non-model plant species (*Amphirrox longifolia*) when 1000+ and 1500+ SNPs were used, respectively. Developing empirically based recommendations for minimum population sampling is important, particularly for genomic approaches, since such knowledge is important for study planning and methodological considerations.

Utility of Knowledge of Minimum Sample Sizes

An obvious utility of knowing a minimum sample size for population genomics studies is the impact it can have on the cost to generate data. Sampling fewer individuals not only saves on usage of library prep reagents and resources, it also reduces the number of NGS lanes that have to be purchased for performing sequencing. Although NGS costs continue to fall year by year, larger sample sizes can still very quickly drive the price of a study up beyond the means of many lab budgets. This, however, may not remain the case for too long. If sequencing costs continue to decline, a point may soon be reached were most labs can easily afford to sample well above minimum sampling recommendations to generate larger and larger datasets for multiple applications, a 'one cost with many benefits' strategy (Rossetto et al, 2021).

A less obvious utility is the implications for conservation of endangered or imperiled species. Management strategies typically call for population genetic assessment to inform conservation decisions, but also often include policies that restrict or prohibit invasive or destructive sampling. To get around this problem, many techniques have been developed for extraction of DNA from non-invasively sampled materials that are left behind by individuals, such as fecal matter, hair or fur, feathers, body mucus, etc. (Waits & Paetkau, 2005; Beja-Pereira et al, 2009). These non-invasive sampling methods, however, typically result in low DNA concentrations, even when highly refined extraction protocols are implemented (Russello et al, 2015). Traditional population genetic approaches utilizing single locus mitochondrial DNA sequencing or microsatellites are achievable (though difficult) using low yield DNA, but population *genomic* approaches like RADseq require comparatively large amounts of high quality DNA in order to succeed. This may be contributing to a lag in the transition from conservation genetics to conservation genomics, relative to many other fields, that has been

continuously observed for nearly 20 years now (Feder & Mitchell-Olds, 2003; McMahon et al, 2014; Shafer et al, 2015; Meek & Larson, 2019; von Thaden et al, 2020).

Frankham (2010) proposed 13 priorities for advancing the field of conservation biology, many of which have been intractable using traditional population genetic approaches, and Allendorf et al (2010) have diagrammed interacting factors in the conservation of natural populations, nearly half of which cannot be directly assessed by conservation genetics but all of which can be assessed by conservation genomics. Rossetto et al (2021) similarly argue a number of advantageous reasons for transitioning conservation efforts into the realm of genomics.

Clearly there are benefits to applying genomics to conservation, but it will remain unfeasible for situations where non-invasive sampling is insisted on due to the amount of sampling that traditional population genetic markers call for. However, further studies corroborating that accurate estimates of population parameters can be achieved through genomic approaches using minimal sampling may aid in convincing management agencies to allow limited amounts of invasive or destructive sampling for assessment of endangered populations, and thus facilitate a more rapid transition to conservation genomics.

Evaluation of Minimum Sample Sizes for Population Genomics in *Webbhelix*

The present study sought to evaluate minimum sample sizes for estimating genetic diversity (observed and expected heterozygosity, H_O and H_S), inbreeding coefficient (G_{IS}), and genetic differentiation (F_{ST}) in the terrestrial gastropod *Webbhelix multilineata* (the striped whitelip snail) using a SNP dataset generated with a ddRADseq approach similar to Peterson et al (2012). A double digest approach was utilized because genome size in *Webbhelix* was estimated by flow cytometry to be very large (19C), and additional complexity reduction was desired. *Webbhelix* is a riparian habitat-restricted species of interest for population genomic assessment owing to recent habitat fragmentation and a broad geographic range that

encompasses previously glaciated and unglaciated regions of the American Midwest. Since no genomic resources existed for this species prior to the present study and no assessment of minimum sample sizes has (to our knowledge) been done before on a molluscan system, four preliminary sample localities of *W. multilineata* (part of a broader sampling of localities for a separate population genomics study; see Chapter 3) were utilized to investigate sample size effects on estimates of demographic parameters. Evaluation of minimum sample sizes was conducted using a resampling approach similar to Nazareno et al (2017), but here mean population parameter estimates of 11 stepwise subsampling treatments were also statistically compared using analysis of variance (ANOVA), Kruskal-Wallis rank sum tests and pairwise comparisons using t-tests with pooled standard deviations.

Materials and Methods

Species Sample Localities and Collection

Four sample localities of *Webbhelix multilineata*, from a broader set of sampled localities for a separate population genomics study, were chosen for use in the present study based on geographical considerations that facilitated *a priori* expectations of pairwise F_{ST}: Timmons Grove County Park (TGCP) along the Iowa River in Marshall County, Iowa; Pikes Timber Park (PTP) along the Iowa River in Wright County, Iowa; the Steamboat Trace Trail (STT) east of Peru, Nebraska, along the Missouri River in Nemaha County; and Mandan Park (MPO) in Omaha, Nebraska, along the Missouri River in Douglas County. Both the TGCP & PTP localities and the STT & MPO localities are separated by ~75 miles of river distance along their respective rivers. The two rivers are an average straight-line distance of 170 miles from one another between their respective sample locality pairs, and the two rivers do not reach a downstream meeting point within the shared Mississippi River watershed for hundreds of miles (an average of ~406 miles from the Iowa River locality pair and ~503 miles from the Missouri

River pair). Assuming strict habitat restriction as is believed to be the case for *Webbhelix*, and assuming that if any gene flow is occurring in this species it is primarily downstream along shared rivers via flooding events, one might reasonably expect that the TGCP and PTP locations would share more genetic diversity with each other than they would with the STT and MPO locations, and vice versa.

Individual *W. multilineata* at these four sample localities were searched for by a combination of carefully walking over and visually scanning the ground and leaf litter for their rounded shell shape and characteristic red-brown striping, and carefully turning over wood logs and pieces of fallen tree bark, as individuals can often be found sheltered on the undersides of these. Individual snails were removed and placed whole into a 50 mL Falcon tube filled with 100% ethanol, killing them via rapid dehydration of soft body tissues and preserving them for later DNA extraction.

DNA Extraction, Genomic Library Preparation and Sequencing

Genetic material was extracted from the soft body tissue using Qiagen's DNeasy® Blood & Tissue Kit, using the Purification of Total DNA from Animal Tissues (Spin-Column) protocol. Tissue used for extractions was taken from the anterior-most tip of each animal's foot (~20 – 30 mg). In the final step of the protocol, 175 uL of elution buffer was used to elute the DNA of each sample in a single elution step. Concentration of eluted DNA ranged from 14 ng/uL to as much as 240 ng/uL, measured using a Thermo Scientific™ NanoDrop 2000 Spectrophotometer, with an average of 59 ng/uL across all samples. A portion of these stock DNA elutions was diluted to ~20 ng/uL for use in preparation of genomic libraries.

Genomic libraries were prepared in-house using the double digest restriction-site associated DNA (ddRAD) protocol described in Peterson et al (2012), with modifications. Digestion of genomic DNA was carried out using 10 uL of ~20 ng/uL DNA extraction, 0.75 uL

of 20 U/uL digestion enzyme PstI, 0.75 uL of 20 U/uL digestion enzyme MspI, 2 uL of 10X CutSmart buffer (New England BioLabs), and 6.5 uL of molecular grade ddH_2O per sample, and incubating at 37˚C for 600 minutes in an eppendorf Mastercycler® ep*gradient* thermocycler. Successful digestion was confirmed by agarose gel electrophoresis. Gels were 1% agarose and wells were loaded with 3 uL of digested sample plus 2 uL of loading dye; electrophoresed samples were compared to an undigested control, a lambda DNA/Hind III ladder (Promega) and a 1 kb DNA ladder (Fisher Scientific). Undigested control samples presented as a bright band at the 23+ kb range of the Hind III ladder, while successfully digested samples presented as a dim smear the length of the 1 kb ladder.

Ligation of adaptor sequences to restriction cut sites was accomplished using 10 uL of ~20 ng/uL digested DNA, 1.5 uL of 1 uM PstI adaptor (with barcode sequence), 0.6 uL of 25 uM MspI adaptor, 0.175 uL of 400,000 U/mL T4 DNA ligase, 2 uL of 10 mM adenosine triphosphate (ATP), 1 uL of 10X CutSmart buffer (New England BioLabs), and 4.725 uL of molecular grade ddH_2O per sample. Prepared samples were then incubated in a thermocycler at 16˚C for 360 minutes, 37˚C for 60 minutes, and a final incubation at 20˚C for 60 minutes. Adaptor ligations were carried out in batches of 48 samples or less, with the intention of submitting no more than 48 individuals per Illumina sequencing lane in order to conserve a reasonable read depth of coverage (at least 10). A total of 16 different barcoded PstI adaptors were used on each batch of 48 or less samples, such that each of the first 16 samples were assigned a unique barcoded adaptor, each of the next 16 samples were assigned a unique barcode from the same pool of 16 available adaptors, and each of the remaining samples were assigned a unique barcode out of those same 16 adaptors (all barcoded adaptor primer sequences can be seen in **Table 2.1**). From there, the 48 or less samples in each batch were further uniquely tagged

Table 2.1 – Ligation site adaptor and PCR primer sequences used for genomic library preparation. PstI and MspI adaptor sequence names designated "F-" are forward and "R-" are reverse sequences. Underlined portion of sequences indicate barcodes (PstI adaptors) and indices (PCR index primers).

Sequence Type	Sequence Name	Sequence
PstI Adaptors with Barcodes	F-PAwB 1	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ATC ACG TGC* A
	F-PAwB 2	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT CGA TGT TGC* A
	F-PAwB 3	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT TTA GGC TGC* A
	F-PAwB 4	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT TGA CCA TGC* A
	F-PAwB 5	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ACA GTG TGC* A
	F-PAwB 6	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GCC AAT TGC* A
	F-PAwB 15	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT CTT GAT GC*A
	F-PAwB 16	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT TCA CCT GC*A
	F-PAwB 17	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT CTA GCT GC*A
	F-PAwB 18	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ACA AAT GC*A
	F-PAwB 19	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT TTC TCT GC*A
	F-PAwB 20	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT AGC CCT GC*A
	F-PAwB 27	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ATGAAAC TGC* A
	F-PAwB 28	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT AAAAGTT TGC* A
	F-PAwB 29	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GAATTCA TGC* A
	F-PAwB 30	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GGACCTA TGC* A
	R-PAwB 1	/5Phos/CGT GAT AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 2	/5Phos/ACA TCG AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 3	/5Phos/GCC TAA AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 4	/5Phos/TGG TCA AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 5	/5Phos/CAC TGT AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 6	/5Phos/ATT GGC AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 15	/5Phos/TCA AGA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 16	/5Phos/GGT GAA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 17	/5Phos/GCT AGA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 18	/5Phos/TTT GTA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 19	/5Phos/GAG AAA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 20	/5Phos/GGG CTA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA A*A
	R-PAwB 27	/5Phos/GTTTCAT AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 28	/5Phos/AACTTTT AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 29	/5Phos/TGAATTC AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
	R-PAwB 30	/5Phos/TAGGTCC AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT AA*A
MspI Adaptors	F-MA	/5Phos/CGA GAT CGG AAG AGC ACA CGT CTG AAC TCC AGT C
	R-MA	GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T
PCR Forward Primer and Index Primers	PCR FP	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC* T
	PCR IP 4	CAA GCA GAA GAC GGC ATA CGA GAT TGG TCA GTG ACT GGA GTT CAG ACG TGT G*C
	PCR IP 6	CAA GCA GAA GAC GGC ATA CGA GAT ATT GGC GTG ACT GGA GTT CAG ACG TGT G*C
	PCR IP 12	CAA GCA GAA GAC GGC ATA CGA GAT TAC AAG GTG ACT GGA GTT CAG ACG TGT G*C

through the use of three index primers in the following polymerase chain reaction (PCR) step of the genomic library prep, such that the first 16 samples (each with a unique barcode) would be amplified using one index primer, the next 16 samples amplified using the second index primer, and the remaining samples amplified using the third index primer.

Amplification of double restriction cut site appended genomic fragments via PCR was executed using 1.8 uL of 10X diluted adaptor ligated DNA, 1 uL of 5 uM forward PCR primer, 1 uL of 5 uM index primer, 7.5 uL of 2X Phusion® High-Fidelity PCR Master Mix w/HF Buffer (New England BioLabs), and 3.7 uL of molecular grade ddH_2O per sample. Polymerase chain reaction was facilitated in a thermocycler with an initial 30 second DNA denaturation step at 98˚C, followed by 25 cycles of a 20 second denaturation step at 98˚C, 30 seconds of annealing at 60˚C, and 40 seconds of extension at 72˚C. At the end of 25 cycles, a final extension occurred at 72˚C for 12 minutes. As described above, three different index primers were used per batch of 48 or less samples such that every individual in a sample batch was tagged with a unique barcode/index combination. In order to distinguish samples *between* batches, each sample batch was slated for sequencing in an individual Illumina sequencing lane, adding a total of three layers of sample identification (1. barcode; 2. index; 3. sequencing lane/submission tracking number) for downstream demultiplexing of raw sequencing data.

After amplification each batch of PCR samples was pooled, and a portion of these pooled samples was fragment size selected (325 – 500 bp) using BluePippin (Sage Science) to remove leftover adapter fragments, unbound primers and primer dimers. Subsequent to the BluePippin protocol, size selected pooled batch samples were confirmed to contain mostly sequence fragments of the desired size via agarose gel electrophoresis, as in the post-digestion confirmation step previously described. Size selected pooled samples were compared to leftover

pooled samples to check for reduced smear band size that roughly fit within the 300 to 500 bp range of a 1 kb ladder.

Pooled samples were submitted for 100-cycle Illumina HiSeq® high-output mode genotyping by sequencing (GBS) at the Iowa State University Office of Biotechnology's DNA Facility. Each sample was submitted to be run in its own Illumina sequencing lane in order to maintain the ability to identify reads belonging to specific individuals within and between pooled samples, as described previously. Samples were quantified by the sequencing facility using Qubit™ quantification prior to sample dilution for initial sequencing. Additionally, a separate sequencing step was requested in order to read the index sequence of the 3' region PCR primers used during library construction.

Stacks Core Parameter Optimization and SNP Discovery

Sequencing data was processed through the STACKS pipeline of software packages (Catchen et al, 2011; Catchen et al, 2013), using STACKS version 1.46. Raw sequence data was first demultiplexed using the program PROCESS_RADTAGS, which also checks the quality score of sequence reads and discards them if the average score drops lower than 90% probability of being correct (raw phred score of 10). Before proceeding through a manual run of the component programs in the pipeline for the sequencing data generated in this study, a set of replicate samples (11 individuals, each from different localities sampled for a separate, broader population genomics study; see Chapter 3) were used to optimize a core set of STACKS *de novo* assembly parameters based on minimizing quantification of SNP error rate, as outlined in Mastretta-Yanes et al (2015). Our GBS protocol (from PCR amplification to Illumina sequencing) was performed twice on these samples, creating a dataset of 11 replicate sample pairs for this procedure.

Mastretta-Yanes et al (2015) identified four core parameters in STACKS that they argue should be optimized specifically for a given genomic data set: '-m' which specifies the minimum

number of identical raw reads required to create a read stack in the program USTACKS (default value 3); '-M' which specifies the number of mismatches allowed between loci when processing a single individual in USTACKS (default value 2); '--max_locus_stacks' specifying the maximum number of stacks at a single *de novo* locus, also in USTACKS (default value 3); and '-n' which specifies the number of mismatches allowed between loci when building a sample reference catalog in the program CSTACKS (default value 0; Catchen et al, 2011; Catchen et al, 2013). Note that in later versions of STACKS the default value for '-n' was changed from 0 to 1 by the developers, though 0 was used in the present study using STACKS v1.46 following Mastretta-Yanes et al (2015).

After demultiplexing with PROCESS_RADTAGS, variation in the aforementioned core parameters to find an optimal setting was tested using the STACKS program DENOVO_MAP. A pipeline-executing program, DENOVO_MAP allows for each component program of STACKS (USTACKS, CSTACKS, SSTACKS and POPULATIONS) to be run sequentially from a single initiating command script, while also permitting specification of parameters across programs. DENOVO_MAP was used iteratively to generate differing sets of SNPs for the sample replicate dataset, varying (with each run) one parameter at a time through a prescribed range (2 – 10 for '-m'; 2 – 8 for '-M'; 2 – 6 for '--max_locus_stacks'; 0 – 7 for '-n') while keeping the others set to their default values (Mastretta-Yanes et al, 2015).

In order to quantify SNP error rate (defined as the proportion of SNP mismatches between sample replicate pairs), a custom R script (Mastretta-Yanes et al, 2015) that reads SNP genotypes coded as a single allele dosage number was used. This necessitated outputting SNP genotypes (via DENOVO_MAP) for sample replicate pairs in PLINK format. PLINK (http://pngu.mgh.harvard.edu/purcell/plink/; Purcell et al, 2007), a program that can perform a

range of basic, large-scale analyses and formatting procedures on genomic data, was used to recode DENOVO_MAP SNP genotype outputs in terms of additive components. The resulting *plink.raw* files were then used to estimate SNP error rate using the custom R script.

Replicate pair error rates for a given varied STACKS core parameter were averaged, and the parameter value that generated the lowest average SNP error rate was designated the optimal. Once an optimal value was determined for each core parameter, the resulting parameter combination was used for all proceeding steps. Three additional STACKS parameters and an additional data correction procedure were also tested after core parameter optimization: '--model_type' specifying use of either a 'snp' or 'bounded' model in USTACKS; '--min_maf' specifying the implementation of a minimum minor allele frequency in POPULATIONS; use of RXSTACKS, another program in the STACKS package that makes corrections to genotype and haplotype calls in individual samples based on data accumulated from a population-wide examination; and '--write_single_snp' for specifying restriction of data analysis in POPULATIONS to only the first SNP per locus. For '--model_type,' DENOVO_MAP runs were performed using the 'snp' model, 'bounded' model with option '--bound_high 0.05,' '--bound_high 0.10,' '--bound_high 0.20' and '--bound_high 0.30,' to test a short range of upper bounds for epsilon (model error rate) in USTACKS. The model that resulted in a sample replicate pair SNP dataset with the lowest SNP error rate was used in subsequent tests. For '--min_maf,' running DENOVO_MAP with a minimum minor allele frequency of 0.05 was compared with running the pipeline with no minimum minor allele frequency set, with the lowest SNP error rate result determining subsequent use of the program option. After '--model_type' and '--min_maf' usage were determined, RXSTACKS was implemented to see if call corrections reduced SNP error rate. Lastly, DENOVO_MAP was run a final time using the '--write_single_snp' option to compare

average SNP error rate between a reduced dataset comprised of the first SNP from near the 5' end of each sequence read, versus an expanded dataset of multiple SNPs from along the full length of each read.

Once all parameters were optimized for minimizing SNP error rate, a proper run of the STACKS pipeline was executed (using optimal parameter settings) by manually running each program in sequence to generate a workable *Webbhelix multilineata* SNP dataset with a minimum population constraint of ~75% and a minimum individual constraint of 80%. The resulting dataset was outputted from the POPULATIONS program in several coding formats, including *.plink.* This output was used in an additional post-processing step performed using PLINK to filter out SNPs in high linkage disequilibrium (LD) pairs and SNPs showing extreme deviation from Hardy-Weinberg Equilibrium (HWE); high LD and extreme HWE deviation can be (in the case of the former) and often is (in the latter) associated with genotyping error (Hosking et al, 2004; Scheet & Stephens, 2008; Chen et al, 2017). Removal of these SNPs also served to reduce the overall number of markers to a more computationally manageable size for downstream population genomic analyses. The PLINK program is limited in SNP coding formats it can output, so post-filtering SNP data was output in *.vcf* format and then converted to *.genepop* format using the program PGDSPIDER version 2.1.1.5 (Lischer & Excoffier, 2012). The *.genepop* format is more flexible in usage since it is more commonly accepted as input by population genetic software packages, and was used in the proceeding examination of sample size effects.

Evaluation of Minimum Sample Size for Population Genomics

To evaluate how many samples per population are needed in order to accurately estimate basic population genomic statistics, the approach of Nazareno et al (2017) was largely followed. A total of 24 individual snails from each of the TGCP, PTP, STT and MPO localities were sequenced, following a 'rule of thumb' of 20+ individuals per population largely carried over

from use with traditional markers like microsatellites. A subsampling scheme was used on the data generated for these four localities, in which individuals would be randomly sampled in decreasing increments to create a series of datasets for estimating basic population genetic statistics (observed heterozygosity, H_O; expected heterozygosity, H_S; inbreeding coefficient, G_{IS}; fixation index, F_{ST}) that could be compared based on population sample size.

In order to efficiently subsample individuals from each sample locality, a custom PYTHON script was created (**https://github.com/jmahguib-ISU/Webbhelix_Project**). The script reads as input into PYTHON a *.genepop* file as outputted by the STACKS v1.46 POPULATIONS program [**NOTE**: POPULATIONS outputs *.genepop* files with all loci names printed on the second line and separated by commas, while PGDSPIDER converts files into *.genepop* format with every locus name printed on a new line starting from line 2; the *.genepop* file used here was manually reformatted to all loci names on line 2 after conversion using PGDSPIDER]. The PYTHON script ignores the first two lines of the file and begins reading on the third, where the word 'pop' denotes the beginning of the first population of samples. The script then adds all sample names from each population into a list (called 'nameslist') in PYTHON; from 'nameslist,' the script then selects a user-specified random number of sample names and adds them into another list (called 'smalllist'). Next, PYTHON writes a plain text output file by first copying the initial header line and loci names line from the *.genepop* input file, then writes the full line of text associated with each sample name on 'smalllist' while maintaining their population assignments in the original file format. The result is a new *.genepop* formatted file with SNP data for a user specified subset number of randomly selected individuals within each population denoted in the input file.

Population subsampling was conducted for 11 stepwise treatments, starting with 22 individuals per sample locality and decreasing in increments of 2 all the way down to 2

individuals per locality. The custom PYTHON script described above was used to resample random individuals from each population 100 times per treatment, producing a total of 1,100 resampled datasets. Each one of these resampled datasets was then individually analyzed using the program GENODIVE version 3.01 (Meirmans, 2020) to generate genetic diversity estimates of H_O, H_S and G_{IS}, and to calculate distances for pairwise F_{ST}. GENODIVE calculates estimates of expected heterozygosity while correcting for sampling bias due to limited sampling numbers per population (Nei, 1987), and the inbreeding coefficient G_{IS} is analogous to the F-statistic F_{IS}. All GENODIVE results were manually collated and formatted for analysis of variance (ANOVA), performed in R, to analyze the differences among mean diversity estimates between subsampling treatments. Kruskal-Wallis rank sum tests were also performed to non-parametrically evaluate whether or not statistically significant differences in diversity estimate means existed between subsampling treatments, and pooled standard deviation t-tests were used for pairwise comparisons of treatments. In addition to evaluating differences between mean diversity estimates, boxplots were generated for each preliminary population, for each genetic diversity parameter, to show the variability in diversity estimates across subsampling treatments. Boxplots were created in R v3.3.3 using RSTUDIO v1.0.136.

Results

Optimal Stacks Core Parameters and SNP Discovery

Mean estimated SNP error rates for sample replicate pairs and values for STACKS core parameters (‘-m’; ‘-M’; ‘--max_locus_stacks’; ‘-n’) resulting in the lowest mean estimated SNP error rates are shown in **Table 2.2**. For the minimum number of identical raw reads required to create a read stack in the program USTACKS (‘-m’; default value 3), mean estimated SNP error

Table 2.2 a – Mean estimated SNP error rates from 11 *Webbhelix* sample replicate pairs used to determine optimal values for STACKS core parameters identified by Mastretta-Yanes et al (2015). The four core parameters include '-m' (the minimum number of identical raw reads required to create a read stack in the program USTACKS; default value 3), '-M' (the number of mismatches allowed between loci when processing a single individual in USTACKS; default value 2), '--max_locus_stacks' (the maximum number of stacks at a single *de novo* locus in USTACKS; default value 3), and '-n' (the number of mismatches allowed between loci when building a sample reference catalog in CSTACKS; default value 0). The 'parameter combo' rows show the numerical values for core parameters inputted (one combo at a time) to the STACKS program DENOVO_MAP to iteratively generate differing sets of SNPs for the sample replicate dataset. With each run of DENOVO_MAP, one parameter at a time was varied through the range of values shown in the table for each parameter, while keeping the others set to their default values. The order of numerical values in all parameter combos correspond to the four core parameters, such that '2_2_3_0' represents -m = 2, -M = 2, --max_locus_stacks = 3, and -n = 0. Each 'mean SNP error rate' is the average value of the 11 SNP error rates calculated from the SNP data of each of the *Webbhelix* sample replicate pairs used. Standard deviations ('St. Dev.') for mean SNP error rates are also shown. The lowest mean SNP error rates for each core parameter are highlighted in green, and all together indicate the optimal core parameter combo for minimizing error in the process of SNP discovery for *Webbhelix* using STACKS is 2_2_5_4.

	SNP Error Rates for Variations of Stacks parameter '-m' from 2 - 10								
Parameter Combo	2_2_3_0	3_2_3_0	4_2_3_0	5_2_3_0	6_2_3_0	7_2_3_0	8_2_3_0	9_2_3_0	10_2_3_0
Mean SNP Error Rate	0.13174	0.13441	0.13891	0.14721	0.15743	0.17028	0.18035	0.19235	0.19778
St. Dev.	0.03372	0.03331	0.03513	0.03402	0.03442	0.03415	0.03199	0.03030	0.02773

	SNP Error Rates for Variations of Stacks parameter '-M' from 2 - 8						
Parameter Combo	3_2_3_0	3_3_3_0	3_4_3_0	3_5_3_0	3_6_3_0	3_7_3_0	3_8_3_0
Mean SNP Error Rate	0.13441	0.15245	0.16333	0.17147	0.17745	0.18496	0.18805
St. Dev.	0.03331	0.04002	0.04254	0.04379	0.04390	0.04416	0.04421

	SNP Error Rates for Variations of Stacks parameter '--max_locus_stacks' from 2 - 6				
Parameter Combo	3_2_2_0	3_2_3_0	3_2_4_0	3_2_5_0	3_2_6_0
Mean SNP Error Rate	0.13490	0.13441	0.13478	0.13323	0.13373
St. Dev.	0.03163	0.03331	0.03451	0.03241	0.03339

	SNP Error Rates for Variations of Stacks parameter '-n' from 0 - 7							
Parameter Combo	3_2_3_0	3_2_3_1	3_2_3_2	3_2_3_3	3_2_3_4	3_2_3_5	3_2_3_6	3_2_3_7
Mean SNP Error Rate	0.13441	0.08113	0.05912	0.05027	0.04769	0.20231	0.20231	0.20231
St. Dev.	0.03331	0.01538	0.01087	0.00962	0.00956	0.21174	0.21174	0.21174

Table 2.2 b – Mean estimated SNP error rates from 11 *Webbhelix* sample replicate pairs used to determine optimal values for an additional set of three secondary STACKS parameters and a data correction procedure, while using the optimal values for the four STACKS core parameters (2_2_5_4). The secondary parameters and procedure tested were '--model_type' (specifying use of either a 'snp' or 'bounded' model in USTACKS), '--min_maf' (specifying the implementation of a minimum minor allele frequency in the STACKS program POPULATIONS), use of RXSTACKS (another program in the STACKS package that makes corrections to genotype and haplotype calls in individual samples based on data accumulated from a population-wide examination), and '--write_single_snp' (specifying restriction of data analysis in POPULATIONS to only the first SNP per locus). Each of these was tested sequentially in the order presented, and after each set of tests the optimal option was implemented into the parameter combo used to test the next secondary option. For '--model_type,' DENOVO_MAP runs were performed using the 'snp' model (designated '00' in parameter combos) and the 'bounded' model with the option '--bound_high 0.05' (designated '05' in parameter combos), '--bound_high 0.10' (designated '10' in parameter combos), '--bound_high 0.20' (designated '20' in parameter combos) and '--bound_high 0.30' (designated '30' in parameter combos) to test a short range of upper bounds for epsilon (model error rate) in USTACKS. For '--min_maf,' running DENOVO_MAP with a minimum minor allele frequency of 0.05 (designated '05' in parameter combos) was compared with running the pipeline with no minimum minor allele frequency set (designated '00' in parameter combos). For RXSTACKS, the DENOVO_MAP run implementing the call corrections (designated 'corr' in parameter combos) was compared to running DENOVO_MAP without call corrections (designated '----' in parameter combos). For the option '--write_single_snp,' enabling it (designated '1' in parameter combos) was compared to the previous DENOVO_MAP runs testing RXSTACKS, since the default setting when the '--write_single_snp' option is not specified is for the program to not enable the option. Overall the lowest mean SNP error rate resulted from using the 'snp' model (00), while not using the '--min_maf' option (00), not using the RXSTACKS correction (----), and using the 'write_single_snp' option (1). In the table this is represented as parameter combo 2_2_5_4_00_00_----_1.

	SNP Error Rates for Variations of Stacks parameter model type 'snp' (default; designated here 00) vs model type '--bound_high' (0.05, 0.10, 0.20 and 0.30)				
Parameter Combo	2_2_5_4_00	2_2_5_4_05	2_2_5_4_10	2_2_5_4_20	2_2_5_4_30
Mean SNP Error Rate	0.04310	0.08117	0.04979	0.04797	0.04523
St. Dev.	0.00861	0.01182	0.01061	0.00976	0.00895

	SNP Error Rates for min_maf 0.05	SNP Error Rates with rxstacks correction	SNP Error Rates w/o rxstacks correction	SNP Error Rates for write_single_snp
Parameter Combo	2_2_5_4_00_05	2_2_5_4_00_00_corr	2_2_5_4_00_00_----	2_2_5_4_00_00_----_1
Mean SNP Error Rate	0.04434	0.05110	0.05017	0.03921
St. Dev.	0.00852	0.00844	0.01116	0.00823

rates for *Webbhelix* sample replicate pairs ranged from 13.2 to 19.8%, with -m = 2 resulting in the lowest error rate. For number of mismatches allowed between loci when processing a single individual in USTACKS ('-M'; default value 2), mean error rates ranged from 13.4 to 18.8%, with -M = 2 resulting in the lowest error rate. For the maximum number of stacks at a single *de novo* locus in USTACKS ('--max_locus_stacks'; default value 3), error rates ranged from 13.3 to 13.5%, with --max_locus_stacks = 5 resulting in the lowest error rate. And finally, for number of mismatches allowed between loci when building a sample reference catalog in the program CSTACKS ('-n'; default value 0), error ranged from 4.8 to 20.2%, with -n = 4 resulting in the lowest error rate.

Testing the USTACKS option '--model_type' resulted in mean error rates ranging from 4.3 to 8.1%, with the default 'snp' model resulting in the lowest mean error. Specifying a minimum minor allele frequency set to 0.05 ('--min_maf 0.05') did not reduce mean SNP error rate, and the mean error rate was actually slightly lower without specifying '--min_maf' (mean SNP error rate of 4.4% with vs. 4.3% without). Similarly, implementation of RXSTACKS on this data set did not reduce mean SNP error rate and was actually slightly lower when not implemented (mean SNP error rate of 5.1% with vs. 5.0% without). Lastly, testing use of the '--write_single_snp' option resulted in a lower mean error using the option vs. not using the option (3.9% vs. 4.3%, respectively).

Parameter values resulting in lower mean estimated SNP error rates were used to perform a proper run of the STACKS pipeline by manually running each program in sequence to generate a *Webbhelix multilineata* SNP dataset for 96 individuals from 4 sample localities (TGCP, PTP, STT and MPO), with a minimum population constraint of ~75% and a minimum individual constraint of 80%. The resulting dataset was comprised of 10,506 SNP loci, which was further

processed using PLINK to filter out SNPs in high LD pairs and SNPs showing extreme HWE deviation. The final dataset after post-processing was comprised of a total of 6,843 SNPs, which was used for downstream analyses.

Minimum Sample Sizes for Population Genomics in *Webbhelix*

Mean population genomic parameter estimations of H_O, H_S, G_{IS} and F_{ST} for each subsampling treatment for assessing minimum sample sizes were analyzed using an ANOVA to compare pairwise treatment estimations. For H_O, ANOVA analyses indicated that there was a statistically significant difference between treatments applied to each of the four sample localities in this study (TGCP, p = 3.79e-10; STT, p = 4.18e-05; MPO, p = 1.94e-10; PTP, p = 2.75e-14). Pairwise comparisons of treatments showed that for all four sample localities, the subsampling treatment of 2 individuals per population (2subsamp) produced estimates of H_O that were significantly different from estimates produced by all other subsampling treatments, except in STT where one pairwise comparison with 2subsamp was not significant (2subsamp vs. 4subsamp, p = 0.85). The 4subsamp treatment in STT was also significantly different than the 8subsamp treatment (p = 0.029).

The results of the ANOVA analyses for H_S were also significant for differences among treatments in all four sample localities (TGCP, p = 7.05e-14; STT, p = 2.33e-09; MPO, p = <2e-16; PTP, p = <2e-16). The pattern of pairwise differences was exactly the same as for H_O above (2subsamp vs. 4subsamp, p = 0.23; 4subsamp vs. 8subsamp, p = 0.0079).

For G_{IS} the ANOVA analyses for TGCP and STT were not statistically significant (p = 0.78 and p = 0.092, respectively), and no pairwise comparisons were significant either. The G_{IS} ANOVA analyses for MPO and PTP were slightly significant (p = 0.039 and p = 0.043, respectively), with the 2subsamp treatment being significantly different than the 6subsamp

treatment in MPO ($p = 0.016$) and 2subsamp being significantly different than 8subsamp in PTP ($p = 0.011$).

The ANOVA analyses assessing differences in mean estimates of F_{ST} among treatments were not significant for most of the sample locality pairs (TGCP-STT, $p = 0.66$; TGCP-MPO, $p = 0.78$; TGCP-PTP, $p = 0.077$; STT-MPO, $p = 0.83$), but were significant for differences in treatments in two pairs (STT-PTP, $p = 0.0019$; MPO-PTP, $p = 0.0018$); for STT-PTP, the 2subsamp estimate of F_{ST} was significantly different than the 14subsamp – 22subsamp treatments ($p = 0.00076$, 0.0022, 0.018, 0.018 and 0.012, respectively), and in MPO-PTP the 2subsamp estimate was significantly different than the 12subsamp estimate ($p = 0.035$).

Results of ANOVA analyses comparing estimates of H_O, H_S, G_{IS} and F_{ST} across subsampling treatments were not sufficient to assess minimum sample sizes for accurately estimating these parameters because these data were not normally distributed, violating an assumption of ANOVA. Therefore, non-parametric Kruskal-Wallis rank sum tests were also performed on estimates of these parameters (since normality of data is not assumed for this test) along with t-tests with pooled standard deviations for pairwise comparisons. **Tables 2.3 and 2.4** show p-values for pairwise comparisons between each subsampling treatment and the best case scenario subsampling treatment (22 individuals per population), for each of the three genetic diversity parameters (H_O, H_S, G_{IS}; **Table 2.3**) and for genetic differentiation (F_{ST}; **Table 2.4**). For H_O, Kruskal-Wallis tests indicated that there was a statistically significant difference between treatments applied to each of the four sample localities (TGCP, $p = 1.46e\text{-}05$; STT, $p = 0.018$; MPO, $p = 0.0061$; PTP, $p = 2.089e\text{-}06$). Pairwise comparisons of treatments using t-tests showed that for all four locations, the 2subsamp treatment produced estimates of H_O that were significantly different from estimates produced by all other subsampling treatments, again except

Table 2.3 – Pairwise comparisons using t-tests with pooled standard deviations of population genetic parameter (H_S, H_O and G_{IS}) estimate means from 11 subsampling treatments (22 individuals per sample locality, 20 per sample locality, 18 per locality, etc.) across four sample localities of *Webbhelix multilineata* (Timmons Grove County Park, TGCP; Pikes Timber Park, PTP; Steamboat Trace Trail, STT; Mandan Park, Omaha, MPO). Subsampling treatments consisted of 100 randomly resampled datasets from a 6,843 SNP loci dataset with 24 individuals per locality. Table shows pairwise comparisons between all subsampling treatments relative to the highest subsampling treatment (22 individuals per locality). Bold p-values with asterisks were statistically significant (* = weakly significant; ** = significant; * = strongly significant).**

T-Test Pairwise P-Values	Localities	TGCP	PTP	STT	MPO
	Treatments	22_SubSamp	22_SubSamp	22_SubSamp	22_SubSamp
Expected Heterozygosity (Hs)	20_SubSamp	0.92	0.93	0.95	0.94
	18_SubSamp	0.93	0.93	0.91	0.94
	16_SubSamp	0.92	0.93	0.91	0.96
	14_SubSamp	0.92	0.93	0.91	0.94
	12_SubSamp	0.92	0.93	0.91	0.87
	10_SubSamp	0.92	0.93	0.91	0.98
	08_SubSamp	0.92	0.67	0.39	0.76
	06_SubSamp	0.92	0.93	0.98	0.7
	04_SubSamp	0.92	0.93	0.062	0.87
	02_SubSamp	**4.80E-10*** **	**1.30E-14*** **	**4.80E-06*** **	**< 2.00E-16*** **
Observed Heterozygosity (Ho)	20_SubSamp	0.96	0.91	0.97	0.99
	18_SubSamp	0.96	0.91	0.97	0.87
	16_SubSamp	0.96	0.9	0.97	0.87
	14_SubSamp	0.96	0.91	0.97	0.87
	12_SubSamp	0.96	0.9	0.97	0.87
	10_SubSamp	0.96	0.96	0.97	0.87
	08_SubSamp	0.43	0.41	0.48	0.87
	06_SubSamp	0.96	0.84	0.97	0.85
	04_SubSamp	0.74	0.84	0.13	0.87
	02_SubSamp	**6.00E-07*** **	**9.70E-10*** **	**0.00138** **	**9.00E-08*** **
Inbreeding Coefficient (G_{IS})	20_SubSamp	0.99	0.99	1	1
	18_SubSamp	0.99	0.99	1	1
	16_SubSamp	0.99	0.99	1	0.96
	14_SubSamp	0.99	0.99	1	1
	12_SubSamp	0.99	0.55	1	1
	10_SubSamp	0.87	0.96	1	0.96
	08_SubSamp	0.87	0.59	1	1
	06_SubSamp	0.99	0.56	1	0.96
	04_SubSamp	0.87	0.99	0.22	0.86
	02_SubSamp	0.87	0.11	0.16	**0.03***

Table 2.4 – Pairwise comparisons using t-tests with pooled standard deviations of fixation index (F_{ST}) estimate means from 10 subsampling treatments (22 individuals per sample locality, 20 per sample locality, 18 per locality, etc.) across six pairings of four sample localities of *Webbhelix multilineata* (Timmons Grove County Park, TGCP; Pikes Timber Park, PTP; Steamboat Trace Trail, STT; Mandan Park, Omaha, MPO). Subsampling treatments consisted of 100 randomly resampled datasets from a 6,843 SNP dataset with 24 individuals per locality. The table shows pairwise comparisons between all subsampling treatments relative to the highest subsampling treatment (22 individuals per locality). Bold p-values with asterisks were statistically significant (* = weakly significant; ** = significant; * = strongly significant).**

T-Test Pairwise P-Values	Localities	TGCP - STT	TGCP - PTP	TGCP - MPO	STT - PTP	STT - MPO	MPO - PTP
	Treatments	22_SubSamp	22_SubSamp	22_SubSamp	22_SubSamp	22_SubSamp	22_SubSamp
Fixation Index (F_{ST})	20_SubSamp	0.94	0.8	0.95	0.95	0.96	1
	18_SubSamp	0.98	0.98	0.95	0.95	0.96	1
	16_SubSamp	0.99	0.8	0.95	0.81	0.96	1
	14_SubSamp	0.94	0.3	0.95	0.72	0.96	0.57
	12_SubSamp	0.94	0.62	0.95	0.72	0.96	1
	10_SubSamp	0.94	0.62	0.95	0.72	0.96	0.60
	08_SubSamp	0.94	0.11	0.95	0.81	0.96	0.09
	06_SubSamp	0.94	0.3	0.95	0.57	0.96	0.069
	04_SubSamp	0.94	0.3	0.95	**0.0046****	0.96	**0.018***

in STT where one pairwise comparison with 2subsamp was not significant (2subsamp vs. 4subsamp, p = 0.26) and the 4subsamp treatment was also significantly different than the 8subsamp treatment (p = 0.0044).

The results of the Kruskal-Wallis tests for H_S were also significant for differences among treatments in all four sample localities (TGCP, p = 1.78e-07; STT, p = 0.00025; MPO, p = 6.12e-06; PTP, p = 6.53e-08). Pairwise comparisons of treatments using t-tests showed that for all four locations, the 2subsamp treatment produced estimates of H_S that were significantly different from estimates produced by all other subsampling treatments. Additionally, for STT the 4subsamp treatment estimate was significantly different than the 8subsamp estimate (p = 0.00094) and the 12subsamp – 18subsamp estimates of H_S (p = 0.031, 0.017, 0.034 and 0.031, respectively), and was nearly significantly different than the 22subsamp treatment (p = 0.062).

For G_{IS} the Kruskal-Wallis tests for TGCP and STT were not statistically significant ($p = 0.55$ and $p = 0.93$, respectively), and no t-test pairwise comparisons were significant either. The G_{IS} Kruskal-Wallis tests for MPO and PTP were also not significant ($p = 0.32$ and $p = 0.13$, respectively), but oddly the t-test comparisons between the 2subsamp treatment estimate for MPO were significantly different than estimates from the 6subsamp ($p = 0.021$), 8subsamp ($p = 0.022$), 12subsamp ($p = 0.022$), 14subsamp ($p = 0.022$), 18subsamp ($p = 0.022$), 20subsamp ($p = 0.044$) and 22subsamp ($p = 0.03$) treatments. In PTP, t-test comparison between the 2subsamp treatment estimate and the 8subsamp estimate was also significant ($p = 0.014$).

For F_{ST}, Kruskal-Wallis tests and t-tests showed no significant results for differences in treatment estimates or pairwise comparisons among treatments for the TGCP-STT, TGCP-MPO or STT-MPO locality differentiation pairs. The Kruskal-Wallis test for differences among treatment estimates of F_{ST} for TGCP-PTP was significant ($p = 0.00032$), but curiously no pairwise t-test comparisons of treatment estimates were significantly different for this locality pair. The converse was true for STT-PTP; the Kruskal-Wallis test was not significant for differences among treatment estimates, but pairwise t-test comparisons between the 4subsamp treatment and 8subsamp – 22subsamp treatment estimates were significantly different ($p = 0.017$, 0.035, 0.035, 0.00082, 0.0012, 0.0046, 0.0046 and 0.0046, respectively). Lastly, the Kruskal-Wallis test for differences among treatments for MPO-PTP was significant ($p = 0.0075$) and t-test pairwise comparisons of the 4subsamp treatment estimate was significantly different than the estimates from the 12subsamp – 22subamp treatments ($p = 0.018$, 0.09, 0.018, 0.018, 0.018 and 0.018, respectively). Based on mean estimates of H_O, H_S, G_{IS} and F_{ST} for resampled localities of *Webbhelix multilineata*, it appears that a minimum of 6 individuals were needed to accurately estimate these population genetic parameters for all four localities assessed.

Boxplots generated from genetic diversity estimates can be seen in **Figures 2.1 – 2.4**. **Figure 2.1** shows boxplots for H_O estimates across subsample treatments for TGCP, PTP, MPO and STT. Median and mean estimates for each treatment are shown within each box as a solid, bold black line and as a red diamond, respectively. For the TGCP sample locality, variability in H_O estimates ranged from 0.152 to 0.156 for the 22 individuals per population subsampling treatment, to as much as 0.133 to 0.18 for the 2 individuals per population subsampling treatment. For PTP, estimates varied from 0.048 to 0.052 for the 22 subsampling treatment, to as much as 0.032 to 0.067 for the 2 subsampling treatment. For MPO, estimates varied from 0.101 to 0.108 for the 22 subsampling treatment, to as much as 0.061 to 0.144 for the 2 subsampling treatment. And for the STT sample locality, estimates of H_O varied from 0.111 to 0.118 for the 22 subsampling treatment, to as much as 0.08 to 0.156 for the 2 subsampling treatment.

Figure 2.2 shows boxplots for H_S estimates across subsample treatments for all four sample localities. For TGCP, variability in H_S estimates ranged from 0.154 to 0.157 for the 22 individuals per population subsampling treatment, to as much as 0.134 to 0.18 for the 2 individuals per population subsampling treatment. For PTP, estimates varied from 0.055 to 0.057 for the 22 subsampling treatment, to as much as 0.042 to 0.07 for the 2 subsampling treatment. For MPO, estimates varied from 0.117 to 0.121 for the 22 subsampling treatment, to as much as 0.091 to 0.142 for the 2 subsampling treatment. And for the STT sample locality, estimates of H_S varied from 0.123 to 0.127 for the 22 subsampling treatment, to as much as 0.097 to 0.152 for the 2 subsampling treatment.

Figure 2.3 shows boxplots for G_{IS} estimates across subsample treatments for sample localities. For TGCP, variability in G_{IS} estimates ranged from 0.004 to 0.017 for the 22 individuals per population subsampling treatment, to as much as -0.086 to 0.087 for the 2

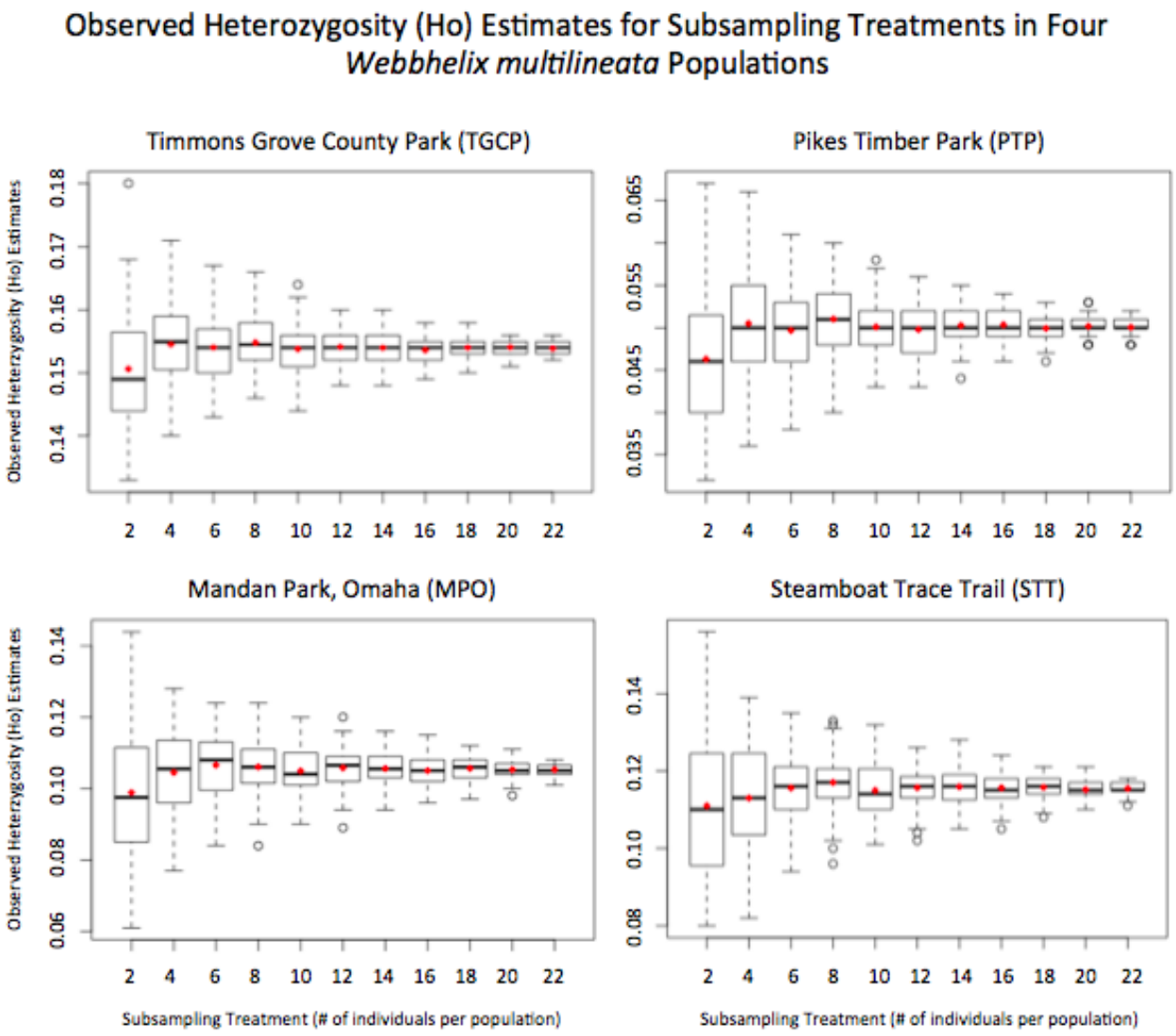


Figure 2.1 – Boxplots showing observed heterozygosity (H_O) estimates generated from 100 resampled *Webbhelix multilineata* datasets of 6,843 SNPs for each of 11 subsampling treatments (2 – 22 individuals per population, in increments of 2), for each of four sample localities. Open circles are outliers; black bar in each box is the median value; red diamond in each box is the mean value. The smallest subsampling treatment with a mean statistically non-different from subsampling treatment 22, applicable to all four sample localities, was subsampling treatment 4 (based on pairwise comparison of treatment means via ANOVA and Kruskal-Wallis rank sum and T-tests; see Table # for pairwise p-values).

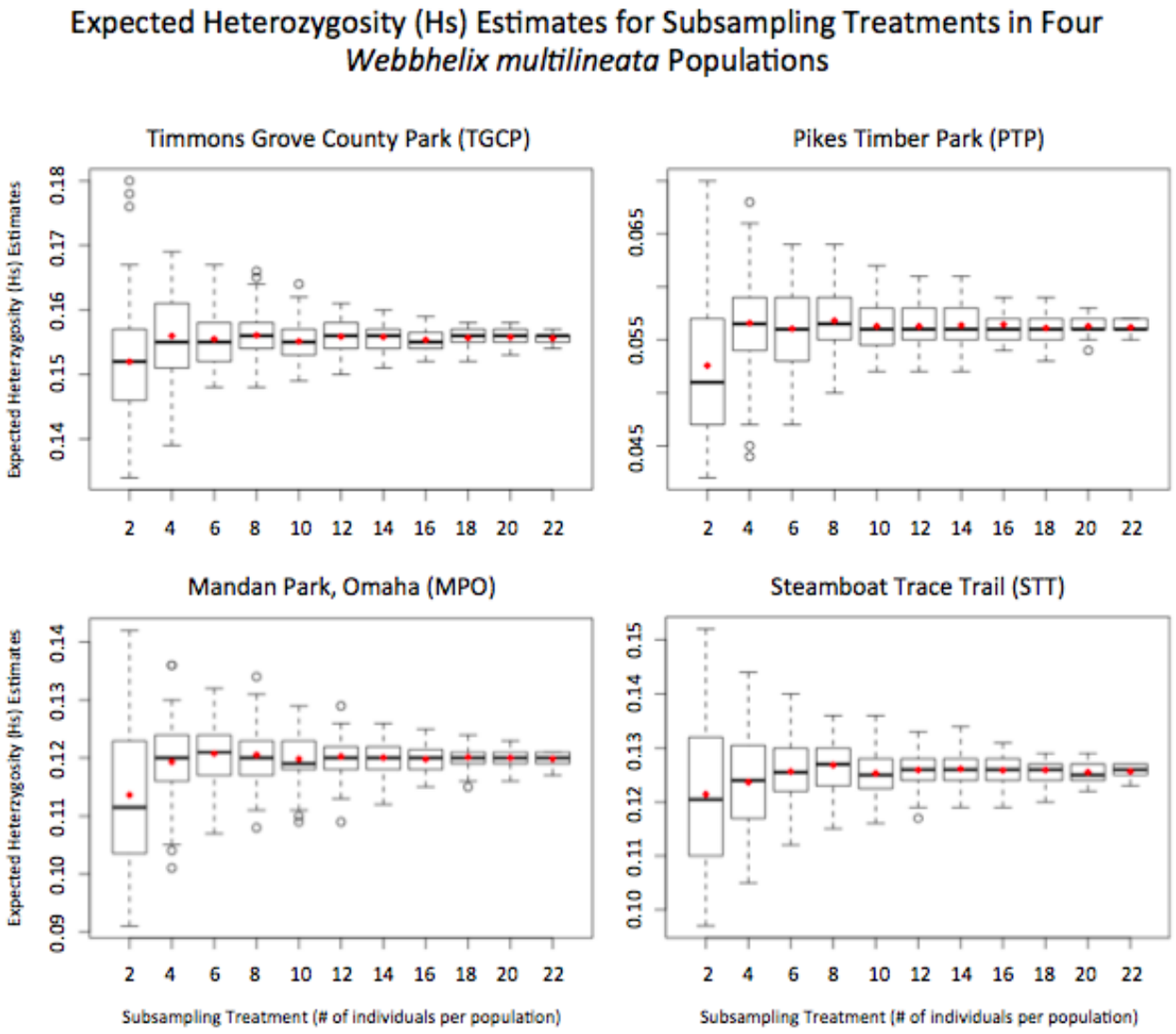


Figure 2.2 – Boxplots showing expected heterozygosity (H_S) estimates generated from 100 resampled *Webbhelix multilineata* datasets of 6,843 SNPs for each of 11 subsampling treatments (2 – 22 individuals per population, in increments of 2), for each of four sample localities. Open circles are outliers; black bar in each box is the median value; red diamond in each box is the mean value. The smallest subsampling treatment with a mean statistically non-different from subsampling treatment 22, applicable to all four sample localities, was subsampling treatment 6 (based on pairwise comparison of treatment means via ANOVA and Kruskal-Wallis rank sum and T-tests; see Table # for pairwise p-values).

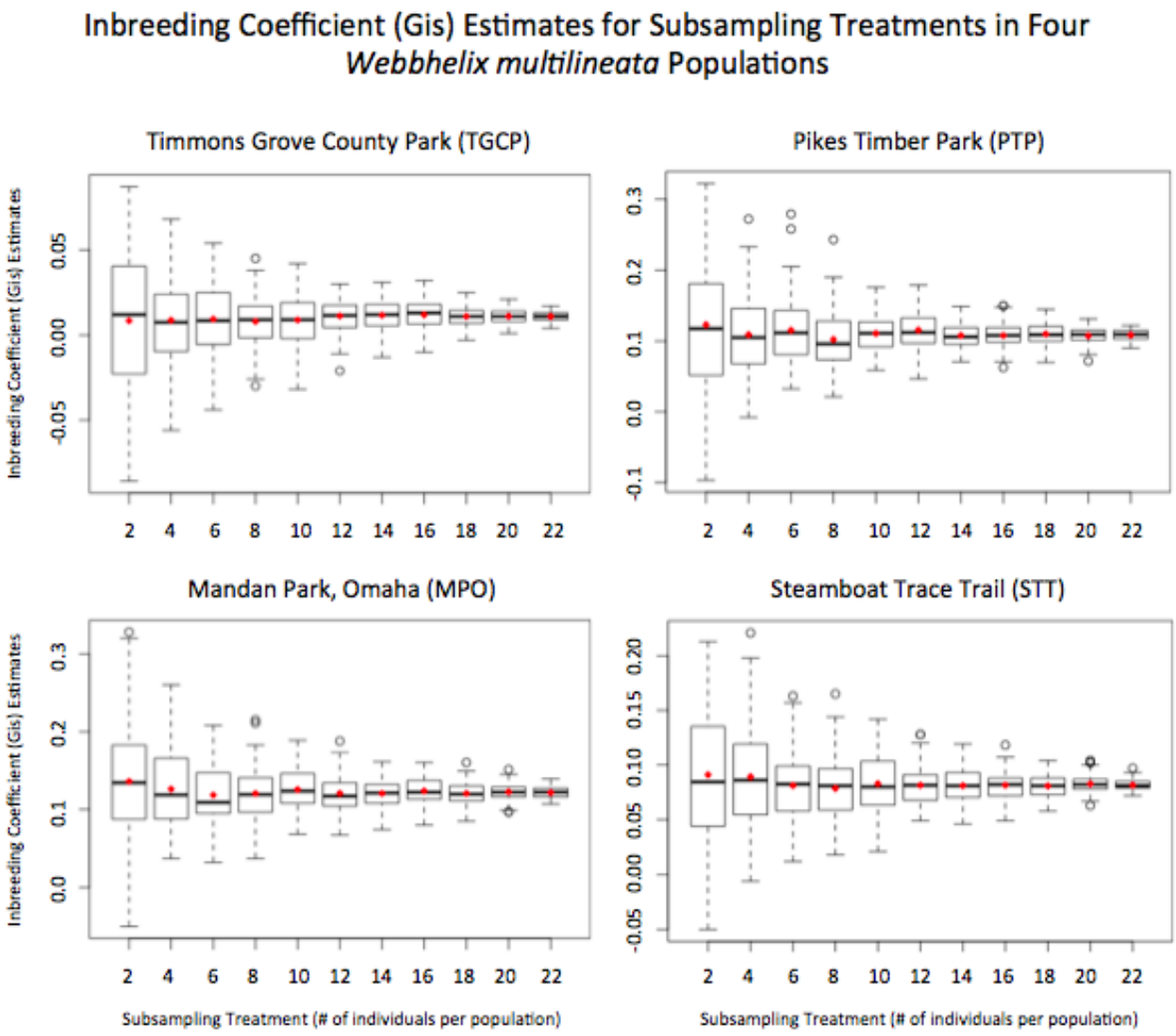


Figure 2.3 – Boxplots showing inbreeding coefficient (G_{IS}) estimates generated from 100 resampled *Webbhelix multilineata* datasets of 6,843 SNPs for each of 11 subsampling treatments (2 – 22 individuals per population, in increments of 2), for each of four sample localities. Open circles are outliers; black bar in each box is the median value; red diamond in each box is the mean value. The smallest subsampling treatment with a mean statistically non-different from subsampling treatment 22, applicable to all four sample localities, was subsampling treatment 4 (based on pairwise comparison of treatment means via ANOVA and Kruskal-Wallis rank sum and T-tests; see Table # for pairwise p-values).

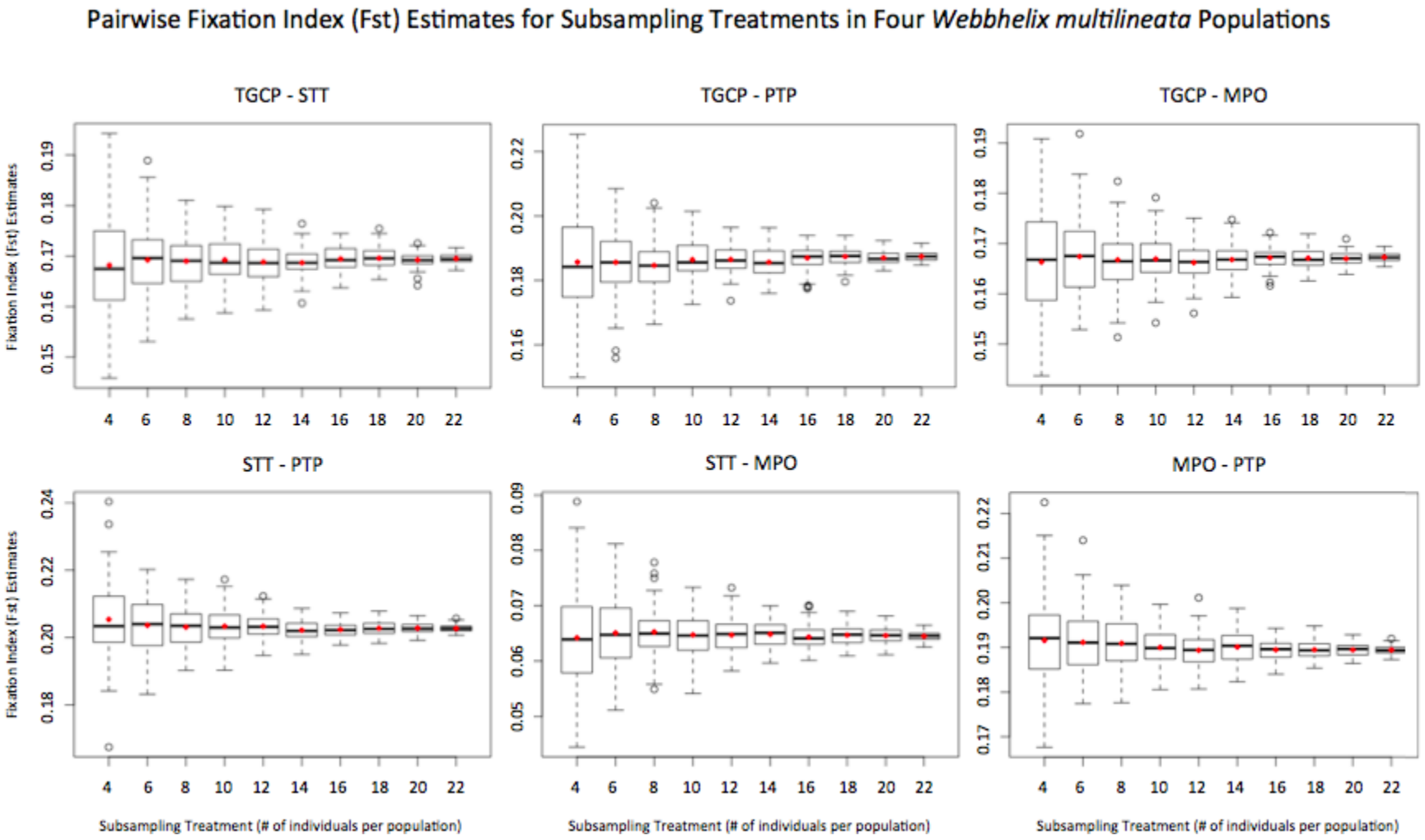


Figure 2.4 – Boxplots showing fixation index (F_{ST}) estimates generated from 100 resampled *Webbhelix multilineata* datasets of 6,843 SNPs for each of 10 subsampling treatments (4 – 22 individuals per population, in increments of 2), for each of six pairwise sample localities. Open circles are outliers; black bar in each box is the median value; red diamond in each box is the mean value. The smallest subsampling treatment with a mean statistically non-different from subsampling treatment 22, applicable to all six pairwise sample localities, was subsampling treatment 6 (based on pairwise comparison of treatment means via ANOVA and Kruskal-Wallis rank sum and t-tests).

individuals per population subsampling treatment. For PTP, estimates varied from 0.09 to 0.122 for the 22 subsampling treatment, to as much as -0.097 to 0.322 for the 2 subsampling treatment. For MPO, estimates varied from 0.107 to 0.139 for the 22 subsampling treatment, to as much as -0.05 to 0.328 for the 2 subsampling treatment. And for the STT sample locality, estimates of G_{IS} varied from 0.072 to 0.097 for the 22 subsampling treatment, to as much as -0.05 to 0.213 for the 2 subsampling treatment.

Figure 2.4 shows boxplots for F_{ST} estimates across subsample treatments for six pairwise combinations of the four sample localities. For the TGCP – STT locality pair, variability in F_{ST} estimates ranged from 0.167 to 0.172 for the 22 individuals per population subsampling treatment, to as much as 0.146 to 0.194 for the 4 individuals per population subsampling treatment. For the TGCP – PTP pair, estimates varied from 0.185 to 0.192 for the 22 subsampling treatment, to as much as 0.15 to 0.225 for the 4 subsampling treatment. For TGCP – MPO, estimates varied from 0.165 to 0.169 for the 22 subsampling treatment, to as much as 0.144 to 0.191 for the 4 subsampling treatment. For STT – PTP, estimates varied from 0.201 to 0.206 for the 22 subsampling treatment, to as much as 0.168 to 0.240 for the 4 subsampling treatment. For STT – MPO, estimates varied from 0.062 to 0.066 for the 22 subsampling treatment, to as much as 0.044 to 0.089 for the 4 subsampling treatment. And for the MPO – PTP sample locality pair, estimates of F_{ST} varied from 0.187 to 0.192 for the 22 subsampling treatment, to as much as 0.168 to 0.222 for the 4 subsampling treatment.

Discussion

Determination of an appropriate sampling scheme is an important aspect of research study design, particularly so for population genetic studies (Nazareno et al, 2017). Sampling adequate numbers of individuals per sample locality to accurately estimate population parameters of interest is critical, but curving data generation costs is also desirable. Although

genome sequencing is currently affordable enough for even modest labs to pursue, and the cost continues to decline each year, oversampling individuals can still quickly drive the price tag up beyond the budget of most studies, especially if they call for sampling many localities. Aside from cost considerations, limiting sampling may be critical for conservation studies on systems with restrictions regarding invasive or destructive sampling. Having knowledge of minimum sample sizes for accurate estimation of population parameters could be used to persuade agencies to allow limited amounts of invasive sampling, opening the door for genomic approaches and circumventing troublesome non-invasive sampling techniques.

The present study contributes to a growing body of evidence informing appropriate sample sizes for population genomic studies (Nazareno et al, 2017; Qu et al, 2020; McLaughlin & Winker, 2020; Li et al, 2020). Our results indicate that on average, estimates of observed and expected heterozygosity (H_O and H_S), inbreeding coefficient (G_{IS}), and pairwise fixation index (F_{ST}) were not statistically significantly different when utilizing 6 or more individuals per population, and for three out of four sample localities (TGCP, PTP and MPO) as few as 4 individuals per population were sufficient to accurately estimate these parameters for *Webbhelix multilineata*. In terms of variability in parameter estimates across subsampling treatments, in general estimate ranges gradually increased from 22 individuals per population down to about 12 to 10 individuals per population, after which they increased more sharply for successively smaller subsampling treatments. The largest deviations from estimate means, for the smallest subsampling treatment (4 individuals per population) with mean estimates statistically non-different than the largest (22 individuals per population), were ~ 0.1 for G_{IS}, which is a fairly substantial amount for an estimate of inbreeding coefficient to be off in terms of accuracy. However for H_O, H_S and F_{ST} the largest deviations from those means were only ~ 0.03 to 0.04,

which, comparatively, is a reasonably small loss of accuracy considering the dramatic cost reduction or expansion of number of sample localities that utilizing only 4 individuals per population affords, and considering that on average estimates will be statistically non-different than using 22 individuals per population.

These results are relatively consistent with previous studies on minimum sample sizes for population genomics using empirical datasets. Nazareno et al (2017), one of the first empirical studies to utilize a large SNP dataset and substantial population sampling to explicitly test the effect of sample size on genetic diversity estimates in a non-model system, found that for an Amazonian tree species (*Amphirrhox longifolia*) 8 individuals per sample locality was sufficient to accurately estimate measures of genetic diversity (number of effective alleles, observed/expected heterozygosity and fixation index) when 1000 or more SNPs were used. More recently, three additional studies have been published that used a similar approach to Nazareno et al to explicitly test minimum sample sizes for population genomics. For a species of whitefly (*Bemisia tabaci*), Qu et al (2020) found that as few as 4 individuals per population was enough to generate accurate estimates of effective number of alleles, heterozygosity and genetic differentiation. Li et al (2020) determined that for the lady beetle *Harmonia axyridis*, 6 individuals was the minimum sample size needed for accurate estimation of the same population genetic diversity parameters; both studies utilized 3000 SNP loci in their assessments of minimum sample size. McLaughlin & Winker (2020) took this process a step further, using SNP datasets from eight diverging Beringian bird lineages with varying demographic and evolutionary histories to investigate sample size effects on coalescent-based estimates of parameters such as effective population size, migration and divergence time. In general they found that their data performed best for estimating these parameters when using at least 6

individuals per population. Overall, the results of these studies and our own suggest that between 4 to 8 individuals per population are minimally necessary to accurately estimate population genetic parameters, depending on study system and parameters of interest.

There are a few perspectives from which to consider the results presented here that we think are important to highlight in terms of informing study designs. First and foremost is that, as previously stated, different study systems do appear to have variable minimum sampling requirements for estimating different population parameters. McLaughlin & Winker (2020) showed that divergence levels in a study system impact the adequacy of sampling numbers for estimating different parameters, with shallower levels of divergence necessitating more individual sampling per population. Given that sample size requirements seem to vary between lineages and that empirical studies on sampling effects are relatively few at present, we recommend that, when possible, first-time population genomic studies on new study systems pursue an assessment of minimum sample sizes for estimating desired demographic parameters. Doing so provides the most effective guideline for study design in terms of sampling for a new system and has the broader benefit of expanding our baseline for best practices in the field.

While this suggestion echoes what is advocated in similar previous studies, it is understandable that pursuing the additional endeavor is not always feasible or desirable; as such, the results provided here and in previous studies do still offer a reasonable guideline for future studies to follow, depending on the circumstance. For example, in studies seeking to work on a protected species with strict invasive or destructive sampling restrictions, an argument could be proposed to management agencies for pursuing an approach that will result in genomic data from high quality DNA sampled using only 4 to 6 individuals per population. As previously argued above and discussed extensively in the literature (Allendorf et al, 2010; Frankham, 2010; Saura

et al, 2013; Toro et al, 2014; Frankham et al, 2017; Li et al, 2020; Rossetto et al, 2021), there are numerous benefits associated with transitioning conservation studies to genomic-based approaches over traditional ones. The results of the present and previous research provide strong support for designing population genomic studies using minimum sample sizes that can be done at relatively low cost and impact to endangered populations, with the caveat that some loss of accuracy in parameter estimates may occur at lower sampling levels.

For studies on non-endangered species, consideration of these results on minimum sample sizes comes down to a three-way tradeoff between spending less money, sampling more localities, and maintaining higher accuracy of estimates. If funding is very limited then obviously sampling at the lowest end per population would be best to limit costs as much as possible. If sufficient funding is available then sampling 20 or more individuals per population all but guarantees accurate estimates. Likely in most cases a balance of the two will be preferred; sampling between 10 to 12 individuals per population in most cases results in a reasonably tight range of possible estimate values while keeping the project price tag at an affordable level. These considerations of course must be further balanced if sampling many different localities is called for in a study, by sacrificing either additional expense or parameter estimate accuracy.

To summarize, a proper sampling scheme for population genomic studies is critical to their success and requires appropriate foresight and planning, to which studies on minimum sampling are incredibly useful. For the terrestrial gastropod mollusk *Webbhelix multilineata*, 6 individuals were sufficient to accurately estimate observed and expected heterozygosity, inbreeding coefficient and fixation index on average compared to all treatments with more individuals per population, for all four sample localities tested. However, minimum sampling requirements appear to be somewhat variable between different taxa, owing to factors such as

level of lineage divergence, and so more studies assessing these requirements across a broader sweep of taxa are needed to expand our view of how sampling impacts estimation of important demographic parameters. Even so, available studies such as the present one provide a valuable guideline for future population genomic studies to consider.

References

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, *11*(10), 697-709.

Beja-Pereira, A. L. B. A. N. O., Oliveira, R., Alves, P. C., Schwartz, M. K., & Luikart, G. (2009). Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources*, *9*(5), 1279-1301.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomics, Genetics*, *1*, 171-182.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*.

Chen, B., Cole, J. W., & Grond-Ginsbach, C. (2017). Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Frontiers in Genetics*, *8*(167).

Feder, M. E., & Mitchell-Olds, T. (2003). Evolutionary and ecological functional genomics. *Nature reviews genetics*, *4*(8), 649-655.

Frankham, R. (2010). Where are we in conservation genetics and where do we need to go?. *Conservation Genetics*, *11*(2), 661-663.

Frankham, R., Ballou, J. D., Ralls, K., Eldridge, M. D. B., Dudash, M. R., Fenster, C. B., Lacy, R. C., & Sunnucks, P. (2017). *Genetic management of fragmented animal and plant populations*. Oxford University Press.

Hale, M. L., Burg, T. M., & Steeves, T. E. (2012). Sampling for Microsatellite-Based Population Genetic Studies: 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies. *PloS one*, *7*(9), e45170.

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I., & Xu, C. (2004). Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *European Journal of Human Genetics*, *12*, 395-399.

Li, H., Qu, W., Obrycki, J. J., Meng, L., Zhou, X., Chu, D., & Li, B. (2020). Optimizing Sample Size for Population Genomic Study in a Global Invasive Lady Beetle, Harmonia axyridis. *Insects*, *11*(5), 290.

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298-299.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*, 28-41.

McLaughlin, J. F., & Winker, K. (2020). An empirical examination of sample size effects on population demographic estimates in birds using single nucleotide polymorphism (SNP) data. *PeerJ*, *8*, e9939.

McMahon, B. J., Teeling, E. C., & Höglund, J. (2014). How and why should we implement genomics into conservation?. *Evolutionary applications*, *7*(9), 999-1007.

Meek, M. H., & Larson, W. A. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, *19*, 795-803.

Meirmans, P. G. (2020). GENODIVE version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, *20*(4), 1126-1131.

Nazareno, A., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*.

Nei, M. (1987). Molecular Evolutionary Genetics. *Columbia University Press*, New York.

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.

Qu, W. M., Liang, N., Wu, Z. K., Zhao, Y. G., & Chu, D. (2020). Minimum sample sizes for invasion genomics: Empirical investigation in an invasive whitefly. *Ecology and evolution*, *10*(1), 38-49.

Rossetto, M., Yap, J. Y. S., Lemmon, J., Bain, D., Bragg, J., Hogbin, P., Gallagher, R., Rutherford, S., Summerell, B., & Wilson, T. C. (2021). A conservation genomics workflow to guide practical management actions. *Global Ecology and Conservation*, *26*, e01492.

Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, *3*, e1106.

Saura, M., Fernández, A., Rodríguez, M. C., Toro, M. A., Barragán, C., Fernández, A. I., & Villanueva, B. (2013). Genome-Wide Estimates of Coancestry and Inbreeding in a Closed Herd of Ancient Iberian Pigs. *PLoS One*, *8*(10), e78314.

Scheet, P., & Stephens, M. (2008). Linkage Disequilibrium-Based Quality Control for Large-Scale Genetic Studies. *PLoS Genetics*, *4*(8), e1000147.

Shafer, A. B., Wolf, J. B., Alves, P. C., Bergstrom, L., Bruford, M. W., Brannstrom, I., Cooling, G., Dalen, L., Meester, L. D., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Hoglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., Krutzen, M., McKay, J. K., Norman, A. J., Ogden, R., Osterling, E. M., Ouborg, N. J., Piccolo, J., Popovic, D., Primmer, C. R., Reed, F. A., Roumet, M., Salmona, J., Schenekar, T., Schwartz, M. K., Segelbacher, G., Senn, H., Thaulow, J., Valtonen, M., Veale, A., Vergeer, P., Vijay, N., Vila, C., Weissensteiner, M., Wennerstrom, L., Wheat, C. W., & Zielinski, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in ecology & evolution*, *30*(2), 78-87.

Toro, M. A., Villanueva, B., & Fernández, J. (2014). Genomics applied to management strategies in conservation programmes. *Livestock Science*, *166*, 48-53.

Tsykun, T., Rellstab, C., Dutech, C., Sipos, G., & Prospero, S. (2017). Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus Armillaria cepistipes. *Heredity*, *119*(5), 371-380.

von Thaden, A., Nowak, C., Tiesmeyer, A., Reiners, T. E., Alves, P. C., Lyons, L. A., Mattucci, F., Randi, E., Cragnolini, M., Galian, J., Hegyeli, Z., Kitchener, A. C., Lambinet, C., Lucas, J. M., Molich, T., Ramos, L., Schockert, V., & Cocchiararo, B. (2020). Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels. *Molecular ecology resources*, *20*(3), 662-680.

Waits, L. P., & Paetkau, D. (2005). Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. *The Journal of Wildlife Management*, *69*(4), 1419-1433.

CHAPTER 3. POPULATION GENOMICS AND COALESCENT SIMULATION OF REFUGIAL EXPANSION IN *WEBBHELIX MULTILINEATA* (MOLLUSCA, POLYGYRIDAE)

Jermaine Mahguib

Iowa State University, Ames, Iowa 50011

Modified from a manuscript to be submitted to *Molecular Ecology*

Abstract

Major climatic shifts such as the repeated glaciations of the Pleistocene in North America have huge consequences on entire assemblages of organisms, evidence of which can persist for thousands of years and be observed in modern populations. The southern extent of the massive Laurentide ice sheet plowing over half of the American Midwest during the last ice age contracted species ranges into southern refugia, dramatically impacting the genetic diversity and structuring of sessile and low dispersal species. The terrestrial gastropod *Webbhelix multilineata*, the striped white-lip snail, is an interesting molluscan system in this regard because it has an incredibly broad geographic range stretching across the American Midwest, parts of the South, and as far east as Pennsylvania. This range encompasses large regions of previously glaciated and unglaciated land, meaning that *Webbhelix* underwent extensive range contraction and expansion as glaciers expanded and receded. For a species with poor vagility and reproductive dispersal like *Webbhelix*, evidence of serial founder effects is expected to be reflected in estimates of current genetic diversity. This species is also contemporarily restricted to moist, wooded riparian habitat which in recent history has largely been lost and become highly fragmented across the range. Here, population genetics of *Webbhelix multilineata* was assessed by utilizing the first and (at present) only genomic dataset generated for this species, and one of only a relatively few genomic datasets generated for terrestrial gastropods. Genomic libraries

were prepared using a ddRADseq approach, and an 8,480 SNP dataset was used to estimate genomic structure and diversity. Additionally, coalescent simulations were used for testing models of refugial expansion to assess the mode and timing of expansion. Structure analysis suggested 4 population clusters: two in the Midwest, one to the south in the Mississippi River region, and one to the east in southern Ontario. Genetic diversity estimates were highest in the south, and generally decreased further north and away from the Mississippi River. Assessment of post-glacial expansion models suggested *Webbhelix* stuck to major river corridors while expanding its range, and that expansion either began as earlier as 20,000 years ago around the LGM, or more northerly refugia existed for this species than initially considered.

Introduction

North American Glaciations during the Pleistocene

The past 700,000 years of Earth's geologic history were dominated by major glacial periods that characterized the late Pleistocene (Martin & Neuner, 1978; Hewitt, 1996; Taberlet et al, 1998; Hewitt, 2000). In the northern hemisphere, continental ice sheets grew south to maximal glacial boundaries and receded north again in oscillating 100,000 year cycles interspersed with shorter, warmer interglacial periods such as at present. In North America during the most recent glacial period (starting ~115,000 years ago), the massive Laurentide ice sheet grew to cover all of central and eastern Canada and extended south into the American Midwest, completely covering Minnesota and Michigan, about three quarters of Wisconsin and North Dakota, and roughly half of South Dakota, Iowa, Illinois, Indiana and Ohio under 2 miles of ice until the last glacial maximum (LGM) some 20,000 years ago (Hewitt, 1996; Clark et al, 2009; Hewitt et al, 2018).

Species Range Contractions and Expansions, and Genetic Consequences

The southward advance of ice sheets during glacial periods decimated previously established biomes and contracted and shifted temperate zones and species ranges toward the equator (Hewitt, 1996). For highly mobile species, these contractions may have amounted to mass redistributions of populations toward southern refugia as northern habitat were bulldozed by ice, but for immobile and slow dispersing species, like terrestrial snails, the advance of ice sheets may have eliminated or reduced the frequency of some alleles as northern populations were wiped out.

When glaciers receded during interglacial periods and northern habitat became available again, species expanded their ranges from refugial populations. Typically, expansion is led by subsets of populations that occupy the northern-most extent of refugial ranges; as these individuals (representing a subset of the total genetic diversity of a species) establish new populations and grow exponentially, genetic diversity is lost via founder effects (Hewitt, 1996; Slatkin & Excoffier, 2012). A series of founder effects are expected to result in a gradation of genetic diversity along expansion routes, with diversity decreasing further away from founding refugia.

Time Frame for Glacial Retreat and Biotic Expansion

Genetic diversity gradients due to founder effects can give us a sense of expansion routes species took post glacial retreat, which can offer an opportunity to model and test scenarios of refugial expansion using coalescent approaches. For this, a reasonable time frame in which refugial expansion could have occurred is needed. The LGM represents roughly the last time period in which the ice sheets were at their maximum southern extent (in the northern hemisphere) and full on ice age conditions existed (Hewitt, 1996; Clark, 2009). Thereafter, ice sheets began to recede and continued to do so through the onset of the current interglacial period,

which began some 18,000 years ago. As the glaciers retreated and new land opened up, organisms would have been able to expand into these areas, depending on their tolerance for colder conditions nearer the ice sheet and how far south their refugia were to begin with. Species with a higher cold tolerance may have occupied refugia directly south of the ice sheet and been able to begin expanding their range northward, following the receding glaciers; species with a lower cold tolerance likely had refugia further south from the ice and pursued the glacial retreat only after temperate zones shifted further north again.

For the former case, it seems reasonable to apply a time frame for refugial expansion that stretches back to the LGM as the earliest possible start time; for the latter case, a later start time would need to be selected. The last 700,000 years of the Pleistocene saw a continuous global trend toward increasing seasonality that ultimately settled at the planet's contemporary seasonality between 12,000 and 8,000 years ago (Martin & Neuner, 1978), corresponding to the time frame that saw the mass extinction of North American mega fauna and the end of the Pleistocene. For a less cold hardy species, around the beginning of this time frame may be a reasonable starting time to apply for parameterization of refugial expansion scenarios.

Webbhelix multilineata and Timing for Range Expansion

In the case of the terrestrial gastropod *Webbhelix multilineata* (the striped whitelip snail), an argument can be made for either start time. *Webbhelix* is an interesting study system, owing to a very broad geographic range that spans nearly the entire American Midwest, south into parts of Arkansas, northeast into parts of Canada, and as far to the east as Pennsylvania where it has recently been rediscovered after more than 70 years. What makes this broad distribution so interesting, apart from the area being so large for such a poor disperser, is that roughly half of it encompasses regions that have experienced cyclical glaciations during the Pleistocene. Additionally, *Webbhelix* has particular habitat requirements; it is restricted to woodland riparian

habitat such as river floodplains, wetlands and marshes, all of which have become highly fragmented throughout the Midwest due to more than a century of human resourcing and agricultural development. These features make the striped whitelip an interesting subject for both population genomic study and investigation of timing for refugial expansion.

Considering that contemporary populations of *Webbhelix* are most active during the warm spring and summer months and that individuals form an epiphragm to seal themselves in their shells and hibernate during the cold winter months, an argument can be made for using a time frame for refugial expansion with a later post-LGM start time. However, Nash et al (2018) have reported finding *Webbhelix multilineata* shells in western Illinois relatively near the maximal extent of the Wisconsinin ice sheet that were radiocarbon dated to around the LGM, suggesting that *Webbhelix* may have occupied colder refugia than previously considered. Given this, an argument could certainly be made for using a time frame with an earlier start at the LGM.

Post-Glacial Habitat, Contemporary Range and Genetic Structure in *Webbhelix*

Refugia south of the Midwestern glaciation was heavily forested both before and after the LGM, but it's thought that with so much water locked up in the ice sheets much of the landscape was relatively dry before the glaciers began to recede (La Rocque, 1966; Martin & Neuner, 1978; Hewitt, 1996). This would have meant that a species like *Webbhelix* would have had to remain restricted to riparian habitat during glacial periods. Once glacial recession was underway, meltwater would have expanded river floodplains, wetlands, and marshes (*Webbhelix's* preferred habitat; Hotopp, 2005; Perez et al, 2014; personal observation). What is uncertain though is to what extent did meltwater from the receding ice sheet expand these habitats? Did most of the water drain away through river drainages, or was there significant runoff across large swathes of land that transformed southern refugia and the post-glacial Midwest into expansive marshes and

swamplands? In the case of the former, *Webbhelix* would have been constrained to river corridors while expanding its range into northern regions; in the case of the latter, *Webbhelix* would presumably have been able to expand outward in all directions, unconstrained by river corridors, to colonize new habitat in the north. Such a question can be addressed through appropriate modeling of testable scenarios of refugial expansion.

In addition to a reasonable time frame in which refugial expansion could have occurred and a question regarding the mode of that expansion to inform the modeling of testable scenarios, modeling and testing said scenarios is simplified when further informed by underlying genetic structure estimated from contemporary data. *Webbhelix multilineata's* restrictive and fragmented habitat, spread out across such a broad contemporary range (from eastern Arkansas up to Minnesota and eastern Nebraska out to southern Ontario) combined with a low individual and reproductive dispersal ability, is likely to have resulted in population genomic structure that can be utilized for this purpose. Population structure informs us of population subdivisions which can also be used to estimate population differentiation and relative rates of migration.

Population Genomics and Refugial Expansion in *Webbhelix*

In this study, single nucleotide polymorphism (SNP) data generated by ddRADseq from populations of *Webbhelix multilineata* sampled across a large portion of the species range were used to estimate measures of genetic diversity (H_O, H_S and G_{IS}) and population differentiation (F_{ST}), assess population genomic structure, and compare scenarios of refugial expansion using a coalescent approach. Genetic diversity measures were estimated to detect a signature of refugial expansion: a directional gradation in genetic diversity resulting from serial founder effects. Genomic structure was assessed to inform models of refugial expansion, as well as to inform estimation of levels of population differentiation and relative rates of migration across the range. Scenarios of refugial expansion were developed to test two hypotheses for the mode of post-

glacial refugial expansion in *W. multilineata*: 1) *Webbhelix* expanded north in all directions utilizing a broad expansion of its habitat resulting from glacial runoff; 2) *Webbhelix* expanded its range north via river corridors, namely the Mississippi and Ohio Rivers and their connecting drainages.

Materials and Methods

Species Populations and Sample Collection

Samples of *Webbhelix multilineata* were collected from across the American Midwest, with some additional population sampling to the south in the Mississippi River region and to the east in the southern-most region of Ontario, Canada (**Figure 3.1**). A total of 21 extant populations were located through the course of this study, all in woodland riparian habitat within various watersheds. Specific locality information can be seen in **Table 3.1**. Ten local populations were identified across central and eastern Iowa; three in the north/west of Illinois; two in southeast Minnesota; two along the southeastern border of Nebraska; one on the border of the southeastern corner of Missouri; one along Arkansas's eastern border; one at the southernmost tip of mainland Ontario and another on Pelee Island in Lake Erie.

Fieldwork to identify and collect samples from these local populations was conducted during the summers of 2016, 2017 and 2018, though most of this work was done in 2018. Due to limited available information on known living population locations, most of the localities identified in this study were found by visiting sites that were selected by visually scanning satellite imagery in the program GOOGLE EARTH PRO for suitable-looking habitat along rivers within the known species range. Once a list of candidate sites was established, location surveys and sampling were carried out at a total of 66 sites. Of the total sites visited, no living specimens of *W. multilineata* could be found in 31 of them, though some of these sites did contain old, faded shells of the species, suggesting past occupation. In addition, 13 sites were found to be

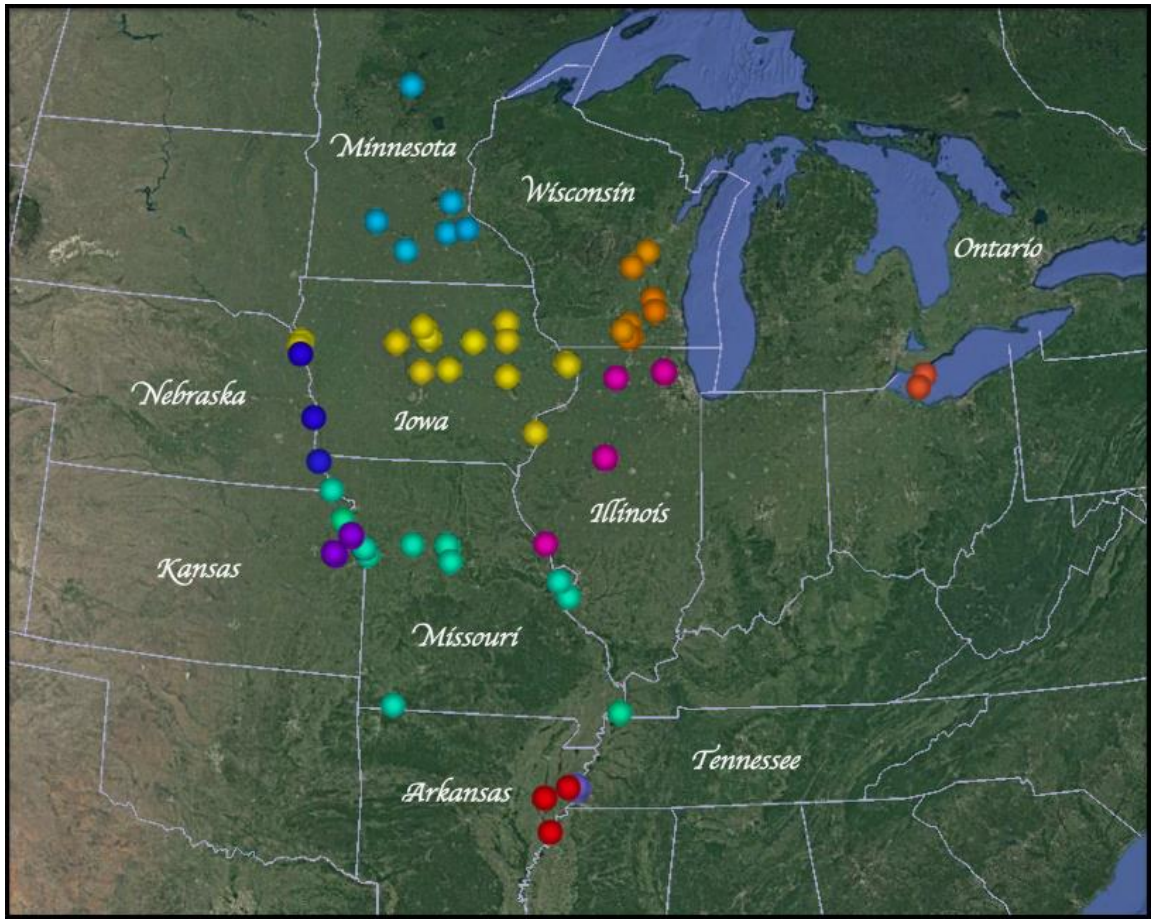


Figure 3.1 a – Map of field sites within the species range of *Webbhelix multilineata* visited for this study. A total of 66 localities (shown on map above) were searched for extant *Webbhelix* populations over the course of the summers of 2016, 2017 and 2018. Living individuals were found at only 21 localities (see Figure 5 continued). Color coding of field sites on map correspond to the state or province they are located within: red for sites in Arkansas; hot pink in Illinois; yellow in Iowa; dark purple in Kansas; sky blue in Minnesota; teal in Missouri; dark blue in Nebraska; dark orange in Ontario; light purple in Tennessee; and light orange in Wisconsin.

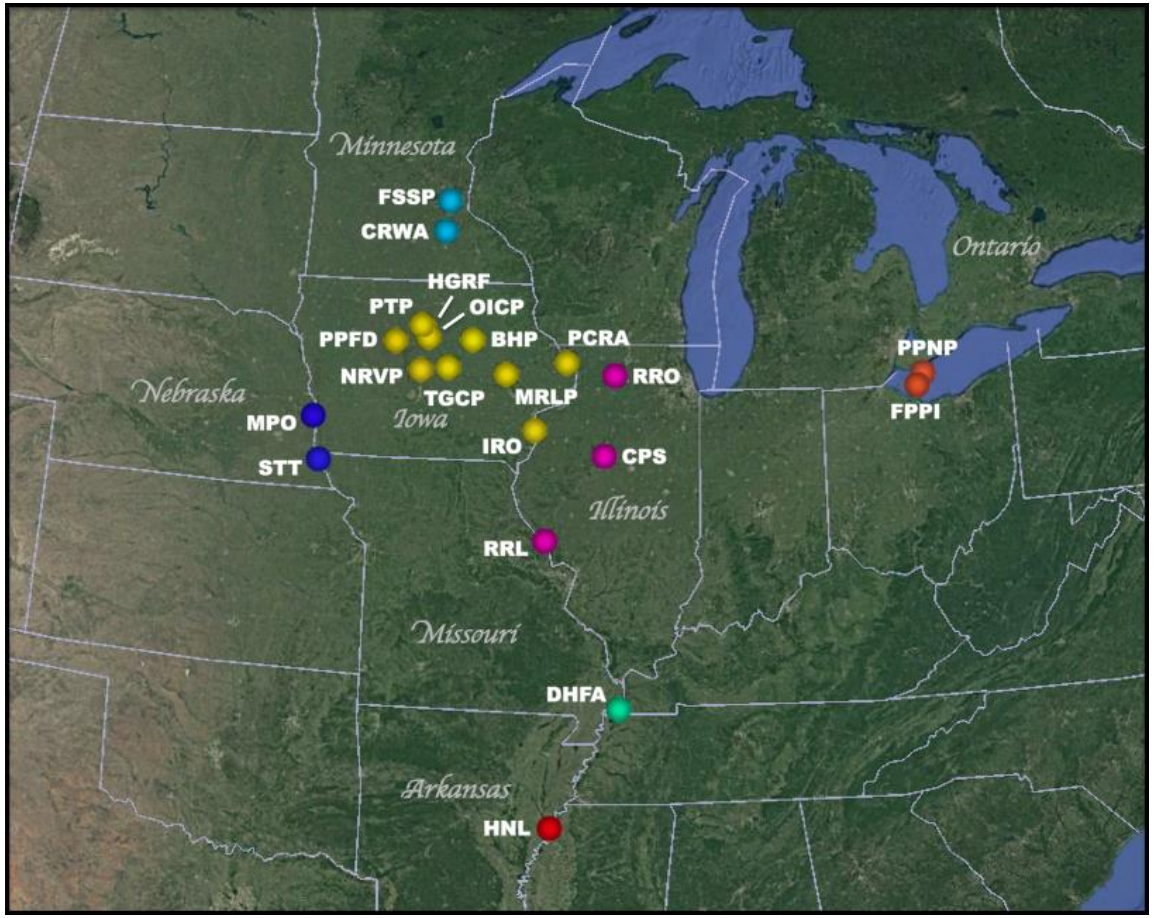


Figure 3.1 b – Map of 21 localities found in this study to contain extant populations of *Webbhelix multilineata*. Localities include: Horner Neck Lake (HNL), in Arkansas; Cooper Park South (CPS), Rip Rap Landing State Fish and Waterfowl Management Area (RRL) and Rock River in Oregon (RRO), all in Illinois; Black Hawk Park (BHP), Horse Grove Rietz Forest Area (HGRF), Iowa River northwest of Oakville (IRO), Manhattan/Robbins Lake Park (MRLP), Carr Woods in North River Valley Park (NRVP), Oakland-Iowa River County Park (OICP), Pleasant Creek Recreation Area (PCRA), Phinney Park in Fort Dodge (PPFD), Pike Timber Park (PTP), and Timmons Grove County Park (TGCP), all in Iowa; south Cannon River Wilderness Area (CRWA) and Fort Snelling State Park (FSSP), in Minnesota; woods east of Dorena-Hickman Ferry Access Point (DHFA), in Missouri; Mandan Park in Omaha (MPO) and Steamboat Trace Trail east of Peru (STT), in Nebraska; Fish Point Pelee Island (FPPI) and Point Pelee National Park (PPNP), in Ontario.

Table 3.1 – Locality information for all 21 field sites at which living populations of *Webbhelix multilineata* were found and sampled from.

Sample Localities	Locality Name	City/Area or County	State/Province	Collection Date(s)	Latitude	Longitude	# Individuals Sampled	# DNA Extractions
FSSP	Fort Snelling State Park	Minneapolis	Minnesota	6/6/2017 6/10/2017	44°53'11.80"N	93°10'47.30"W	32	24
CRWA	Cannon River Wilderness Area (south)	Northeast of Faribault	Minnesota	7/25/18	44°22'5.00"N	93°13'37.90"W	27	25
PTP	Pikes Timber Park	Southwest of Belmond	Iowa	5/20/18	42°46'58.08"N	93°38'39.17"W	24	24
HGRF	Horse Grove Rietz Forest Area	Southwest of Rowan	Iowa	5/22/18	42°43'42.78"N	93°34'3.72"W	25	10
OICP	Oakland-Iowa River County Park	Southwest Franklin County	Iowa	5/24/18	42°34'55.32"N	93°27'53.72"W	25	10
BHP	Black Hawk Park	Cedar Falls	Iowa	6/27/16	42°33'50.77"N	92°28'50.55"W	33	3
PPFD	Phinney Park	Fort Dodge	Iowa	7/7/17	42°30'18.37"N	94°12'11.52"W	37	1
PCRA	Pleasant Creek Recreation Area	Pleasant Creek	Iowa	6/2/18	42°12'51.69"N	90°22'54.78"W	24	24
TGCP	Timmons Grove County Park	Albion	Iowa	6/20/16	42° 5'45.62"N	93° 0'7.19"W	50	32
NRVP	Carr Woods, North River Valley Park	Ames	Iowa	7/31/17	42° 2'32.68"N	93°36'8.15"W	25	1
RRO	Rock River	Oregon	Illinois	7/28/18	42° 0'38.99"N	89°19'24.35"W	24	24
MRLP	Manhattan/Robbins Lake Park	Cedar Rapids	Iowa	7/12/16	42° 0'30.52"N	91°42'25.22"W	19	10
PPNP	Point Pelee National Park	Southeast of Leamington	Ontario	7/19/18	41°55'46.75"N	82°30'43.89"W	29	11
FPPI	Fish Point	Pelee Island	Ontario	7/18/18	41°43'56.37"N	82°40'23.71"W	26	24
MPO	Mandan Park	Omaha	Nebraska	7/9/17	41°11'38.03"N	95°55'46.66"W	38	25
IRO	Iowa River	Northwest of Oakville	Iowa	6/15/18	41° 6'11.07"N	91° 3'6.52"W	24	24
CPS	Cooper Park South	Peoria	Illinois	8/11/2018 8/18/2018	40°41'51.85"N	89°32'33.79"W	26	24
STT	Steamboat Trace Trail	Peru	Nebraska	7/19/16	40°28'58.36"N	95°43'38.69"W	57	26
RRL	Rip Rap Landing State Fish and Waterfowl Management Area	South of Belleview	Illinois	8/8/18	39°18'34.46"N	90°46'50.74"W	28	25
DHFA	Woods East of Dorena-Hickman Ferry Access Point	Mississippi County	Missouri	7/7/18	36°35'7.86"N	89°13'9.22"W	25	25
HNL	Hornor Neck Lake, St Francis National Forest	Northeast of West Helena	Arkansas	7/20/17	34°35'46.40"N	90°35'45.80"W	40	25

inaccessible on arrival, either due to physical impassability of the terrain or private ownership. Locality information of sites where no living specimens were found can be found at **https://github.com/jmahguib-ISU/Webbhelix_Project**.

On entering a wooded riverbank at a potential site, searches for *W. multilineata* were conducted using a combination of carefully scanning the ground and leaf litter for their rounded shell shape and characteristic red-brown striping, and carefully turning over wood logs and pieces of fallen tree bark, as individuals can often be found sheltered on the undersides of these. Individual snails were removed and placed whole into a 50 mL Falcon tube filled with 100% ethanol, killing them via rapid dehydration of soft body tissues and preserving them for later DNA extraction. A total of 638 individuals were collected in this manner from 21 sites. The number of specimens collected from each site varied from as few as 14 to as many as 56, depending on observed population abundance. At some sites it took as long as 5 or 6 hours to find 20 or fewer individuals, whereas at others up to 40 or 50 could be found in an hour or two. Low abundance sites were sampled more conservatively to lessen the impact on the population, while some sites with high abundance were sampled more liberally. Specific number of individuals taken from each location can be seen in **Table 3.1**.

DNA Extraction, Genomic Library Preparation and Sequencing

Genetic material was extracted from the soft body tissue of a total of 255 individual *Webbhelix multilineata* with Qiagen's DNeasy® Blood & Tissue Kit, using the Purification of Total DNA from Animal Tissues (Spin-Column) protocol (see **Table 3.1** for specific number of individuals DNA was extracted from per locality). Tissue used for extractions was taken from the anterior-most tip of each animal's foot (~20 – 30 mg). In the final step of the protocol, 175 uL of elution buffer was used to elute the DNA of each sample in a single elution step. Concentration of eluted DNA ranged from 14 ng/uL to as much as 240 ng/uL, measured using a

Thermo Scientific™ NanoDrop 2000 Spectrophotometer, with an average of 59 ng/uL across all samples. A portion of these stock DNA elutions was used to make ~20 ng/uL dilutions (where applicable) for use in preparation of genomic libraries.

Genomic libraries were prepared in-house using the double digest restriction-site associated DNA (ddRAD) protocol described in Peterson et al (2012), with modifications. Digestion of genomic DNA was carried out using 10 uL of ~20 ng/uL DNA extraction, 0.75 uL of 20 U/uL digestion enzyme PstI, 0.75 uL of 20 U/uL digestion enzyme MspI, 2 uL of 10X CutSmart buffer (New England BioLabs), and 6.5 uL of molecular grade ddH_2O per sample, and incubating at 37˚C for 600 minutes in an eppendorf Mastercycler® ep*gradient* thermocycler. Successful digestion was confirmed by agarose gel electrophoresis. Gels were 1% agarose and wells were loaded with 3 uL of digested sample plus 2 uL of loading dye; electrophoresed samples were compared to an undigested control, a lambda DNA/Hind III ladder (Promega) and a 1 kb DNA ladder (Fisher Scientific). Undigested control samples presented as a bright band at the 23+ kb range of the Hind III ladder, while successfully digested samples presented as a dim smear the length of the 1 kb ladder.

Ligation of adaptor sequences to restriction cut sites was accomplished using 10 uL of ~20 ng/uL digested DNA, 1.5 uL of 1 uM PstI adaptor (with barcode sequence), 0.6 uL of 25 uM MspI adaptor, 0.175 uL of 400,000 U/mL T4 DNA ligase, 2 uL of 10 mM adenosine triphosphate (ATP), 1 uL of 10X CutSmart buffer (New England BioLabs), and 4.725 uL of molecular grade ddH_2O per sample. Prepared samples were then incubated in a thermocycler at 16˚C for 360 minutes, 37˚C for 60 minutes, and a final incubation at 20˚C for 60 minutes. Adaptor ligations were carried out in batches of 48 samples or less, with the intention of submitting no more than 48 individuals per Illumina sequencing lane in order to conserve a

reasonable read depth of coverage (at least 10). A total of 16 different barcoded PstI adaptors were used on each batch of 48 or less samples, such that each of the first 16 samples were assigned a unique barcoded adaptor, each of the next 16 samples were assigned a unique barcode from the same pool of 16 available adaptors, and each of the remaining samples were assigned a unique barcode out of those same 16 adaptors (all barcoded adaptor primer sequences can be seen in **Table 2.1**). From there, the 48 or less samples in each batch were further uniquely tagged through the use of three index primers in the following polymerase chain reaction (PCR) step of the genomic library prep, such that the first 16 samples (each with a unique barcode) would be amplified using one index primer, the next 16 samples amplified using the second index primer, and the remaining samples amplified using the third index primer.

Amplification of double restriction cut site appended genomic fragments via PCR was executed using 1.8 uL of 10X diluted adaptor ligated DNA, 1 uL of 5 uM forward PCR primer, 1 uL of 5 uM index primer, 7.5 uL of 2X Phusion® High-Fidelity PCR Master Mix w/HF Buffer (New England BioLabs), and 3.7 uL of molecular grade ddH_2O per sample. Polymerase chain reaction was facilitated in a thermocycler with an initial 30 second DNA denaturation step at 98˚C, followed by 25 cycles of a 20 second denaturation step at 98˚C, 30 seconds of annealing at 60˚C, and 40 seconds of extension at 72˚C. At the end of 25 cycles, a final extension occurred at 72˚C for 12 minutes. As described above, three different index primers were used per batch of 48 or less samples such that every individual in a sample batch was tagged with a unique barcode/index combination. In order to distinguish samples *between* batches, each sample batch was slated for sequencing in an individual Illumina sequencing lane, adding a total of three layers of sample identification (1. barcode; 2. index; 3. sequencing lane/submission tracking number) for downstream demultiplexing of raw sequencing data. **Table 3.2** shows the

Table 3.2 – Sequencing order information and composition of sample batches per submission. All sequencing was carried out at the Iowa State University DNA Facility using Illumina HiSeq in high output mode, with read type SE and read length of 100 nucleotide bases.

ISU Sequencing Facility Submission Tracking #	ISU Pooled Sample ID #	Sequencing Lane #	Average Fragment Size (bp)	Sample Locality Composition	# of Individuals per Locality	# of Barcoded PstI Adaptors Used	# of Indexed PCR Primers Used
1083	11039	1	325 - 500	TGCP	24	8	3
1282	22330	1	325 - 500	STT	24	16	3
				MPO	24		
	22331	2	325 - 500	PTP	24	16	3
				TGCP	2		
				BHP	2		
				MRLP	2		
				NRVP	2		
				PPFD	2		
				MPO	2		
				HNL	2		
				RRL	2		
				RRO	2		
				PPNP	2		
				CRWA	2		
				DHFA	2		
1524	32794	1	325 - 500	MRLP	7	16	3
				FSSP	10		
				HNL	10		
				OICP	10		
				IRO	9		
	32795	2	325 - 500	IRO	1	16	3
				HGRF	10		
				RRL	10		
				RRO	10		
				CPS	10		
				PPNP	5		
	32796	3	325 - 500	PPNP	5	15	3
				FPPI	10		
				CRWA	10		
				PCRA	10		
				DHFA	10		

composition of all sample batches, including how many sample localities and individuals per locality comprised each batch, and how many barcode adapters and index primers were utilized per batch, as well as additional details relevant to sequencing submission.

After amplification each batch of PCR samples was pooled, and a portion of these pooled samples was fragment size selected (325 – 500 bp) using BluePippin (Sage Science) to remove leftover adapter fragments, unbound primers and primer dimers. Subsequent to the BluePippin protocol, size selected pooled batch samples were confirmed to contain mostly sequence fragments of the desired size via agarose gel electrophoresis, as in the post-digestion confirmation step previously described. Size selected pooled samples were compared to leftover pooled samples to check for reduced smear band size that roughly fit within the 300 to 500 bp range of a 1 kb ladder.

A total of 6 library prepped and size selected pooled samples (detailed in **Table 3.2**) were submitted for 100-cycle Illumina HiSeq® high-output mode genotyping by sequencing (GBS) at the Iowa State University Office of Biotechnology's DNA Facility. Each sample was submitted to be run in its own Illumina sequencing lane in order to maintain the ability to identify reads belonging to specific individuals within and between pooled samples, as described previously. Samples were quantified by the sequencing facility using Qubit™ quantification prior to sample dilution for initial sequencing. Additionally, a separate sequencing step was requested in order to read the index sequence of the 3' region PCR primers used during library construction.

Stacks Core Parameter Optimization and SNP Discovery

Sequencing data was processed through the STACKS pipeline of software packages (Catchen et al, 2011; Catchen et al, 2013), using STACKS version 1.46. Raw sequence data was first demultiplexed using the program PROCESS_RADTAGS, which also checks the quality score of sequence reads and discards them if the average score drops lower than 90% probability of

being correct (raw phred score of 10). Before proceeding through a manual run of the component programs in the pipeline for the bulk of the sequencing data generated in this study, a small subset of replicate samples (see details for pooled sample ID # 22331 in **Table 3.2**) were used to optimize a core set of STACKS *de novo* assembly parameters based on minimizing quantification of SNP error rate, as outlined in Mastretta-Yanes et al (2015).

Out of the 21 populations of *Webbhelix multilineata* sampled, 11 populations (BHP, MRLP, NRVP, PPFD, MPO, HNL, RRL, RRO, PPNP, CRWA, and DHFA) were chosen at random and a single individual was drawn from each. Our GBS protocol (from PCR amplification to Illumina sequencing) was performed twice on these samples, creating a dataset of 11 replicate sample pairs that were used for STACKS parameter optimization. Mastretta-Yanes et al (2015) identified four core parameters in STACKS that they argue should be optimized specifically for a given genomic data set: '-m' which specifies the minimum number of identical raw reads required to create a read stack in the program USTACKS (default value 3); '-M' which specifies the number of mismatches allowed between loci when processing a single individual in USTACKS (default value 2); '--max_locus_stacks' specifying the maximum number of stacks at a single *de novo* locus, also in USTACKS (default value 3); and '-n' which specifies the number of mismatches allowed between loci when building a sample reference catalog in the program CSTACKS (default value 0; Catchen et al, 2011; Catchen et al, 2013). Note that in later versions of STACKS the default value for '-n' was changed from 0 to 1 by the developers, though 0 was used in the present study using STACKS v1.46 following Mastretta-Yanes et al (2015).

After demultiplexing with PROCESS_RADTAGS, variation in the aforementioned core parameters to find an optimal setting was tested using the STACKS program DENOVO_MAP. A pipeline-executing program, DENOVO_MAP allows for each component program of STACKS

(USTACKS, CSTACKS, SSTACKS and POPULATIONS) to be run sequentially from a single initiating command script, while also permitting specification of parameters across programs. DENOVO_MAP was used iteratively to generate differing sets of SNP's for the sample replicate dataset, varying (with each run) one parameter at a time through a prescribed range (2 – 10 for '-m'; 2 – 8 for '-M'; 2 – 6 for '--max_locus_stacks'; 0 – 7 for '-n') while keeping the others set to their default values (Mastretta-Yanes et al, 2015).

In order to quantify SNP error rate (defined as the proportion of SNP mismatches between sample replicate pairs), a custom R script (Mastretta-Yanes et al, 2015) that reads SNP genotypes coded as a single allele dosage number was used. This necessitated outputting SNP genotypes (via DENOVO_MAP) for sample replicate pairs in PLINK format. PLINK (http://pngu.mgh.harvard.edu/purcell/plink/; Purcell et al, 2007), a program that can perform a range of basic, large-scale analyses and formatting procedures on genomic data, was used to recode DENOVO_MAP SNP genotype outputs in terms of additive components. The resulting *plink.raw* files were then used to estimate SNP error rate using the custom R script.

Replicate pair error rates for a given varied STACKS core parameter were averaged, and the parameter value that generated the lowest average SNP error rate was designated the optimal. Once an optimal value was determined for each core parameter, the resulting parameter combination was used for all proceeding steps. Three additional STACKS parameters and an additional data correction procedure were also tested after core parameter optimization: '--model_type' specifying use of either a 'snp' or 'bounded' model in USTACKS; '--min_maf' specifying the implementation of a minimum minor allele frequency in POPULATIONS; use of RXSTACKS, another program in the STACKS package that makes corrections to genotype and haplotype calls in individual samples based on data accumulated from a population-wide

examination; and '--write_single_snp' for specifying restriction of data analysis in POPULATIONS to only the first SNP per locus. For '--model_type,' DENOVO_MAP runs were performed using the 'snp' model, 'bounded' model with option '--bound_high 0.05,' '--bound_high 0.10,' '--bound_high 0.20' and '--bound_high 0.30,' to test a short range of upper bounds for epsilon (model error rate) in USTACKS. The model that resulted in a sample replicate pair SNP dataset with the lowest SNP error rate was used in subsequent tests. For '--min_maf,' running DENOVO_MAP with a minimum minor allele frequency of 0.05 was compared with running the pipeline with no minimum minor allele frequency set, with the lowest SNP error rate result determining subsequent use of the program option. After '--model_type' and '--min_maf' usage were determined, RXSTACKS was implemented to see if call corrections reduced SNP error rate. Lastly, DENOVO_MAP was run a final time using the '--write_single_snp' option to compare average SNP error rate between a reduced dataset comprised of the first SNP from near the 5' end of each sequence read, versus an expanded dataset of multiple SNPs from along the full length of each read.

Once all parameters were optimized for minimizing SNP error rate, a proper run of the STACKS pipeline was executed (using optimal parameter settings) by manually running each program in sequence to generate a workable *Webbhelix multilineata* SNP dataset with a minimum population constraint of ~75% and a minimum individual constraint of 80%. The resulting dataset was outputted from the POPULATIONS program in several coding formats, including *.plink*. This output was used in an additional post-processing step performed using PLINK to filter out SNPs in high linkage disequilibrium (LD) pairs and SNPs showing extreme deviation from Hardy-Weinberg Equilibrium (HWE); high LD and extreme HWE deviation can be (in the case of the former) and often is (in the latter) associated with genotyping error

(Hosking et al, 2004; Scheet & Stephens, 2008; Chen et al, 2017). Removal of these SNPs also served to reduce the remaining number of markers (8,480 SNPs) to a more computationally manageable size for downstream population genomic analyses. The PLINK program is limited in SNP coding formats it can output, so post-filtering SNP data was output in *.vcf* format and then converted to *.genepop* format using the program PGDSPIDER version 2.1.1.5 (Lischer & Excoffier, 2012). The *.genepop* format is more flexible in usage since it is more commonly accepted as input by population genetic software packages, and was used in the proceeding genomic analyses.

Population Genomic Structure and Estimation of Genetic Diversity & Differentiation

Genomc structure in *Webbhelix multilineata* was examined to detect any distinct population clusters using the program STRUCTURE v2.3.4 (Pritchard et al, 2000). The *.genepop* SNP dataset above was converted into a *.structure* formatted file using GENODIVE, and STRUCTURE was run with a burn-in of 100,000 generations followed by 500,000 MCMC iterations for k = 1 – 18 with 20 replicates per k. The logarithm of the data probability and estimates of Δk were evaluated using the program STRUCTURE HARVESTER web v0.6.94 (Earl & vonHoldt, 2012) to determine what the most likely number of population clusters was. The program DISTRUCT v1.1 (Rosenberg, 2004) was used to produce a STRUCTURE bar plot to display the results.

The program GENODIVE was used to estimate observed (H_O) and expected (H_S) heterozygosity, inbreeding coefficient (G_{IS}), and pairwise population differentiation (F_{ST}) for an 8,480 SNP dataset formatted as a *.genepop* file. The dataset was comprised of a total of 178 individual *Webbhelix multilineata* sampled from 18 localities (TGCP, MRLP, STT, FSSP, HNL, MPO, OICP, IRO, HGRF, PTP, RRL, RRO, CPS, PPNP, FPPI, CRWA, PCRA, DHFA; see **Table 3.1**). All sample localities contained 10 individuals, except MRLP, which contained 8

individuals. Diversity indices (H_O, H_S, G_{IS}) were calculated using Analysis > Genetic Diversity… and selecting the option to Calculate indices separately for every population. The program calculates indices according to Nei (1987). Pairwise population differentiation (F_{ST}) was calculated using Analysis > Calculate Distances > F_st in the program; GENODIVE calculates F_{ST} via Analysis of Molecular Variance (AMOVA; Excoffier et al, 1992; Michalakis & Excoffier, 1996).

Testing for Isolation by Distance

The GENODIVE program was also used to test the above dataset for isolation by distance (IBD) via Mantel test (Mantel, 1967). The pairwise F_{ST} distance matrix outputted in the analysis from the previous section was transformed into a linearized F_{ST} matrix using Data > Transformation…, then selecting Standard mode and method For Isolation by Distance: D/(1-D); testing for IBD using pairwise F_{ST} distances against pairwise geographical distances via Mantel test calls for use of linearized F_{ST} ($F_{ST}/(1-F_{ST})$). A pairwise geographical distance matrix was manually created by measuring straight-line distances between all population pairs using the program GOOGLE EARTH PRO. Straight-line geographical distances between all sample locality pairs can be seen in **Table 3.3**. Linearized F_{ST} values were plotted against geographic distance values to check if the data showed a linear relationship and determine what R statistic to calculate when performing the Mantel test; for data that are roughly linear, Mantel's R statistic is appropriate, while Spearman's R statistic is more appropriate for data that are strongly non-linear. Mantel test was performed running 1,000,000 permutations.

Estimation of Relative Migration Rates

Relative rates of migration between population clusters identified in the STRUCTURE analysis were estimated using the program DIVMIGRATE-ONLINE (Keenan, 2012; Sundqvist et al, 2016). Analysis was done using G_{ST} to calculate relative migration, with 1,000 bootstrap replicates for calculating statistical significance of directional migration using alpha 0.05. A

Table 3.3 – Straight line geographical distances between all sample locality pairs. Distances are in miles, and were measured using the ruler feature in the program Google Earth Pro. Pairwise geographical distances were initially plotted against linearized pairwise genetic distances (linearized F_{ST}) to observe whether or not the plotted data points were roughly linear; this would determine the appropriate R statistic to use for performing a Mantel test to see how geographic distance correlates with estimates of linearized genetic distance. If the scatter plot of geographic distance vs. linearized genetic distance turns out to be at least roughly linear, then Mantel's R statistic would be appropriate to use in performing a Mantel test; if the scatter plot is strongly non-linear, then Spearman's R would be the appropriate statistic to use. Once the appropriate statistic was determined, the geographic distance matrix below was used as an input file for performing a Mantel test in the program GENODIVE, along with a matrix of pairwise linearized genetic distances.

Sample Localities	TGCP	MRLP	STT	FSSP	HNL	MPO	OICP	IRO	HGRF	PTP	RRL	RRO	CPS	PPNP	FPPI	CRWA	PCRA	DHFA
TGCP	0.0	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
MRLP	65.7	0.0	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
STT	180.0	235.1	0.0	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
FSSP	192.6	211.8	331.3	0.0	--	--	--	--	--	--	--	--	--	--	--	--	--	--
HNL	533.8	514.9	494.2	723.2	0.0	--	--	--	--	--	--	--	--	--	--	--	--	--
MPO	164.7	227.4	51.1	290.8	541.2	0.0	--	--	--	--	--	--	--	--	--	--	--	--
OICP	42.3	98.7	187.1	160.6	571.7	159.4	0.0	--	--	--	--	--	--	--	--	--	--	--
IRO	122.1	70.7	249.3	282.1	451.3	254.2	160.7	0.0	--	--	--	--	--	--	--	--	--	--
HGRF	52.4	109.5	191.1	150.8	584.4	161.6	11.4	171.5	0.0	--	--	--	--	--	--	--	--	--
PTP	57.8	114.1	192.8	147.1	589.7	160.5	16.6	177.0	5.4	0.0	--	--	--	--	--	--	--	--
RRL	224.9	193.5	275.5	403.7	325.7	301.5	266.7	123.5	276.6	282.3	0.0	--	--	--	--	--	--	--
RRO	188.8	123.3	350.8	276.5	517.7	345.8	216.2	109.8	222.5	226.8	201.2	0.0	--	--	--	--	--	--
CPS	203.9	143.6	326.7	342.4	425.3	335.0	240.4	83.1	250.0	256.7	115.9	92.0	0.0	--	--	--	--	--
PPNP	538.9	472.6	695.6	573.4	669.6	696.2	561.9	446.3	568.0	572.7	471.4	350.7	375.4	0.0	--	--	--	--
FPPI	533.1	465.7	686.3	571.1	653.8	687.6	556.4	436.8	562.2	566.6	458.4	343.1	364.9	16.0	0.0	--	--	--
CRWA	156.9	180.2	296.6	37.3	689.8	258.0	123.1	251.6	115.1	110.8	370.3	255.9	316.0	566.8	564.2	0.0	--	--
PCRA	134.3	69.9	303.0	232.3	526.2	294.8	159.9	85.0	166.2	171.8	201.2	56.4	111.9	404.6	398.7	207.4	0.0	--
DHFA	430.8	397.2	443.6	609.5	158.2	481.6	471.3	326.3	482.9	488.8	207.2	375.0	284.2	515.7	499.6	577.8	393.3	0.0

subset of the 8,480 SNP dataset described above was used to estimate relative migration rates, due to computational limits of DIVMIGRATE-ONLINE. The web browser application cannot handle computing estimates for large SNP datasets or datasets with many pairwise estimates before being disconnected from the server. For computing pairwise rates between four population clusters, the application could handle no more than ~3000 SNP loci; therefore, loci with more than 30% missing data within any of the four population clusters from the STRUCTURE analysis were filtered out from the entire dataset, and the remaining 2,760 SNP's were used for estimating migration rates.

The web application was also used to calculate relative migration rates between 8 *Webbhelix* sample localities: PTP, HGRF, OICP, TGCP, IRO, RRL, DHFA and HNL. Each of these sampling sites is located downstream of prior sites in the order listed here (such that HGRF is downstream of PTP; OICP is downstream of HGRF and PTP; etc.), being sites along a continuous route along the Iowa River starting in central north Iowa, following that river till it merges with the Mississippi River, and continuing down the Mississippi to eastern Arkansas. The first three sites along this route (PTP, HGRF and OICP) are within 15 miles river-distance from each other; the next site, TGCP, is nearly 50 miles river-distance away from the first three sites; the remaining four sites (IRO, RRL, DHFA and HNL) are more than 100 miles downstream from the first four sites and from each other. Relative migration between these 8 sites was calculated to assess at what scale downstream migration is relevant in this riparian species (here, downstream migration refers to migrants being washed downstream during river flooding events). As with the previous DIVMIGRATE-ONLINE analysis, a subset of the 8,480 SNP dataset was needed in order to successfully run the application. For computing pairwise rates between 8 localities, the application could handle no more than ~2000 SNP loci; therefore, loci

with more than 15% missing data across all 8 sample localities were filtered out of the entire dataset, and the remaining 1,762 SNP's were used for migration rate estimates.

Coalescent Simulation of Refugial Expansion Scenarios

Population clusters determined by STRUCTURE analysis were used to model scenarios of refugial expansion for *Webbhelix multilineata* after and during glacial recession following the LGM roughly 20,000 years ago. Expansion models and accompanying tree scenarios are shown in **Figure 3.2**. Coalescent simulations based on prior distributions of effective population size and number of *Webbhelix* generations, as well as model checking analyses and comparisons of scenario posterior probabilities, where done using the program DIYABC v2.1.0 (Cornuet et al, 2014). The *.genepop* SNP dataset described in the genetic diversity section above was converted into a *.DIYABC.snp* file format using a python script accompanying the program; subsequently, SNP loci for which more than 30% of individuals within a STRUCTURE population cluster had missing data were removed from the entire dataset, for two reasons: 1) DIYABC does not accept datasets with any loci for which all individuals in a designated population have missing data; 2) DIYABC is not able to computationally handle running datasets larger than ~4,000 SNPs. Filtering loci with more than 30% missing data within designated populations reduced the 8,480 *Webbhelix* SNP dataset down to 2,760 SNPs, which the program could handle running.

Two runs of DIYABC were performed, restricting scenario simulations to within two time frames: one with prior distribution for T1 and T2 set from 10 to 5000 generations (restricting scenario simulations to later than 12,500 years ago, given a 2.5 year generation time for *Webbhelix*); and a second run with prior distribution for T1 and T2 set from 10 to 8000 generations (restricting scenario simulations to later than 20,000 years ago). The condition that T1<T2 was set for both runs of the program, and each was run for 3,000,000 simulations. After simulations were completed, model checking analyses were performed in DIYABC to visualize

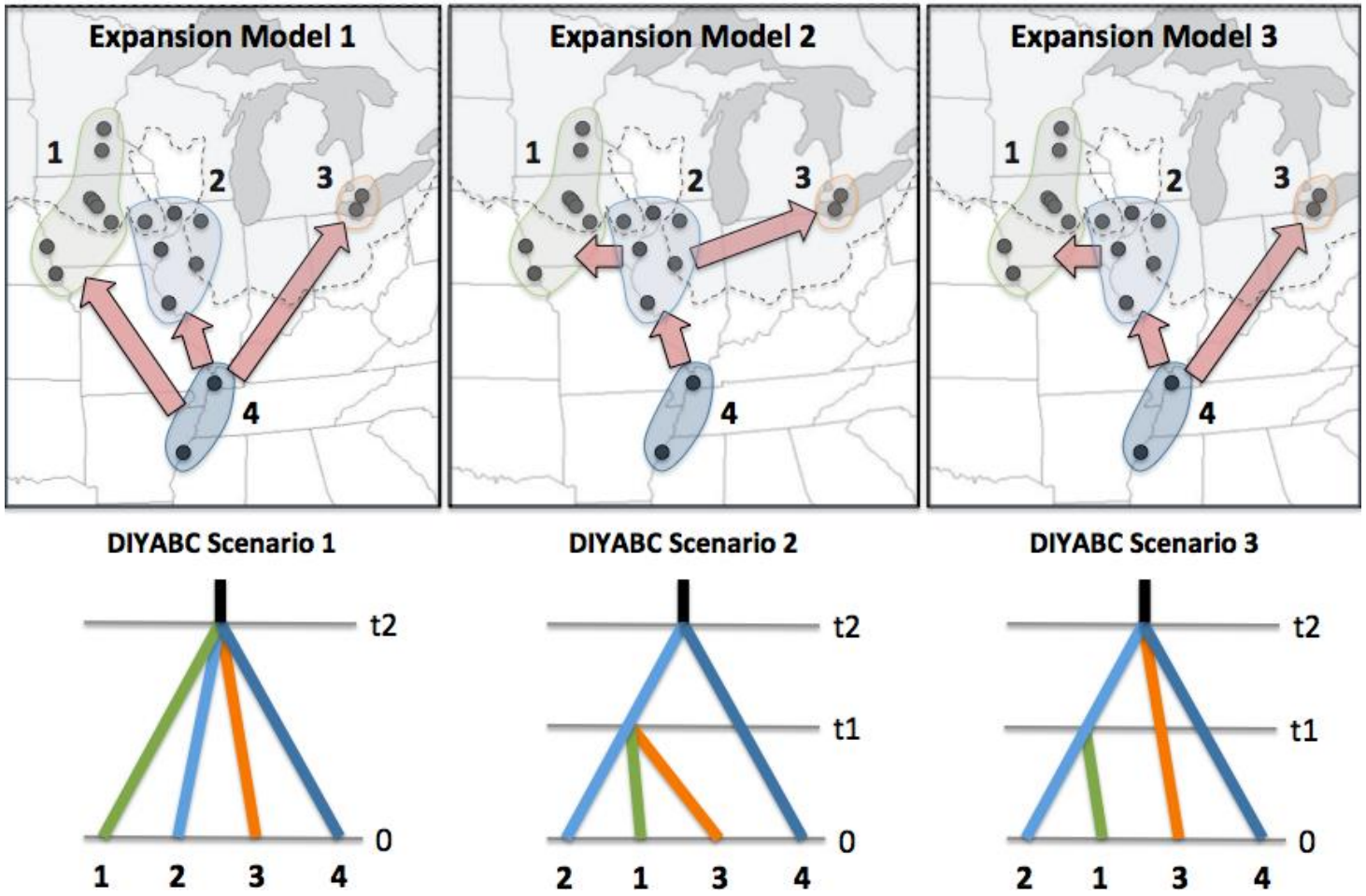


Figure 3.2 – Refugial expansion models and tree scenarios used for coalescent simulations in the program DIYABC. Scenario 1 is meant to model refugial expansion where *Webbhelix* left southern refugia and expanded their range north in a multidirectional pattern unconstrained by river corridors; the arrow from Cluster 4 to Cluster 1 in particular is what distinguishes that aspect of the model, since expanding from 4 to 1 directly would require broad availability of *Webbhelix* habitat (marsh, wetland, flood plain, etc) beyond availability along river corridors. Arrows from 4 to 2 and 4 to 3 suggest migration following habitat along river corridors (the Mississippi and Ohio in particular). Expansion from 2 to 1 could be achieved following the Missouri River and the Iowa and Cedar Rivers. Expansion from 2 to 3 could be achieved following the Illinois River and various interconnected smaller rivers that stretch across northern Illinois and Indiana, and southern Michigan.

how well data sets simulated from posterior distributions for each scenario fit the observed data. Additionally, the three scenarios for each run were compared using estimates of the posterior probability of each scenario via a direct approach and a logistic regression approach.

Results

Population Genomic Structure and Diversity Estimates

Results of the STRUCTURE analysis on all 18 sample localities of *Webbhelix multilineata* using the 8,480 SNP dataset can be seen in **Figure 3.3**. The most likely number of population clusters was determined to be 4, based on observation of the highest Δk and a smaller standard deviation of lnP(k) (model choice criterion) than the next highest Δk from the STRUCTURE HARVESTER results. The bar plot shown in **Figure 3.3** indicates that four Iowa sample localities (TGCP, OICP, HGRF and PTP), the two Minnesota sample localities (FSSP and CRWA) and the two Nebraska sample localities (STT and MPO) form one population cluster (Cluster 1); the three remaining Iowa localities (MRLP, IRO and PCRA) and the three Illinois localities (RRL, RRO and CPS) form a second population cluster (Cluster 2); the two southern Ontario localities (PPNP and FPPI) formed a third population cluster (Cluster 3); and the two southern-most sample localities (DHFA in Missouri and HNL in Arkansas) formed a fourth population cluster (Cluster 4).

Estimates of H_O, H_S and G_{IS} for all *Webbhelix multilineata* sample localities based on 8,480 SNPs can be seen in **Figure 3.4**. Observed heterozygosity values ranged from 0.006 (PTP) to 0.050 (DHFA). Expected heterozygosity ranged from 0.006 (PTP) to 0.053 (DHFA). Inbreeding coefficient ranged from -0.018 (HGRF) to 0.176 (MRLP). A full summary of all genetic diversity estimates can be viewed in **Table 3.4**. Estimates of pairwise F_{ST} ranged from 0.049 (between MRLP-IRO) to 0.23 (between OICP-PPNP); all F_{ST} estimates are shown as a pairwise distance matrix in **Table 3.5**.

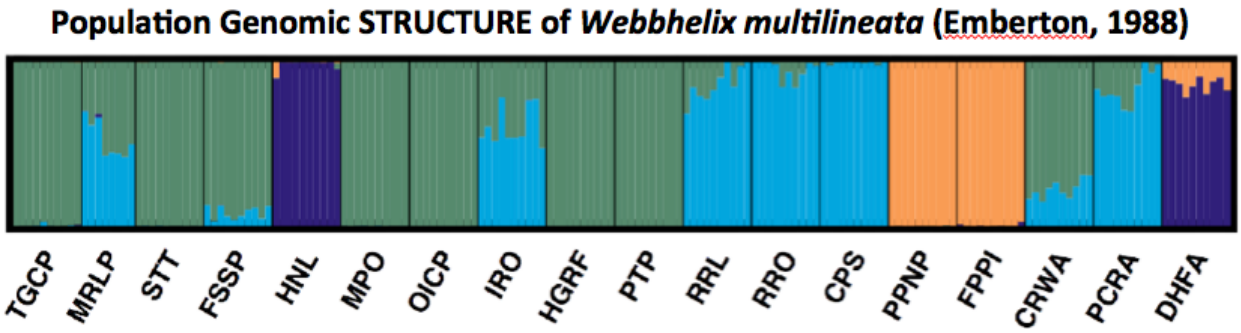
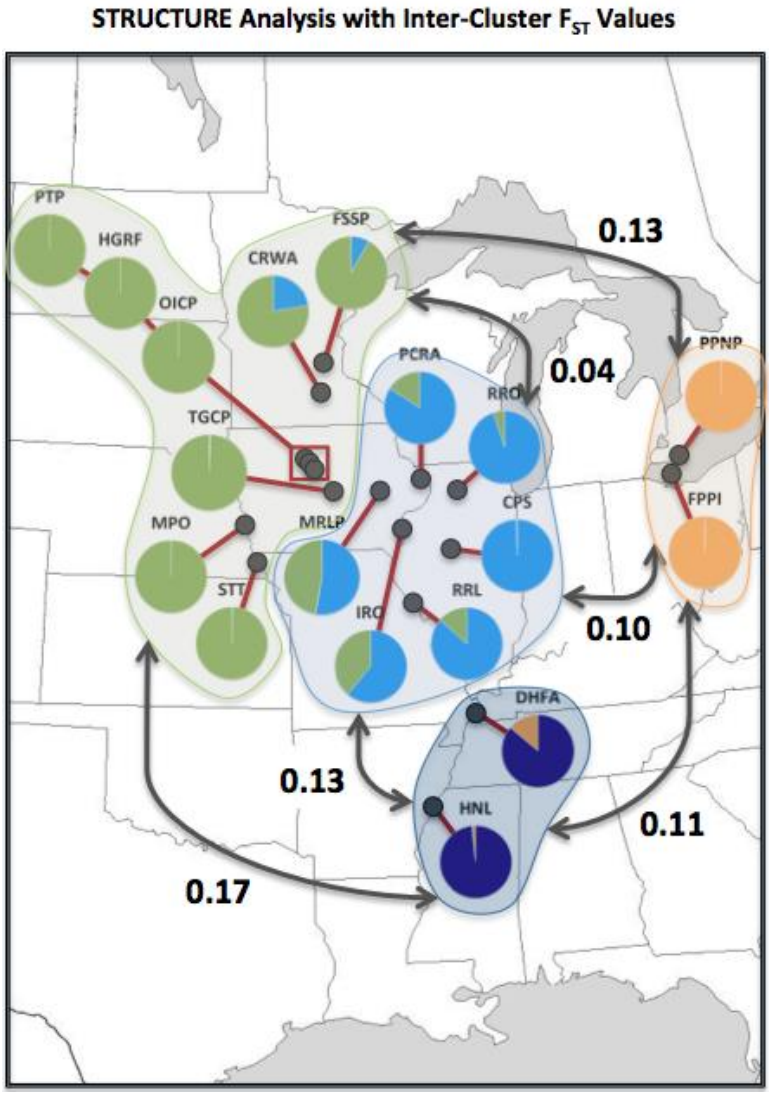


Figure 3.3 – Population genomic structure in *Webbhelix multilineata* from 18 sample localities based on 8,480 SNPs analyzed in the program STRUCTURE.



Analysis outputs were processed using STRUCTURE HARVESTER, and number of population clusters (4) was assessed based on the highest Δk and smaller standard deviation of lnP(K) (model choice criterion) than the next highest Δk. The bar plot shown above was produced using the program DISTRUCT. The map on the left shows the geographic location of all 18 sample localities in the American Midwest and South (TGCP, MRLP, OICP, IRO, HGRF, PTP and PCRA in the state of Iowa; STT and MPO in Nebraska; FSSP and CRWA in Minnesota; HNL in Arkansas; RRL, RRO and CPS in Illinois; DHFA in Missouri), and southern Canada (PPNP and FPPI in the province of Ontario). The smaller black circles are the sample localities; colored pie charts for each locality on the map reflect the genetic variation shown in the STRUCTURE bar plot above. Colored shapes encircling pie charts reflect the four population clusters inferred from the STRUCTURE analysis. Double arrows indicate cluster pairs for which pairwise F_{ST} was calculated (F_{ST} estimates are shown next to arrows).

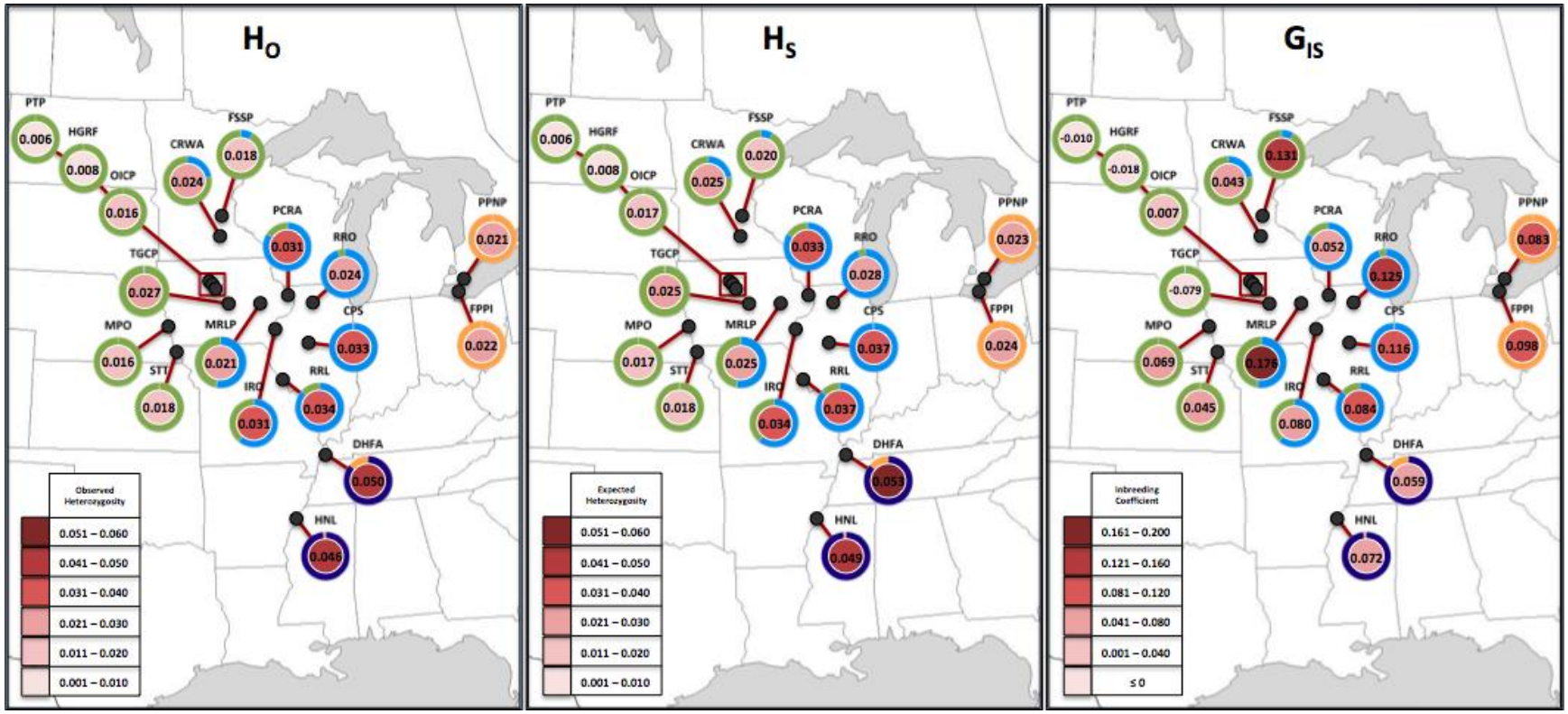


Figure 3.4 – Estimates of observed heterozygosity (H_O), expected heterozygosity (H_S) and inbreeding coefficient (G_{IS}) for 18 population sample localities of *Webbhelix multilineata* based on a dataset of 8,480 SNPs. Red color gradients indicate increasing ranges of estimates for each parameter (increasing from light red to dark red). Filled-in red circles on maps contain parameter estimates for each population sample locality, and colored rings around red circles indicate population genomic structure from a STRUCTURE analysis of the same dataset. In general, observed and expected heterozygosity appear to be highest in southern populations and decrease to the north and again to the northwest. For inbreeding coefficient there did not appear to be any general geographic pattern, as sample localities had varying amounts of inbreeding across the species range.

Table 3.4 – Summary of genetic diversity estimates for 18 sample localities of *Webbhelix multilineata* based on 8,480 SNP loci. Genetic diversity estimates were calculated using the program GENODIVE.

Genetic Diversity: Nei 1987.
178 individuals included, 8480 SNP loci included

– Summary of indices of genetic diversity

Statistic	Value	Std.Dev.	c.i.2.5%	c.i.97.5%	Description
Num	2	0	2	2	Number of alleles
Eff_num	1.027	0	1.027	1.028	Effective number of alleles
Ho	0.025	0	0.025	0.026	Observed Heterozygosity
Hs	0.027	0	0.027	0.028	Heterozygosity Within Populations
Ht	0.032	0	0.031	0.033	Total Heterozygosity
H't	0.032	0	0.032	0.033	Corrected total Heterozygosity
Gis	0.076	0.005	0.067	0.085	Inbreeding coefficient

Standard deviations of F-statistics were obtained through jackknifing over loci.
95% confidence intervals of F-statistics were obtained through bootstrapping over loci.

– Indices of genetic diversity per population

Population	Num	Eff_num	Ho	Hs	Gis
TGCP	1.102	1.035	0.027	0.025	-0.079
MRLP	1.113	1.032	0.021	0.025	0.176
STT	1.078	1.025	0.018	0.018	0.045
FSSP	1.087	1.027	0.018	0.02	0.131
HNL	1.181	1.07	0.046	0.049	0.072
MPO	1.072	1.023	0.016	0.017	0.069
OICP	1.058	1.024	0.016	0.017	0.007
IRO	1.171	1.043	0.031	0.034	0.08
HGRF	1.028	1.011	0.008	0.008	-0.018
PTP	1.024	1.009	0.006	0.006	-0.01
RRL	1.172	1.048	0.034	0.037	0.084
RRO	1.109	1.039	0.024	0.028	0.125
CPS	1.156	1.051	0.033	0.037	0.116
PPNP	1.087	1.033	0.021	0.023	0.083
FPPI	1.095	1.034	0.022	0.024	0.098
CRWA	1.104	1.034	0.024	0.025	0.043
PCRA	1.147	1.045	0.031	0.033	0.052
DHFA	1.206	1.074	0.05	0.053	0.059

Table 3.5 – Estimates of pairwise F_{ST} between 18 sample localities of *Webbhelix multilineata* based on an 8,480 SNP dataset. The dataset was comprised of a total of 178 individuals, 10 individuals per sample locality with the exception of MRLP, which included only 8 individuals. Values of F_{ST} ranged from as low as 0.049 between MRLP and IRO, to as high as 0.23 between OICP and PPNP. Pairwise F_{ST} values are shown below and to the left of the main diagonal of the matrix; p-values for F_{ST} estimates based on 1000 permutations are shown above and to the right of the main diagonal. Analysis of molecular variance was carried out using the 'Pairwise Differentiation…' option in the program GENODIVE. Some of the estimates of genetic differentiation in this table are relatively high for within-species sample locality comparisons. For example, Funk et al (2007) assessed genetic differentiation in three subspecies of snowy plovers and concluded that there was strong evidence for maintaining that species complex based on F_{ST} values higher than 0.2 for some of their pairwise comparisons. Here, at least a dozen pairwise comparisons of *Webbhelix* sample localities have estimated values of F_{ST} above 0.2, and a couple dozen more comparisons with values nearly as high as 0.2. This suggests that *Webbhelix multilineata* may actually itself be a species complex, though this level of genetic differentiation alone is not enough to confidently say. A species delineation analysis of *Webbhelix* sample localities is needed.

\ p-value FST \	TGCP	MRLP	STT	FSSP	HNL	MPO	OICP	IRO	HGRF	PTP	RRL	RRO	CPS	PPNP	FPPI	CRWA	PCRA	DHFA
TGCP	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MRLP	0.136	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
STT	0.177	0.12	--	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
FSSP	0.146	0.105	0.145	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
HNL	0.205	0.165	0.187	0.189	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MPO	0.192	0.123	0.056	0.139	0.193	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
OICP	0.103	0.142	0.186	0.163	0.207	0.183	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
IRO	0.109	0.049	0.103	0.097	0.163	0.112	0.12	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
HGRF	0.147	0.14	0.187	0.165	0.205	0.185	0.155	0.12	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
PTP	0.149	0.135	0.177	0.142	0.187	0.178	0.159	0.117	0.116	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
RRL	0.133	0.096	0.135	0.119	0.175	0.137	0.154	0.081	0.155	0.154	--	0.001	0.001	0.001	0.001	0.001	0.001	0.001
RRO	0.168	0.12	0.162	0.152	0.192	0.167	0.194	0.115	0.201	0.191	0.106	--	0.001	0.001	0.001	0.001	0.001	0.001
CPS	0.149	0.116	0.147	0.147	0.181	0.157	0.168	0.102	0.173	0.171	0.109	0.121	--	0.001	0.001	0.001	0.001	0.001
PPNP	0.201	0.161	0.188	0.182	0.176	0.197	0.23	0.155	0.228	0.209	0.171	0.199	0.187	--	0.001	0.001	0.001	0.001
FPPI	0.199	0.157	0.186	0.178	0.166	0.197	0.222	0.152	0.22	0.205	0.17	0.199	0.182	0.138	--	0.001	0.001	0.001
CRWA	0.167	0.104	0.149	0.12	0.192	0.155	0.174	0.097	0.187	0.183	0.124	0.16	0.146	0.198	0.195	--	0.001	0.001
PCRA	0.155	0.1	0.138	0.134	0.185	0.146	0.17	0.087	0.166	0.15	0.098	0.125	0.129	0.182	0.183	0.133	--	0.001
DHFA	0.19	0.155	0.18	0.173	0.112	0.18	0.201	0.16	0.194	0.18	0.171	0.184	0.176	0.158	0.151	0.187	0.181	--

Isolation by Distance

A scatter plot of linearized F_{ST} values versus straight-line geographic distances between sample localities can be seen in **Figure 3.5**. The data were roughly linear, indicating that use of Mantel's R statistic was appropriate to calculate in a Mantel test for correlation between genetic distance (linearized F_{ST}) and geographic distance (straight-line pairwise distances) in sampled individuals of *Webbhelix multilineata*. The result of the Mantel test showed that there is a highly statistically significant (one-tailed p = 0.000) strong positive correlation between genetic distance and geographic distance (Mantel's R = 0.641; R^2 = 0.411; error SS = 3073144.656; total SS = 5216834.078). This indicates that populations of *Webbhelix* are under strong influence of IBD.

Relative Migration Rates

A relative migration network, generated with DIVMIGRATE-ONLINE, showing migration rates between four population clusters from the STRUCTURE analysis is shown in **Figure 3.6**. In addition to numerical rates, the network features arrows of varying thickness, color and length indicating the relative strength of gene flow detected from one population cluster to another (with thicker, darker and shorter arrows conveying a stronger signal). A useful consequence of this feature is that populations with high rates of gene flow will cluster closer together in the network space and further away from populations with which they share low or no gene flow, providing an additional visualization of genetic structuring pattern to compare with the STRUCTURE analysis result. The migration network shows that population clusters 1 and 2 landed closer together in the network space, reflecting the shared genetic variation shown by STRUCTURE and the low estimate of F_{ST} between the pair. In general, the spacing between clusters in the network closely reflected estimates of pairwise F_{ST} among them. The DIVMIGRATE-ONLINE analysis (using 2,760 SNP loci) detected statistically significant relative rates of unidirectional migration from Cluster 1 to Clusters 2, 3 and 4; from Cluster 2 to Clusters 3 and 4;

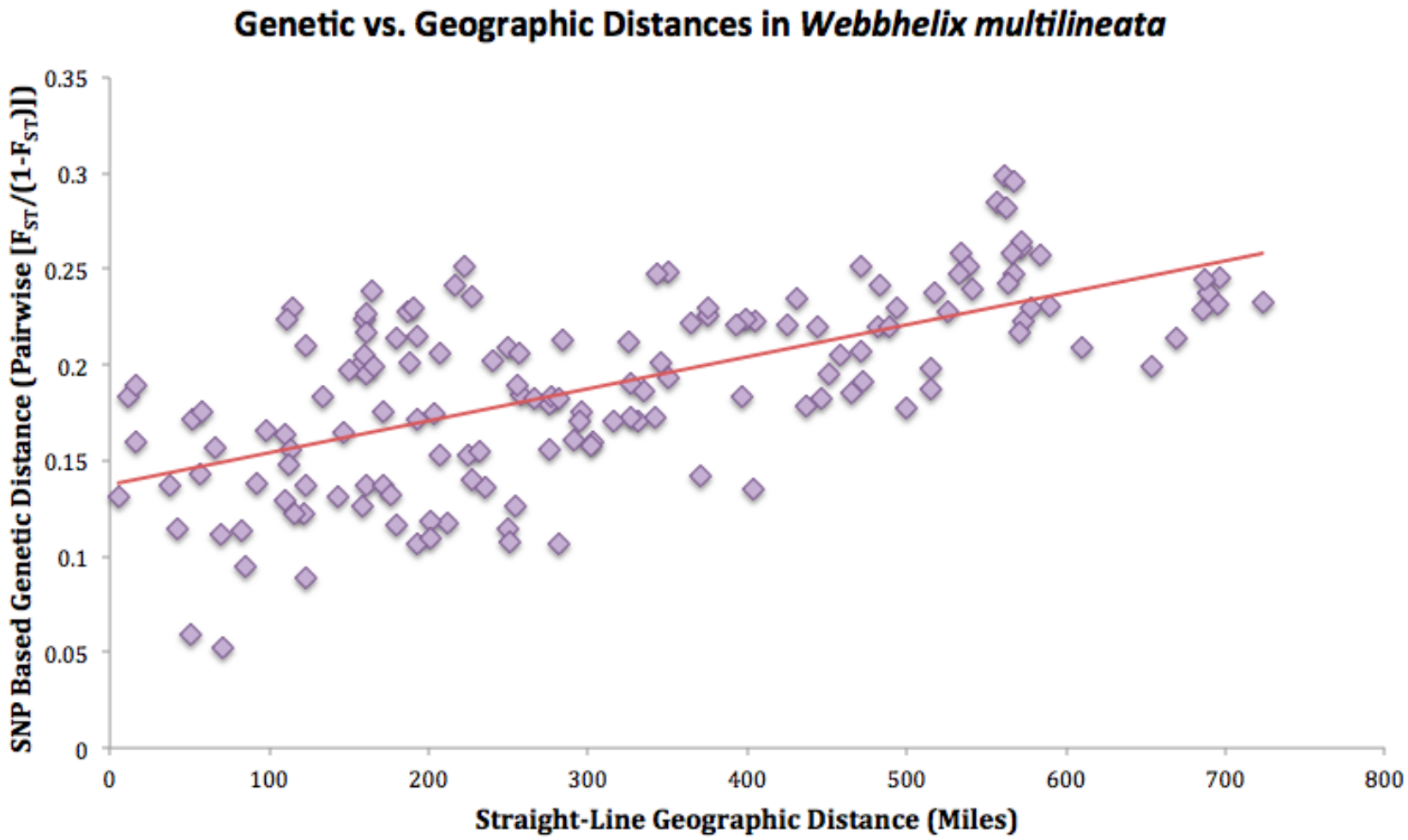


Figure 3.5 – Scatter plot of linearized F_{ST} values (from pairwise comparison of 18 *Webbhelix* sample localities) vs. straight-line geographic distances (pairwise geographic distances between sample localities). Data are roughly linear, indicating use of Mantel's R statistic is appropriate for performing a Mantel test for correlation between a linear F_{ST} distance matrix and a geographic distance matrix constructed from *Webbhelix* SNP data (8,480) and sample localities.

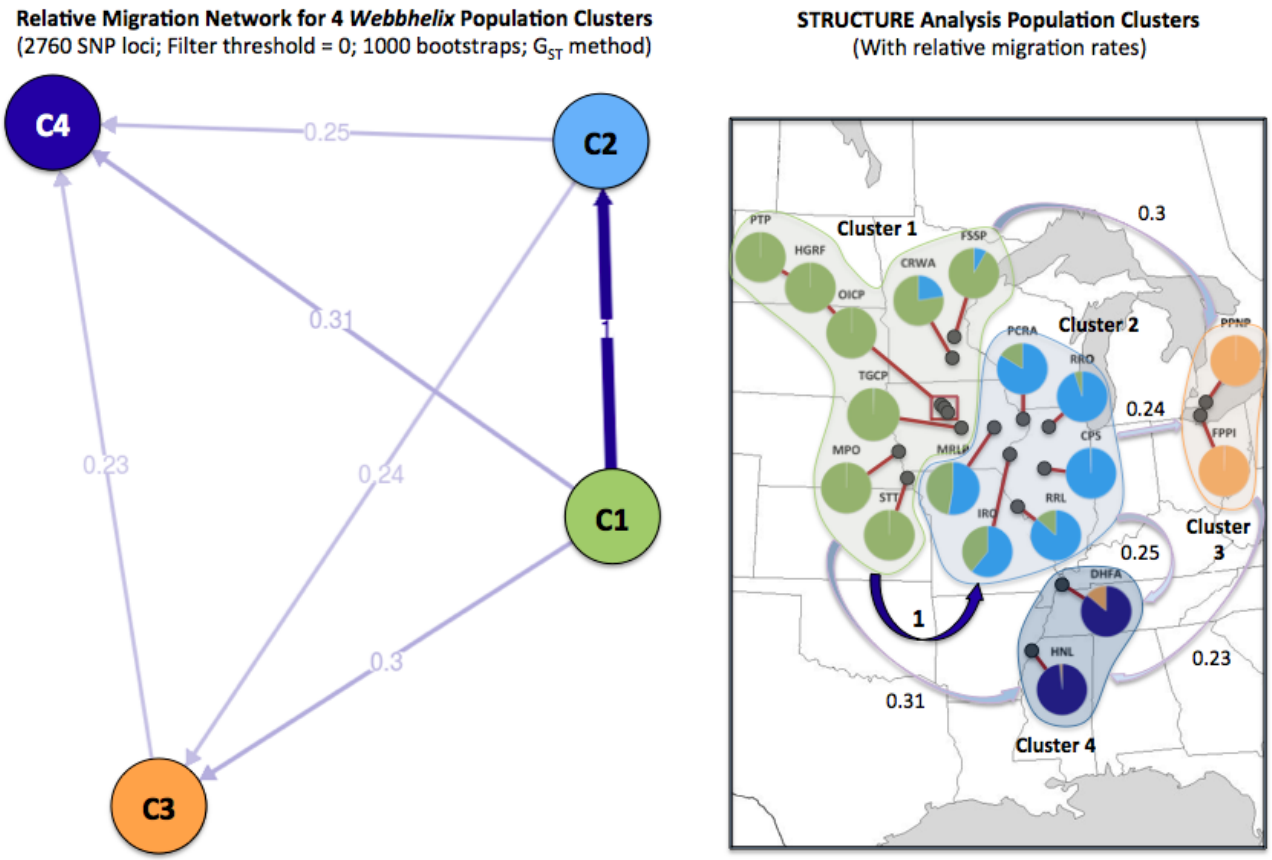


Figure 3.6 – Network from DIVMIGRATE-ONLINE showing relative rates of migration (left) between population clusters of *Webbhelix multilineata* inferred from the STRUCTURE analysis (right) described in Figure 7. The relative migration network was generated using a subset (2,760 SNP's) of the full 8,480 SNP dataset used in the STRUCTURE analysis; the subset were the remaining SNP's after filtering out (from the full dataset) loci with more than 30% missing data within any of the four population clusters from the STRUCTURE analysis. The network features numerical migration rates and arrows of different thickness, color and length conveying relative strength of gene flow (thicker, darker and shorter indicate stronger signal). A consequence of this is that population clusters with stronger gene flow between them cluster closer together in the network space, providing a visualization of genetic structuring to compare with the STRUCTURE analysis results. The map on the right shows geographic locations of each sample locality (black circles), plus arrows and numerical values matching relative migration estimates from the network. Migration rates shown were statistically significant based on 1000 bootstraps (alpha = 0.05), with the strongest signal indicating unidirectional gene flow from Cluster 1 to 2. Weak unidirectional gene flow was also detected between other cluster pairs and in general occurs downstream, though some signal was detected from Clusters 1 & 2 going to 3.

and from Cluster 3 to Cluster 4, based on 1,000 bootstraps (alpha = 0.05). The highest relative migration rate (RMR) was estimated to be from Cluster 1 to Cluster 2 (RMR = 1). The relative migration rates from Cluster 1 to Clusters 3 and 4 were RMR = 0.3 and RMR = 0.31, respectively. Relative migration rates from Cluster 2 to Clusters 3 and 4 were RMR = 0.24 and RMR = 0.25, respectively. Lastly, relative migration rate from Cluster 3 to Cluster 4 was 0.23.

Relative migration rates estimated between 8 sample localities (PTP, HGRF, OICP, TGCP, IRO, RRL, DHFA and HNL) along a continuous stretch of the Iowa-Mississippi Rivers from northern central Iowa down to the eastern boarder of Arkansas, can be seen in **Figure 3.7**. The analysis (using 1,762 SNP loci) detected statistically significant migration from the HGRF sample locality to TGCP (RMR = 1), ~45 miles downstream along the Iowa River, and no migration rate between any other locality pairs regardless of river distance (including PTP to HGRF, localities that are only ~5 miles river distance apart). Localities in the network space were distributed in a (roughly) evenly spread out ring, a pattern that is outputted when there is no gene flow detected between populations. For the single detection of gene flow from HGRF to TGCP, it should be noted that the arrow in the figure is thin and light grey not because the signal is weak but rather because it is the only signal detected from the HGRF locality; arrow thickness, color and length only indicate signal strength *relative* to other gene flow signals from the same population/locality, and default to thin and grey when there is only one gene flow signal.

Coalescent Simulation of Refugial Expansion Scenarios

Data sets simulated with prior distributions of parameters in DIYABC were used to assess the goodness-of-fit of model parameter posterior combinations in three refugial expansion scenarios proposed for *Webbhelix multilineata*. Principal component analysis (PCA) plots from these model-checking analyses, for both time frames for which simulations were run, can be seen in **Figure 3.8**. Approximately 1,000,000 simulated data sets were used to compute posterior

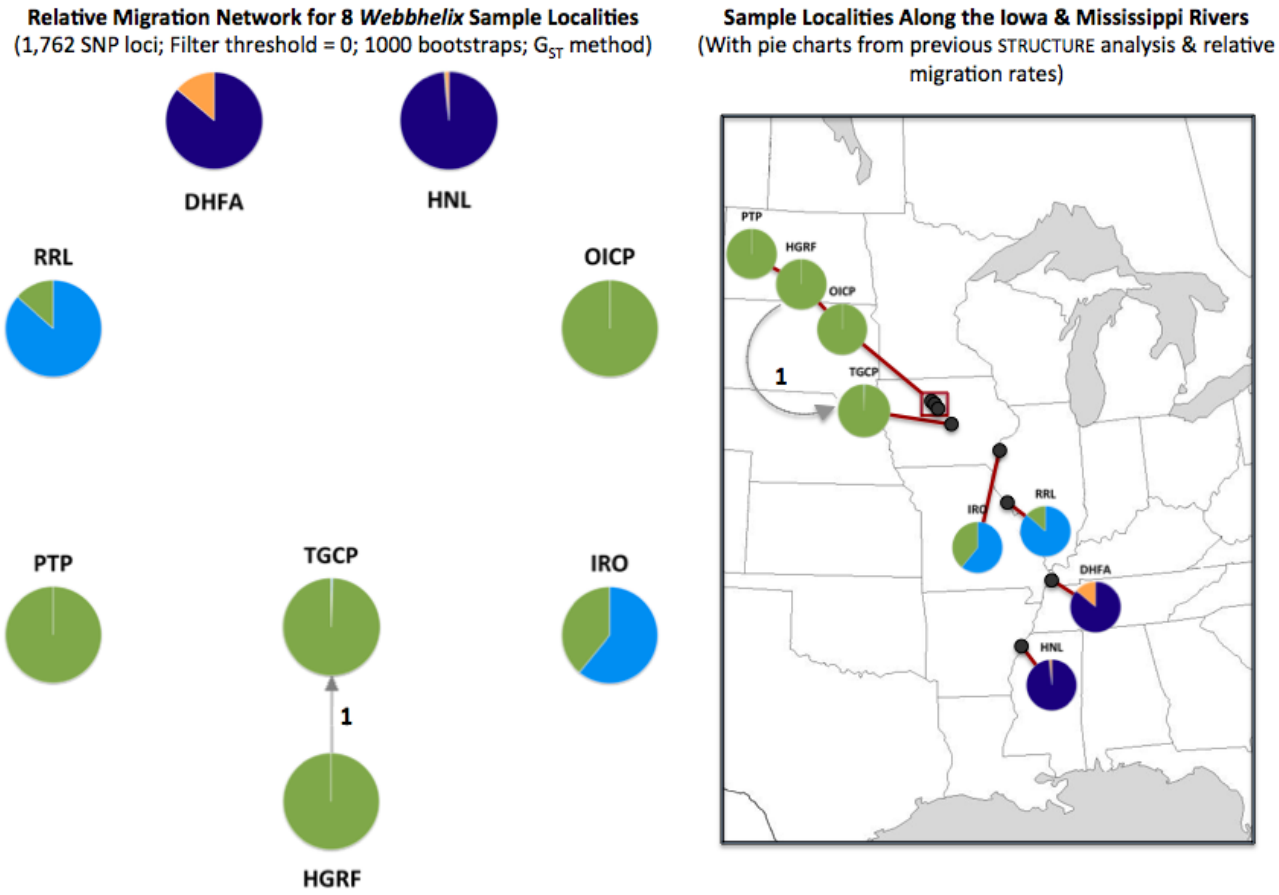


Figure 3.7 – Relative migration network generated using DIVMIGRATE-ONLINE, showing migration rates between 8 sample localities of *Webbhelix multilineata* (10 individuals per locality) along a continuous downstream route from the northern Iowa River, down to its drainage into the Mississippi River in southeast Iowa, and along the Mississippi to eastern Arkansas. These sample localities were used to see at what scales migration occurs between localities, to explore the relevance of river flooding to downstream gene flow in this riparian species. The analysis was performed using 1,762 SNP loci remaining after filtering (from the entire 8,480 SNP dataset) loci with more than 15% missing data across all 80 individuals. The map on the right shows geographic locations of each sample locality (black circles), and both the map and the migration network feature colored pie charts from the previously described STRUCTURE analysis of all 18 *Webbhelix* sample localities. The analysis resulted in only one statistically significant signal of unidirectional gene flow from HGRF to TGCP (based on 1000 bootstraps with an alpha = 0.05). Significant gene flow was not detected between any other locality pairs in the analysis.

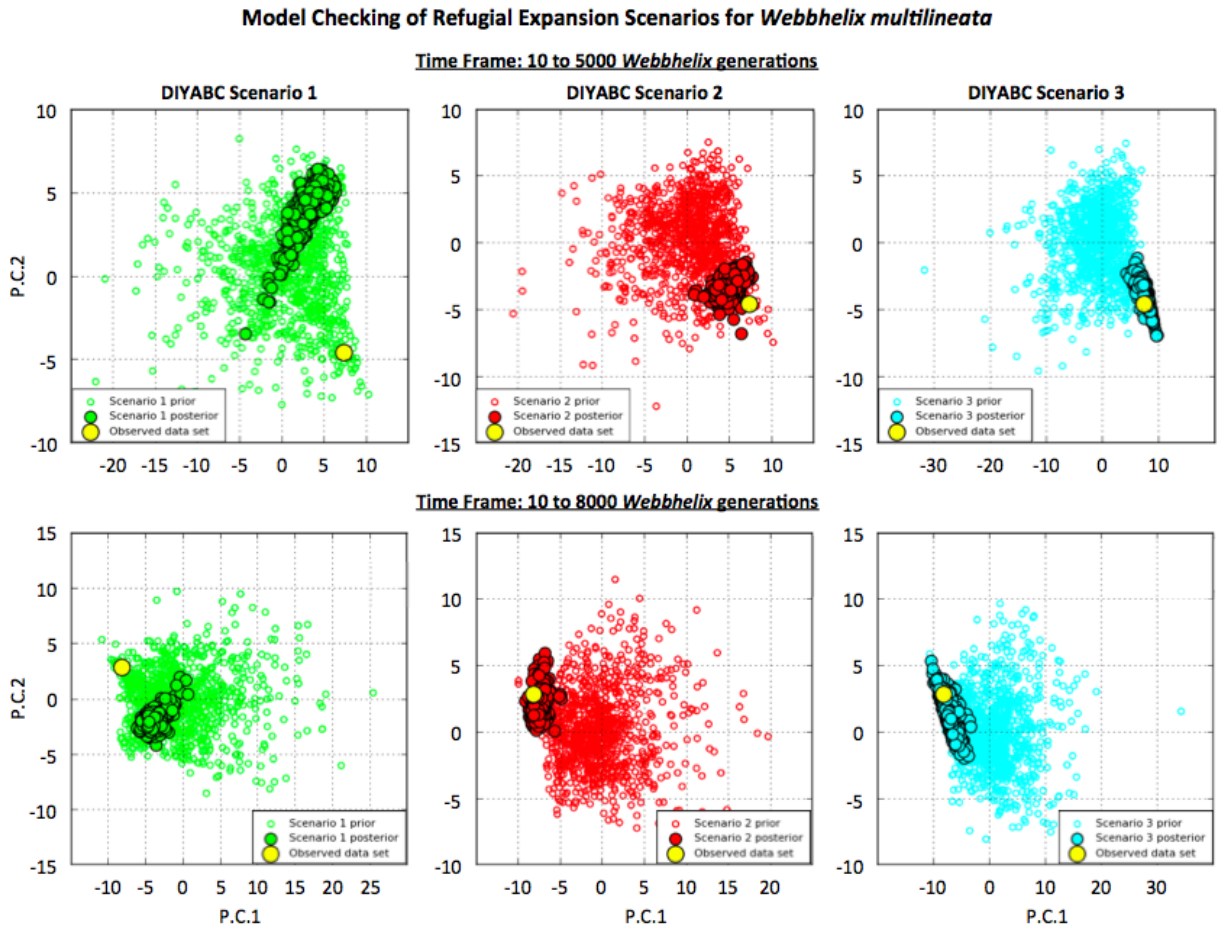


Figure 3.8 – Results of model checking against prior and posterior distributions of three scenarios for refugial expansion in *Webbhelix multilineata*, from the program DIYABC. Open circles represent parameter values drawn from prior distributions for each of the three scenarios (scenario 1 = green; scenario 2 = red; scenario 3 = blue). Filled circles represent posterior distributions of simulated parameters for each scenario. Large yellow circles represent the observed data. A generation time of 2.5 years was used for *Webbhelix*, so simulations under the 10 to 5,000 generations time frame operated within the hypothesis that refugial expansion occurred sometime later than 12,500 years ago; simulations under the 10 to 8,000 generations time frame allowed for expansion to have occurred sometime later than 20,000 years ago. For both time frames, data sets simulated from the posterior (1,000) clustered closer to the observed data under scenarios 2 & 3 compared to scenario 1, and closer overall to the observed data under scenario 2 than scenario 3.

distributions for each scenario; 1% of those were used in the local regression for each model checking analysis, and 1000 data sets were simulated from the posterior. The plots of principal component 1 (PC 1; 35.7% of data variance in the 10 to 5000 generation time frame; 36.1% in the 10 to 8000 generation time frame) and principal component 2 (PC 2; 13.7% of data variance in the 10 to 5000 generation time frame; 12.9% in the 10 to 8000 generation time frame) show a wide cloud of datasets simulated from the prior (open colored circles), a cluster of datasets from the posterior predictive distribution (filled color circles), and the observed data (single yellow circle). For scenario 1 (green plots), the posterior datasets did not cluster near or around the observed data for either time frame. For scenarios 2 and 3, however, posterior datasets filled the summary statistic space much closer to the observed data.

For the 10 to 5000 generations time frame, posterior datasets for scenarios 2 and 3 similarly fit the observed data when looking at the plots of PC 1 vs. PC 2 in **Figure 3.8**; however, plots of PC 1 vs. PC 3 – 5, and PC 2 vs. PC 3 – 5 consistently showed that posterior datasets for scenario 3 fit the observed data better than posterior datasets for scenario 2 (PC 3 – 11.5% of data variance; PC 4 – 10.8%; PC 5 – 9.9%). For the 10 to 8000 generations time frame, the opposite was the case. The plot of PC 1 vs PC 2 in **Figure 3.8** shows clearly that posterior datasets for scenario 2 fit the observed data better than posterior datasets for scenario 3, and nearly all other PCA plots showed the same thing for this time frame (PC 3 – 11.2% of data variance; PC 4 – 10.9%; PC 5 – 10.1%). The exception was in the PC 1 vs. PC 5 plots, in which scenario 3 fit the observed data better.

Comparisons of posterior probabilities for each refugial expansion scenario, for each time frame, are shown in **Figure 3.9**. Two different approaches for comparing posterior probabilities are shown: the graphs on the left show comparisons of posterior probabilities for the three

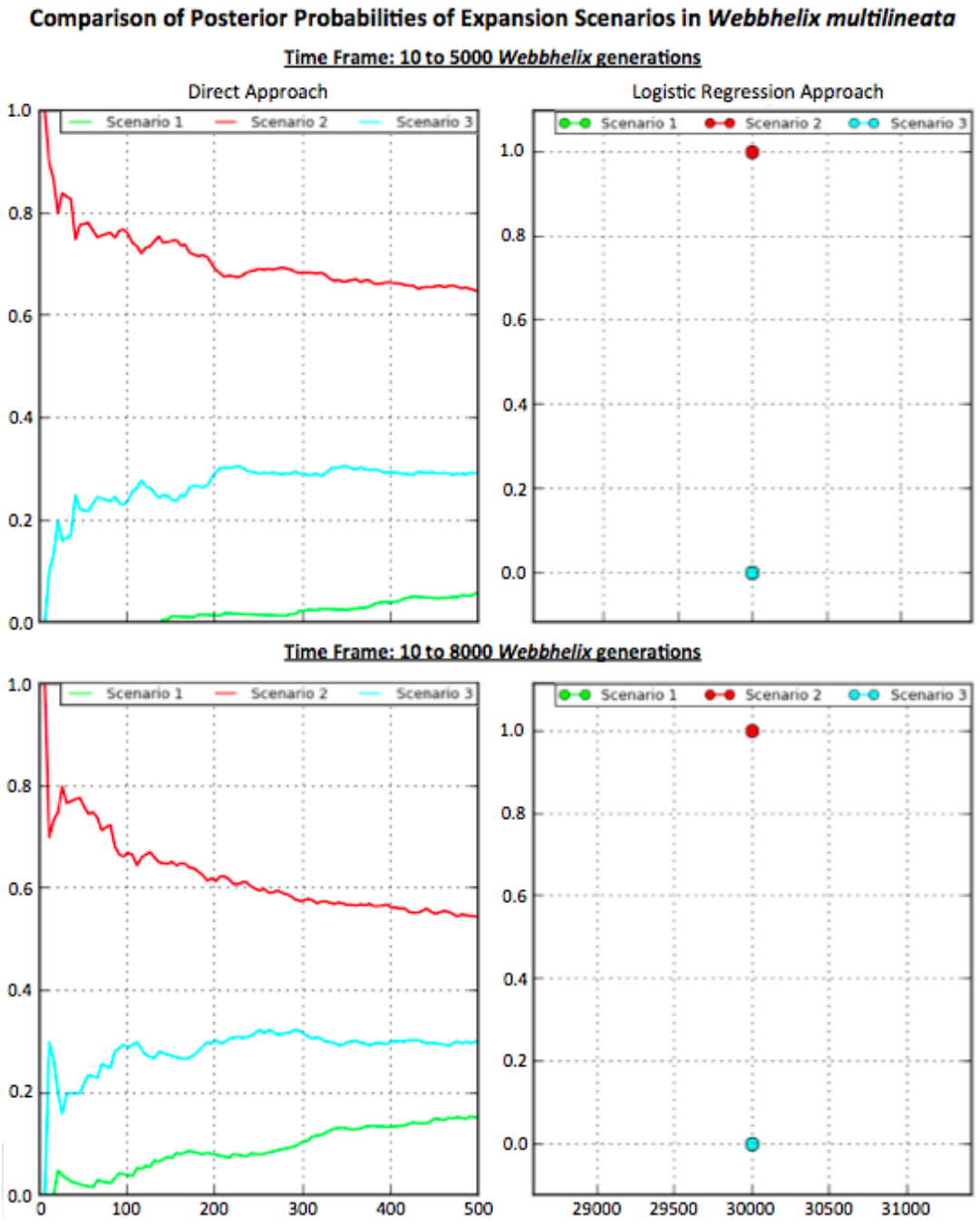


Figure 3.9 – Comparison of posterior probabilities for three refugial expansion scenarios in *Webbhelix multilineata*, from the program DIYABC. Comparisons for each of the two sets of coalescent simulations based on different time frames are shown, with the 10 to 5,000 *Webbhelix* generations time frame comparisons shown in the top two graphs and the 10 to 8,000 generations time frame comparisons shown in the bottom two graphs (1 generation = 2.5 years). The two left graphs show the probability (y-axis) of each scenario using a direct approach, by simply considering the 500 closest simulated data sets (x-axis) to the observed data. The two right graphs show the probability of each scenario using a logistic regression approach, with the x-axis showing the number of simulations used (1% of total simulations). In both regression comparisons, scenarios 1 & 3 are overlapping at 0.0.

scenarios estimated using a direct approach in DIYABC, with the probability of each scenario (y-axis) based on the distance to the observed data of the 500 closest posterior simulated datasets (x-axis) to the observed data for each scenario; the graphs on the right also show comparisons of posterior probabilities for the three scenarios but estimated using a logistic regression approach in DIYABC, with the probability of each scenario (y-axis) based on a polychotomic weighted logistic regression using 1% of the total number of simulations (x-axis) and taking the intercept of the regression as the point estimate.

Based on the direct approach, scenario 2 was consistently the most likely scenario (for both time frames tested) across all 500 of the closest simulated datasets to the observed data in each scenario. For the 10 to 5000 generations time frame, the scenario probabilities based on the 100 closest simulated datasets to the observed data were 76% for scenario 2, followed by 24% for scenario 3 and 0% for scenario 1; based on the 500 closest simulated datasets, the probabilities were 64.6% for scenario 2, followed by 29.4% for scenario 3 and 6% for scenario 1. For the 10 to 8000 generations time frame, the scenario probabilities based on the 100 closest simulated datasets to the observed data were 67% for scenario 2, followed by 29% for scenario 3 and 4% for scenario 1; lastly, based on the 500 closest simulated datasets, the probabilities were 54.6% for scenario 2, followed by 30% for scenario 3 and 15.4% for scenario 1.

The logistic regression approach also indicated that scenario 2 was the most likely scenario, even more so than the direct approach. For both time frames tested, the logistic regression estimated a 100% probability for scenario 2 and 0% probability for scenarios 1 and 3. Note that the blue dots representing scenario 3 in the logistic regression plots are overlapping the green dots for scenario 1, since they are both 0.0 on the probability axis.

Discussion

Population Structure and Diversity in *Webbhelix multilineata*

Assessment of a large SNP dataset generated from 178 *Webbhelix multilineata* sampled from 18 localities across a broad portion of the species range suggests 4 population clusters: a 'western cluster' (Cluster 1) consisting of localities from central Iowa, southern Minnesota and along the Missouri River in Nebraska; a 'central cluster' (Cluster 2) consisting of localities from northern and central Illinois, and eastern Iowa; an 'eastern cluster' (Cluster 3) consisting of the two localities from Ontario; and a 'southern cluster' (Cluster 4) consisting of two localities along the Mississippi in southeastern Missouri and eastern Arkansas. Clusters 1 & 2 share the most genetic variation, as evidenced by both the STRUCTURE analysis and the low estimation of pairwise F_{ST}. But while the genetic structuring between them does show substantial shared variation, it also shows extreme differentiation between some localities (SNPs from PTP, HGRF, OICP, TGCP, MPO and STT are nearly fully differentiated from RRO and CPS). Additionally, the estimated F_{ST} for this cluster pair is somewhat artificial because of the MRLP locality, which is almost evenly split between the two (green and blue variation from Figure 7). The locality was ultimately included in Cluster 2 because the proportion of membership from the STRUCTURE analysis was slightly higher for it compared to Cluster 1 (52.7% vs. 47% respectively), but the split is so even that inclusion in either population cluster may be questionable and lowers the estimate of F_{ST} in either case despite the localities with extreme proportions of membership in each cluster. Nevertheless, it is interesting to note (for now) that localities in these two clusters found along or near the Mississippi River tended to have more shared variation than those further away from it.

Interpretation of the genetic structuring between population Clusters 3 & 4 is more speculative owing to a significant gap in sampling from Kentucky, Indiana, Ohio and Michigan,

but one might imagine a similar pattern across that stretch of the species range as seen for Clusters 1 & 2 (where localities are more admixed toward the center and more differentiated out toward the edges). This gap in sampling has left the current estimate of F_{ST} between Clusters 3 & 4 relatively high since the genetic variation of their respective localities are so different, possibly being outer edge localities of two population clusters that span the geographical gap between them. Though the differentiation between Clusters 3 & 4 in this study is somewhat high, the differentiation between each of Clusters 3 & 4 and each of Clusters 1 & 2 is as high or higher. In fact, the estimate of pairwise F_{ST} for Clusters 1 & 4 (0.17) is as high as estimates between subspecies that have been obtained in studies for other systems (Carreras-Carbonell et al, 2006; Marion & Le Gentil, 2006; Funk et al, 2007; Kim & Roe, 2021).

While some estimates of F_{ST} were unexpected in their magnitude, observed and expected heterozygosity (H_O and H_S) estimates roughly followed an anticipated pattern: the highest estimates were found to be for the two southern-most localities in southeast Missouri and eastern Arkansas (Cluster 4), decreased for localities to the north (Cluster 2), and then decreased further from there for localities to the west and northwest (Cluster 1) and northeast (Cluster 3). This pattern is characteristic of serial founder effects, a consequence of species range expansions from refugia into previously unavailable habitat whereby subsets of refugial populations lead colonization fronts that found new colonies with low effective population sizes that experience genetic bottlenecking as their populations grow (Slatkin & Excoffier, 2012). These new populations end up having lowering genetic diversity then their founding populations, and also become the source of new colonization fronts that continue to expand the species range and found new colonies that experience even more bottlenecking and further reduced genetic diversity as they grow. This pattern has been observed for many species in North America with

ranges that include previously glaciated regions (Hewitt, 2000), particularly so for species with low motor and reproductive dispersal like *Webbhelix*. It is also suggestive of Model 2 for refugial expansion in *Webbhelix* from **Figure 3.2**, though alone does not rule out Model 3 as a possibility (discussed further in a proceeding section).

Relative to estimates of H_O and H_S, estimates of inbreeding coefficient (G_{IS}) seemed curiously misaligned from expectation. For populations with higher levels of inbreeding we typically expect to see lower levels of genetic diversity as a result of processes such as selective sweeps or background selection (Liu & Charlesworth, 1998). But for the present dataset some localities (particularly in Cluster 2) with high inbreeding also had higher genetic diversity than other localities with little or no signal of inbreeding and very low diversity (particularly in Cluster 1). However this is not entirely unusual, as inbreeding populations can occasionally have high diversity due to population mixture or local adaptation between populations with subdivision that is not apparent (Bonnin et al, 1996; Charlesworth et al, 1997; Liu & Charlesworth, 1998).

Isolation by Distance and Migration

The results of the Mantel test performed on linearized estimates of F_{ST} from the full 8,480 SNP dataset indicate that populations of *Webbhelix multilineata* at sampled localities are under strong influence of isolation by distance. This means that genetic divergences among sample localities are being driven by geographic distance restricting gene flow, rather than driven by selective forces. This is supported by the results of the DIVMIGRATE-ONLINE analysis for estimating relative migration rates between 8 sample localities along an approximately 670 mile continuous stretch of the Iowa and Mississippi Rivers, which showed almost no gene flow between locality pairs. Even the PTP and HGRF locality pair had no signal of gene flow despite being only 5 miles river distance apart. While it is perhaps not surprising that no gene flow was

detected between pairs that are 100+ miles apart, it is odd that the one pair of localities for which a strong and significant signal of downstream gene flow was detected (HGRF to TGCP) are ~52 miles river distance apart; yet no gene flow was detect from HGRF to OICP, a site only ~11 miles downstream from HGRF.

One explanation could be that population size of *Webbhelix* at some of these localities is simply too small for any meaningful gene flow to be occurring via flooding events washing migrants downstream. The three sample localities furthest upstream (PTP, HGRF, OICP) all had very low estimates of within-population genetic diversity, and the next locality downstream (TGCP) also had lower diversity compared to the remaining four downstream localities. This could be due to small population sizes and would explain the lack of migration signal despite the relative proximity of the upstream localities to each other. Another possible explanation could be that successful migration of terrestrial gastropods downriver via flooding events is highly dependent on floodplain arrangement (incline, position at river bends vs. straightaways, etc.) at the locality being emigrated from and the locality being immigrated to, along with the right fluid dynamics between them to consistently deposit debris at the receiving locality. Also the presence/absence of islands, and their size and number relative to the river's width, might be a relevant factor for river-facilitated gene flow of riparian terrestrial gastropods. Sinclair (2010) demonstrated that 14 populations of the land snail *Ventridens ligera* along a 5 mile stretch of the Potomac River, sampled from both riverbanks, multiple sites across the large, elongated Bear Island, and sites on a secondary island, had low estimates of F_{ST}, no genetic substructure, and low levels of IBD compared to 4 land-locked populations, suggesting the river facilitates gene flow. Bear Island, as well as several smaller islands, fill much of that stretch of the Potomac and likely break up the river flow and provide more opportunity for deposition of debris from the

riverbanks to islands and vice versa. None of the three short-river-distance *Webbhelix* sample localities (PTP, HGRF, OICP) in Iowa feature islands or large depositions of silt to break up the flow of the Iowa River, but there is quite a bit of silt buildup at the TGCP locality and the incline of the floodplain there is fairly level with the river, resulting in a marshy habitat that can flood easily. These features may help explain why TGCP is able to receive upstream migrants. As an aside, the RRL locality in western Illinois is on a 5+ mile stretch of the Mississippi that does have several large and small islands that occupy the width and length of the river; it would be interesting to model a follow-up study on *Webbhelix* there based on the sampling approach of Sinclair (2010), to see if multiple collection sites on the islands and along the adjacent riverbanks function as a single population with gene flow facilitated by the river.

As for estimates of relative migration between the population clusters determined from the STRUCTURE analysis, several statistically significant signals of unidirectional gene flow were detected. However, all signals were weak compared to the only strong signal detected: gene flow from Cluster 1 to Cluster 2. This makes sense considering the overall genetic structuring for all four clusters and the fact that Clusters 3 & 4 are underrepresented in terms of number of localities included in the analyses. That only gene flow from Cluster 1 to Cluster 2 was detected, and not the reciprocal, would be suggestive of downstream river-mediated gene flow considering that most Cluster 1 localities are upstream of the Cluster 2 localities that share variation with them. But the results of the migration analysis on the 8 continuous downstream localities along the Iowa and Mississippi Rivers showed no detectable gene flow from any of the Cluster 1 localities (PTP, HGRF, OICP, TGCP) to either of the Cluster 2 localities (IRO, RRL), suggesting river-mediated gene flow at the scale of 100+ miles river distance may not be realistic.

There are other random passive dispersal mechanisms of terrestrial snails that are thought to occur, including attachment to and dispersal by mammals (Fischer et al, 1996; Beinlich & Plachter, 2010; Maciorowski et al, 2012) and birds (Spencer & Patchett, 1997; Green & Figuerola, 2005; Kawakami et al, 2008; Maciorowski et al, 2012). Bird-mediated dispersal in particular may help explain these long-distance migration patterns between regional populations, as a study by Kawakami et al (2008) suggests that small snails can sometimes survive being ingested whole, passing through the digestive system, and being defecated out at a distant location by birds. But even if hitchhiking on (or in) birds was realistically enough to account for the signal of gene flow from Cluster 1 to Cluster 2 absent river-facilitated migration, there would still be the question of why the signal is unidirectional if the dispersal mechanism is a random passive one? If random displacement by birds was a significant component of snail migration, it would be reasonable to expect signals of gene flow to be reciprocal between population clusters.

While the results of the migration analyses in the present study were not able to provide a clear insight into the mechanism by which gene flow in *Webbhelix multilineata* populations operates across a large portion of the contemporary species range, the weak signals of downstream gene flow may be indicative of a historical migratory pattern. The last glacial recession reopened fresh habitat that drove the expansion of species ranges, the availability of newly opened resource niches enough to pull even low dispersal organisms like terrestrial gastropods hundreds of square miles across the continent. Once expansion ends and new habitat becomes occupied, however, gene flow for land snails likely remains relatively localized even across hundreds of miles of continuous riparian habitat as there existed for *Webbhelix* before broad deforestation and agricultural development by humans. But if large enough numbers of *Webbhelix* existed continuously along hundreds of miles of river corridors across the range, then

the likelihood of flooding events successfully and consistently moving migrants downstream increases, and the weak signals of regional gene flow in a generally downstream direction shown in this study may be reflective of that history.

Post-Glacial Refugial Expansion

Of the three models proposed in this study for refugial expansion of *Webbhelix multilineata*, for both the time frame starting as early as the LGM and the time frame starting after recession of the Laurentide Ice Sheet, model 2 was the most likely scenario, followed by model 3 and lastly by model 1. Model checking analyses performed in DIYABC showed that posterior simulations based on scenarios 2 and 3 both fit the observed data fairly well, with posterior simulations based on scenario 2 for the time frame starting at the LGM appearing to fit the observed data the most closely. Posterior simulations based on scenario 1 did not fit the observed data for either time frame tested. When posterior probabilities of each scenario were compared using a direct approach and a logistic regression, scenario 2 was clearly the most likely, for both time frames.

The DIYABC analyses support the hypothesis that post-glacial species range expansion in *Webbhelix* occurred following river corridors, rather than broadly across a transformed landscape of extended river floodplains and marshland due to glacial runoff. Expansion likely primarily followed the Mississippi River corridor, from which localities of population Cluster 2 were presumably first established. Additional expansion continued further north along the Mississippi, to the west along the Missouri River, and to the northwest along the Iowa River, establishing localities of population Cluster 1 from Cluster 2. Interpretation of the results of this study to infer the expansion route to population Cluster 3, however, is less clear.

Model 2 was the most probabilistic scenario here, suggesting an expansion route from population Cluster 2 to Cluster 3 likely involving the Illinois River up into the region south of

Lake Michigan, spreading east across multiple networks of smaller river corridors that characterize northern Indiana (such as are part of the upper Wabash watershed) and Ohio (the Maumee watershed), and southern Michigan (the St. Joseph watershed), then up into eastern Michigan and into southern Ontario. But there are some considerations that throw this proposed route to Cluster 3 into question. First is that unlike scenario 1, posterior simulations for scenario 3 actually clustered around the observed data in the probability space of the model-checking principal components analysis, almost fitting it as well as simulations for scenario 2. So discounting the possibility of scenario 3 based on the present data would be a mistake, particularly when one considers how unlikely it seems that expansion of *Webbhelix* via river corridors would proceed north along the Mississippi and not also to the northeast along the Ohio River corridor as scenario 3 proposes. Second is that the STRUCTURE analysis showed some shared variation between Cluster 3 and Cluster 4, and virtually no shared variation between Cluster 3 and Cluster 2. It is possible that filtering loci with too much missing data to reduce the size of the STRUCTURE dataset (8,480 SNPs) to a computationally manageable size for DIYABC (2,760 SNPs) resulted in this discrepancy; had the full STRUCTURE SNP dataset been used for coalescent simulations, scenario 3 may have been determined to be the most likely scenario. Related to this is the gap in sampling across the geographic area between Clusters 3 & 4, previously discussed. Additional sampled localities from Kentucky, Indiana, and Ohio (and particularly localities along or near the Ohio River across that region) may reveal more of a gradient of shared genetic structuring between Clusters 3 & 4, compensate for the loss of variation in the reduced SNP dataset size used for coalescent simulations in DIYABC, and result in scenario 3 being the more likely model for refugial expansion of *Webbhelix*.

Lastly, assessment in DIYABC of two time frames within which refugial expansion could have occurred for *Webbhelix* has led us to consider the 10 to 8000 generations time frame (corresponding to a possible start time for expansion around 20,000 years ago at the LGM) to be more suitable than the 10 to 5000 generations time frame (corresponding to an expansion start time around 12,500 years ago at the end of the Pleistocene and the ramping up of global seasonality), based on consistency between model checking results and posterior probability comparisons between scenarios within each time frame. Model checking analyses for the 10 to 5000 generations time frame showed scenario 3 fitting the observed data better than scenario 2, across comparisons of PC 1 with PC 2 – 5 and PC 2 with PC 3 – 5, but the posterior probability comparison indicated that scenario 2 was the most likely scenario. For the 10 to 8000 generations time frame the two results were more consistent; scenario 2 both fit the observed data best across major PC comparisons in the model checking analyses, and was the most likely scenario determined by posterior probability comparison with the other two scenarios.

Additionally, a study by Nash et al (2018) found 2 *Webbhelix multilineata* shells in the New Cottonwood School Peoria Silt loess section (approximately 150 km north of St. Louis) that were radiocarbon dated to 21,400 ± 300 and 20,400 ± 300 calibrated years before present (cal yr BP). This either supports our inference of an expansion north from southern refugia as early as or earlier than the LGM, or suggests a more northerly refugia for *Webbhelix* , much closer to the maximal extent of the Wisonsinin ice sheet, than we initially considered. That *Webbhelix* may have occupied colder refugia near the ice sheet means there could have been two populations (a northern ice sheet-adjacent refugial population and a southern refugial population) of this species occupying climatically disparate regions for hundreds or thousands of years prior to the LGM, enough for a substantial amount of species differentiation to have begun to occur. If this were the

case, it would help explain the pattern of genetic structuring and high estimates of F_{ST} seen between northern population Clusters 1 & 2 and the southern population Cluster 4. More localities from the sampling gap in this study previously discussed would not only help to better determine what expansion scenario is most likely, but also allow for an investigation into species delimitation between possible northern refugial populations and post-glacial offshoots (Clusters 2 and 1, respectively) and southern refugial populations and (likely) post-glacial offshoots (Clusters 4 and 3, respectively).

Summary and Future Directions

This study presents the first genomic dataset for *Webbhelix multilineata* and is one of only a relatively small number of genomic studies for a terrestrial gastropod available. Assessment of genomic structuring based on 18 localities sampled from across a broad span of the species range revealed four population clusters: two occupying the central Midwest on either side of the upper Mississippi (Cluster 1 in Iowa, Minnesota and Nebraska; Cluster 2 in Illinois and eastern Iowa), and showing admixed localities along or near the river; one in the south (Cluster 4) consisting of localities along the lower Mississippi River; and one to the east (Cluster 3) on the southwestern shores of Lake Erie. Genetic differentiation between Clusters 1 & 2 was low but substantial differentiation was estimated for all other pairwise clusters, with differentiation between Clusters 1 & 4 as high as between some subspecies. Estimates of genetic diversity showed a rough gradient of decreasing diversity from south to north along the Mississippi, and away from it to the west and east. Estimates of relative migration rates, both between population clusters and between a set of continuous downstream localities, suggest a historical pattern of downstream migration, while coalescent simulations suggest post-glacial expansion of the species range along major river corridors from Cluster 4 to Cluster 2, and from Cluster 2 to Clusters 1 & 3, with expansion beginning as early as the LGM.

Additional sampling of localities in the geographic region between Clusters 3 & 4, and particularly along the Ohio River, would greatly enhance the present assessment of genetic structure across the species range of *Webbhelix multilineata*, as well as allow for further testing of post-glacial expansion routes to better distinguish between scenarios 2 & 3 proposed in this study. This additional sampling would also help to facilitate a species delimitation study to investigate whether southern refugial populations of *Webbhelix* may have diverged substantially from populations stemming from a more northerly refugium adjacent to the ice sheet at the LGM.

References

Beinlich, B., & Plachter, H. (2010). Sheep: a functional corridor system. *Largescale livestock grazing. A management tool for nature conservation. Springer Verlag, Berlin*, 281-288.

Bonnin, I., Huguet, T., Gherardi, M., Prosperi, J. M., & Olivieri, I. (1996). High level of polymorphism and spatial structure in a selfing plant species, Medicago truncatula (Leguminosae), shown using RAPD markers. *American Journal of Botany*, *83*(7), 843-855.

Carreras-Carbonell, J., Macpherson, E., & Pascual, M. (2006). Population structure within and between subspecies of the Mediterranean triplefin fish Tripterygion delaisi revealed by highly polymorphic microsatellite loci. *Molecular Ecology*, *15*(12), 3527-3539.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomics, Genetics*, *1*, 171-182.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*.

Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetics Research*, *70*(2), 155-174.

Chen, B., Cole, J. W., & Grond-Ginsbach, C. (2017). Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Frontiers in Genetics*, *8*(167).

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., Mitrovica, J. X., Hostetler, S. W., & McCabe, A. M. (2009). The Last Glacial Maximum. *Science*, *325*(5941), 710-714.

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J. M., & Estoup, A. (2014). DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, *30*(8), 1187-1189.

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.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics*, *131*(2), 479-491.

Fischer, S. F., Poschlod, P., & Beinlich, B. (1996). Experimental Studies on the Dispersal of Plants and Animals on Sheep in Calcareous Grasslands. *Journal of Applied Ecology*, 1206-1222.

Funk, W. C., Mullins, T. D., & Haig, S. M. (2007). Conservation genetics of snowy plovers (Charadrius alexandrinus) in the Western Hemisphere: population genetic structure and delineation of subspecies. *Conservation Genetics*, *8*(6), 1287-1309.

Green, A. J., & Figuerola, J. (2005). Recent advances in the study of long-distance dispersal of aquatic invertebrates via birds. *Diversity and Distributions*, *11*(2), 149-156.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological journal of the Linnean Society*, *58*(3), 247-276.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907-913.

Hewitt, T. L., Bergner, J. L., Woolnough, D. A., & Zanatta, D. T. (2018). Phylogeography of the freshwater mussel species Lasmigona costata: testing post-glacial colonization hypotheses. *Hydrobiologia*, *810*(1), 191-206.

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I., & Xu, C. (2004). Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *European Journal of Human Genetics*, *12*, 395-399.

Hotopp, K. (2005). Webbhelix multilineata. Carnegie Museum of Natural History. <http://www.carnegiemnh.org/science/mollusks/pa_webbhelix_multilineata.html>

Kawakami, K., Wada, S., & Chiba, S. (2008). Possible dispersal of land snails by birds. *Ornithological Science*, *7*(2), 167-171.

Keenan, K. (2012). divMigrate-online: visualise and test gene flow patterns among populations. https://popgen.shinyapps.io/divMigrate -online/

Kim, K. S., & Roe, K. J. (2021). Genome-wide SNPs redefines species boundaries and conservation units in the freshwater mussel genus Cyprogenia of North America. *Scientific reports*, *11*(1), 1-16.

La Rocque, A. (1966). *Pleistocene Mollusca of Ohio*. State of Ohio Department of Natural Resources Division of Geological Survey, Bulletin 62 Part 1 + 2

Liu, F., Zhang, L., & Charlesworth, D. (1998). Genetic diversity in Leavenworthia populations with different inbreeding levels. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1393), 293-301.

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298-299.

Maciorowski, G., Urbanska, M., & Gierszal, H. (2012). An example of passive dispersal of land snails by birds-short note. *Folia Malacologica*, *20*(2).

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer research*, *27*(2 Part 1), 209-220.

Marion, L., & Le Gentil, J. (2006). Ecological segregation and population structuring of the Cormorant Phalacrocorax carbo in Europe, in relation to the recent introgression of continental and marine subspecies. *Evolutionary Ecology*, *20*(3), 193-216.

Martin, L. D., & Neuner, A. M. (1978). The End of the Pleistocene in North America. *Transactions of the Nebraska Academy of Sciences, Volume VI*, 117-126.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*, 28-41.

Michalakis, Y., & Excoffier, L. (1996). A Generic Estimation of Population Subdivision Using Distances Between Alleles With Special Reference for Microsatellite Loci. *Genetics*, *142*(3), 1061-1064.

Nash, T. A., Conroy, J. L., Grimley, D. A., Guenthner, W. R., & Curry, B. B. (2018). Episodic deposition of Illinois Valley Peoria silt in association with Lake Michigan Lobe fluctuations during the last glacial maximum. *Quaternary Research*, *89*(03).

Nei, M. (1987). Molecular Evolutionary Genetics. *Columbia University Press*, New York.

Perez, K. E., Defreitas, N., Slapcinsky, J., Minton, R. L., Anderson, F. E., & Pearce, T. A. (2014). Molecular Phylogeny, Evolution of Shell Shape, and DNA Barcoding in Polygyridae (Gastropoda: Pulmonata), an Endemic North American Clade of Land Snails. *American Malacological Bulletin*, *32*(1), 1-31.

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.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945-959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, *81*.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular ecology notes*, *4*(1), 137-138.

Scheet, P., & Stephens, M. (2008). Linkage Disequilibrium-Based Quality Control for Large-Scale Genetic Studies. *PLoS Genetics*, *4*(8), e1000147.

Sinclair, C. S. (2010). Surfing Snails: Population Genetics of the Land Snail Ventridens ligera (Stylommatophora: Zonitidae) in the Potomac Gorge. *American Malacological Bulletin*, *28*(2), 105-112.

Slatkin, M., & Excoffier, L. (2012). Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift. *Genetics*, *191*(1), 171-181.

Spencer, J. E., & Jonathan Patchett, P. (1997). Sr isotope evidence for a lacustrine origin for the upper Miocene to Pliocene Bouse Formation, lower Colorado River trough, and implications for timing of Colorado Plateau uplift. *Geological Society of America Bulletin*, *109*(6), 767-778.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P., & Kleinhans, D. (2016). Directional genetic differentiation and relative migration. *Ecology and evolution*, *6*(11), 3461-3475.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular ecology*, *7*(4), 453-464.

CHAPTER 4. ECOGEOGRAPHIC PATTERNS OF SHELL SIZE AND BRIGHTNESS VARIATION IN WEBBHELIX MULTILINEATA (MOLLUSCA, POLYGYRIDAE)

Jermaine Mahguib

Iowa State University, Ames, Iowa 50011

Modified from a manuscript to be submitted to *Journal of Molluscan Studies*

Abstract

Interpretive biogeographical studies often test established hypotheses that describe patterns that associate biological traits with ecological, environmental or geographic variables. Two well-known examples are Bergmann's Rule and Gloger's Rule, relating body size to environmental temperature and body pigmentation to temperature and humidity, respectively. Bergmann's hypothesizes that larger body sized congeners ought to be found in colder environments, while Gloger's hypothesizes more heavily pigmented forms ought to be found in warmer and more humid environments. Many such 'rules,' including these two, were originally intended to describe patterns specific to endotherms, but have since been employed more broadly in studies on ectotherms as well. The extent to which popular study patterns like Bergmann's applies in different ectothermic groups has been found to vary dramatically, with some conforming at very low rates and others at fairly high rates. In some groups though, like mollusks, the applicability is less clear-cut, with patterns like Bergmann's seeming to apply at about a 50/50 rate along with its inverse pattern. Both Bergmann's and Gloger's are often tested using a geographic variable, like latitude or elevation, as a proxy for temperature; thus, these rules tend to be tested on systems with a broad range extending over multiple degrees of these variables, such as seen in the terrestrial gastropod *Webbhelix multilineata*. This snail has a large range across much of the American Midwest and parts of the South that spans more than 10° of latitude; additionally, the typical size and brightness of individual shells has been observed to

vary noticeably from one locality to another. Here, we had the opportunity to generate morphological size and brightness data for up to 21 sampled localities of *Webbhelix* from across the species range and explicitly test the patterns described by two ecogeographic hypotheses, Bergmann's Rule and Gloger's Rule. A moderately strong negative correlation was found between shell size and latitude in *Webbhelix*, indicating it follows a reverse Bergmann cline. A weak positive correlation was found for shell brightness and latitude, suggestive of the pattern described by Gloger's, though brightness data generated was lacking compared to size data.

Introduction

Ecological Biogeographic Hypotheses

The field of biogeography encompasses a range of studies that seek to describe the geographical distribution of organisms (Posadas et al, 2006; Monge-Nájera, 2008). Such studies are typically either descriptive or interpretive; descriptive studies are concerned with uncovering information regarding biotic distributions, while interpretive studies (such as in ecological biogeography) look for correlative relationships between biological traits and geographic variables, and often test hypotheses about such relationships. Two examples of ecological biogeographic (ecogeographic) hypotheses are Bergmann's Rule and Gloger's Rule, both of which relate variation in morphological traits to an environmental variable or set of variables.

Bergmann's Rule is the most studied ecogeographic rule, the original form of which states that intraspecific variation in body size of homeothermic (endothermic) animals varies based on climate, with larger body sizes being found in colder climates and smaller body sizes found in warmer ones, owing to thermoregulatory mechanics (Bergmann, 1847; Salewski & Watt, 2017). Over time the form of this 'rule' (more of a hypothesized trend or pattern of ecological clines) has been altered due to shifting use of verbiage or use of surrogate variables

(like latitude for temperature) to describe it (Meiri & Dayan, 2003; Meiri, 2011; Salewski & Watt, 2017).

Gloger's Rule, another classic ecogeographical hypothesis, was originally defined by Rensch (1929) in two forms, a simplistic version and a complex version. The simplistic version states that endothermic animals will be darker in warmer and more humid climates due to heavier melanin pigmentation (Delhey, 2019). The complex version makes two statements that segregate the relationship between each type of melanin and environmental factors: the first is that more of the dark brown and black eumelanins will be present for animals in higher humidity climates, and less for conspecifics or congeners in extreme low temperature climates; the second is that the red-brown-yellowish pheomelanins are more prominent for individuals in dryer and warmer climates, with a dramatic reduction expected for individuals residing in lower temperature regions. As with Bergmann's Rule, Gloger's has seen its definition shifted across studies over the years (Delhey, 2019).

Bergmann's Rule and Gloger's Rule in Ectotherms

The validity of these ecogeographic hypotheses has been hotly contested in the scientific literature for over 160 years now, even among studies on endotherms for whom these hypotheses were originally intended (Meiri & Dayan, 2003; Meiri, 2011; Riemer et al, 2018; Delhey, 2019). This nevertheless has not stopped researchers from routinely testing these hypotheses (or variations of them) on poikilotherms (ectotherms) as well. Support for these hypotheses among studies on ectotherms varies between different groups of organisms (Vinarski, 2014). For example, in mollusks it's been shown that Bergmann's and converse Bergmann's clines occur at about the same rates, though converse Bergmann's appears to be more common, and frequently no correlation between mollusk size and latitude is found at all.

Much of the contention over the validity of ecogeographic hypotheses like Bergmann's and Gloger's revolves around the inapplicability or lack of generalizability of the mechanistic explanations that have been assigned to them, when trying to apply these hypotheses broadly across taxa. Meiri (2011), however, argues that mechanistic explanations are irrelevant to the validity of using ecogeographic hypotheses as a valuable toolset of *patterns* to explore for distributions in different taxa (though elucidating adaptive mechanisms for observed patterns *is* of course a logical and valuable next step). From this perspective, testing ecogeographic hypotheses can be viewed as a valid jumping off point for exploring population distributions, particularly in poorly studied taxa.

Shell Size and Brightness Ecogeographical Variation in a Terrestrial Gastropod

In this study, shell size and shell brightness data were measured for adult individuals of the terrestrial gastropod mollusk *Webbhelix multilineata*, the striped whitelip snail, sampled from localities across the species range to test two ecogeographic hypotheses: Bergmann's Rule and Gloger's Rule. *Webbhelix* has a broad geographical distribution across most of the American Midwest and parts of the American South, encompassing a seasonal north-south temperature cline. The shell of individuals in this species exhibits characteristic reddish-brown color bands formed by pigmentation in the periostracum, the thin outer layer of organic tissue covering the shell. Individuals can vary in the number and thickness of bands present, making some individuals (and some local populations) lighter or darker than others (personal observation). Additionally, *Webbhelix* is a low dispersal species and has habitat requirements restricting it to woodland riparian zones, a habitat type that has become highly fragmented across the Midwest over the past century due to human activity. As a result, morphological parameters such as body size and pigmentation at any given locality are influenced predominantly by local genetic diversity and environmental conditions, without much (if any) influence from migration. All of

these features of *Webbhelix multilineata* make it an ideal system for testing Bergmann's and Gloger's Rules to elucidate ecogeographical patterns that may be present in this poorly studied terrestrial gastropod.

Measures of shell size were used as a proxy for animal body size to assess whether sample localities of *Webbhelix* conform to the pattern described by Bergmann's Rule, a converse Bergmann's Rule, or show no correlation between body size and latitude of sample locality. Shell brightness (digitally measured from lateral view photo images of shells) was used as a proxy for amount of pigmentation in the shell periostracum to assess whether sample localities of *Webbhelix* conform to the pattern described by the simplistic version of Gloger's Rule.

Materials and Methods

Sample Localities and Collection of *Webbhelix multilineata*

A total of 21 sample localities were used in this study: Fort Snelling State Park (FSSP) and Cannon River Wilderness Area (CRWA) in southern Minnesota; Timmons Grove County Park (TGCP), Pikes Timber Park (PTP), Horse Grove Rietz Forest (HGRF), Oakland-Iowa River County Park (OICP), Phinney Park (PPFD) and Carr Woods, North River Valley Park (NRVP) in central Iowa; Black Hawk Park (BHP), Manhattan/Robbins Lake Park (MRLP), Pleasant Creek Recreation Area (PCRA) and an Iowa River woodland bank northwest of Oakville (IRO) in eastern Iowa; Mandan Park, Omaha (MPO) and Steamboat Trace Trail, Peru (STT) on the southeastern border of Nebraska; Rock River, Oregon (RRO) and Cooper Park South (CPS) in north central Illinois; Rip Rap Landing State Fish and Waterfowl Management Area (RRL) in western Illinois; Point Pelee National Park (PPNP) and Fish Point, Pelee Island (FPPI) in southern Ontario; Dorena-Hickman Ferry Access Point (DHFA) on the southeast border of Missouri; and Hornor Neck Lake, St. Francis National Forest (HNL) on the eastern border of Arkansas. Additional locality information, including number of individuals sampled per locality,

can be seen in **Table 3.1**. Samples from these localities were collected over the course of the summers of 2016, 2017 and 2018, with most localities sampled in 2018. Collection year for each locality is also shown in **Table 3.1**. Live snails sampled from localities were placed whole into a 50 mL Falcon tube filled with 100% ethanol, killing them via rapid dehydration of soft body tissues and preserving them for later use.

Assessment of Animal Size Among Sample Localities of *Webbhelix*

Three linear shell dimensions (shown in **Figure 4.1A**) were measured using a Grainger® stainless steel digital caliper with 0-150 mm range and 0.0005”/0.01 mm resolution. The first linear dimension was a long shell width (shell width 1), measuring a line starting from the most distant point of the aperture relative to the columella and running through the umbilicus all the way to the far opposite end of the shell. The second dimension was a short shell width (shell width 2), measuring a line starting on the body whorl near the point where the aperture lip connects dorsally to it, and running through the umbilicus to the far opposite end of the shell. The third linear dimension was shell height, measuring a line starting from the tip of the protoconch down to the center of the callus covering the umbilicus opening. A total of 410 specimens were measured in this manner, accounting for all fully mature adult snails collected; here, ‘fully mature adult’ refers to individuals who when collected had already grown to their maximum body/shell size. Achievement of max individual growth was determined based on the presence of a reflected aperture lip, as members of this species do not continue to grow their shells once this lip has been formed. Adolescent individuals (showing no reflected lip) were not measured in order to eliminate age as a variable contributing to size.

Additionally, the number of shell whorls per adult individual was also observed and recorded as a supplemental measure of size. Whorls were counted by eye using the scheme diagramed in **Figure 4.1B**. From a dorsal view, a shell was rotated in hand until the terminal end

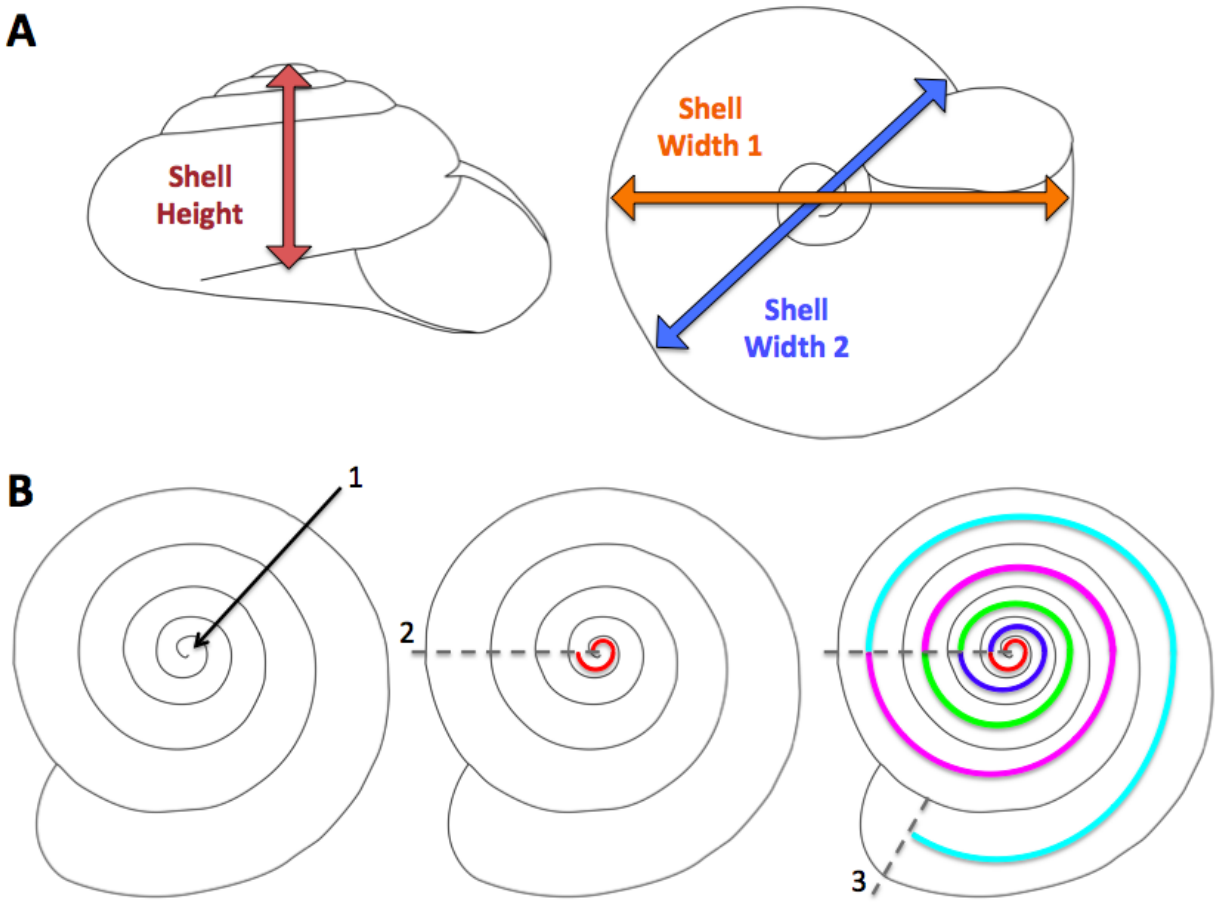


Figure 4.1 – Illustrations showing shell size measures used in this study. The top images (A) show shell height (left), measuring a line starting from the tip of the protoconch down to the center of the callus covering the umbilicus opening, and two shell width measures (right). Shell Width 1 measures a line starting from the most distant point of the aperture relative to the columella and running through the umbilicus all the way to the far opposite end of the shell. Shell Width 2 measures a line starting on the body whorl near the point where the aperture lip connects dorsally to it, and running through the umbilicus to the far opposite end of the shell. The bottom images (B) diagram how shell whorls were counted. Facing the dorsal view of the shell, it was rotated in hand until the terminal end of the whorl suture on the protoconch pointed up (1); a perpendicular line extending from the vertically aligned suture terminus (2) was used to mark the beginning and end of a single whorl, a full 360 degree turn around the line. Whorls were counted starting from the protoconch and ending at the aperture (3); incomplete final whorls (as shown here) were approximately rounded to the nearest quarter whorl based on position of the aperture relative to the whorl start/end line. The diagramed shell in (B) has 4 complete whorls (red, blue, green and purple) and 1 incomplete whorl (teal) that would be rounded down to and recorded as ¾ of a whorl, for a total of 4.75 whorls.

of the whorl suture on the protoconch pointed up; a perpendicular line extending left and right from the suture terminus point was used to mark the beginning and end of a single whorl, being a full 360 degree turn around the line. Whorls were counted starting from the protoconch and ending at the aperture; incomplete final whorls were approximately rounded to the nearest quarter whorl based on position of the aperture relative to the whorl start/end line, as shown in **Figure 4.1B**.

Once all shell size measurements were obtained, two sets of ANOVA's were conducted in R: the first set were performed using the data for each shell measure to calculate within-locality means and compare them pairwise across all 21 sample localities; the second set of ANOVA's were also performed using the data for each shell measure, but instead of sample localities the data were grouped into four population clusters from a separate population genomics study on *Webbhelix* (see Chapter 3). These population clusters were determined using the program STRUCTURE and a dataset of 8,480 SNP's from 18 of the 21 sample localities described for the present study. Population Cluster 1 consisted of sample localities FSSP, CRWA, PTP, HGRF, OICP, TGCP, MPO and STT; Cluster 2 consisted of sample localities PCRA, MRLP, IRO, RRO, CPS and RRL; Cluster 3 consisted of PPNP and FPPI; and Cluster 4 consisted of DHFA and HNL. Analysis of variance was used to statistically compare group means for each of the shell measures taken. Non-parametric Kruskal-Wallis rank sum tests were also performed on these data, both as a supplemental assessment of statistical differences of group means and as an alternative assessment when the data did not meet the assumptions of ANOVA.

These data were also tested for correlation between shell size measurements and latitude. Each shell size measure (long and short shell width, shell height) was plotted against latitude and

simple linear regression analyses were performed using the Excel-based program STATPLUS. Analyses of variance were also done to test for significance of regression models and slopes of trend lines for each size measure. Normality tests were conducted to test for normal distribution of data sets, and histogram and normal Q-Q plots for residuals were plotted to visually assess normality. Residuals were also plotted against predicted values and latitude to assess homogeneity of variance.

Assessment of Shell Brightness Among Sample Localities of *Webbhelix*

In order to assess shell color banding variation in *W. multilineata*, a basic morphometric approach was taken using shell brightness as a proxy. This involved analyzing high quality photographic images of shell transects (reverse-apertural view) to obtain a quantitative measure of overall brightness. To start, a couple of steps were taken to eliminate material that could absorb or reflect light other than the shell and pigmentation of the periostracum. The first was removing the soft body tissue from shells as carefully as could be managed. In some cases, the animal body could be maneuvered out completely intact; in most cases, however, the first quarter of the animal's soft body length would rip clean off, leaving the remainder stuck coiled within the spiral. To remove the remainder of the tissue, a pair of tweezers was used to punch and widen a roughly rectangular opening in the inner shell wall entering into the aperture, extending about equidistant into and out of the aperture plane. This opening was used to pull on the body tissue within the shell to help loosen the coil around the inner spiral and break some of the connective tissue fixing it to the inner shell walls. Once an animal was removed, the body tissues were placed into an individual glass vial filled with 100% ethanol for long-term storage. The shell interior was then painstakingly cleaned by inserting a thin, angled metal dissection pick through the aperture opening and rectangular opening made in the shell wall in order to gently scrape excess tissue that remained affixed to the interior walls. The fragile shell material of the inner

whorls was gently broken apart to gain access to most of the inner surface of the outer-most shell wall. For the purposes of long-term dry storage of the shells, all residual tissue within needed to be scraped and rinsed out with water; otherwise, any remaining tissue rots and generates a repulsive odor.

The second step for eliminating extra material that could absorb or reflect light was cleaning shell exteriors, as many specimens retained varying amounts of dirt and mud packed in between the small ribbing lining the outer shell surface. Dirt impaction on the outer shell makes them appear darker than they are based solely on color banding, so it was meticulously removed using a medium firmness toothbrush and brush strokes parallel to shell ribs, with intermittent water rinses. A total of 268 individual shells, representing 12 out of the 21 sample localities used in this study, were cleaned in this manner to prepare them for photographing. The 12 sample localities used were CRWA, PTP, HGRF, OICP, PCRA, RRO, IRO, CPS, RRL, PPNP, FPPI and DHFA.

It should be noted at this point that development of this project began in fall of 2019 and proceeded through 2020, which impacted the number of samples that could be included for examination as it was not previously known by the investigator that extended storage in ethanol has a couple of undesirable effects on gastropod shell integrity. For one, ethanol appears to weaken the calcium carbonate structure of shells over time, thinning and making them more brittle; this can make cleaning unmanageable, as weakened shells can break apart or collapse entirely when handled for brushing. Second, submersion in ethanol eventually begins to degrade the periostracum, which as previously mentioned is responsible for much of the color definition of the shell. Degradation of the periostracum results in faded coloration or, worse, outright peeling of the tissue when the shell is brushed. On careful inspection of all stored samples, it was

observed that those collected in 2016 and 2017 had become degraded as described above, while the samples collected in the summer of 2018 still had firm shells and only minor color fading relative to a few fresh samples collected in fall of 2019 for comparison. Since the 12 sample localities collected in 2018 were all stored in ethanol for roughly the same amount of time, and still spanned a broad north-south seasonal temperature cline, those samples were used to assess differences in shell brightness among localities along a latitudinal gradient.

Once these samples were cleaned and dried, they were photographed using the setup shown in **Figure 4.2**. Shells were held upright for presenting a reverse-apertural view to an overhead camera by mounting them on a set of custom-made, size-variable grey molding clay stands (**Figure 4.2A**). Each shell was individually mounted on top of a Pixel Perfect 4 x 6 inch 18% grey card to provide a neutral color background, and set next to a 24 standard color calibration chart (of the same brand and dimensions) for a consistent color reference in each image (**Figure 4.2B**). The color card backgrounds were fixed to a wooden block and centered on a Smith-Victor 36" pro-duty copy stand with twin adjustable LED copy light arms (**Figure 4.2C**). To help reduce glare from light reflecting off the shells, each LED was fixed with a plastic linear polarization sheet wrapped around the front of the light; to further reduce glare, each LED was also covered in an opaque plastic shopping bag (**Figure 4.2D**).

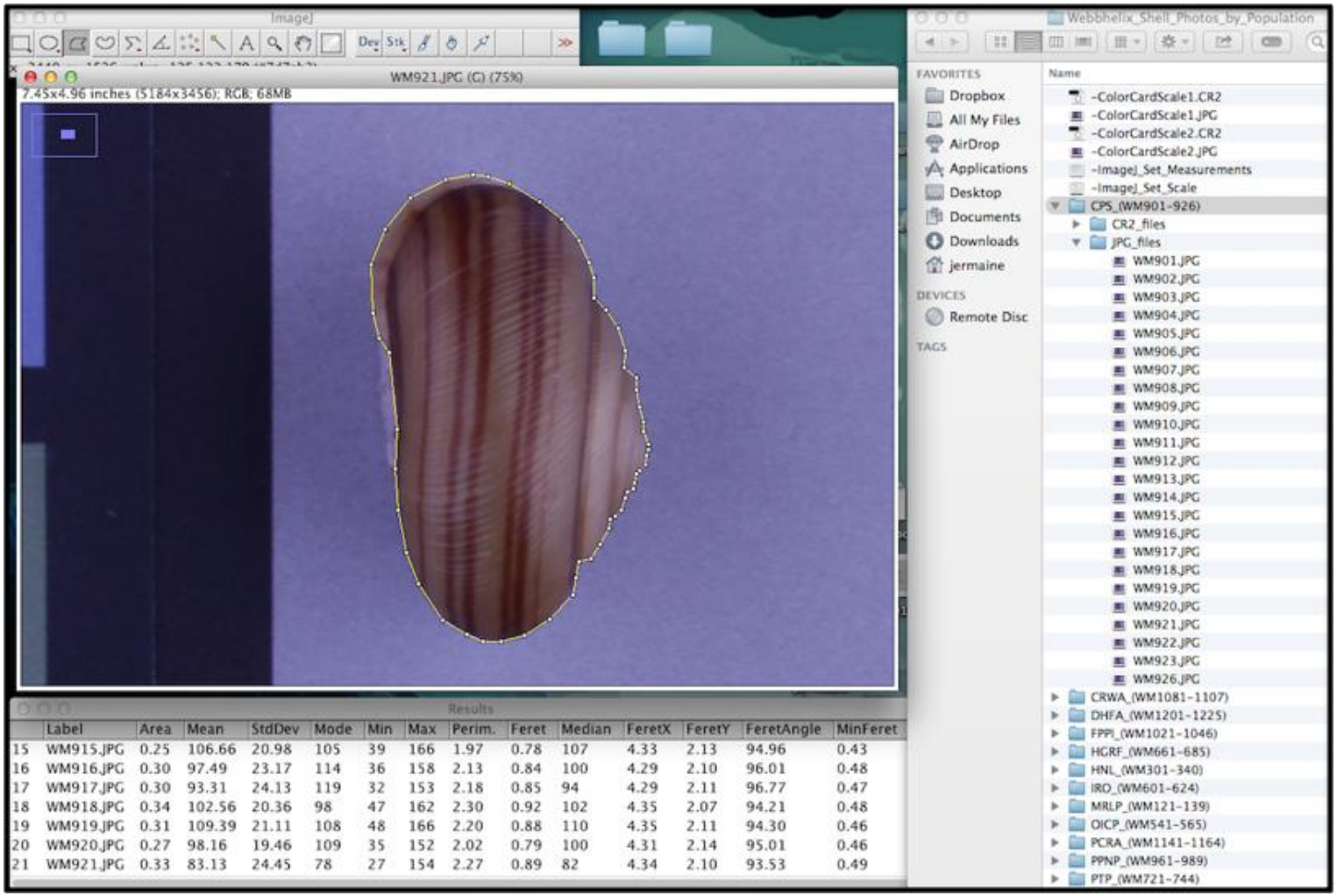
Shell photographs were taken using a Canon EOS Rebel T6 DSLR camera with 18 – 55 mm lens, mounted to the copy stand overhead. The camera was used in manual mode so that the same light and image quality settings would be applied to every shell photo. The aperture setting of the camera (ratio of focal length to entrance pupil diameter) was set to f/18 to capture as much detail as possible. A low ISO speed (image sensor's sensitivity to light) of 100 was used to further preserve image details. Camera shutter speed was set to ¼ of a second to increase image

Figure 4.2 – Images showing shell photography setup for capturing shell brightness. (A) shows how shells were held upright for presenting a reverse-apertural view to an overhead camera by mounting them on grey molding clay stands. (B) shows how shells were mounted on top of an 18% grey card to provide a neutral color background, and set next to a 24 standard color calibration chart for consistent color referencing in each image. (C) shows how the camera and lighting were mounted on a copy stand, and plastic linear polarization sheets were wrapped in front of the lights to reduce glare. (D) shows how plastic shopping bags were also used to further reduce glare.

exposure, since the steps described above to reduce glare on shells also reduced light availability. Lastly, the digital file setting on the camera was set to the maximum number of pixels and image quality. Each image was manually focused on the shell specimen to produce as clear an object as could be managed, and care was taken not to contact any part of the copy stand setup when initiating the camera shutter to minimize camera shake. All images were captured in a dark room with the only light source being the copy stand LED's. As an aside, positioning of mounted shells on the grey card was shifted toward the image center rather than the center of the grey card to make it easier to focus on them.

Shell photos were loaded into IMAGEJ version 1.50i (Rasband, 1997-2018; Abramoff et al, 2004; Schneider et al, 2012), a commonly used image processing and analysis program that can generate descriptive statistics for select portions of an image. After setting a global image scale using the 0.75 x 0.75 inch color squares of the calibration chart, images were zoomed in on the shell and a segmented line was drawn around the perimeter (as shown in **Figure 4.3**) to designate the area for image analysis. A set of measurements was calculated to describe the brightness of shell transects, including mean, median and modal gray value, and gray value standard deviation. Mean gray value (the sum of the gray values of all the pixels within the image selection divided by the number of pixels) is calculated in IMAGEJ by converting each pixel in the selection to grayscale using the formula gray = (red + green + blue)/3. A set of supplemental size measurements were also calculated, including area of the image selection in square inches, perimeter length, and Feret's diameter.

These data were then formatted for ANOVA in R to compare sample locality means, to see primarily if degree of shell color banding (proxied by shell brightness) varied significantly between any localities sampled. As with the manually measured size data from the previous



	Label	Area	Mean	StdDev	Mode	Min	Max	Perim.	Feret	Median	FeretX	FeretY	FeretAngle	MinFeret
15	WM915.JPG	0.25	106.66	20.98	105	39	166	1.97	0.78	107	4.33	2.13	94.96	0.43
16	WM916.JPG	0.30	97.49	23.17	114	36	158	2.13	0.84	100	4.29	2.10	96.01	0.48
17	WM917.JPG	0.30	93.31	24.13	119	32	153	2.18	0.85	94	4.29	2.11	96.77	0.47
18	WM918.JPG	0.34	102.56	20.36	98	47	162	2.30	0.92	102	4.35	2.07	94.21	0.48
19	WM919.JPG	0.31	109.39	21.11	108	48	166	2.20	0.88	110	4.35	2.11	94.30	0.46
20	WM920.JPG	0.27	98.16	19.46	109	35	152	2.02	0.79	100	4.31	2.14	95.01	0.46
21	WM921.JPG	0.33	83.13	24.45	78	27	154	2.27	0.89	82	4.34	2.10	93.53	0.49

Figure 4.3 – Image showing how shell photographs were digitally analyzed using the program IMAGEJ to measure mean gray value as a proxy for shell brightness. Shell photos were zoomed in on the shell and a segmented line was drawn around the perimeter to designate the area for image analysis. IMAGEJ calculates mean gray value (the sum of the gray values of all the pixels within the image selection divided by the number of pixels) by converting each pixel in the selection to grayscale using the formula gray = (red + green + blue)/3.

section, brightness data were additionally grouped into population clusters and tested for differences of means among clusters. For the brightness data, Population Cluster 1 consisted of sample localities CRWA, PTP, HGRF, OICP; Cluster 2 consisted of sample localities PCRA, IRO, RRO, CPS and RRL; Cluster 3 consisted of PPNP and FPPI; and Cluster 4 consisted of only the DHFA locality. Non-parametric Kruskal-Wallis rank sum tests were also performed on these data as a supplementary assessment of group means and alternative assessment when the data did not meet the assumptions of ANOVA.

These data were also tested for correlation between shell brightness measurements and latitude. Shell brightness data were plotted against latitude the same as described previously for the manually measured shell size data, using the program STATPLUS. A regression model was fit to the data, ANOVA was used to test for significance of the model, normality tests were performed, and residuals were plotted to assess homogeneity.

Results

Sample Locality and Population Cluster Comparisons of Shell Size

The shell size data grouped by sample localities met assumptions of ANOVA except in the case of the shell whorls data, which were shown by the Shapiro-Wilk normality test to be non-normally distributed ($W = 0.986$, $p = 0.001$). Both the shell width measures and the shell height measure were statistically significant for differences between sample locality means via ANOVA (shell width 1: $df = 20$, $F = 30.58$, $p < 2e\text{-}16$; shell width 2: $df = 20$, $F = 31.96$, $p < 2e\text{-}16$; shell height: $df = 20$, $F = 37.75$, $p < 2e\text{-}16$) and all four shell size measures were statistically significant for differences between sample locality means via non-parametric Kruskal-Wallis rank sum tests (shell width 1: $df = 20$, $chi^2 = 239.63$, $p < 2.2e\text{-}16$; shell width 2: $df = 20$, $chi^2 = 247.21$, $p < 2.2e\text{-}16$; shell height: $df = 20$, $chi^2 = 260.31$, $p < 2.2e\text{-}16$; shell whorls: $df = 20$, $chi^2 = 117.67$, $p = 7.68e\text{-}16$). All p-values for pairwise comparisons of sample localities using T-tests

with pooled standard deviations can be viewed in **Tables 4.1 – 4.4** for shell width 1, shell width 2, shell height and shell whorls, respectively. As a preliminary assessment of Bergmann's ecogeographic hypothesis, sample localities in these tables were arranged in order by latitude from north to south. However, no overarching pattern could be interpreted from pairwise comparisons of sample localities across the latitudinal range. What did stand out was that the PTP and HGRF locality means were statistically non-different from each other, and statistically different (smaller) than all other sample localities. The PTP and HGRF localities were on average ~2 – 2.5 mm smaller in width and height than all other sample localities, and have approximately a quarter fewer whorls.

The shell width 1 & 2 size data grouped by population clusters met the assumptions of ANOVA, but the shell height and shell whorls data were non-normally distributed (shell height: $W = 0.985$, $p = 0.001$; shell whorls: $W = 0.968$, $p = 5.69e\text{-}07$). Both shell width measures were statistically significant for differences between population cluster means via ANOVA (shell width 1: $df = 3$, $F = 63.15$, $p < 2e\text{-}16$; shell width 2: $df = 3$, $F = 72.41$, $p < 2e\text{-}16$) and all four shell size measures were statistically significant for differences between population cluster means via non-parametric Kruskal-Wallis rank sum tests (shell width 1: $df = 3$, $chi^2 = 133.58$, $p < 2.2e\text{-}16$; shell width 2: $df = 3$, $chi^2 = 144.9$, $p < 2.2e\text{-}16$; shell height: $df = 3$, $chi^2 = 99.18$, $p < 2.2e\text{-}16$; shell whorls: $df = 3$, $chi^2 = 45.55$, $p = 7.08e\text{-}10$). **Table 4.5** shows data summary statistics, p-values for pairwise comparisons of group means using T-tests with pooled standard deviations, and p-values for Tukey multiple comparison tests of group means, for shell size measures grouped by population clusters. Cluster 1 means for all four size measures were statistically significantly smaller compared to Clusters 2, 3 & 4. No significant difference was found for any of the pairwise comparisons between Clusters 2, 3 & 4.

Table 4.1 – Shell width 1 pairwise comparisons of sample locality means using T-tests with pooled standard deviations. Statistically significant p-values ($p < 0.05$) are shown in bold and highlighted in yellow. Sample localities are arranged by latitude on the left, from the northern-most locality (FSSP) down to the southern-most locality (HNL).

Latitude	Sample Localities	FSSP	CRWA	PTP	HGRF	OICP	BHP	PPFD	PCRA	TGCP	NRVP	RRO	MRLP	PPNP	FPPI	MPO	IRO	CPS	STT	RRL	DHFA
44°53'11.80"N	FSSP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44°22'5.00"N	CRWA	0.955	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°46'58.08"N	PTP	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°43'42.78"N	HGRF	**0.000**	**0.000**	0.505	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°34'55.32"N	OICP	**0.002**	**0.002**	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°33'50.77"N	BHP	0.702	0.663	**0.000**	**0.000**	**0.005**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°30'18.37"N	PPFD	0.287	0.244	**0.000**	**0.000**	**0.000**	0.683	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°12'51.69"N	PCRA	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.008**	**0.004**	-	-	-	-	-	-	-	-	-	-	-	-	-
42° 5'45.62"N	TGCP	0.263	0.227	**0.000**	**0.000**	**0.000**	0.627	0.896	**0.012**	-	-	-	-	-	-	-	-	-	-	-	-
42° 2'32.68"N	NRVP	**0.041**	**0.041**	**0.000**	**0.000**	0.376	**0.042**	**0.001**	**0.000**	**0.002**	-	-	-	-	-	-	-	-	-	-	-
42° 0'38.99"N	RRO	**0.047**	**0.035**	**0.000**	**0.000**	**0.000**	0.244	0.329	0.082	0.470	**0.000**	-	-	-	-	-	-	-	-	-	-
42° 0'30.52"N	MRLP	**0.001**	**0.001**	**0.000**	**0.000**	**0.000**	**0.017**	**0.015**	0.953	**0.030**	**0.000**	0.127	-	-	-	-	-	-	-	-	-
41°55'46.75"N	PPNP	0.657	0.618	**0.000**	**0.000**	**0.003**	0.955	0.721	**0.008**	0.657	**0.030**	0.256	**0.017**	-	-	-	-	-	-	-	-
41°43'56.37"N	FPPI	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.006**	**0.000**	**0.000**	**0.000**	**0.021**	**0.000**	-	-	-	-	-	-	-
41°11'38.03"N	MPO	0.133	0.136	**0.000**	**0.000**	0.127	0.114	**0.006**	**0.000**	**0.008**	0.592	**0.000**	**0.000**	0.088	**0.000**	-	-	-	-	-	-
41° 6'11.07"N	IRO	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.187	**0.000**	**0.000**	**0.002**	0.282	**0.000**	0.194	**0.000**	-	-	-	-	-
40°41'51.85"N	CPS	0.657	0.614	**0.000**	**0.000**	**0.000**	0.972	0.580	**0.001**	0.512	**0.010**	0.129	**0.004**	0.935	**0.000**	**0.042**	**0.000**	-	-	-	-
40°28'58.36"N	STT	0.206	0.181	**0.000**	**0.000**	**0.000**	0.479	0.657	0.082	0.757	**0.003**	0.769	0.113	0.505	**0.000**	**0.011**	**0.004**	0.374	-	-	-
39°18'34.46"N	RRL	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.084	**0.000**	**0.000**	**0.000**	0.174	**0.000**	0.215	**0.000**	0.855	**0.000**	**0.001**	-	-
36°35'7.86"N	DHFA	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.086	**0.000**	**0.000**	**0.000**	0.181	**0.000**	0.181	**0.000**	0.903	**0.000**	**0.001**	0.947	-
34°35'46.40"N	HNL	**0.007**	**0.004**	**0.000**	**0.000**	**0.000**	0.112	0.113	0.145	0.206	**0.000**	0.663	0.206	0.113	**0.000**	**0.000**	**0.004**	**0.029**	0.513	**0.000**	**0.000**

Table 4.2 – Shell width 2 pairwise comparisons of sample locality means using T-tests with pooled standard deviations. Statistically significant p-values ($p < 0.05$) are shown in bold and highlighted in yellow. Sample localities are arranged by latitude on the left, from the northern-most locality (FSSP) down to the southern-most locality (HNL).

Latitude	Sample Localities	FSSP	CRWA	PTP	HGRF	OICP	BHP	PPFD	PCRA	TGCP	NRVP	RRO	MRLP	PPNP	FPPI	MPO	IRO	CPS	STT	RRL	DHFA
44°53'11.80"N	**FSSP**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44°22'5.00"N	**CRWA**	0.405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°46'58.08"N	**PTP**	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°43'42.78"N	**HGRF**	**0.000**	**0.000**	0.939	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°34'55.32"N	**OICP**	**0.032**	**0.002**	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°33'50.77"N	**BHP**	0.718	0.776	**0.000**	**0.000**	**0.039**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°30'18.37"N	**PPFD**	0.321	0.866	**0.000**	**0.000**	**0.001**	0.707	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°12'51.69"N	**PCRA**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-
42° 5'45.62"N	**TGCP**	0.129	0.445	**0.000**	**0.000**	**0.000**	0.394	0.533	**0.002**	-	-	-	-	-	-	-	-	-	-	-	-
42° 2'32.68"N	**NRVP**	0.265	**0.043**	**0.000**	**0.000**	0.352	0.207	**0.028**	**0.000**	**0.008**	-	-	-	-	-	-	-	-	-	-	-
42° 0'38.99"N	**RRO**	**0.000**	**0.003**	**0.000**	**0.000**	**0.000**	**0.010**	**0.004**	0.227	**0.049**	**0.000**	-	-	-	-	-	-	-	-	-	-
42° 0'30.52"N	**MRLP**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.003**	**0.002**	0.826	**0.015**	**0.000**	0.424	-	-	-	-	-	-	-	-	-
41°55'46.75"N	**PPNP**	0.099	0.321	**0.000**	**0.000**	**0.001**	0.295	0.387	**0.016**	0.752	**0.009**	0.179	0.058	-	-	-	-	-	-	-	-
41°43'56.37"N	**FPPI**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.003**	**0.000**	**0.000**	**0.000**	**0.004**	**0.000**	-	-	-	-	-	-	-
41°11'38.03"N	**MPO**	0.403	0.082	**0.000**	**0.000**	0.195	0.303	0.054	**0.000**	**0.016**	0.752	**0.000**	**0.000**	**0.016**	**0.000**	-	-	-	-	-	-
41° 6'11.07"N	**IRO**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.101	**0.000**	**0.000**	**0.004**	0.105	**0.000**	0.195	**0.000**	-	-	-	-	-
40°41'51.85"N	**CPS**	0.195	0.630	**0.000**	**0.000**	**0.000**	0.523	0.741	**0.000**	0.760	**0.013**	**0.015**	**0.005**	0.555	**0.000**	**0.026**	**0.000**	-	-	-	-
40°28'58.36"N	**STT**	**0.039**	0.158	**0.000**	**0.000**	**0.000**	0.158	0.195	0.059	0.460	**0.003**	0.388	0.146	0.711	**0.000**	**0.005**	**0.001**	0.311	-	-	-
39°18'34.46"N	**RRL**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.128	**0.000**	**0.000**	**0.004**	0.136	**0.000**	0.085	**0.000**	0.760	**0.000**	**0.001**	-	-
36°35'7.86"N	**DHFA**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.300	**0.000**	**0.000**	**0.014**	0.282	**0.000**	**0.024**	**0.000**	0.432	**0.000**	**0.004**	0.581	-
34°35'46.40"N	**HNL**	**0.000**	**0.003**	**0.000**	**0.000**	**0.000**	**0.013**	**0.004**	0.097	0.061	**0.000**	0.752	0.272	0.233	**0.000**	**0.000**	**0.001**	**0.016**	0.493	**0.000**	**0.002**

Table 4.3 – Shell height pairwise comparisons of sample locality means using T-tests with pooled standard deviations. Statistically significant p-values ($p < 0.05$) are shown in bold and highlighted in yellow. Sample localities are arranged by latitude on the left, from the northern-most locality (FSSP) down to the southern-most locality (HNL).

Latitude	Sample Localities	FSSP	CRWA	PTP	HGRF	OICP	BHP	PPFD	PCRA	TGCP	NRVP	RRO	MRLP	PPNP	FPPI	MPO	IRO	CPS	STT	RRL	DHFA
44°53'11.80"N	FSSP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44°22'5.00"N	CRWA	0.369	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°46'58.08"N	PTP	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°43'42.78"N	HGRF	**0.000**	**0.000**	0.324	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°34'55.32"N	OICP	**0.000**	**0.003**	**0.000**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°33'50.77"N	BHP	0.599	0.876	**0.000**	**0.000**	**0.015**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°30'18.37"N	PPFD	0.710	0.599	**0.000**	**0.000**	**0.000**	0.804	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°12'51.69"N	PCRA	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**	**0.005**	**0.001**	-	-	-	-	-	-	-	-	-	-	-	-	-
42° 5'45.62"N	TGCP	0.080	**0.006**	**0.000**	**0.000**	**0.000**	**0.049**	**0.026**	0.349	-	-	-	-	-	-	-	-	-	-	-	-
42° 2'32.68"N	NRVP	0.273	0.822	**0.000**	**0.000**	**0.009**	0.759	0.461	**0.000**	**0.005**	-	-	-	-	-	-	-	-	-	-	-
42° 0'38.99"N	RRO	**0.004**	**0.000**	**0.000**	**0.000**	**0.000**	**0.004**	**0.001**	0.991	0.331	**0.000**	-	-	-	-	-	-	-	-	-	-
42° 0'30.52"N	MRLP	**0.008**	**0.001**	**0.000**	**0.000**	**0.000**	**0.006**	**0.002**	0.797	0.282	**0.000**	0.797	-	-	-	-	-	-	-	-	-
41°55'46.75"N	PPNP	0.271	0.710	**0.000**	**0.000**	0.054	0.668	0.419	**0.001**	**0.010**	0.848	**0.000**	**0.001**	-	-	-	-	-	-	-	-
41°43'56.37"N	FPPI	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	0.378	0.076	**0.000**	0.375	0.609	**0.000**	-	-	-	-	-	-	-
41°11'38.03"N	MPO	**0.017**	0.145	**0.000**	**0.000**	0.151	0.211	**0.039**	**0.000**	**0.000**	0.253	**0.000**	**0.000**	0.470	**0.000**	-	-	-	-	-	-
41° 6'11.07"N	IRO	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.007**	**0.000**	**0.031**	**0.000**	-	-	-	-	-
40°41'51.85"N	CPS	**0.000**	**0.005**	**0.000**	**0.000**	0.870	**0.022**	**0.001**	**0.000**	**0.000**	**0.015**	**0.000**	**0.000**	0.075	**0.000**	0.209	**0.000**	-	-	-	-
40°28'58.36"N	STT	0.114	**0.017**	**0.000**	**0.000**	**0.000**	0.065	0.051	0.510	0.894	**0.012**	0.497	0.407	**0.017**	0.157	**0.000**	**0.000**	**0.000**	-	-	-
39°18'34.46"N	RRL	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.002**	**0.000**	**0.011**	**0.000**	0.875	**0.000**	**0.000**	-	-
36°35'7.86"N	DHFA	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.161	**0.012**	**0.000**	0.152	0.405	**0.000**	0.822	**0.000**	**0.019**	**0.000**	0.055	**0.004**	-
34°35'46.40"N	HNL	**0.006**	**0.000**	**0.000**	**0.000**	**0.000**	**0.007**	**0.001**	0.671	0.541	**0.000**	0.655	0.518	**0.001**	0.169	**0.000**	**0.000**	**0.000**	0.719	**0.000**	**0.030**

Table 4.4 – Shell whorls pairwise comparisons of sample locality means using T-tests with pooled standard deviations. Statistically significant p-values ($p < 0.05$) are shown in bold and highlighted in yellow. Sample localities are arranged by latitude on the left, from the northern-most locality (FSSP) down to the southern-most locality (HNL).

Latitude	Sample Localities	FSSP	CRWA	PTP	HGRF	OICP	BHP	PPFD	PCRA	TGCP	NRVP	RRO	MRLP	PPNP	FPPI	MPO	IRO	CPS	STT	RRL	DHFA
44°53'11.80"N	FSSP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44°22'5.00"N	CRWA	0.643	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°46'58.08"N	PTP	**0.002**	**0.013**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°43'42.78"N	HGRF	**0.000**	**0.000**	0.057	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°34'55.32"N	OICP	0.467	0.855	**0.031**	**0.000**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°33'50.77"N	BHP	0.822	0.953	**0.045**	**0.000**	0.822	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°30'18.37"N	PPFD	0.467	0.136	**0.000**	**0.000**	0.083	0.358	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42°12'51.69"N	PCRA	0.733	0.989	**0.019**	**0.000**	0.830	0.989	0.204	-	-	-	-	-	-	-	-	-	-	-	-	-
42° 5'45.62"N	TGCP	0.389	0.119	**0.000**	**0.000**	0.078	0.304	0.888	0.175	-	-	-	-	-	-	-	-	-	-	-	-
42° 2'32.68"N	NRVP	0.663	0.997	**0.017**	**0.000**	0.855	0.953	0.159	0.990	0.138	-	-	-	-	-	-	-	-	-	-	-
42° 0'38.99"N	RRO	0.297	0.083	**0.000**	**0.000**	0.053	0.246	0.822	0.124	0.953	0.094	-	-	-	-	-	-	-	-	-	-
42° 0'30.52"N	MRLP	0.247	0.087	**0.000**	**0.000**	0.060	0.197	0.650	0.117	0.795	0.094	0.830	-	-	-	-	-	-	-	-	-
41°55'46.75"N	PPNP	0.247	0.087	**0.000**	**0.000**	0.060	0.197	0.650	0.117	0.795	0.094	0.830	1.000	-	-	-	-	-	-	-	-
41°43'56.37"N	FPPI	**0.044**	**0.009**	**0.000**	**0.000**	**0.005**	**0.045**	0.162	**0.017**	0.267	**0.012**	0.309	0.591	0.591	-	-	-	-	-	-	-
41°11'38.03"N	MPO	0.491	0.878	**0.025**	**0.000**	0.992	0.844	0.087	0.855	0.083	0.878	0.057	0.063	0.063	**0.005**	-	-	-	-	-	-
41° 6'11.07"N	IRO	0.189	0.060	**0.000**	**0.000**	**0.039**	0.159	0.591	0.085	0.759	0.066	0.815	0.997	0.997	0.583	**0.042**	-	-	-	-	-
40°41'51.85"N	CPS	0.326	0.087	**0.000**	**0.000**	0.057	0.267	0.856	0.138	0.997	0.104	0.966	0.795	0.795	0.253	0.060	0.759	-	-	-	-
40°28'58.36"N	STT	0.997	0.769	**0.019**	**0.000**	0.614	0.855	0.583	0.810	0.492	0.779	0.417	0.323	0.323	0.082	0.643	0.278	0.466	-	-	-
39°18'34.46"N	RRL	0.066	**0.012**	**0.000**	**0.000**	**0.006**	0.074	0.319	**0.026**	0.498	**0.017**	0.583	0.856	0.856	0.700	**0.006**	0.856	0.488	0.140	-	-
36°35'7.86"N	DHFA	0.411	0.822	**0.033**	**0.000**	0.989	0.806	0.066	0.807	0.063	0.822	**0.043**	0.053	0.053	**0.004**	0.953	**0.031**	**0.045**	0.583	**0.004**	-
34°35'46.40"N	HNL	0.087	**0.015**	**0.000**	**0.000**	**0.007**	0.095	0.467	**0.035**	0.675	**0.021**	0.779	0.997	0.997	0.481	**0.007**	0.997	0.667	0.197	0.822	**0.004**

Table 4.5 – Results of shell size data analyses for size data grouped by population clusters, determined by a STRUCTURE analysis from a separate population genomics study on *Webbhelix multilineata*. The top nested table shows data summary statistics for each of the four shell size measures (shell widths 1 & 2, shell height and shell whorls), including number of sample individuals (Count), average size measures (Mean), and standard deviation (SD) per cluster. The middle nested table shows pairwise comparisons of population cluster means for shell size measures using T-tests with pooled standard deviations. The bottom nested table shows Tukey multiple comparisons of population cluster means for shell sizes, including differences in means (Diff), lower and upper confidence levels (Lwr / Upr), and adjusted p-values (p Adj). Significant p-values are bold and highlighted in yellow.

Data Summary Statistics

Population Clusters	Shell Width 1			Shell Width 2			Shell Height			Shell Whorls		
	Count	Mean	SD	Count	Mean	SD	Count	Mean	SD	Count	Mean	SD
Cluster1	166	19.2	1.39	166	16.5	1.2	166	10.1	0.941	166	4.9	0.197
Cluster2	105	21.1	1.27	105	18.2	1.11	105	11.3	1.04	105	5.05	0.183
Cluster3	22	21.3	1.56	22	18.5	1.34	22	11	0.875	21	5.11	0.203
Cluster4	55	21.3	1.21	58	18.3	1.05	57	11.4	0.741	58	5.01	0.169

Pairwise Comparisons of Means using T-Tests with Pooled SD

	Shell Width 1			Shell Width 2			Shell Height			Shell Whorls		
	Cluster1	Cluster2	Cluster3	Cluster1	Cluster2	Cluster3	Cluster1	Cluster2	Cluster3	Cluster1	Cluster2	Cluster3
Cluster2	**0.000**	-	-	**0.000**	-	-	**0.000**	-	-	**0.000**	-	-
Cluster3	**0.000**	0.650	-	**0.000**	0.490	-	**0.000**	0.180	-	**0.000**	0.205	-
Cluster4	**0.000**	0.610	0.980	**0.000**	0.910	0.490	**0.000**	0.580	0.130	**0.000**	0.236	0.062

Tukey Multiple Comparisons of Means Tests

	Shell Width 1			Shell Width 2			Shell Height			Shell Whorls		
	Diff	Lwr / Upr	p Adj	Diff	Lwr / Upr	p Adj	Diff	Lwr / Upr	p Adj	Diff	Lwr / Upr	p Adj
Cluster2-Cluster1	1.898	1.466 / 2.330	**0.000**	1.785	1.412 / 2.158	**0.000**	1.196	0.893 / 1.498	**0.000**	0.149	0.088 / 0.210	**0.000**
Cluster3-Cluster1	2.091	1.305 / 2.876	**0.000**	2.047	1.368 / 2.726	**0.000**	0.879	0.328 / 1.429	**0.000**	0.211	0.098 / 0.324	**0.000**
Cluster4-Cluster1	2.083	1.544 / 2.621	**0.000**	1.806	1.350 / 2.262	**0.000**	1.280	0.908 / 1.653	**0.000**	0.113	0.038 / 0.187	**0.001**
Cluster3-Cluster2	0.192	-0.620 / 1.004	0.928	0.261	-0.440 / 0.963	0.771	-0.317	-0.886 / 0.251	0.475	0.062	-0.054 / 0.178	0.517
Cluster4-Cluster2	0.184	-0.392 / 0.761	0.843	0.021	-0.469 / 0.510	1.000	0.085	-0.314 / 0.484	0.947	-0.037	-0.116 / 0.043	0.636
Cluster4-Cluster3	-0.008	-0.882 / 0.865	1.000	-0.241	-0.990 / 0.509	0.841	0.402	-0.207 / 1.010	0.323	-0.099	-0.222 / 0.025	0.172

Sample Locality and Population Cluster Comparisons of Shell Brightness

The shell mean gray value data grouped by sample localities was non-normally distributed (W = 0.940, p = 6.32e-09) and statistically significant for differences between sample locality means via non-parametric Kruskal-Wallis rank sum test (df = 11, chi^2 = 40.09, p = 3.45e-05). All p-values for pairwise comparisons of sample localities using T-tests with pooled standard deviations can be viewed in **Table 4.6** for the mean gray value data. As with the T-test tables for the shell size data, sample localities in **Table 4.6** were arranged in order from north to south based on latitude. Notably, only two sample localities (RRL and OICP) had statistically significant differences of means with more than one other sample locality. The RRL locality had an average mean gray value (93.3) that was statistically significantly smaller than the average of the CRWA (102), FPPI (103), HGRF (103), IRO (105), and OICP (107) localities. The OICP locality had an average mean gray value that was statistically significantly larger than the average of the PPNP (97.3), RRO (97.4) and RRL localities. Smaller mean gray values equate to a darker image (and in the case of the present study, darker shells by proxy) and larger values equate to a lighter or brighter image (lighter shells).

The shell mean gray value data grouped by population clusters were also non-normally distributed (W = 0.957, p = 4.31e-07), and were weakly statistically significant for differences between population cluster means via non-parametric Kruskal-Wallis rank sum test (df = 3, chi^2 = 7.91, p = 0.048). **Table 4.7** shows data summary statistics, p-values for pairwise comparisons of group means using T-tests with pooled standard deviations, and p-values for Tukey multiple comparison tests of group means, for shell mean gray values grouped by population clusters. The Cluster 1 average mean gray value was statistically significantly larger than the average mean gray value of Cluster 2, meaning individuals from Cluster 1 on average were significantly brighter (lighter colored) than the average individual from Cluster 2.

Table 4.6 – Shell mean gray value pairwise comparisons of sample locality means using T-tests with pooled standard deviations. Statistically significant p-values ($p < 0.05$) are shown in bold and highlighted in yellow. Sample localities are arranged by latitude on the left, from the northern-most locality (CRWA) down to the southern-most locality (DHFA).

Latitude	Sample Localities	CRWA	PTP	HGRF	OICP	PCRA	RRO	PPNP	FPPI	IRO	CPS	RRL
44°22'5.00"N	CRWA	-	-	-	-	-	-	-	-	-	-	-
42°46'58.08"N	PTP	0.530	-	-	-	-	-	-	-	-	-	-
42°43'42.78"N	HGRF	0.953	0.484	-	-	-	-	-	-	-	-	-
42°34'55.32"N	OICP	0.258	0.079	0.385	-	-	-	-	-	-	-	-
42°12'51.69"N	PCRA	0.530	0.997	0.481	0.079	-	-	-	-	-		-
42° 0'38.99"N	RRO	0.213	0.560	0.179	**0.012**	0.560	-	-	-	-	-	-
41°55'46.75"N	PPNP	0.179	0.530	0.154	**0.012**	0.530	0.997	-	-	-	-	-
41°43'56.37"N	FPPI	0.980	0.528	0.987	0.346	0.524	0.213	0.179	-	-	-	-
41° 6'11.07"N	IRO	0.530	0.216	0.617	0.656	0.215	0.079	0.061	0.577	-	-	-
40°41'51.85"N	CPS	0.481	0.953	0.389	0.061	0.953	0.633	0.595	0.456	0.179	-	-
39°18'34.46"N	RRL	**0.012**	0.104	**0.012**	**0.000**	0.104	0.336	0.313	**0.012**	**0.003**	0.132	-
36°35'7.86"N	DHFA	0.530	0.997	0.481	0.079	0.997	0.560	0.530	0.524	0.213	0.953	0.104

Correlation Between Shell Size and Latitude

Webbhelix shell size measures plotted against latitude can be seen in **Figure 4.4**. Each scatter plot contains a trend line, 95% confidence intervals for slopes (orange dashed lines) and 95% prediction intervals (green dashed lines) from simple linear regression analyses performed on the data. Summaries of regression statistics and ANOVA tables can be seen in **Table 4.8** and **Table 4.9**, respectively. Analyses of variance for the regression models of all three plots indicated there is a highly statistically significant linear relationship between size measures and latitude of sample locality (shell width 1: df = 1, F = 57.77, p = 2.09e-13; shell width 2: df = 1, F = 60.25, p = 6.85e-14; shell height: df = 1, F = 56.32, p = 3.94e-13). The regression analyses indicated a moderately strong negative linear correlation between size variables and latitude (shell width 1: R = 0.354; shell width 2: R = 0.359; shell height: R = 0.349). Approximately 12% of the variance in size measurements can be accounted for by latitude (shell width 1: $R^2 = 0.125$; shell width 2: $R^2 = 0.129$; shell height: $R^2 = 0.122$). Plot trend line intercepts and slopes were all

Table 4.7 – Results of shell brightness (mean gray value) data analyses for the data grouped by population clusters determined via STRUCTURE analysis from a separate population genomics study on *Webbhelix multilineata*. The top nested table shows data summary statistics for shell mean gray value, including number of sample individuals (Count), average mean gray value measures (Mean), and standard deviation (SD) per cluster. The middle nested table shows pairwise comparisons of population cluster means for shell mean gray value using T-tests with pooled standard deviations. The bottom nested table shows Tukey multiple comparisons of population cluster means for shell mean gray value, including differences in means (Diff), lower and upper confidence levels (Lwr / Upr), and adjusted p-values (p Adj). All statistically significant p-values are shown in bold and highlighted in yellow.

Population Clusters	Data Summary Statistics Shell Mean Gray Value		
	Count	Mean	SD
Cluster1	90	103	12.3
Cluster2	108	98.6	8.66
Cluster3	47	99.6	9.61
Cluster4	22	99.7	6.03

Pairwise Comparisons of Means using T-Test with Pooled SD

	Cluster1	Cluster2	Cluster3
Cluster2	**0.010**	-	-
Cluster3	0.160	0.750	-
Cluster4	0.310	0.750	0.970

Tukey Multiple Comparisons of Means Test

	Diff	Lwr	Upr	p Adj
Cluster2-Cluster1	-4.535	-8.238	-0.831	**0.009**
Cluster3-Cluster1	-3.492	-8.162	1.177	0.217
Cluster4-Cluster1	-3.384	-9.556	2.787	0.489
Cluster3-Cluster2	1.042	-3.492	5.577	0.934
Cluster4-Cluster2	1.151	-4.919	7.220	0.961
Cluster4-Cluster3	0.108	-6.595	6.812	1.000

highly statistically significantly non-zero (shell width 1: intercept p = 0, slope p = 2.09e-13; shell width 2: intercept p = 0, slope p = 6.85e-14; shell height: intercept p = 0, slope p = 3.94e-13).

Correlation Between Shell Brightness and Latitude

Shell brightness data for *Webbhelix* shells is plotted against latitude in **Figure 4.5**. The scatter plot contains a trend line, 95% confidence intervals for the slope (orange dashed lines) and 95% prediction intervals (green dashed lines) from the simple linear regression analysis

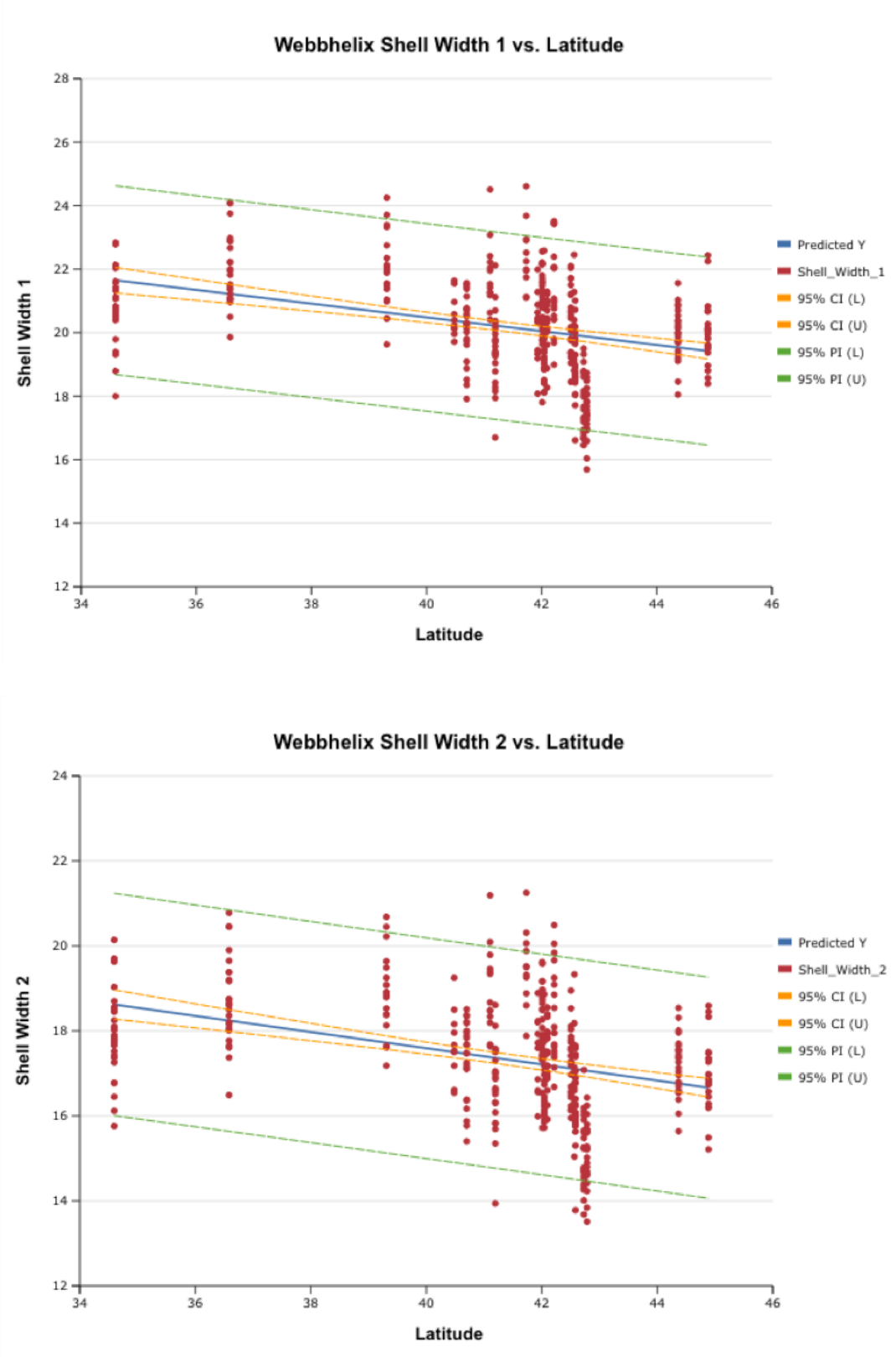


Figure 4.4 – *Webbhelix* shell size measures plotted against latitude (continued on next page)…

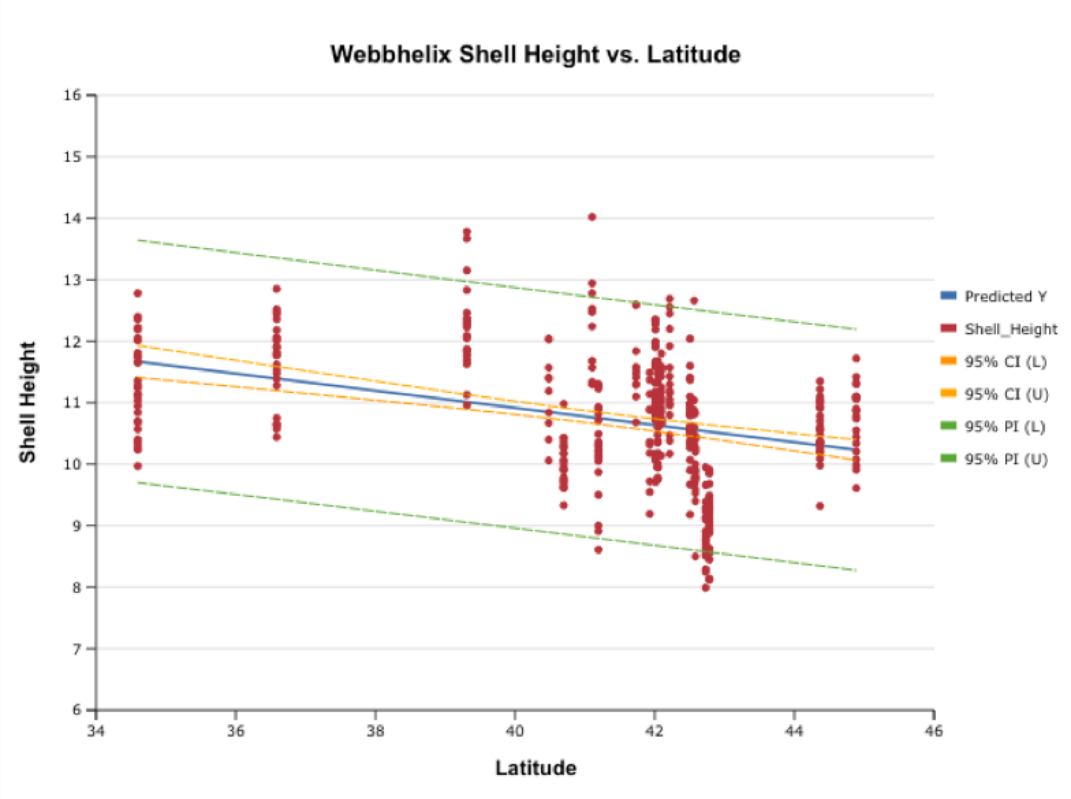


Figure 4.4 continued – *Webbhelix* shell size measures plotted against latitude. Plots contain trend lines for the data, 95% confidence intervals for slopes (orange dashed lines) and 95% prediction intervals (green dashed lines) from simple linear regression analyses. ANOVA analyses for the regression models of all three plots indicated there is a highly statistically significant linear relationship between size measures and latitude, and the regression analyses indicated a moderately strong negative linear correlation between size variables and latitude (Width 1: p = 2.1e-13, R = 0.35; Width 2: p = 6.8e-14, R = 0.36; Height: p = 3.9e-13, R = 0.35).

performed on the data. A summary of the regression statistics and ANOVA table can be seen in **Table 4.8** and **Table 4.9**, respectively. Analysis of variance for the regression model of the plot indicated there is a statistically significant linear relationship between measure of brightness (mean gray value) and latitude of sample locality (p = 0.020). The regression analysis indicated a weak positive linear correlation between the shell brightness variable and latitude (mean gray value: R = 0.142). Only 2% of the variance in the shell brightness measurement can be accounted for by latitude (mean gray value: R^2 = 0.020). Plot trend line intercept and slope were statistically significantly non-zero (mean gray value: intercept p = 4.16e-07, slope p = 0.020).

Table 4.8 – Summary table of regression statistics from all linear regression analyses performed on morphological data (shell size measures and shell brightness measure) that were plotted against latitude of sample localities to test ecogeographical hypotheses.

Morphologic Trait vs. Latitude	Regression Statistics					
	R	R-Squared	S	DW	AIC	BIC
Shell Width 1	0.35408	0.12537	1.49958	0.85867	3.65317	3.67294
Shell Width 2	0.35948	0.12922	1.31912	0.84858	3.3967	3.41636
Shell Height	0.34942	0.12209	0.99368	0.77917	2.83011	2.84981
Shell Brightness	0.14243	0.02029	10.09215	1.73054	7.46885	7.49573

Table 4.9 – Summary table of ANOVA statistics on all linear regression models performed on morphological data (shell size measures and shell brightness measure) that were plotted against latitude of sample localities to test ecogeographical hypotheses.

Morphologic Trait vs. Latitude		ANOVA Tables					
Shell Width 1		d.f.	SS	MS		F	p-value
	Regression	1	129.90	129.90		57.77	2.09E-13
	Residual	403	906.24	2.25			
		Coefficients	Std Err	LCL	UCL	t Stat	p-value
	Intercept	29.15	1.18	26.83	31.47	24.72	0
	Latitude	-0.22	0.029	-0.27	-0.16	-7.60	2.09E-13
Shell Width 2		d.f.	SS	MS		F	p-value
	Regression	1	104.84	104.84		60.25	6.85E-14
	Residual	406	706.47	1.74			
		Coefficients	Std Err	LCL	UCL	t Stat	p-value
	Intercept	25.20	1.01	23.21	27.19	24.90	0
	Latitude	-0.19	0.025	-0.24	-0.14	-7.76	6.85E-14
Shell Height		d.f.	SS	MS		F	p-value
	Regression	1	55.62	55.62		56.32	3.94E-13
	Residual	405	399.90	0.99			
		Coefficients	Std Err	LCL	UCL	t Stat	p-value
	Intercept	16.50	0.77	14.99	18.01	21.47	0
	Latitude	-0.14	0.019	-0.18	-0.10	-7.50	3.94E-13
Shell Brightness		d.f.	SS	MS		F	p-value
	Regression	1	558.86	558.86		5.49	0.020
	Residual	265	26990.63	101.85			
		Coefficients	Std Err	LCL	UCL	t Stat	p-value
	Intercept	69.20	13.33	42.95	95.44	5.19	4.16E-07
	Latitude	0.75	0.32	0.12	1.38	2.34	0.020

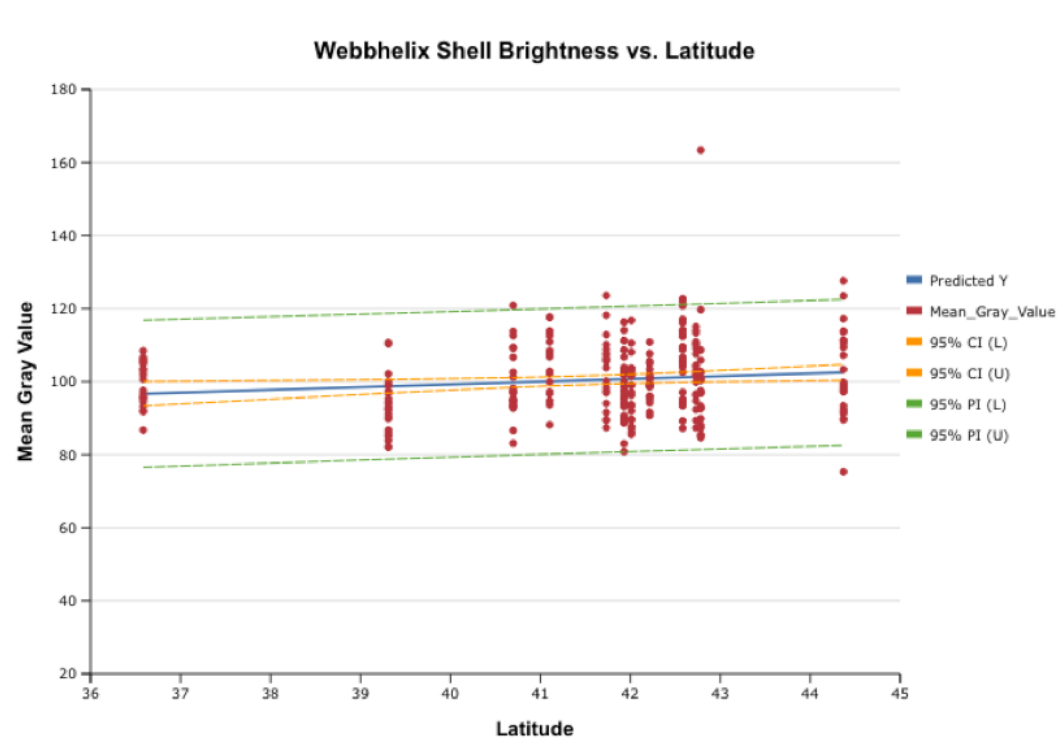


Figure 4.5 – *Webbhelix* shell brightness (mean gray value) plotted against latitude. Plot contains trend line for the data, 95% confidence interval for slope (orange dashed lines) and 95% prediction interval (green dashed lines) from a simple linear regression analysis. ANOVA analysis for the regression model of the plot indicated there is a statistically significant linear relationship between measure of brightness and latitude, and the regression analysis indicated a weak positive linear correlation between brightness and latitude (p = 0.02, R = 0.14).

Discussion

Locality and Population Comparisons for Shell Size and Brightness in *Webbhelix*

Statistical comparisons of sample locality means alone were not suggestive of any geographic pattern based on shell size measures or shell brightness. If the data were indicative of a strong north-south cline for shell size or brightness, then it would be reasonable to expect to see a pattern in the matrix distribution of significant p-values in **Tables 4.1 – 4.4** and **Table 4.6** that reflects it. Regardless of whether the shell size measures followed a Bergmann ecogeographic pattern or a reverse Bergmann's (Gloger's or reverse Gloger's for shell pigmentation), we should expect to see significant p-values concentrated toward the bottom-left corner of each table,

owing to the ordering of the sample localities by latitude. Pairs of sample localities with the smallest difference in latitude between them have p-values that are along or near the main diagonal of the matrix, while pairs with larger differences in latitude between them are further from the main diagonal. But for shell width and height measures the significant comparisons are distributed evenly throughout the matrices, with many of the significant p-values found along the main diagonal between sample localities that are very similar in latitude. This suggests that shell width and height in *Webbhelix* varies randomly among sample localities. In contrast, pairwise comparisons among sample localities for the shell whorls data showed far fewer significant comparisons across most locality pairs, suggesting that the number of shell whorls in *Webbhelix* does not vary much among localities sampled.

One notable pattern, however, did stand out across all four shell size measures with regard to two sample localities: Pikes Timber Park (PTP) and Horse Grove Rietz Forest (HGRF), both in Iowa along the Iowa River, with HGRF roughly five miles river distance downstream of PTP. Locality means for all size measures were statistically non-different between these two localities, but both were statistically significantly smaller than all other sampled localities. While this study did not investigate why individuals from these two localities were smaller than the rest, any number of environmental factors could be at play. For example at the HGRF site the east riverbank from which individuals were collected was a broad, wooded mudflat that appeared to be flooded regularly, and the population density was low enough that it took several hours of searching to find little more than a dozen individuals. The reduced average size could have been due to genetic bottlenecks from repeated flooding events, or it could have been an effect of low sampling relative to other localities.

While average shell sizes varied among sample localities, average shell brightness (as proxied by mean gray value of lateral shell transect images) was statistically non-different among most pairwise comparisons. Only two sample localities, Rip Rap Landing State Fish and Waterfowl Management Area (RRL) and Oakland-Iowa River County Park (OICP), had average mean gray values that were statistically significantly different from more than one other locality. These two localities were on opposite ends of the range of average brightness for the localities tested in this study, with RRL having the darkest individuals on average and OICP having the lightest or brightest individuals on average. Even so, individuals from the RRL locality were only statistically significantly darker on average than five other localities (CRWA, HGRF, OICP, FPPI and IRO) and individuals from OICP were only significantly lighter or brighter than three other localities (RRO, PPNP and RRL). To the question of whether the degree of shell banding (and thus shell brightness) in *W. multilineata* varies with respect to latitude, based on statistical comparisons of average mean gray values it would appear that it does not, since the overwhelming majority of localities sampled were statistically non-different from one another.

Assessment of the shell size data grouped by population clusters, determined from a previous population genomics study on *Webbhelix*, indicated that sampled individuals from Cluster 1 were on average statistically smaller than each of the other three population clusters across all four shell size measures. And this was the case for both shell widths and shell height even when the PTP and HGRF localities were removed from Cluster 1. The shell brightness data grouped by population clusters indicated, unsurprisingly, that shells from individuals in Cluster 2 were significantly darker than those from Cluster 1; this was due to the RRL locality (the darkest shelled locality) being included in Cluster 2, and the OICP locality (the lightest/brightest shelled

locality) being included in Cluster 1. Otherwise, shell brightness did not differ among the population clusters used in this study.

Correlation Between Latitude and Shell Size & Brightness in *Webbhelix*

Although pairwise statistical comparisons of sample locality means for shell size and brightness measures did not suggest a pattern of variation associated with latitude, they were shown to be statistically correlated via simple linear regression. The data for both shell width measures and for shell height all showed moderately strong negative correlations with latitude of sample localities, indicating that *Webbhelix multilineata* actually follows a reverse Bergmann ecogeographic cline. This means that more northerly localities tend to have smaller individuals and more southerly localities tend to have larger ones. While an explicit mechanistic explanation of this correlation was beyond the exploratory scope of the present study, we can at least begin to speculate as to why an organism like *Webbhelix* might show a reverse Bergmann cline for body size. In the northern parts of the species range, the encroaching onset of harsh, snow-packed winter conditions cause *Webbhelix* to seal the aperture of their shell with an epiphragm in late fall and hibernate through the winter, during which they are unable to feed and grow. So one explanation for the latitudinal size cline observed could be that populations in more southerly localities enter hibernation later and emerge from it sooner, since they experience milder or even snowless winters, and thus can spend a larger proportion of the year feeding and growing. Or perhaps the negative correlation between size and latitude could be due to latitude covarying with an environmental factor like calcium availability in soils, since snails use calcium to make and grow their shells. Follow-up studies are certainly needed to elucidate underlying mechanistic reasons for the ecogeographic pattern observed in these data, but also to fill in sampling gaps and re-assess the strength of the correlation signal. Sampling of southern localities is spares in the present dataset, both in terms of number of localities and in longitudinal representation,

compared to northern localities. There is also a significant northern sampling gap in the region encompassing northern Indiana, Ohio, and southern Michigan, which has historically been part of the species range. The paucity of southern localities plus the northern sampling gap, as well as low individual sampling for some of the presently sampled localities, could be leaving us with a misleading ecogeographic pattern, particularly when considering that only 12% of the variation in size measures could be attributed to latitude. Further research is needed to strengthen our understanding of the ecogeography of body size in this poorly studied terrestrial gastropod.

In contrast to the negative correlation found for shell size and latitude, regression analysis of the shell brightness data showed that mean gray value of lateral shell transect images correlated positively (albeit weakly) with latitude, suggesting *Webbhelix* adheres to the ecogeographic pattern described by the simplistic version of Gloger's Rule. Based on the present data, it appears that individuals tend to be slightly darker (in terms of shell pigmentation) at localities the further south they are and lighter the further north they are. However, this result comes with even heavier caveats than for the shell size data. On top of the sampling issues described above for the shell size datasets, the unfortunate impact of long-term storage of shells in ethanol (as described in the Methods section) left the extent of the shell brightness dataset at just over half that of the datasets for shell size (only 12 out of 21 sample localities used). Only 2% of the variation in mean gray value could be attributed to changes in latitude; while this could be reflective of the real correlation between shell brightness and latitude, we cannot be confident without filling in the sampling gaps and reassessing the correlation to see if the same pattern emerges. Nevertheless the results presented here provide a first look into how shell pigmentation varies geographically in *Webbhelix multilineata*, a low-mobility species with a broad range comprised of regions varying in seasonal extremes and pedology that is ripe for

further basic research. This study also details a relatively inexpensive method for generating color or brightness data for gastropod shells using basic tools and a free and easy to use software package.

The Contention Over Bergmann's Rule in Mollusks and Other Animal Groups

While each of the two hypotheses tested in this study were originally intended to describe ecogeographic patterns in endotherms, their popular use as a jumping off point for investigating such patterns in a broad range of animals, including those characterized as ectotherms, is apparent in the sheer number of studies that have utilized them over the past century (Avise, et al, 1987; Vinarski, 2014). By far most research studies on ecogeographic patterns have been done on Bergmann's Rule or similar variations related to body size, with numerous reviews published on studies within particular taxonomic groups. Therefore the remainder of this discussion will focus on some of the contentions surrounding the use of Bergmann's Rule, to what extent it has been found to apply to different animal groups, and highlighting some recommended best practices regarding reporting results when broadly applying Bergmann's and other ecogeographic hypotheses.

The main contention regarding the use of Bergmann's Rule to test for the ecogeographic pattern it describes is over whether or not the mechanism for the pattern ought also be tested. Watt et al (2010) argued that inherent to Bergmann's Rule (as put forward by Carl Bergmann in his original 1847 German text) is a mechanistic explanation for the proposed pattern that states larger body sizes are adaptive for cold environments because body surface area decreases relative to mass; therefore, larger bodies lose less heat relative to mass and are an energy conservation adaptation. They contended that studies claiming to test Bergmann's Rule by observing the pattern in a system without also explicitly testing the hypothesized mechanism for it are by definition not testing Bergmann's Rule. Along that same vein, Watt et al (2010) also

contended that Bergmann's Rule ought only be tested on endotherms, since the original definition was explicit about whom the rule applied to.

Conversely, Meiri (2011) contended that the pattern which Bergmann's Rule describes is on its own valuable enough to warrant use in exploratory studies on systems in which body size is a variable of interest, and that failing to experimentally verify the mechanism(s) of an observed pattern does not discount or invalidate it. They argue that ecological 'rules' such as Bergmann's are first and foremost descriptions of observable patterns; that something like a temperature-size relationship *is a pattern* and definitionally independent of mechanism, especially when one considers that the same pattern may arise due to different mechanisms under different circumstances. And so much of the tension surrounding the use of Bergmann's Rule (and other hypotheses like it) appears to stem from semantic disagreement. One thing that seems clear though in the debate on whether such 'rules' should be described in terms of pattern or pattern *and* mechanism, is that in practice many more researchers have leaned toward the former over the latter.

For over a century researchers have performed thousands of studies testing Bergmann's Rule across endotherms *and* ectotherms, with only a relatively few of them having attempted to evaluate a mechanism (the present study included; Watt et al, 2010). Vinarski (2014) did a meta-review summarizing the results of hundreds of these studies for different animal groups. They showed that according to different reviews, studies on endotherms have shown fairly consistently high rates of confirmation to the pattern described by Bergmann's Rule. In mammals, between 65% and 71% of study systems in the group conform to Bergmann's cline. In birds, it was between 72% and 76% of study systems. On the other hand, for ectothermic groups results have been less consistent. Conformation to Bergmann's within different reptile groups has ranged

from 26% to 83%; in amphibians results have been contradictory among different reviews, with Bergmann's being found to apply at rates ranging from 7% to 68% and some groups showing different ecogeographic patterns all together, like U-shaped clines. For fish and arthropods, fewer studies and reviews summarizing their results have been done. One review found fewer than 30% of fish species examined showed a Bergmann cline, while others have shown that insects sometimes exhibit Bergmann's cline, reverse Bergmann's cline, and U-shaped clines; but relative to the species richness of insects, these patterns are barely detectable given the number of studies available (Vinarski, 2014). Across terrestrial, marine and fresh water mollusks, reverse Bergmann's (decreased body size with increased latitude and decreased temperature) has actually been found to be common to the phylum, though both Bergmann's and reverse Bergmann's occur at almost the same rate.

Despite the lack of generalizability of an ecogeographic pattern such as Bergmann's Rule across most animal groups, it's clear that a broad diversity of different animal forms can end up exhibiting the pattern. And as Meiri (2011) alluded to, it seems unlikely that the diversity of species that exhibit this pattern do so as a result of the same underlying mechanism(s), but rather a variety of different ones or different combinations of mechanisms. With that in mind, the assertion by Meiri (2011) and others that ecogeographic rules ought to be described simply in terms of patterns, rather than necessarily in terms of patterns plus mechanisms, is the most straightforward and unconfused approach for communicating these concepts. Otherwise, every combination of temperature-size relationship plus underlying mechanism(s) would need to be given its own name, rather than all temperature-size relationships falling under the common umbrella of Bergmann's Rule. That said, communicators of ecological science could do a couple of things to help promote clarity and appease concerns over mechanism when discussing

ecogeographical rules: (1) instead of reporting that a study system conforms to or follows (for instance) Bergmann's Rule, it could be reported that a study system conforms to *the pattern described by Bergmann's Rule* or *exhibits a Bergmann cline*; (2) whenever no mechanism for observed ecogeographical patterns is tested, make an explicit statement that the mechanism(s) for an observed pattern remain as of yet unknown.

Conclusions and Future Directions

As is commonly found in terrestrial mollusks, *Webbhelix multilineata* appears to conform to the reverse pattern described by Bergmann's Rule and weakly follows the pattern described by Gloger's Rule. Individuals tend to be smaller in size and slightly lighter in color at more northerly latitudes, and larger and slightly darker at more southerly latitudes within the species range. Bergmann's and Gloger's Rule have provided useful jumping-off points for investigating observed morphological differences in the field, and the present study adds to the body of work that has been uncovering the prevalence of these patterns across diverse taxa for over a century now. Sampling across the species range of *Webbhelix* was extensive, and has provided insight into ecogeographic patterns present in this poorly studied terrestrial gastropod, but was not comprehensive and has left the door open for follow-up studies to reveal a clearer view of how size and color banding vary in this species. Future studies should focus on filling in the sampling gaps in the southern and northeastern parts of the range, reassessing the correlation between morphological traits and latitude, and developing tests for possible mechanisms underlying observed ecogeographic patterns.

References

Abramoff, M. D., Magalhaes, P. J., & Ram, S. J. (2004). Image Processing with ImageJ. *Biophotonics International*, *11*(7), 36-42.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual review of ecology and systematics*, *18*(1), 489-522.

Bergmann, C. (1847). Über die Ver- hältnisse der wärmeokönomie der Thiere zu ihrer Grösse. *Göttinger Studien*, *3*, 595-708.

Delhey, K. (2019). A review of Gloger's rule, an ecogeographical rule of colour: definitions, interpretations and evidence. *Biological Reviews*, *94*(4), 1294-1316.

Meiri, S., & Dayan, T. (2003). On the validity of Bergmann's rule. *Journal of biogeography*, *30*(3), 331-351.

Meiri, S. (2011). Bergmann's Rule–what's in a name?. *Global Ecology and Biogeography*, *20*(1), 203-207.

Monge-Nájera, J. (2008). Ecological Biogeography: A Review With Emphasis On Conservation and the Neutral Model. *Gayana*, *72*(1), 102-112.

Posadas, P., Crisci, J. V., & Katinas, L. (2006). Historical biogeography: A review of its basic concepts and critical issues. *Journal of Arid Environments*, *66*(3), 389-403.

Rasband, W. S. (1997-2018). ImageJ. *U. S. National Institutes of Health*, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/

Rensch, B. (1929). Das Prinzip geographischer Rassenkreise und das Problem der Artbildung. *Gebrueder Borntraeger*, Berlin.

Riemer, K., Guralnick, R. P., & White, E. P. (2018). No general relationship between mass and temperature in endothermic species. *Elife*, *7*, e27166.

Salewski, V., & Watt, C. (2017). Bergmann's rule: a biophysiological rule examined in birds. *Oikos*, *126*(2).

Schneider, C. A., Rasband, W. S., Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*, 671-675.

Vinarski, M. V. (2014). On the Applicability of Bergmann's Rule to Ectotherms: The State of the Art. *Biology Bulletin Reviews*, *4*(3), 232-242.

Watt, C., Mitchell, S., & Salewski, V. (2010). Bergmann's rule; a concept cluster?. *Oikos*, *119*(1), 89-100.

CHAPTER 5. GENERAL CONCLUSIONS

The dissertation presented here had three major aims to achieve: (1) test for a minimum sample size to accurately estimate basic population genomic parameters for the terrestrial gastropod *Webbhelix multilineata*; (2) generate a genomic dataset of thousands of SNPs from sampled localities of *Webbehlix* for population genomic assessment; (3) test variation in *Webbhelix* shell size and brightness within the context of the ecogeographic patterns described by Bergmann's Rule and Gloger's Rule. Addressing the first major aim revealed that as few as 6 individuals per sample locality could be used to accurately estimate observed and expected heterozygosity, inbreeding coefficient and fixation index between four sampled localities of *Webbhelix*. This result was relatively consistent with an assessment of minimum sample sizes for population genomics in an Amazonian plant system (8 individuals per locality; Nazareno et al, 2017), a species of whitefly (4 individuals per locality; Qu et al, 2020), and a species of lady beetle (6 individuals per locality; Li et al, 2020). The variability in minimum sample size appears to depend on the taxon of interest and the set of parameters of interest, as each study focused on different combinations of the two factors. But what seems to be clear at this point is that fewer than 10 individuals per population are needed to conduct population genomic studies, which positively impacts the financial feasibility of pursuing genomic methods on non-model organisms.

Insight into the minimum number of individuals needed to estimate population genomic parameters led here to the generation of the first genomic dataset for *Webbhelix multilineata*, and one of a relatively few current genomic datasets available for terrestrial gastropod mollusks. Assessment of population genomic structure of 18 sample localities identified four population clusters: two Midwestern clusters, one along and east of the upper Mississippi River (6

localities), one west of the upper Mississippi River (8 localities); a southern Mississippi River cluster (2 localities); and an eastern cluster on the shores of Lake Erie (2 localities). Genetic differentiation between the two Midwestern clusters was low, but as high between them and the southern Mississippi and Lake Erie clusters as is sometimes found between subspecies, indicating that *Webbhelix* could possibly be a species complex. A general pattern of decreasing genetic diversity from south to north along the Mississippi, and away from the northern Mississippi to the east and west, was observed and suggestive of serial founder effects indicative of post-glacial range expansion, likely along major river corridors. Coalescent simulations were also suggestive of post-glacial range expansion along major river corridors, and together with shells carbon dated to the LGM that have been found near the maximal extent of the Wisconsinan ice sheet (Nash et al, 2018) indicate the possibility of a more northerly refugium for *Webbhelix* than previously considered.

On the morphological side, *Webbhelix multilineata* shells from 21 sample localities across the species range showed a moderate negative correlation between size and latitude, conforming to a reverse Bergmann cline as has been commonly found in terrestrial mollusks (Vinarski, 2014). Additionally, shells from 12 sample localities showed a weak positive correlation between shell brightness (as a proxy for pigmentation) and latitude, suggesting that *Webbhelix* may conform to the ecogeographic pattern described by the simplistic version of Gloger's Rule. In more straightforward terms, *W. multilineata* individuals tend to be smaller in size and slightly brighter in color at more northerly localities, larger and slightly darker in more southerly localities. The observation of these biogeographical patterns in *Webbhelix* opens the door for further research to elucidate mechanistic reasons for why these traits may be distributed geographically as the present data suggests.

The studies that comprise this dissertation have sought to identify natural patterns associated with populations of the wide-ranging and morphologically charming striped white-lip snail. They have provided a look into the species' genomic structuring across much of the known range, afforded insight into how the species range expanded after the last ice age, and begun to show how variation in morphological traits is geographically distributed. Future studies on this system should focus on expanding the sampling distribution to more comprehensively encompass the known species range. Doing so would improve assessment of population genomic structure, and allow for more robust modeling of post-glacial refugial expansion. More comprehensive sampling would also improve analysis of correlations between morphological traits and environmental factors. Lastly, future studies should attempt to devise ways to experimentally test for mechanisms to explain ecogeographic patterns present in this terrestrial gastropod mollusk.

References

Li, H., Qu, W., Obrycki, J. J., Meng, L., Zhou, X., Chu, D., & Li, B. (2020). Optimizing Sample Size for Population Genomic Study in a Global Invasive Lady Beetle, Harmonia axyridis. *Insects*, *11*(5), 290.

Nash, T. A., Conroy, J. L., Grimley, D. A., Guenthner, W. R., & Curry, B. B. (2018). Episodic deposition of Illinois Valley Peoria silt in association with Lake Michigan Lobe fluctuations during the last glacial maximum. *Quaternary Research*, *89*(03).

Nazareno, A., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*.

Qu, W. M., Liang, N., Wu, Z. K., Zhao, Y. G., & Chu, D. (2020). Minimum sample sizes for invasion genomics: Empirical investigation in an invasive whitefly. *Ecology and evolution*, *10*(1), 38-49.

Vinarski, M. V. (2014). On the Applicability of Bergmann's Rule to Ectotherms: The State of the Art. *Biology Bulletin Reviews*, *4*(3), 232-242.