

**PATTERNS OF POPULATION STRUCTURE AND GENETIC DIVERSITY AMONG
WESTERN RATTLESNAKES (*CROTALUS OREGANUS*) IN THE PACIFIC
NORTHWEST**

by

Danielle Aimery Schmidt

B.Sc. [Hon], North Carolina State University, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

August 2019

© Danielle Aimery Schmidt, 2019

The following individuals certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis/dissertation entitled:

PATTERNS OF POPULATION STRUCTURE AND GENETIC DIVERSITY AMONG
WESTERN RATTLESNAKES (*CROTALUS OREGANUS*) IN THE PACIFIC NORTHWEST

submitted by Danielle Aimery Schmidt in partial fulfillment of the requirements of
the degree of Master of Science .

Dr. Michael Russello, Department of Biology, Irving K. Barber School of Arts and Sciences

Supervisor

Dr. Karl Larsen, Department of Natural Resource Sciences, Thompson Rivers University

Co-Supervisor

Dr. Purnima Govindarajulu, British Columbia Ministry of Environment and Climate Change
Strategy

Supervisory Committee Member

Dr. Mathieu Bourbonnais, Department of Earth, Environmental, and Geographic Sciences, Irving
K. Barber School of Arts and Sciences

University Examiner

Dr. Robert Lalonde, Department of Biology, Irving K. Barber School of Arts and Sciences

Additional Examiner

Abstract

Advances in DNA sequencing technologies have enabled the collection of genome-wide data, facilitating a shift from population genetics to population genomics. This transition has greatly augmented the amount of data that can be collected from DNA samples as well as increased the scope of ecological and conservation studies. Historically, non-invasive or minimally-invasive sampling (MIS) has been widespread in wildlife studies, especially for those species of elevated conservation concern. Despite the benefits of MIS, the utility of these samples for collecting genomic information is limited. Low DNA quantity, degradation, and contamination by exogenous DNA make MIS challenging to use with modern genotyping-by-sequencing approaches, which have traditionally been developed for high quality DNA sources. Here, we demonstrate the utility of Genotyping-in-Thousands by sequencing (GT-seq), a targeted, multiplex amplicon approach, for genotyping minimally-invasive DNA samples from the Western Rattlesnake (*Crotalus oreganus*), a species-at-risk in British Columbia, Canada. Using a panel of 362 single nucleotide polymorphisms (SNPs) generated from high quality blood samples identified via restriction site-associated DNA sequencing (RADseq), we genotyped a total of 96 blood (n=68), minimally-invasive cloacal swab (n=9), and opportunistically collected roadkill tissue samples (n=19) across the distribution of Western Rattlesnakes in BC and Washington (USA). The targeted GT-seq panel yielded comparable estimates of within- and among-population variation to that of the larger RADseq dataset (n=9568 SNPs). We then applied this approach to a set of blood (n=68) and minimally-invasive samples (n =690) collected from across species the distribution. Hierarchical STRUCTURE analyses found evidence for population structure within and among all five geographic regions of *C. oreganus* occurrence in BC and Washington, as well as low levels of migration among groups of den sites across the distribution, with no evidence for a pattern of isolation-by-distance. These results provide evidence for population discreteness, as well as barriers to gene flow across the distribution suggesting that the single, recognized designatable unit for conservation for Western Rattlesnakes in BC should be re-assessed. More broadly, this thesis demonstrates the efficacy of combining genomic tools with MIS to investigate ecological and evolutionary processes impacting wild populations and better inform conservation management.

Lay Summary

Minimally invasive sampling (MIS) is widespread in wildlife studies involving species of conservation concern. This type of sample collection is advantageous as it limits the need to directly handle or interact with a species of interest. However, DNA from MIS is often of low quantity and poor quality. The potentially degraded nature of these samples and contamination with DNA from other sources, such as bacteria, make obtaining genetic data challenging with traditional methods. Here, we adapt and apply a targeted genetic approach to gather data from minimally invasive DNA samples collected from the Western Rattlesnake, a species at-risk in British Columbia, Canada. Through evaluating the structure and connectivity of rattlesnake populations in BC, we detected genetic differentiation that suggests that the designation of a single conservation unit for this species warrants re-examination. We also demonstrate the utility of combining new DNA sequencing technologies with MIS to inform conservation management.

Preface

Multiple individuals have contributed to Chapters 2 and 3 of this thesis. The study design for Chapter 2 is a shared effort between myself and Dr. Michael Russello, along with Nathan Campbell who conducted primer design and assisted with primer optimization. Chapter 3 also shares study design between myself and Dr. Michael Russello, as well as Dr. Purnima Govindarajulu and Dr. Karl Larsen. For both chapters I was primarily responsible for all lab work, data collection, data analysis and manuscript preparation.

A manuscript version of the methods and results presented in Chapter 2 has been submitted to *Molecular Ecology Resources*, with Nathan Campbell, Dr. Purnima Govindarajulu, Dr. Karl Larsen, and Dr. Michael Russello as co-authors for publication.

For both chapters, sample collection was a combined effort among many individuals and organizations. In BC samples were collected by, Jeff Brown, Kyle Horner, Owain M. McKibbin, Dr. Christine Bishop, Orville Dyer, Jared Hobbs, Lindsay Anderson, Karl Larsen, Mike Dunn, Stephanie Winton, and Jared Maida, and myself. All sample collection took place under Thompson Rivers University Animal Ethics File No. 101547, BC Ministry of Environment Park Use Permit No. 108794, and British Columbia Wildlife Act (PE15- 171661). Samples from Washington State were provided by Daniel Beck (University of Central Washington) and John Rohrer (US Forest Service).

Table of Contents

Abstract.....	iii
Lay Summary	iv
Preface.....	v
Table of Contents	vi
List of Tables	ix
List of Figures.....	xi
Acknowledgements	xiii
Dedication	xv
Chapter 1: Introduction	1
1.1 Habitat Fragmentation	1
1.1.1 Impacts of Habitat Fragmentation on Reptiles	2
1.1.2 Influences of Roads on Reptile Behavior and Genetics.....	2
1.2 Defining Units for Species Conservation	3
1.3 Rattlesnakes in British Columbia.....	5
1.3.1 Current Species Management	6
1.4 Study Objectives	7
Chapter 2: Genotyping-in-Thousands by sequencing (GT-seq) panel development and application to minimally-invasive DNA samples to support studies in molecular ecology	9
2.1 Background.....	9
2.2 Methods.....	12
2.2.1 Sample Collection.....	12
2.2.2 DNA Extraction	12
2.2.3 Restriction Site-Associated DNA Sequencing (RADseq)	13
2.2.4 SNP Discovery.....	13
2.2.5 SNP Filtering and Panel Selection.....	14
2.2.6 GT-seq Primer Design	15
2.2.7 GT-seq Test Library Preparation	16
2.2.8 GT-seq Genotyping and Primer Pool Optimization.....	17
2.2.9 Genotype Error/Concordance	18
2.2.10 Population Diversity Estimates.....	18

2.3	Results.....	19
2.3.1	RADseq, SNP Discovery and Panel Selection	19
2.3.2	GT-seq SNP Genotyping	19
2.3.3	Genotype Concordance Between RADseq and GT-seq	20
2.3.4	Population Diversity Estimates and Individual Ancestry Estimation.....	20
2.4	Discussion.....	20
Chapter 3: Genotyping-in-Thousands by sequencing reveals marked population structure in Western Rattlesnakes to inform conservation status		29
3.1	Background.....	29
3.2	Methods.....	31
3.2.1	Study area/ sample collection and distribution	31
3.2.2	DNA Extraction	32
3.2.3	GT-Seq Library Preparation	32
3.2.4	GT-seq Genotyping and Genotyping Error.....	33
3.2.5	Regional Population Differentiation and Connectivity.....	33
3.2.6	Fine Scale Population Structure and Connectivity	34
3.2.7	Genetic Diversity Within and Among Dens	35
3.2.8	Sex Bias in Dispersal	36
3.3	Results.....	37
3.3.1	GT-seq Genotyping.....	37
3.3.2	Inferred Regional Population Structure and Connectivity.....	37
3.3.3	Within-region Population Structure and Connectivity	38
3.3.4	Den Level Genetic Diversity	39
3.3.5	Sex Bias in Dispersal	40
3.4	Discussion.....	40
3.4.1	Population Structure and Connectivity	40
3.4.2	Genetic Diversity	44
3.4.3	Sex Bias in Dispersal	45
3.4.4	Implications for the Conservation of Snakes	47
Chapter 4: Conclusions		61
4.1	Findings and Significance.....	61

4.1 Future Directions	62
References	65
Appendices	79
Appendix A	79
Chapter 2 Supplementary Material	79
Appendix B	84
Chapter 3 Supplementary Material	84

List of Tables

Table 2.1. Population genetic diversity estimates within and among sampled populations across methods and dataset sizes, including estimates of mean effective number of alleles (Eff_{Ar}), mean observed heterozygosity (H_o), mean expected heterozygosity within populations (H_s), and mean inbreeding coefficient (G_{is}).	27
Table 2.2. Pairwise estimates of θ (Weir & Cockerham 1984) for all four sampled populations of Western Rattlesnakes across all three data sets (GT-seq_311, RAD_311, RAD_9568). All pairwise comparisons were significant ($p < 0.05$) with 2000 permutations. Population acronyms as in Figure 1.....	28
Table 3.1 Estimates of pairwise genetic distance (F_{ST}) between all 6 geographic regions across the distribution. All comparisons were significant ($p < 0.001$) with 1000 replicates.....	53
Table 3.2 Estimates of pairwise F_{ST} and straight-line distance (km) between the 11 defined den groups. Values above the diagonal represent F_{ST} and those below are straight-line distances. All pairwise comparisons of F_{ST} were significant $p < 0.001$ with 1000 replicates.....	54
Table 3.3 Estimates of contemporary migration among den groups as calculated in BAYESASS. Numbers represent the mean proportion of migrants moving from one group to another per generation averaged across 5 iterations. Italicized values along the diagonal indicate the mean proportion of non-migrants within a den group. Bolded values indicate significant estimates of migration where calculated 95% credible sets did not include zero.....	55
Table 3.4 Estimates of Queller and Goodnight's pairwise relatedness (r_{QG}) between pairs of individuals within each den group and within each den.....	56
Table 3.5 Estimates of mean observed and expected heterozygosity, proportion of polymorphism (P), nucleotide diversity (N_g), and inbreeding coefficients (F_{IS}) for each den group and dens within each den group. * indicates $p < 0.05$, ** indicates $p < 0.001$. Bolded values indicate a significant excess of heterozygosity.....	58
Table 3.6 Comparison of $mAIC$ and $vAIC$ for adult male and female snakes among all den groups and within den groups containing at least 2 den sites with at least 2 males and females at each. * denotes $p < 0.05$ with 10,000 replicates.....	60
Table S2.1. The number of individual samples that were collected within each region by sample type (region acronyms as in Figure 1).	81

Table S2.2. The average, minimum, and maximum DNA concentrations per sample type employed in this study across all sampled regions. 82

Table S2.3. STRUCTURE HARVESTER output for all K values tested ($K=1-7$) for the RADseq_9568, GT-seq_311, and RADseq_311 datasets. The posterior log-probability for each K (Mean $\text{LnP}(K)$), standard deviation in $\text{LnP}(K)$ (Stdev $\text{LnP}(K)$), as well as the first order rate of change $\text{Ln}'(K)$ and absolute value of the second order rate of change $|\text{Ln}''(K)|$ used to estimate ΔK are presented. Bolded rows represent the optimal K value selected for each dataset ($K=3$) where a plateau in the log probability across all K values was identified, variation was low, and where ΔK was maximized..... 83

Table S3.1 Counts of individual samples used in broad scale population genetic analyses ($n=588$) divided by region and sample type. Sample types include cloacal swab samples (CL), blood samples (BL), roadkill tissue samples (T), buccal swab samples (BU), and shed skins (SH). 86

Table S3.2 Results of AMOVAs across various groupings. * indicates $p<0.05$, while ** denotes $p<0.01$ after 10000 permutations. 87

List of Figures

Figure 1.1 A) Map of the distribution of *Crotalus oreganus* in Canada, the United States, and Mexico shown in green (spatial data from NatureServe and IUCN (International Union for the Conservation of Nature), 2007) B) The 5 main geographic regions of *C. oreganus* occurrence in British Columbia, shown in red (spatial data provided by O. Dyer, British Columbia Ministry of Environment and Climate Change Strategy). 8

Figure 2.1. Western Rattlesnake species range in British Columbia, Canada (shaded) and Washington, USA (approximate; dotted line), including sample distribution from sites in: Thompson-Nicola (KAM; n = 14), Vernon (VER; n = 30), Okanagan-Similkameen (SOK; n = 40), Grand Forks (GF; n = 4), and Washington (WA; n = 16). Within region sample locations are represented by circles with the various colors indicating sample type noted in legend. 24

Figure 2.2. Library designs for A) RADseq and B) GT-seq. Included samples selected to facilitate within and among method genotype comparisons..... 25

Figure 2.3. STRUCTURE barplots displaying inferred clustering and individual ancestry estimation of Western Rattlesnakes (n = 41) for all 3 datasets (GT-seq, RADseq_311, RADseq_9568) across all sampled regions (population acronyms as in Figure 1). Each different color in the plots represents a distinct genetic cluster and each vertical bar represents the proportion of ancestry of a single individual to the different genetic clusters..... 26

Figure 3.1 The distribution of den site sampling locations of Western Rattlesnakes in BC across the five main geographic regions: Thompson-Nicola (KAM), Vernon (VER), Okanagan-Similkameen (SOK), Midway (RC), Grand Forks (GF) and in Washington, USA (WA). Red dotted circles encompass each defined geographic region, while solid black circles designate each of the 11 den groups defined within this study. Coloured points indicate the 36 den sites sampled across den groups: KAM-W (n=2), KAM-C (n=3), KAM-E (n=1), VER-CB (n=2), VER-KL (n=4), WL (n=6), OS (n=3), RC (n=1), GF (n=3), WA-N (n=7), WA-S (n=4). 49

Figure 3.2 STRUCTURE bar plot displaying inferred clustering and ancestry estimation for individuals assigned to dens (n=461), showing K=3 for iteration 1, K=3 for iteration 2, and K=3, 2, and 2, respectively for populations across iteration 3 (population acronyms as defined in Figure 3.1). Each color represents a unique genetic cluster and each vertical bar represents the proportion of ancestry to each cluster for each individual..... 50

Figure 3.3 Mantel’s correlation between straight-line distance (km) and genetic distance (F_{ST}) for all 6 geographic regions ($r = -0.1863916$, $p= 0.6588$) (A) and all 11 den groups (B) ($r = -0.03812$, $p= 0.4989$).....	51
.....	52
Figure 3.4 A) Correlation between latitude and gene diversity (N_g) ($r^2= 0.365$, $p= 0.03$), and latitude and proportion of polymorphism (P) ($r^2= 0.196$, $p= 0.1$) across den groups. B) Correlation between latitude and gene diversity (N_g) ($r^2= 0.311$, $p<0.001$), and latitude and proportion of polymorphism (P) ($r^2= 0.026$, $p= 0.75$) across individual dens.....	52
Figure S2.1. Scatterplot displaying the relationship between percent amplification across 311 loci for samples remaining after quality filtering ($n=103$) and initial DNA sample concentration in ng/ul. Samples with concentrations larger than 30 ng/uL were diluted to 15 ng/ul to avoid preferential amplification of samples with relatively high concentrations.	79
Figure S2.2. Boxplot displaying the mean and range of percent amplification for all samples remaining after quality filtering ($n=103$) across 311 loci for all sample types including: blood (BL; $n=61$), cloacal swab (CL; $n=26$), shed (SH; $n=2$), and roadkill tissue (T; $n=14$).	80
Figure S3.1 Barplot of STRUCTURE results for $K=3$ for both iterations of all samples ($n=588$) regardless of assignment to den.	84
Figure S3.2 Boxplots displaying corrected assignment probabilities for Male and Female adult snakes in the VER-KL den group. The variance of assignment probabilities ($vAIC$) was significantly higher ($p<0.05$) for females than males as evidenced by the range of values for each sex in this plot.	85

Acknowledgements

The completion of this thesis truly would not have been possible without the help and support of so many people. Primarily, I would like to thank my supervisor Dr. Michael Russello for his constant guidance and patience throughout the entirety of this project. I would also like to thank my co-supervisor Dr. Karl Larsen for providing me with the chance to get out into the field and for his insight and endless enthusiasm. Likewise, I am tremendously thankful for the encouragement and feedback provided throughout by my committee member Dr. Purnima Govindarajulu.

Given the hundreds of samples used in this study, I am especially grateful to everyone that has collected or supplied me with DNA samples, including Karl Larsen, Orville Dyer, Jeff Brown, Kyle Horner, Owain M. McKibbin, Dr. Christine Bishop, Jared Hobbs, Lindsay Anderson, Mike Dunn, John Rohrer, Daniel Beck, Jared Maida, Stephanie Winton, Dana Eye, Cole Hooper, Marcus Atkins, Robyn Reudink, JoAnne Hales Brodie (BC Parks), and Shawn Klassen. Special thanks to Stephanie Winton and Jared Maida for being amazing guides out in the field and for teaching me how to collect samples, and to Luke, Brett, Kristine, Bryson, and Kirstie for all your field assistance.

I must also extend my gratitude to Nate Campbell at GTSeek for helping me adapt GT-seq for use in my thesis work. His primer design expertise and generous advice have been essential to the success of this project. I also would like to thank Dr. Karen Hodges for her wisdom, advice, and encouragement throughout my degree.

Additionally, I would like to thank the other members of the ECGL, both past and present, for being absolutely wonderful colleagues and friends. Thank you to Bryson and Luke for (sometimes) sharing an office with me; I truly appreciate you both allowing me to pester you with questions and bioinformatic problems. Thanks to Sarah, Farida, and Brock for being great friends and sources of support.

Last, but not least, I'd like to thank my friends and family for their overwhelming support throughout this adventure. Especially my parents, I could not have made it to this point without you. Thank you for always encouraging me to pursue my ambitions, wherever they may take me.

Funding for this study was provided by the Habitat Conservation Trust Fund (PG, KL, MR), and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant # 341711 (MR).

Dedication

For JB.

Chapter 1: Introduction

Biological diversity, often referred to as biodiversity, includes the total variation within and among living organisms on Earth that make up the myriad of ecosystems that exist (Swingland, 2001). Preserving this diversity is necessary to maintain natural processes, such as nutrient cycles, as well as the ecosystem services provided by them. Humans rely upon natural resources as sources of food, water, fuels, textiles, and even sources of recreation. The loss of even a single species can impact not only ecosystem functioning, but can have severe socioeconomic consequences on society as a whole (Díaz et al., 2006). Although the exacerbation of global biodiversity loss is not solely anthropogenic, human impacts have played a significant role in species decline (Ceballos et al., 2015). Main anthropogenic threats to biodiversity include habitat degradation, environmental pollution, climate change, invasive species, and overexploitation.

While each of the aforementioned threats has a negative impact on biodiversity, habitat loss and degradation pose one of the biggest threats to the persistence of species (Magin et al., 1994; Purvis et al., 2000). The main drivers of habitat loss include harvesting of natural resources, development of infrastructure, and land conversion for agriculture. Over time, continual degradation leads to habitat fragmentation, the process by which continuous habitat is broken down into a heterogeneous patchwork of habitats (Lehmkuhl & Ruggiero, 1991). As a result, patches of suitable habitat often are isolated across a landscape matrix (Todd et al., 2010).

1.1 Habitat Fragmentation

Habitat fragmentation is responsible for losses of up to 75% of biodiversity in some areas (Haddad et al., 2015). This process alters both the spatial structure of habitats and the metapopulation dynamics of species that exist among them. Fragmentation also influences the spatial distribution of species and therefore, the functioning of ecosystems. The interaction of subpopulations of a species largely depends on habitat patch size, shape, and the distance between neighboring patches (reviewed in Fahrig & Merriam, 1994). The more isolated patches are, the more difficult it becomes for species to function as if they still existed in once continuous habitat. The level of connectivity among habitat patches across a fragmented landscape can have quite severe implications for population genetic structure and species persistence (Fahrig & Merriam, 1985; Dixon et al., 2007).

When populations across a landscape become isolated and separated, individuals often cannot move among patches of suitable habitat due to limited connectivity. In turn, this alters patterns of individual dispersal, and therefore patterns of gene flow among metapopulations (Gerlach & Musolf, 2000). Gene flow is necessary for maintaining adequate levels of genetic diversity within populations; without proper genetic diversity, populations become unable to adapt to their environment, which puts them at risk of local extinction (Frankham, 2003). Over time, the extirpation of multiple populations of a species can contribute to overall species decline.

1.1.1 Impacts of Habitat Fragmentation on Reptiles

Reptiles and amphibians, collectively known as herpetofauna, are particularly vulnerable to the effects of habitat loss and fragmentation. These species often are sensitive to environmental changes (Gibbons et al., 2000), as they are ectothermic and rely on their environment to regulate their body temperatures. Herpetofauna also often exhibit low vagility, moving infrequently, and over relatively short distances. Habitat fragmentation is considered one of the most significant threats to herpetofauna worldwide (Lesbarrères et al., 2014; Mittermeier et al., 1992). As of 2019, 1311 of all assessed reptile species were considered threatened (IUCN, 2019).

In particular, habitat fragmentation has been shown to reduce the persistence of reptile species by 90 percent in areas of previously occupied habitat (Driscoll, 2004). It has been suggested that undisturbed habitat supports over 2 times more reptiles than disturbed habitat (Brown et al., 2008). Previous studies suggest that the abundance of reptiles also is significantly impacted by the isolation of habitat fragments (Luiselli & Capizzi, 1997; Watling & Donnelly, 2007) as well as the size of fragments. In the Western Australia wheatbelt, Sarre (1998) found a significant correlation between habitat area and population size of the Tree Dtella gecko (*Gehyra variegata*). At a regional scale, land conversion for agriculture also was found to have the most impact on reptile diversity deficit in Catalonia (Ribeiro et al., 2009).

1.1.2 Influences of Roads on Reptile Behavior and Genetics

One of the main mechanisms of habitat fragmentation affecting reptile species is the development of roads. Road fragmentation has been shown to have both behavioural and genetic effects on populations, even when roads are up to 2 km away from core populations (Beebee,

2013). As demonstrated by both the Eastern Massasauga rattlesnake (*Sistrurus catenatus*) and the Red-Sided Garter snake (*Thamnophis sirtalis parietalis*), reptiles may exhibit road avoidance behaviors, where instead of crossing roads, animals move parallel alongside them (Andrews & Gibbons, 2005; Clark et al., 2010; Shepard et al., 2008; Shine et al., 2004). This avoidance behavior can greatly limit animal dispersal and prohibit gene flow between populations, creating an increased risk for inbreeding depression and population decline (Shepard et al., 2008). Roads also have been shown to increase the chance of mortality for crossing animals. Those reptiles that do cross roads, usually cross for thermoregulatory purposes, as road surfaces are often warmer than the surrounding vegetation (Shine et al., 2004). This not only puts individuals in direct danger of being hit by vehicles, but also increases their exposure to predation. In particular, those species that have delayed sexual maturity are often vulnerable to adult road mortality as it takes longer to replace adults lost from a population (Brooks et al., 1991). As a result, populations can become increasingly susceptible to decline (Row et al., 2007).

In addition to limited dispersal and mortality, gene flow and population connectivity between reptile populations also are reduced due to roads. Several studies show how roadways act as a barrier to gene flow and create population genetic structure among populations. In New York state, a population genetic analysis of Timber rattlesnakes (*Crotalus horridus*) found significant levels of genetic differentiation and structure both among snake hibernacula in a region and between regions that were separated by roads. These patterns of population structure were directly attributed to the disruption of snake dispersal and vehicular mortality (Clark et al., 2010). Similarly, genetic differentiation in subpopulations of Desert Tortoises (*Gopherus agassizii*) was significantly associated with roads in California (Latch et al., 2011). Overall, it is clear that the various anthropogenic causes of habitat fragmentation have substantial influences on the behavior, genetic structure, population connectivity, and persistence of reptile species.

1.2 Defining Units for Species Conservation

Before conservation management decisions can be made for species at risk of extinction, units for conservation must be defined. Without a clear idea of what is being conserved, targeted legislation and policy protections cannot be enforced (Funk et al., 2012; Green, 2005). This can be especially important in cases where multiple populations of a species vary in their viability and status. As conservation funds often are limited, this information can be used to prioritize conservation units deemed critical to overall species persistence (Ando et al., 1998). By defining

units targeted for conservation, conservationists can protect and maintain the evolutionary potential of species (Moritz, 1994).

Although there currently are no set criteria for defining conservation units, several concepts have been proposed and adopted. First proposed by Ryder (1986), Evolutionary Significant Units (ESUs) refer to groups of organisms below the species designation that warrant specific conservation attention, namely in terms of significant genetic characteristics. Over time, this concept has been revisited multiple times (Moritz, 1994; Waples, 1991), each suggesting various considerations for delimiting units. Waples (1991) defined ESUs based upon groups of a species that are a) reproductively isolated and b) considered an important component to the evolutionary trajectory of that species based upon historical patterns of genetic variation, while Moritz (1994) focused upon measures of significant divergence in allele frequencies and reciprocal monophyly in mitochondrial DNA among populations of a species. Further, Moritz (1994) distinguished between ESUs and management units (MUs). While ESUs tend to examine more historical and phylogeographic divergences in allele frequencies of populations, MUs consider patterns of genetic divergence that rely mostly on contemporary genetic differentiation between populations. This allows for the existence of multiple MUs within ESUs that focus upon short-term conservation goals, nested within ESUs that have a more long-term focus (Moritz, 1994). Both ESUs and MUs are equally important for conservation, as the effective conservation of MUs is essential for the persistence of ESUs (Paquette et al., 2007). These conservation units often align with the United States Endangered Species Act (ESA) (Crandall et al., 2000; Waples, 1991), as well as the Endangered Species Protection Act in Australia.

In Canada, Green (2005) proposed designatable units (DUs) that were adopted by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) as a way to prioritize units for conservation. Specifically, this designation addresses the need for biologically-based units that account for species conservation status, especially when one species distinction is not applicable for all populations within a species (Green, 2005). In light of the various approaches used for establishing ESUs, a two-facet system exists for discerning putative DUs that aligns well with the process of ecological risk assessments used to determine conservation status. This ensures that the establishment of units for conservation coincides with conservation concerns. Much like evolutionarily significant units, the two facets of designatable units are 1) population discreteness and 2) the evolutionary significance of a given population. These units coincide with

the Species at Risk Act (SARA) that was implemented in Canada as part of the Canadian Biodiversity Strategy (Environment and Climate Change Canada, 2009).

To determine the discreteness and evolutionary significance of populations, studies can assess the level of genetic differentiation and gene flow between populations and infer patterns of population connectivity, as well as adaptive divergence (reviewed in Allendorf et al., 2010). For instance, populations of landbird species including the Northern Saw-whet Owl (*Aegolius acadicus*), Hairy Woodpecker (*Picoides villosus*), Steller's Jay (*Cyanocitta stelleri*), Chestnut-backed Chickadees (*Poecile rufescens*) and Pine Grosbeaks (*Pinicola enucleator*) in the Queen Charlotte Islands (QCI) were found to exhibit significant genetic differentiation from other populations of these species in Alaska and Washington-Oregon. The observed genetic differences supported the existence of previously hypothesized endemic subspecies of these birds in QCI that were based upon phenotypic differences. Therefore, each of the four subspecies found in QCI called for separate DUs and ESUs (Topp & Winker, 2008). In combination with ecological, environmental and phenotypic information, genetic data can be further analyzed to look for evidence of adaptation at loci within the genome, which can give further support for the delimitation of evolutionarily significant units (Allendorf et al., 2010; Funk et al., 2012).

1.3 Rattlesnakes in British Columbia

The Western Rattlesnake (*Crotalus oreganus*) is a venomous, squamate reptile that occurs from Southern Canada to California and northern Mexico (Fig 1.1a). This snake is one of 4 species of rattlesnake native to Canada (one now being extirpated); in BC, it is represented by the Northern Pacific subspecies (*C.o. oreganus*). The range of Western Rattlesnakes in Canada is restricted to south-central British Columbia, where it occurs in five so-called regions, namely the Okanagan-Similkameen, Midway, Grand Forks, Vernon, and Thompson-Nicola. Given the mountainous terrain of BC, climate and ecotypes change abruptly: the range of this snake in the province is restricted to warmer valleys, with fragmentation of habitats occurring naturally as well as through anthropogenic factors (COSEWIC, 2015) (Figure 1.1b).

Like many other snake species existing at higher latitudes, *C. oreganus* overwinters in communal dens from October to late March/early April. Individuals show high site fidelity, returning to overwinter at the same den each year (Macartney, 1985). Dens are located on south-facing talus slopes, where there is greater availability of basking areas for thermoregulation (Gienger & Beck, 2011). During the active season between April and October, snakes emerge

from their dens to seek food and reproduce. Some snakes have been found to migrate upwards of 4 km (Gomez, 2008). These rattlesnakes are specialized rodent predators that use facial organs to sense heat to locate prey. Their typical diet includes pocket mice (*Perognathus parvus*), deer mice (*Peromyscus maniculatus*), and other small mammals (Macartney, 1989; McAllister et al., 2016).

Western Rattlesnakes reach sexual maturity around 7-8 years old (Macartney et al., 1990). Female snakes will give birth to live pups during the active season, usually at rookery sites located within 300-500 m of the dens. Gravid females have been found to abstain from hunting while pregnant, losing almost all their body fat to developing young. After parturition, non-gravid females may not mate for one or more years; at which time they begin the process of vitellogenesis (egg yolk production) when accumulated fat stores are enough to sustain another pregnancy (Macartney & Gregory, 1988). Therefore, females will only give birth every 2 to 3 years. This creates a lengthy generation time of about 15 years for this species (COSEWIC, 2015).

Currently, *C. oreganus* is of conservation concern in British Columbia, which encompasses the northernmost extent of this species' range. The Western Rattlesnake is provincially protected under the BC Wildlife Act and classified as 'threatened' by the federal Species at Risk Act (COSEWIC, 2015). Main threats to the survival of these rattlesnakes include road mortality (Winton et al., 2018), direct human persecution, and habitat loss/fragmentation due to land conversion for agriculture and urbanization. According to Hobbs (2013), 86 percent of all known rattlesnake dens in BC are within 2 km of a road. Given that roads can influence population declines at distances of up to 2 km away (Beebee, 2013), the majority of animals in the province may be at risk of decline. Additionally, as land is converted for agricultural development in BC, not only is suitable habitat for snakes lost, but snakes and people are brought closer together - enhancing opportunities for conflict (Gomez, 2008). Because there is a general lack of knowledge about these animals, and as people are often afraid of snakes, this closer proximity between people and rattlesnakes may increase the risks of direct persecution.

1.3.1 Current Species Management

At present, the Western Rattlesnake is managed as one designatable unit: Terrestrial Amphibians and Reptiles Faunal Province: Intermountain (COSEWIC, 2015). In part, this designation implies that gene flow is occurring across all populations as if they are functionally

panmictic. A lack of genetic data has precluded a formal assessment regarding the number and distribution of DUs across the Canadian range of this species (COSEWIC, 2015).

1.4 Study Objectives

Currently, there is little known information about the distribution of genetic variation across the species range in British Columbia. To inform conservation management, this thesis aims to fill this knowledge gap by:

1. Adapting and optimizing a targeted genotyping approach (Genotyping-in-thousands by sequencing (GT-seq) for use with minimally invasive DNA samples collected from the Western Rattlesnake in BC (Chapter 2)
2. Examining the distribution of population structure and extent of connectivity among populations of Western Rattlesnakes in British Columbia and Washington at multiple spatial scales to infer potential barriers to gene flow (Chapter 3)
3. Quantifying levels of genetic diversity and testing for a pattern of adult sex biased dispersal across populations range wide (Chapter 3)

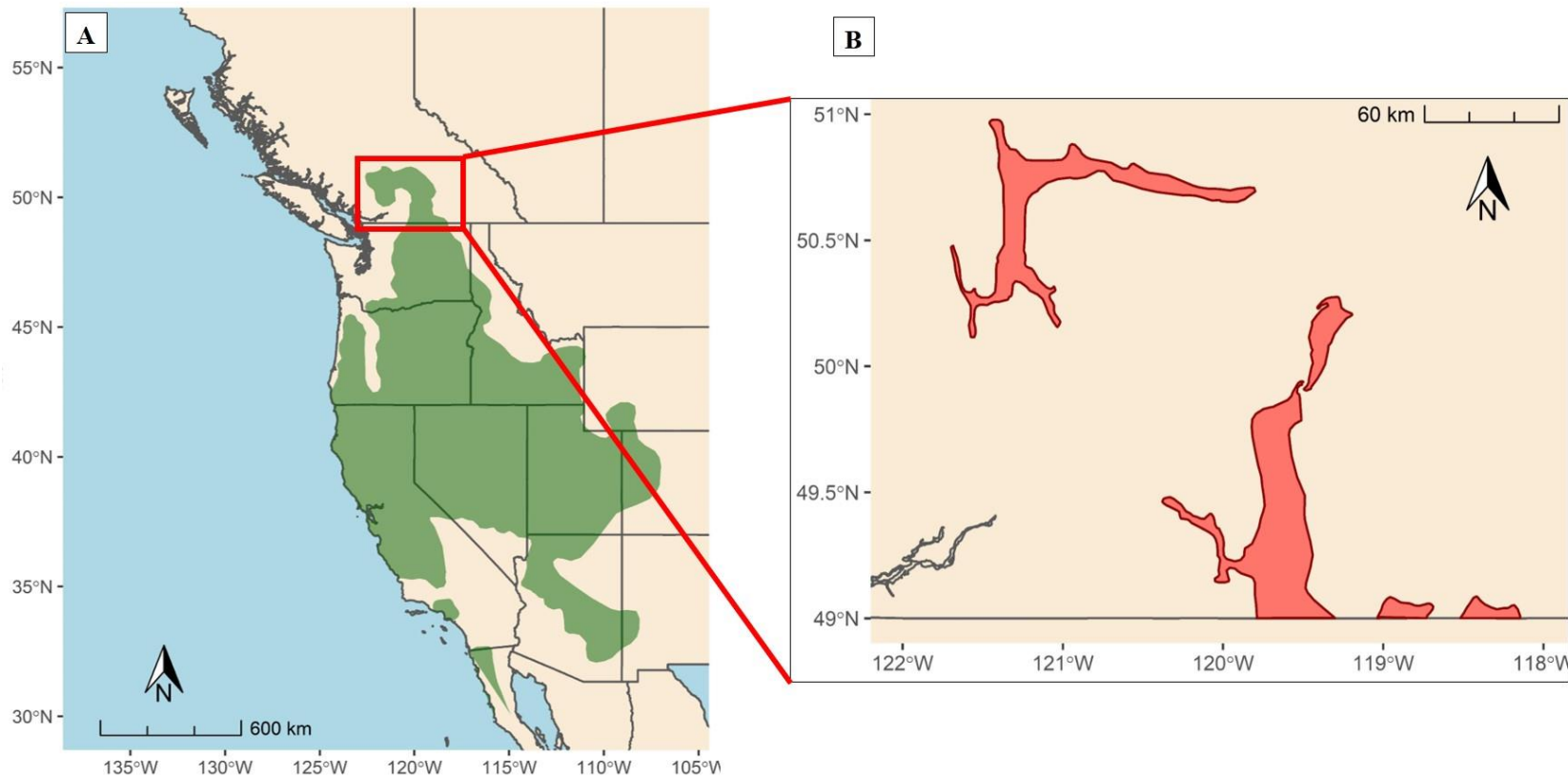


Figure 1.1 A) Map of the distribution of *Crotalus oreganus* in Canada, the United States, and Mexico shown in green (spatial data from NatureServe and IUCN (International Union for the Conservation of Nature), 2007) B) The 5 main geographic regions of *C. oreganus* occurrence in British Columbia, shown in red (spatial data provided by O. Dyer, British Columbia Ministry of Environment and Climate Change Strategy).

Chapter 2: Genotyping-in-Thousands by sequencing (GT-seq) panel development and application to minimally-invasive DNA samples to support studies in molecular ecology

2.1 Background

Massively parallel DNA sequencing (MPS) technologies (also referred to as next-generation or high-throughput sequencing) have made possible the ability to collect genome-wide data in model and non-model organisms alike (Hunter et al., 2018). These tools have enabled a transition from population genetics to population genomics, which has expanded the scope and improved the precision of studies in ecology, evolution and conservation (Luikart et al., 2003). For example, MPS has facilitated the identification and genotyping of single nucleotide polymorphisms (SNPs), allowing researchers to investigate evidence of local adaptation in natural populations (Savolainen et al., 2013), use environmental DNA (eDNA) samples to detect the presence of species in an ecosystem (Bohmann et al., 2014), and incorporate genome-wide and locus-specific data into conservation unit delimitation (Funk et al., 2012).

Historically, minimally-invasive and non-invasive DNA sampling has been used extensively across wildlife studies (Adams et al., 2003; Barba et al., 2010; Bellemain et al., 2005; Kersey & Dehnhard, 2014; Lukacs & Burnham, 2005). These types of sample collection allow researchers to limit contact and avoid disturbing species under investigation, while still collecting valuable genetic and ecological information. Minimally-invasive sampling is especially useful for studying organisms that are elusive and dangerous, as well as those where direct handling is unethical, such as species with elevated conservation status (Taberlet et al., 1999). Depending on the species being studied, many types of minimally or non-invasive samples have been used in genetic studies including hair, scat, and feathers (Beja-Pereira et al., 2009).

While there are merits to employing minimally-invasive sampling, its use in population genomic studies presents challenges (Russello et al., 2015; Andrews et al., 2018). Minimally-invasive samples typically provide DNA of poor quality and lower quantity relative to more traditional starting materials (e.g. tissue, blood). DNA degradation begins shortly after sampling, shedding, or death and can be exacerbated by environmental conditions, including temperature,

pH, humidity, and exposure to UV radiation (Poinar et al., 1996). Moreover, minimally- and non-invasive samples are typically contaminated by exogenous DNA from non-target organisms (Taberlet et al., 1999; Carroll et al., 2018); for example, contemporary fecal samples may be comprised of >95% microbial DNA (Perry et al., 2010; Chiou & Bergey, 2018).

Given these limitations, minimally-invasive samples often are problematic for use with many MPS approaches. Modern genotyping-by-sequencing approaches, such as restriction site-associated DNA sequencing (RADseq; Baird et al. 2008), work most efficiently with large amounts of intact starting material, which is often unattainable from minimally-invasive samples. The potentially degraded nature of DNA extracted from minimally-invasive samples can further disrupt the functionality of a restriction digest, as they may lack intact enzyme recognition sites, and thus, fail to create fragments with the restriction site as expected (Graham et al., 2015). Fragments without a restriction site present can then increase the number of low-quality sequencing reads in a dataset, which can reduce the number of SNPs identified, lower site coverage after quality filtering, and ultimately render data unusable (Graham et al., 2015). Moreover, any exogenous DNA present within minimally-invasive samples will also be digested and sequenced along with target DNA, further diminishing efficiency.

To increase the viability of minimally-invasive samples and similar lower quality DNA samples for use with MPS, several approaches exist that target species-specific DNA for sequencing. These methods include DNA capture, SNP genotyping assays (e.g. Taqman™) and amplicon sequencing (reviewed in Meek & Larson 2019), all of which selectively capture and enrich targeted DNA sequences prior to sequencing and/or genotyping (Ekblom & Galindo, 2011; Carroll et al., 2018). By selecting a limited number of regions in the genome to target, resulting genotypic data are typically of high coverage and informativeness for population genetic analyses. As PCR-based enrichment does not require intact restriction-sites or large amounts of starting material, such methods are useful for sequencing minimally-invasive samples (Carroll et al., 2018).

One such PCR-based targeted capture method for library preparation is Genotyping-in-Thousands by sequencing (GT-seq; Campbell et al. 2015). GT-seq is a multiplexed amplicon sequencing method that allows for the simultaneous genotyping of hundreds of SNPs across thousands of individuals in a single library, making library preparation simple and cost-effective. Although amplification success may vary among samples and loci in multiplex PCR reactions

(Andrews et al., 2018), GT-seq has been found to amplify consistently, with very little genotyping error as compared to genotypes collected for the same samples with TaqMan assays (Campbell et al., 2015). In conjunction with overall amplification success, GT-seq has great potential for use with minimally-invasive samples given that it uses species-specific primers that can help circumvent potential sample contamination, as well as degradation.

Here we tested the efficacy of GT-seq when applied to minimally-invasive samples, using as a case study, cloacal swab DNA collected for the Western Rattlesnake (*Crotalus oreganus*). This venomous species is distributed through out western North America from southern British Columbia (BC), Canada to northern Mexico; in Canada the Western Rattlesnake is classified as ‘threatened’ by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC). Major threats to this species’ persistence include road mortality and direct persecution, as well as habitat loss and fragmentation due to urban and agricultural development. The Canadian range of Western Rattlesnakes is restricted to five distinct geographical regions in BC separated by stretches of unsuitable habitat (Figure 1; COSEWIC 2015). The absence of information on the genetic diversity and population connectivity of this species has limited the effectiveness of wildlife management strategies.

As a venomous snake of elevated conservation status, minimally-invasive cloacal swab sampling is a preferred method for DNA collection from Western Rattlesnakes, as it reduces handling time and is less invasive than drawing blood. Whereas blood sampling requires specific training and procedures (Stephens et al., 2010), cloacal swabs can be obtained more readily in the field and by a single individual (Lanci et al., 2012). Previous studies have validated the use of cloacal swab samples and shed skins as alternative DNA sources to blood for collecting genotypic data in reptile species (Miller, 2006; Stephens et al., 2010; Lanci et al., 2012; Jones et al., 2008). In particular, cloacal swab DNA samples from Western Rattlesnakes have been found to provide microsatellite data consistent to that of high-quality blood samples (Ford et al., 2017).

The primary objective of this study was to assess the utility of GT-seq for collecting genomic data from minimally-invasive samples. As no previous genomic information was available for the Western Rattlesnake, RADseq data was collected from high-quality, geographically-representative blood samples for initial SNP discovery. From these data, an optimized panel of markers for use in GT-seq was identified and the resulting genotypic data were compared to RADseq data generated for the same set of individuals. This design enabled us

to directly evaluate the quality of data from our GT-seq panel relative to RADseq data in terms of genotyping error rate and consistency, as well as to investigate potential causes for discordance. Furthermore, we examined congruence between population-level metrics of genetic diversity calculated both genome-wide with RADseq and with an optimized GT-seq panel.

2.2 Methods

2.2.1 Sample Collection

Blood (n = 65) and cloacal swab (n = 18) samples were collected at den sites, while snake shed (n=4) and roadkill tissue (n = 17) samples were collected opportunistically within four major regions across the Western Rattlesnake range in BC, Canada (Thompson-Nicola (KAM), Vernon (VER), Okanagan-Similkameen (SOK) Grand Forks (GF)), as well as in Washington (WA), USA (Figure 1; TableS1). Blood samples were obtained by inserting a sterile syringe into the caudal vein of a snake, drawing 0.1 mL of blood, and then directly transferring the sample to 1 mL of Longmire lysis buffer (Longmire et al., 1997) for preservation at 4°C. Cloacal samples were collected by gently inserting a sterile cotton-tipped swab (Puritan Medical Products, USA) into the cloacal vent, located on the ventral side of the snake near the tail, and then slowly rotating it 5-10 times. As with blood samples, cloacal swabs were then preserved in 1 mL of Longmire lysis buffer (Longmire et al., 1997) at 4°C. Both shed and roadkill tissue samples were collected and stored at -80°C and -20°C, respectively. Sample collection took place under Thompson Rivers University Animal Ethics File No. 101547, BC Ministry of Environment Park Use Permit No. 108794, and the British Columbia Wildlife Act (PE15- 171661)

2.2.2 DNA Extraction

Whole genomic DNA was extracted from the blood and roadkill tissue samples using a DNeasy Blood and Tissue Kit (Qiagen) following manufacturer's protocols (including RNase A treatment) with only slight variations for the shed and cloacal swab samples. For snake shed DNA samples, ~160–180 mg of each shed underwent a surface clean with a 3% bleach solution. After bleach treatment, sheds were then washed with ethanol and left to air dry completely. Once dry, sheds were manually pulverized and placed into 2 mL tubes. Proteinase K and lysis buffer were added to each tube and samples were incubated for 48 hours at 56°C. The rest of the extraction procedure followed the manufacturer's protocol.

For the cloacal swab samples, the entire cotton tip of each cloacal swab sample was removed from the wood applicator and placed into 2 mL tubes. Samples were incubated with Proteinase K and lysis buffer for 30 minutes at 56°C. The rest of the extraction process followed the manufacturer's protocol. Once eluted, all DNA extracts were quantified using a Qubit 3.0 Fluorometer and the Qubit dsDNA High Sensitivity DNA Quant Kit (Invitrogen).

2.2.3 Restriction Site-Associated DNA Sequencing (RADseq)

We conducted RADseq for 51 blood (48 individuals with 3 in duplicate) and 4 shed DNA samples collected from across the BC distribution to identify and genotype SNPs following (Baird et al., 2008) with modifications reported in Lemay & Russello (2015). Briefly, 200-500 ng of genomic DNA was digested with *SbfI* for each sample, and then P1 adapters containing unique barcodes were ligated. After ligation, samples were pooled and sheared using a Biorupter (Diagenode). Once DNA was fragmented, it underwent size selection for a range of 300-600 bp with a Pippin Prep (Sage Science). P2 adapters were then ligated, and the final pool was amplified and size selected for 300-650 bp fragments. The resulting library was sequenced using one full lane of Illumina HiSeq 2500 paired-end 125-bp sequencing at the McGill University and Génome Québec Innovation Centre.

2.2.4 SNP Discovery

Resulting raw sequencing reads were processed using a workflow implemented in STACKS version 2.0 beta 8 (Catchen et al., 2013). First, overall data quality was assessed using FastQC (Andrews, 2010). Then, raw reads were trimmed to 100 bp and separated according to the P1 barcodes using *process_radtags*. This process also discarded low quality reads (phred score of <10) and those reads without the *SbfI* cut site present. Next, the *clone_filter* module was used to eliminate PCR duplicates by comparing paired end reads and removing additional sets of identical matching reads. After filtering for clones, six parameter sets were tested to determine the sensitivity of the data to parameter choices for the *ustacks* and *cstacks* modules of the STACKS pipeline ($M = 2,3$; $n = 1,2,3$). As the number of variant sites, inbreeding coefficients, and nucleotide diversity measures were similar across all tests (data not shown), it was determined that the data were insensitive to parameter choice and a middling set was used for data analysis (described below).

The STACKS *m* parameter was set such that a minimum of 3 reads ($m = 3$) was required to create a locus in the *ustacks* module. At this stage, only the forward reads were run to identify RAD tag loci within individual samples, and a maximum of 3 mismatching nucleotides were allowed between stacks ($M = 3$). Then, using the *cstacks* module, a *de novo* catalog of RAD tags was created with a maximum of 2 mismatches between sampled loci ($n = 2$). After catalog formation, sets of stacks for each individual were matched to the catalog with the *sstacks* module. To incorporate the paired-end reads for later analysis, the *tsv2bam* module transposed the data so that *gstacks* could align the reverse reads to existing RAD tags and assemble paired-end contigs. The *populations* module was run to further filter SNPs across all 55 samples genotyped. We retained the first SNP per locus (`--write_single_SNP`) and required all loci to be present in at least 80% of individuals (`-r 0.8`) in all five populations (`-p 5`), with a minor allele frequency greater than 0.05 (`--min-maf 0.05`) and $\ln l$ cutoff of -10 (`--lnl_lim -10`). SNPs were output into VCF and GENEPOP (Rousset, 2008) files for further filtering and analysis, and the resulting SNP dataset was filtered for quality using VCFtools (Danecek et al., 2011). Quality parameters aimed to retain loci with exactly two alleles, a minimum site depth of 10 (`--min-meanDP 10`), a maximum site depth of 100 (`min-meanDP 100`), and no more than 30 percent missing data (`--max-missing 0.3`).

2.2.5 SNP Filtering and Panel Selection

Once SNPs were called using STACKS and subsequently filtered for site quality, the entire dataset was then filtered for outliers, or those SNPs potentially under selection. Using the F_{st} -outlier detection method as implemented in Bayescan 2.0 (Foll & Gaggiotti, 2008), we identified and removed those SNP loci with q -values < 0.20 . To avoid introducing any bias in this analysis due to small sample size, we excluded the samples from the Grand Forks region ($n = 4$) (Figure 1). Next, we tested the resulting dataset for SNPs that deviated from Hardy-Weinberg equilibrium (HWE) with the (`--hardy`) function employed in VCFtools (Danecek et al., 2011). This was completed for each population independently (Thompson-Nicola, Vernon, Okanagan-Similkameen, and Washington), again excluding Grand Forks. Those SNPs with p values < 0.05 in two or more populations were excluded from our dataset. These filtering parameters were intentionally stringent and allowed for the retention of SNPs representative of neutral genetic variation.

After filtering for deviation from HWE, we calculated genotyping error using the three blood samples that each had two replicates. Genotyping error, in this case, was calculated as the proportion of discordant genotypes across all loci where both replicate samples were genotyped. SNP loci where the genotypes were either discordant or had missing data in one or both replicates for each pair of repeated blood samples were then removed from the dataset. The removal of these SNPs allowed for increased confidence that the SNPs chosen for our GT-seq panel would consistently provide genotypes in future sequencing runs.

To facilitate primer design, our SNP dataset was then filtered for locus position on the associated paired end contigs as assembled in STACKS. Due to the variation and incomplete overlap of our 100 bp forward and reverse sequencing reads, only those SNPs located between the 25th and 75th base pairs of the contigs were retained in order to increase confidence in the DNA sequence surrounding the SNPs, and to allow for sufficient flanking regions to ultimately design primers. To verify that selecting SNPs by position would not bias estimates of population genetic parameters, we compared average values of the inbreeding coefficient (G_{is}), observed and expected heterozygosity, and effective number of alleles for included populations as calculated with GENODIVE (Meirmans & Van Tienderen, 2004) relative to those found based on the full RADseq dataset, with no effect found (data not shown). We also compared mean minor allele frequency and the proportion of missing data across all SNP sites within populations as calculated with VCFtools (Danecek et al., 2011) (data not shown).

After comparing the genetic diversity estimates across the various SNP datasets, we sorted our dataset by base pair position and selected the first 500 SNPs starting at the first SNP located at base pair 40. Then, all pairs of loci were assessed for evidence of linkage disequilibrium (LD) using a Fisher's exact test as implemented in GENEPOP version 4.7 (Rousset, 2008). We removed one locus from each pair exhibiting LD ($p < 0.05$) after performing a Benjamini-Hochberg correction using the *p.adjust* function in R (R Core Team, 2019). The full RAD tag sequences associated with the resulting subset of SNPs were then sent to GTseek LLC (<https://gtseek.com/>) for locus-specific primer design.

2.2.6 GT-seq Primer Design

Upon receipt of the locus-specific forward and reverse primer sequences, we incorporated Illumina sequencing primer binding sites from the IDT unique dual-index adapters to create the complete PCR1 primers. For PCR2, custom primers containing Illumina adapter sequences were

designed based upon the IDT for Illumina Truseq unique dual index adapter sequences and the six base-pair barcode sequences from the original GT-seq paper (Campbell et al., 2015). Ninety-six i5 and fifteen i7 primer sequences were subsequently designed, allowing for the simultaneous genotyping of up to 1440 individual samples (Supplementary Data File 1).

2.2.7 GT-seq Test Library Preparation

To test the efficacy of using GT-seq for genotyping minimally-invasive samples, an initial test library was prepared for a total of 96 samples, including blood (n = 68), cloacal (n = 9), and roadkill tissue (n = 19) samples. The purpose of this first test was to determine how well samples amplified with the GT-seq method relative to RADseq and to optimize our panel of SNPs to avoid unwanted PCR artefacts and amplification biases. This library had a multi-layered design, facilitating comparisons of sequencing data on multiple scales including within and among sequencing methods and sample types (Figure 2). To assess genotyping error within GT-seq, a total of six samples had 2 replicates each (3 blood, 2 tissue, and 1 cloacal). Of these samples, the 3 blood samples that were repeated within the RADseq library also were replicated within the GT-seq library, allowing for a direct comparison of genotyping error across sequencing methods. To compare the quality of sequencing data across sample types both within GT-seq and across methods, we included seven pairs of blood and cloacal samples collected from the same individual in our GT-seq library. Additionally, 44 of the 68 blood samples we included in our GT-seq test library were previously sequenced using RADseq, allowing us to directly compare genotype consistency across methods. We further assessed genotype consistency between the cloacal samples genotyped with GT-seq and corresponding blood samples genotyped with RADseq.

Library preparation generally followed the protocol of Campbell et al. (2015) with some modifications to the original procedure. For PCR1, DNA samples with concentrations over ~30 ng/ul were diluted to 15ng/ul to help avoid the preferential amplification of samples with high DNA concentrations. One plate was then run per primer pool and 3 uL of each resulting product was pooled together and diluted 1:10 for use in PCR2. Resulting PCR2 products were normalized using a SequelPrep Normalization kit (Thermofisher) and 10 uL of sample from each individual was pooled into a plate library. The plate library was purified using a MinElute PCR Purification Kit (Qiagen) and eluted into a final volume of 22 uL. The library was sequenced using a single lane of Illumina MiSeq paired-end 150 bp sequencing at the McGill University

and Génome Québec Innovation Centre. Whole genomic DNA was extracted from the blood and roadkill tissue samples using a DNeasy Blood and Tissue Kit (Qiagen) following manufacturer's protocols (including RNase A treatment) with only slight variations for the shed and cloacal swab samples. For snake shed DNA samples, ~160–180 mg of each shed underwent a surface clean with a 3% bleach solution. After bleach treatment, sheds were then washed with ethanol and left to air dry completely. Once dry, sheds were manually pulverized and placed into 2 mL tubes. Proteinase K and lysis buffer were added to each tube and samples were incubated for 48 hours at 56°C. The rest of the extraction procedure followed the manufacturer's protocol.

For the cloacal swab samples, the entire cotton tip of each cloacal swab sample was removed from the wood applicator and placed into 2 mL tubes. Samples were incubated with Proteinase K and lysis buffer for 30 minutes at 56°C. The rest of the extraction process followed the manufacturer's protocol. Once eluted, all DNA extracts were quantified using a Qubit 3.0 Fluorometer and the Qubit dsDNA High Sensitivity DNA Quant Kit (Invitrogen).

2.2.8 GT-seq Genotyping and Primer Pool Optimization

Raw sequencing data were demultiplexed and genotyped using the GT-seq pipeline available on GitHub (<https://github.com/GTseq/GTseq-Pipeline>). The *GTseq_Genotyper_v3.pl* script was used to call genotypes for each individual sample and then *GTseq_GenoCompile_v3.pl* was used to compile all the individual genotype files into one. Locus-specific primer optimization was completed by running the *GTseq_SeqTest.pl* and *GTseq_Primer-Interaction-Test.pl* scripts on the entire set of raw sequencing data. The *SeqTest* script counted the number of raw sequencing reads that contained the forward primer associated with each locus to identify those loci that were overrepresented and preferentially amplified during PCR. The *Primer-Interaction-Test* script counted the occurrence of various primer interactions creating PCR artefacts or primer dimers. Once identified, those primers that were either overrepresented or demonstrated significant interactions were removed from the PCR1 primer pool. After primer pool optimization, a second GT-seq library was prepared for the same 96 samples, as described above, as well as for the 4 shed samples previously genotyped with RADseq and 10 cloacal swab samples each with 2 technical replicates (n =120). Library preparation followed the same procedure as the first test library, however the PCR1 primer pools were re-made to only include the optimized panel of primers.

2.2.9 Genotype Error/Concordance

Prior to making comparisons between genotyping methods and calculating genotyping error, SNP loci and samples with >30% missing data and minor allele frequencies <0.05 were removed from the GT-seq dataset. These parameters were used to mirror the quality parameters initially employed to select SNPs for the GT-seq panel from the RADseq data. The loci retained after this filtering ($n = 311$ loci; see “Results” section) were then extracted from the RADseq dataset (hereafter referred to as “RADseq_311”). This allowed for direct comparisons of genotyping error across the same subset of SNPs genotyped with both sequencing methods using the same approaches described above.

2.2.10 Population Diversity Estimates

To further evaluate the utility of using GT-seq for population genetic analysis, estimates of genetic diversity within and among populations of Western Rattlesnakes in BC were compared across the filtered GT-seq and RADseq_311 datasets, as well as the full RADseq dataset of 9568 loci (see “Results” section; hereafter referred to as “RADseq_9568”). Only those individuals genotyped with both methods, were included in this analysis ($n = 41$). For each dataset, we estimated the effective number of alleles (Eff_{Ar}), observed/expected heterozygosity (H_o/H_s), and inbreeding coefficients (G_{is}) as implemented in GENODIVE (Meirmans & Van Tienderen, 2004). Average minor allele frequency and missing data were also calculated across all loci using PLINK version 1.90 beta 4.9 (Purcell et al., 2007). To assess among population genetic differentiation, estimates of pairwise θ (Weir & Cockerham, 1984) were calculated using GENETIX v 4.05 (Belkhir et al., 2004) with 2000 permutations to test for significance. We compared the inferred number and discreteness of genetic units in our study region across datasets using the Bayesian clustering method implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000). We used an admixture model with correlated allele frequencies and a run length of 1,000,000 MCMC replicates after a burn-in period of 500,000. The most likely number of clusters (K) was determined by varying K from 1 to 7 with 10 iterations per value of K and implementing the ΔK method (Evanno et al., 2005) using STRUCTURE HARVESTER (Earl & vonHoldt, 2012) keeping in mind the bias towards $K = 2$ (Janes et al., 2017). We also plotted the log probability of the data $\ln \Pr(X|K)$ (Pritchard et al., 2000) across the range of K values, with optimal K selected where $\ln \Pr(X|K)$ plateaued as suggested in the STRUCTURE manual.

Results for the identified optimal values of K were summarized using CLUMPP (Jakobsson & Rosenberg, 2007) and visualized using DISTRUCT (Rosenberg, 2004).

2.3 Results

2.3.1 RADseq, SNP Discovery and Panel Selection

Our RADseq library yielded 550,266,898 reads that were retained after demultiplexing and discarding reads with ambiguous barcodes and low quality. Each sample ($n = 55$) had an average of 10,004,853 reads. *De novo* assembly in STACKS generated 157,008 RAD tags, from which 9,933 SNPs were identified after filtering with the *populations* module (average read depth = 35.2). From the initial set of SNPs called, 196 failed to meet site quality criteria, 104 were flagged as outliers, and 65 deviated from Hardy-Weinberg, resulting in 9568 SNPs for further filtering. Overall, this dataset had an average of 4.72% missing data ($n=55$) with an average of 0.00% missing data for the shed samples ($n=4$) and 5.09% for the blood samples ($n=51$). The average proportion of genotyping error across the 3 repeated blood samples in the RADseq dataset was 0.80%. A total of 1287 SNPs that exhibited either genotype discordance or missing data in any of the replicate comparisons were then removed from the dataset, leaving 8281 markers for subsequent filtering and GT-seq panel selection. Of the 500 candidate SNPs selected after filtering and evaluation, 488 loci were retained after tests for LD and primers were successfully designed for 393 loci (Supplementary Data Files 1-2). Following the results from the initial MiSeq run and primer pool optimization, we removed 31 loci, leaving a final GT-seq panel of 362 SNPs (Supplementary Data Files 1-2).

2.3.2 GT-seq SNP Genotyping

After filtering for missing data and minor allele frequency, a total of 311 SNPs and 103 samples remained in our GT-seq dataset (average read depth = 185.1) with an average of 4.49 % missing data overall. Genotyping success was neither dependent on initial DNA sample concentration nor DNA sample type (e.g. blood, roadkill tissue, shed, or cloacal swab (Figures S1-S2; Table S2). Of the 16 repeated samples included in the GT-seq library design, three samples had one of two replicates fail to amplify likely due to PCR error, leaving three samples (2 blood and 11 cloacal) to be used for calculating genotyping error. For those SNPs where both replicates were successfully genotyped, the mean percent genotyping error was 0.5%. No

genotyping error was detected for the replicated blood samples, while genotyping error was 0.62% for the replicated cloacal samples. Six of the seven pairs of cloacal and blood samples collected from the same individuals amplified successfully and met quality filtering criteria; the average genotyping discordance across these comparisons was 1.37%.

2.3.3 Genotype Concordance Between RADseq and GT-seq

After quality filtering, 49 of the 55 samples included in the final GT-seq library for direct genotype comparisons to RADseq were retained in our dataset (41 blood, 6 cloacal, 2 sheds). Overall, the mean proportion of genotype discordance between genotyping methods for all 49 samples was 2.53%. For the blood samples genotyped with both sequencing methods ($n = 41$), the average genotype discordance was 2.63%, and for the shed samples ($n = 2$) the average discordance was 1.68%. For the cloacal samples genotyped with GT-seq that had corresponding data collected from blood samples genotyped with RADseq ($n = 6$), the average genotype discordance was 2.10%.

2.3.4 Population Diversity Estimates and Individual Ancestry Estimation

All three datasets (GT-seq, RADseq_311, RADseq_9568) yielded similar estimates of within population diversity across all sites (Thompson-Nicola, Okanagan-Similkameen, Vernon, Washington) (Table 1). In general, the GT-seq dataset tended to have slightly higher values of observed/expected heterozygosity, as well as more negative inbreeding coefficients (Table 1). Estimates of pairwise θ across all four populations were also similar between the GT-seq and RADseq_9568 datasets (Table 2). The clustering analysis implemented in STRUCTURE revealed an optimal value of $K = 3$ (ΔK : GT-seq = 303.3, RADseq_311 = 424.1, RADseq_9568 = 1049.9; Table S3) and produced congruent patterns of individual ancestry and population structure across all three datasets (Figure 3).

2.4 Discussion

Here, we demonstrated for the first time that GT-seq, a genotyping by multiplexed amplicon sequencing approach, can be effectively applied to minimally-invasive samples. The genotypic data generated with our optimized GT-seq panel were consistent with data generated with RADseq for the same subset of loci (i.e. RADseq_311). Moreover, estimates of genetic diversity and population structure based on both reduced SNP datasets (GT-seq, RADseq_311)

were representative of the signal in the full RADseq_9568 dataset. While most genotypes collected were identical across methods for the same individuals, there was a small proportion of genotype discordance (2.57%). Virtually all cases of genotype mismatching (86.0%) across methods were due to allelic dropout in the RADseq genotypes relative to GT-seq at the same loci.

Given the genotyping error rates within each method (GT-seq = 0.50%; RADseq = 0.80%), it is unlikely that the genotype discordance found across datasets was due to sequencing error. Rather, the observed discrepancies between GT-seq and RADseq genotypes may be attributed to differences in how the genotypes were called by each method (Nielsen et al., 2011; Hess et al., 2015). For the data generated with RADseq, SNPs were simultaneously identified and genotyped using the STACKS pipeline, which uses a multinomial maximum-likelihood framework to estimate the sequencing error rate at a nucleotide and then to assign the most likely genotype at that locus (Hohenlohe et al., 2010; Catchen et al., 2013). For loci where one genotype is not significantly more likely than the other based upon a likelihood ratio test, no genotype is inferred. This inherently removes sites with insufficient sequencing depth from the dataset on a per site basis, allowing for genotypes to be called at variable levels of read coverage. For GT-seq, genotypes were called by calculating the allele ratios at a given SNP locus (Campbell et al., 2015), rather than through a maximum-likelihood model as used in STACKS. Loci with allele ratios >10 , <0.1 , or between 2 and 5 were classified as homozygous for allele 1, homozygous for allele 2, or heterozygous, respectively. Sites with less than 10x sequencing depth were not assigned a genotype, requiring a greater depth of coverage for calling genotypes than the model employed by STACKS.

Accordingly, the pattern of allelic dropout observed in the genotypes from the RADseq dataset may be due to the specific SNP calling model used for identifying and genotyping SNPs with STACKS. The default SNP calling model allows the estimated sequencing error parameter to vary such that the maximum-likelihood of a genotype can have an unrealistically high error rate (Catchen et al., 2013). This can cause sites that have alternative alleles present to be identified as homozygous with an excessive error rate, rather than called as heterozygous. Therefore, those loci genotyped as heterozygous in GT-seq and homozygous in RADseq may actually have been heterozygous in the latter; but were misidentified as homozygous due to the SNP calling model treating the alternative alleles as error rather than true polymorphism. In fact,

preliminary comparisons of the RADseq data genotyped using the GT-seq workflow revealed a lower percent allelic dropout rate (73.2%) relative to the GT-seq genotypes than exhibited by the RADseq data genotyped using the STACKS workflow (86.0%) (data not shown). Interestingly, genotypic discordance (2.27%) was observed in the same RADseq data genotyped by the two different workflows (GTseq v. STACKS), further suggesting that variation in the SNP calling models may be playing a role (data not shown). Although slight, the discrepancies found between genotypes in this study highlight the importance of carefully considering the models selected for calling SNPs, especially for use in targeted genotyping.

When it comes to designing GT-seq studies for use with minimally-invasive sampling, there are several important considerations for SNP panel selection and optimization. First, genomic information for the organism under investigation is needed from which SNPs can be selected. Ideally, a reference genome would be used to select SNPs for a GT-seq panel to facilitate assembly and map targeted loci to enable connectivity among studies. However, as many non-model organisms do not yet have high-quality reference genomes, *de novo* SNP identification can be suitable for panel selection (Davey et al., 2011; Kraus et al., 2015; Senn et al., 2013; Lew et al., 2015; Bayerl et al., 2018). In particular, *de novo* reference assemblies should be generated from a set of high-quality samples from the species of interest, such as blood or tissue. This allows for the selection of a subset of species-specific markers based upon stringent parameters for high coverage and limited missing data. Given that high-quality DNA samples are not always accessible for many organisms, pre-established SNP marker datasets or panels from other studies or databases for the same organism can also be adapted for use with GT-seq (Hess et al., 2015). Additionally, SNP panel selection is dependent upon the questions being addressed by a particular study. In cases where the resulting data are intended for population genetic analysis, it may be ideal to choose SNP markers that encompass neutral genetic variation in the context of the sampled populations. For other studies where more diagnostic panels are appropriate, those loci with high θ values or flagged as candidates under selection may be preferentially selected for a SNP panel (Russello et al., 2012; Hess et al., 2015). In this study, we chose SNPs that were variable in all sampled populations and genotyped in at least 80 percent of individuals per population, excluding those that were candidates for selection and that deviated from Hardy Weinberg and linkage equilibria. These filtering parameters were

used to minimize ascertainment bias in our panel selection (Carlson et al., 2003) and help ensure the neutrality of selected SNPs.

Overall, we have shown that GT-seq provides an effective approach for genotyping minimally-invasive samples, providing accurate and precise estimates of within- and among population diversity metrics relative to a companion RADseq dataset. Although initial SNP discovery and panel optimization are required, GT-seq provides a scalable approach that effectively harnesses the various throughput capacities of MPS. In that regard, thousands of individuals can be genotyped at hundreds of SNPs using a single lane of Illumina HiSeq to support a large-scale study (Campbell et al., 2015), or alternatively, smaller numbers of samples can be added to an existing database to support long-term population monitoring or wildlife forensic applications taking advantage of a MiSeq spike or other low-throughput MPS methods. Importantly, the targeted, multiplex amplicon sequencing approach harnessed by GT-seq can be effectively applied to low quality DNA samples, minimizing the inefficiencies presented by exogenous DNA typically found in minimally-invasive samples and continuing the expansion of molecular ecology and conservation genetics in the genomics era.

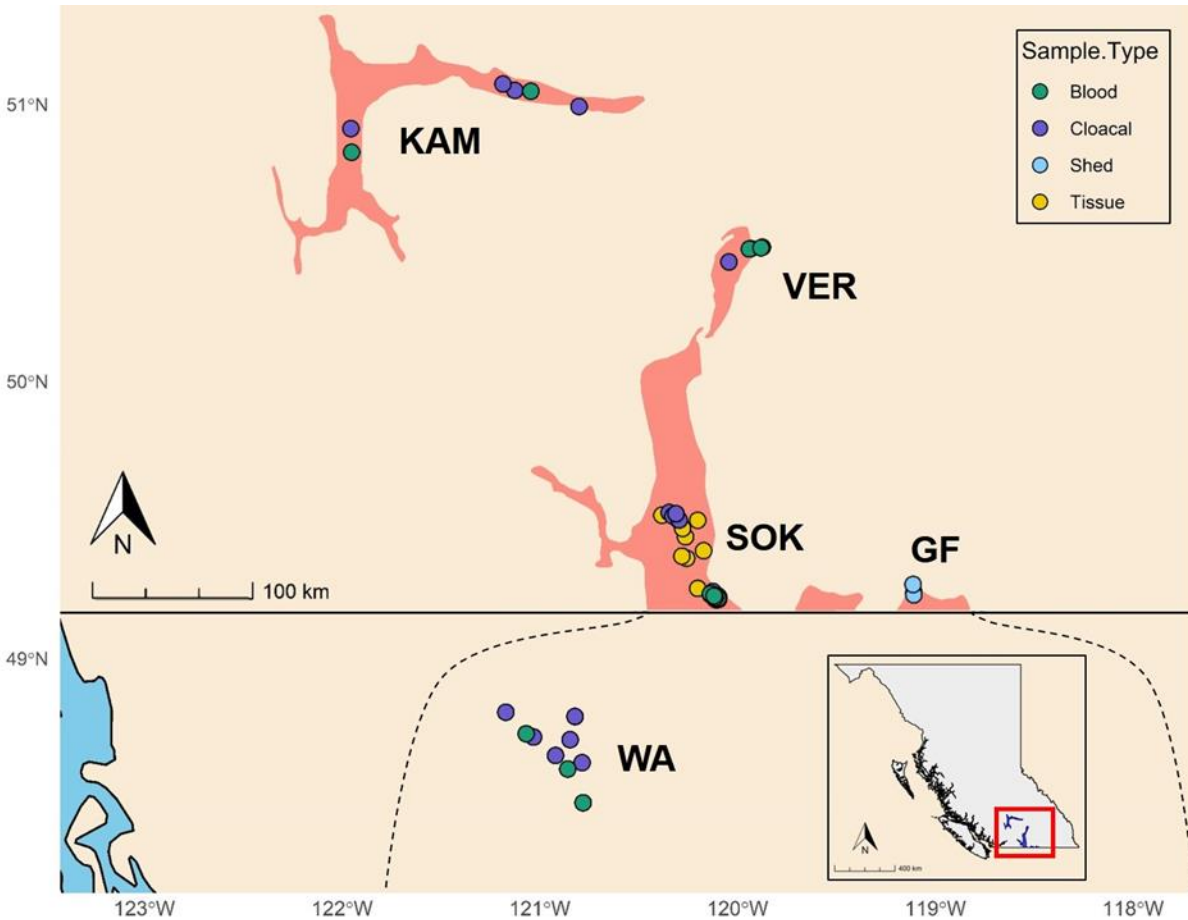


Figure 2.1. Western Rattlesnake species range in British Columbia, Canada (shaded) and Washington, USA (approximate; dotted line), including sample distribution from sites in: Thompson-Nicola (KAM; n = 14), Vernon (VER; n = 30), Okanagan-Similkameen (SOK; n = 40), Grand Forks (GF; n = 4), and Washington (WA; n = 16). Within region sample locations are represented by circles with the various colors indicating sample type noted in legend.

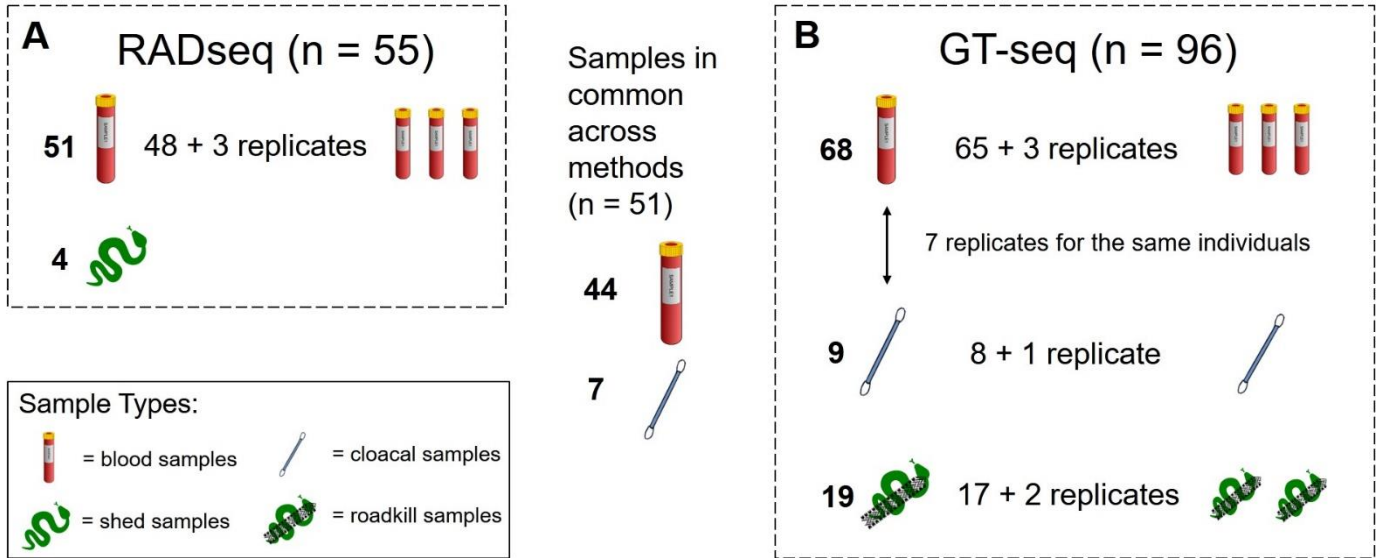


Figure 2.2. Library designs for A) RADseq and B) GT-seq. Included samples selected to facilitate within and among method genotype comparisons.

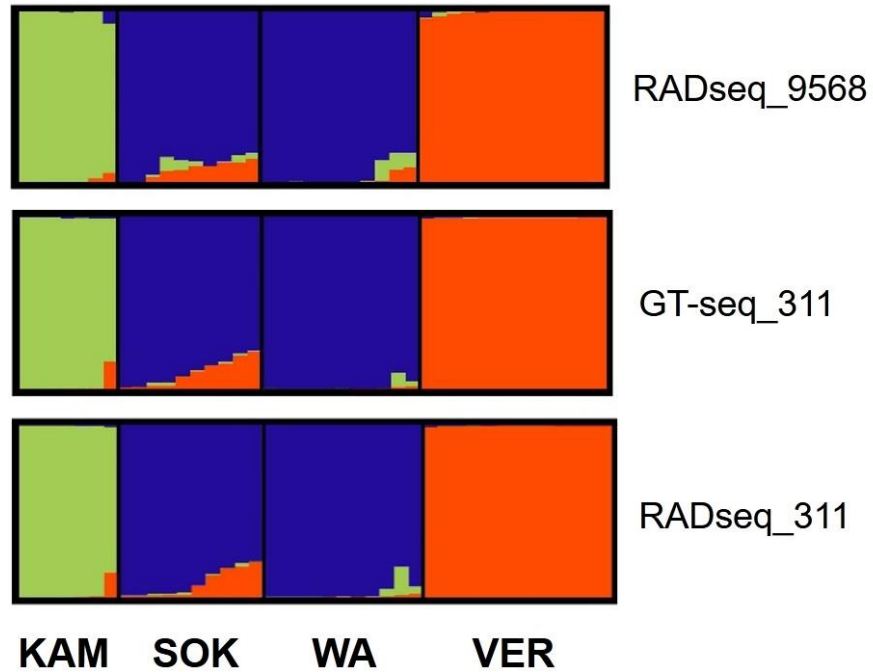


Figure 2.3. STRUCTURE barplots displaying inferred clustering and individual ancestry estimation of Western Rattlesnakes ($n = 41$) for all 3 datasets (GT-seq, RADseq_311, RADseq_9568) across all sampled regions (population acronyms as in Figure 1). Each different color in the plots represents a distinct genetic cluster and each vertical bar represents the proportion of ancestry of a single individual to the different genetic clusters.

Table 2.1. Population genetic diversity estimates within and among sampled populations across methods and dataset sizes, including estimates of mean effective number of alleles (Eff_Ar), mean observed heterozygosity (H_o), mean expected heterozygosity within populations (H_s), and mean inbreeding coefficient (G_{is}).

Population	Genotyping Method	No. SNPs	Eff_Ar	H_o	H_s	G_{is}
Thompson- Nicola (n = 7)	RADseq	9568	1.376	0.223	0.239	0.068
	RADseq	311	1.381	0.209	0.239	0.125
	GT-seq	311	1.393	0.233	0.248	0.06
Okanagan-Similkameen (n = 10)	RADseq	9568	1.53	0.324	0.328	0.012
	RADseq	311	1.527	0.318	0.327	0.028
	GT-seq	311	1.55	0.34	0.338	-0.006
Washington (n = 11)	RADseq	9568	1.541	0.318	0.332	0.044
	RADseq	311	1.566	0.346	0.344	-0.005
	GT-seq	311	1.567	0.355	0.344	-0.033
Vernon (n = 13)	RADseq	9568	1.426	0.253	0.263	0.035
	RADseq	311	1.455	0.271	0.28	0.032
	GT-seq	311	1.467	0.289	0.284	-0.017
Overall (n = 41)	RADseq	9568	1.411	0.279	0.291	0.039
	RADseq	311	1.423	0.286	0.298	0.04
	GT-seq	311	1.436	0.304	0.304	-0.001

Table 2.2. Pairwise estimates of θ (Weir & Cockerham 1984) for all four sampled populations of Western Rattlesnakes across all three data sets (GT-seq_311, RAD_311, RAD_9568). All pairwise comparisons were significant ($p < 0.05$) with 2000 permutations. Population acronyms as in Figure 1.

Dataset	Population Comparisons					
	KAM v SOK	KAM v WA	KAM v VER	SOK v WA	SOK v VER	VER v WA
GT-seq_311	0.208	0.225	0.317	0.079	0.149	0.20
RADseq_311	0.225	0.227	0.330	0.084	0.162	0.204
RADseq_9568	0.212	0.210	0.306	0.083	0.179	0.193

Chapter 3: Genotyping-in-Thousands by sequencing reveals marked population structure in Western Rattlesnakes to inform conservation status

3.1 Background

Habitat loss and degradation are two of the main threats affecting the persistence of global biodiversity (Magin et al., 1994; Purvis et al., 2000). Anthropogenic disturbances such as residential and commercial development, roads, agriculture, and forestry can lead to large areas of formerly intact habitat becoming isolated and disconnected over time. As a result, populations often are reduced in size and exhibit low population connectivity with a decreased capacity for individuals to disperse (Pease et al., 1989). Gene flow among subpopulations is then interrupted, which limits the ability for isolated subpopulations to maintain levels of genetic diversity found in connected populations. Without sufficient genetic variation, populations of a species are unable to adapt in the face of stochasticity, making them more vulnerable to extinction (Frankham, 2003). Over time, the extinction of multiple isolated sub-populations can contribute to overall decline in population size and erosion of genetic diversity. As a result, species distributed across isolated populations often are of elevated conservation concern.

By evaluating patterns of genetic diversity and gene flow, genetic studies play an important role in understanding population connectivity among geographically-isolated populations, especially for rare species or those that are at-risk. In particular, genetic assessments can help delineate conservation units that are integral for enforcing legislation and prioritizing conservation resources (Funk et al., 2012; Green, 2005). Multiple concepts for designating conservation units have been proposed across various contexts, including evolutionary significant units (ESUs), that are based upon historical patterns of genetic divergence between populations (Crandall et al., 2000; Ryder, 1986; Waples, 1991), and management units (MUs) that consider contemporary patterns of genetic differentiation among populations (Moritz, 1994). In Canada, designatable units (DUs) have been adopted by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) to define biologically relevant units for conservation below the species level (Green, 2005). Designatable units are determined based upon population discreteness and evolutionary significance, two criteria that can be examined using genetic data. According to COSEWIC, population(s) may be considered discrete with evidence for genetic distinctiveness, natural geographic disjunction, or occurrence

in geographically-varied ecological zones. Once populations are deemed discrete, they can then be evaluated for significance, meaning that these populations are integral to the evolutionary trajectory of the species and would not be naturally be restored if lost (Environment and Climate Change Canada, 2017). Measures of genetic variation and gene flow among populations can provide evidence for discreteness, while evolutionary significance of populations can be assessed by estimating adaptive divergence (Allendorf et al., 2010).

Here, we conducted the first genetic evaluation of the at-risk Western Rattlesnake (*Crotalus oreganus*) across its distribution in British Columbia, which encompasses the northernmost extent of the species range. The Western Rattlesnake is a venomous snake found along the western portion of North America from southern British Columbia to northern Mexico. In Canada, this species is classified as ‘threatened’ by the federal Species At Risk Act and protected under the BC Wildlife Act (COSEWIC, 2015). The Canadian range of this species is estimated to encompass 5% of the global population, occurring among five main geographic regions in BC (COSEWIC, 2015). These five regions are the Thompson-Nicola, Vernon, Okanagan-Similkameen, Midway, and Grand Forks, among which the rattlesnake populations are isolated by stretches of unsuitable habitat (Figure 1.1b). The current population size of Western Rattlesnakes in the province is not definitively known, but is estimated to be about 10,000 individuals (COSEWIC, 2015). Moreover, populations are suspected to be declining; about 12 percent of all known rattlesnake dens have become extirpated since 1980, and with current threats to survival the overall population size is predicted to decline by 30 percent over the next 45 years (3 generations) (Southern Interior Reptile Recovery Team, 2016).

Major threats to Western Rattlesnake persistence in BC include road mortality (Winton et al., 2018), direct persecution, and habitat fragmentation due to agricultural development (COSEWIC, 2015). At present, all regions of rattlesnake occurrence in BC are managed as a single designatable unit under the COSEWIC Terrestrial Amphibians and Reptiles Faunal Province: Intermountain (COSEWIC, 2015). Due to a lack of evidence for distinct behaviour and morphology among populations, as well as an absence of genetic information, there has been no further investigation into the current delimitation of designatable units for this species in Canada.

To help fill this knowledge gap, we used genotypic data collected at an optimized panel of single-nucleotide polymorphisms (SNPs) to determine the distribution of genetic variation within and among Western Rattlesnake populations across the species range in Canada and

Washington (USA). By examining range-wide genetic differentiation among geographically isolated populations of rattlesnakes in BC, we tested the hypothesis that patterns of genetic structure would align with *a priori* defined geographic regions. We also inferred fine-scale patterns of population structure and connectivity, and quantified genetic diversity within and among den sites, allowing us to assess populations for signals of inbreeding and heterozygosity loss. Further, we tested for a pattern of sex-biased dispersal among adult rattlesnakes to investigate the movement ecology of this species, which can influence population structure and connectivity. We then discuss the implications of these results for informing future conservation status assessments and conservation management for this species in BC.

3.2 Methods

3.2.1 Study area/ sample collection and distribution

Between 2015 and 2017, blood (n = 68), cloacal swab (n = 566), and buccal swab (n = 5) samples were collected from 36 den sites from all five regions across the Western Rattlesnake range in BC, as well as in Washington State, USA, while snake shed (n = 4) and roadkill tissue (n = 115) samples were collected opportunistically (total n = 758) (Figure 3.1). In this study, our sampling regions were defined as follows: Thompson-Nicola (KAM), Vernon (VER), the Okanagan-Similkameen (SOK), Washington (WA), Midway (RC), and Grand Forks (GF). All samples obtained from live snakes were taken with the posterior portion of the animals restrained in a plexiglass handling tube (Murphy, 1971). Blood samples were drawn by inserting a sterile syringe into the caudal vein of a snake, drawing 0.1mL of blood, and then directly transferring the sample to 1mL of Longmire lysis buffer (Longmire et al., 1997) for preservation at 4°C. Cloacal samples, which have been shown to provide genotypic data comparable to that of blood samples (Ford et al., 2017; Chapter 2), were collected by gently inserting a sterile cotton-tipped swab into the cloacal vent, located on the ventral side of the snake near the tail, and then slowly rotating it 5-10 times. To sample from juvenile and neonate snakes that were too small for cloacal sampling, buccal swabs were employed. To collect buccal swabs, a cotton swab was gently inserted into the mouth of a snake and slowly rotated for 5-10 seconds, with care taken to avoid getting the swab stuck on the fangs of the snake. As with blood samples, buccal and cloacal swabs were then preserved in 1mL of Longmire lysis buffer (Longmire et al., 1997) at 4°C. Shed skins and roadkill tissue samples were collected and stored at -80°C and -20°C,

respectively. Sample collection took place under Thompson Rivers University Animal Ethics File No. 101547, BC Ministry of Environment Park Use Permit No. 108794, and British Columbia Wildlife Act (PE15- 171661).

3.2.2 DNA Extraction

Whole genomic DNA was extracted from the blood and roadkill tissue samples using a DNeasy Blood and Tissue Kit (Qiagen) following manufacturer's protocols (including RNase A treatment). Only slight variations were applied for the shed, buccal and cloacal swab samples. For DNA samples from shed skins, ~160-180 mg of each shed underwent a surface clean with a 3% bleach solution. After bleach treatment, sheds were then washed with ethanol and left to air dry completely. Once dry, sheds were manually pulverized and placed into 2 mL tubes. Proteinase K and lysis buffer were added to each tube and samples were incubated for 48 hours at 56°C. The rest of the extraction procedure followed the manufacturer's protocol.

For the cloacal and buccal swab samples, DNA was extracted from the lysis buffer in which the swabs were stored. 180 uL of buffer from each sample was incubated with 20 uL of Proteinase K for 1 hour at 56°C. The rest of the extractions followed the manufacturer's protocol. Once eluted, all DNA extracts were quantified using a Qubit 3.0 Fluorometer and Qubit dsDNA High Sensitivity DNA Quant Kit (Invitrogen).

3.2.3 GT-Seq Library Preparation

A Genotyping-in-Thousands by sequencing (GT-seq) library was used to genotype all 758 samples at a targeted panel of 362 single-nucleotide polymorphism (SNP) markers (see Chapter 2). A total of 22 samples were repeated to calculate within-library genotyping error. Library preparation generally followed the protocol of Campbell et al. (2015) with some modifications to the original procedure. For PCR1, one plate was run per primer pool and 3uL of each resulting product was pooled together and diluted 1:10 (a total dilution of 1:20) for use in PCR2. After each plate of 96 PCR2 products were normalized using a SequelPrep Normalization kit (Thermofisher), 10uL of sample from each individual was pooled into a plate library. Plate libraries were then quantified and an equimolar amount of each was combined in a single tube. This library was purified using a MinElute PCR Purification Kit (Qiagen) and eluted into a final volume of 22uL. Library sequencing was completed on a single lane of Illumina HiSeq4000 paired end 100 sequencing at McGill University and Génome Québec Innovation Centre.

3.2.4 GT-seq Genotyping and Genotyping Error

A Genotyping-in-Thousands by sequencing (GT-seq) library was used to genotype all 758 samples at a targeted panel of 362 single-nucleotide polymorphism (SNP) markers (see Chapter 2). A total of 22 samples were repeated to calculate within-library genotyping error. Library preparation generally followed the protocol of Campbell et al. (2015) with some modifications to the original procedure. For PCR1, one plate was run per primer pool and 3uL of each resulting product was pooled together and diluted 1:10 (a total dilution of 1:20) for use in PCR2. After each plate of 96 PCR2 products were normalized using a SequelPrep Normalization kit (Thermofisher), 10uL of sample from each individual was pooled into a plate library. Plate libraries were then quantified and an equimolar amount of each was combined in a single tube. This library was purified using a MinElute PCR Purification Kit (Qiagen) and eluted into a final volume of 22uL. Library sequencing was completed on a single lane of Illumina HiSeq4000 paired end 100 sequencing at McGill University and Génome Québec Innovation Centre.

Raw sequencing data were separated into one file per sample and then genotyped using the GT-seq pipeline available on GitHub (<https://github.com/GTseq/GTseq-Pipeline>). The *GTseq_Genotyper_v3.pl* script was used to call genotypes for each individual sample and then *GTseq_GenoCompile_v3.pl* was used to compile all the individual genotype files into one. This resulting output file was then converted into PED file format (Purcell et al., 2007) for downstream genetic analysis.

Prior to conducting genetic analyses, the entire dataset of 758 samples was quality filtered and those individuals and SNP loci with >30% missing data and a minor allele frequency <0.05 were removed. Genotyping error was calculated for repeated samples where both replicates amplified and passed data quality filters. Genotyping error was defined as the proportion of discordant genotypes across all loci successfully genotyped for both replicates. To reduce bias due to sampling potentially highly related individuals at den sites, ML-RELATE (Kalinowski et al., 2006) was used to infer relationships between all individuals. Those individuals with an inferred parent as a result of this analysis were subsequently removed from the dataset for all population genetic analyses, except those assessing patterns of relatedness.

3.2.5 Regional Population Differentiation and Connectivity

To infer broad scale patterns of population structure across the known geographic regions, Bayesian clustering analyses were conducted with STRUCTURE v 2.3.4 (Pritchard et

al., 2000). Ten iterations were run for K 1-8 (the number of *a priori* geographic regions +2), each with a burn-in of 500,000 and a run length of 1,000,000 MCMC steps. The resulting output was then summarized using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). The optimal number of genetic clusters (K) was inferred using a combination of the ΔK method (Evanno et al., 2005) and by plotting the log probability of the data (Pritchard et al., 2000) to select the optimal K where $\ln \Pr(X|K)$ plateaued (see STRUCTURE manual). For the optimal value of K , results were summarized with CLUMPP (Jakobsson & Rosenberg, 2007) and then plotted with DISTRUCT v.1.1 (Rosenberg, 2004). Additionally, an Analysis of Molecular Variance (AMOVA) was conducted as implemented in ARLEQUIN v.3.5 (Excoffier & Lischer, 2010) to investigate the broad scale partitioning of genetic variation within and among detected genetic clusters with 10,000 permutations to test for significance. Weir and Cockerham's estimates of pairwise F_{ST} and inbreeding coefficients (F_{IS}) also were calculated with GENETIX 4.05 (Belkhir et al., 2004) to quantify the amount of genetic differentiation within and among den groups, as these estimators are unbiased with regard to sample size (Weir & Cockerham, 1984).

To examine the extent and pattern of connectivity across the range we tested for patterns of isolation-by-distance (IBD) among geographic regions and den groups using a Mantel test (Mantel, 1967) as implemented in the R package *ade4* (Dray & Dufour, 2007) with 10,000 replicates to test for significance. A significant pattern of IBD would indicate that pairs of populations that have a smaller geographic distance between them will be less genetically different from one another than those separated by a larger geographic distance as a result of short-range dispersal of individuals among populations (Wright, 1943). To conduct this test, first, the centroid location of each region or den group was determined using the *rgeos* (Bivand & Rundel, 2014) package in R (R Core Team, 2019). Straight-line distance between centroids was then calculated in kilometers with GeographicDistanceMatrixGenerator_v1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg/download.php) and correlated with estimates of pairwise F_{ST} .

3.2.6 Fine Scale Population Structure and Connectivity

Subsequent fine-scale analyses included only those samples that were associated with known dens ($n=461$), as these sites were the basis for DNA sampling locations for this study. To evaluate the pattern of genetic structure among den sites sampled within each region, we ran STRUCTURE as described previously, using a hierarchical approach (Vähä et al., 2007) based

upon the genetic clusters inferred from the range-wide analysis. Groups of dens (hereafter referred to as den groups) were then defined based upon resulting genetic structure that was significant, as well as geographic distribution, as sets of den sites sampled were generally isolated throughout each region (Figure 3.1). To further examine the hierarchical partitioning of genetic variation within and among den groups encompassed within each geographic region AMOVAs were conducted and pairwise F_{ST} were calculated as described above.

To evaluate patterns of connectivity among den groups, we tested for patterns of IBD using a Mantel test (Mantel, 1967) as outlined for the broad scale analysis. Mantel tests also were used to assess the pattern of den site connectivity within and among geographic regions that encompassed >1 den site. For these tests, straight-line distance was calculated between individual den sites within a given region and then correlated with estimates of pairwise F_{ST} . Additionally, the amount and direction of contemporary migration among den groups were assessed with BAYESASS 3.04 (Wilson & Rannala, 2003). This Bayesian approach uses multilocus genotypes to estimate movement among populations over the past three generations. We estimated levels of migration using 10,000,000 iterations after a burn-in of 1,000,000 MCMC steps, sampling after every 100 steps. Five runs were conducted, each starting with a different random seed and convergence was analyzed by comparing mean posterior estimates of migration. To determine if migration was significant between populations, we calculated 95% credible sets by calculating the mean migration rate $\pm 1.96x$ the standard deviation, as suggested in the user manual. Those migration estimates whose credible sets that did not include zero were deemed significant.

Pairwise estimates of Queller and Goodnight's relatedness coefficient were also calculated for individuals within and among den sites using SPAGeDI 1.5 (Hardy & Vekemans, 2002). This estimator was chosen because it is unbiased towards small sample sizes (Queller & Goodnight, 1989). Reference allele frequencies for each den group were calculated by pooling individuals in each of the geographic regions defined by broad scale STRUCTURE analyses. For this analysis, the 27 individuals removed with presumed parents in the dataset were added back into the dataset, leaving 488 individuals for analysis.

3.2.7 Genetic Diversity Within and Among Dens

We evaluated genetic diversity within each den group and within each hibernaculum by calculating estimates of observed heterozygosity (H_o), expected heterozygosity (H_e), gene

diversity (N_g) and the proportion of variable sites (P) as implemented in ARLEQUIN (Excoffier & Lischer, 2010). Global tests for heterozygote deficiency or excess were run using Fisher's exact tests in GENEPOP 4.5, with 10,000 dememorization steps, 100 batches, and 10,000 iterations. Those groups with a p-value <0.05 were considered to significantly deviate from the level of heterozygosity expected if they were in Hardy-Weinberg equilibrium. Inbreeding coefficients (F_{IS}) also were calculated for each den group with GENETIX (Belkhir et al., 2004), using 1000 permutations to test for significance. To determine if measures of genetic diversity were correlated with latitude, linear regressions with nucleotide diversity (N_g) and the proportion of polymorphic SNP loci in a population (P) were conducted in R (R Core Team, 2019) for both den groups and individual den sites.

3.2.8 Sex Bias in Dispersal

As patterns of dispersal can influence genetic structuring across a metapopulation (Hansson, 1991), we also tested for a pattern of adult sex-biased dispersal in this species. For this analysis, only adult snakes that had been reliably sexed upon sampling were retained in the dataset ($n=337$). Adult snakes were defined as either those snakes identified as adults at the time of sampling or those with an snout-to-vent length (SVL) measurement of $>540\text{mm}$, based upon recent estimates for sexually mature Western Rattlesnakes (Maida et al., 2018). To test for evidence of a sex bias in dispersal among den groups, we compared estimates of mean assignment index (mAIc) and the variance of assignment indices (vAIc) between male and female snakes using the R package *hierfstat* (Goudet, 2005). The assignment index (AI) statistic implemented in this analysis estimates the probability of a given individual's genotype across loci appearing in any of the sampled populations, and then assigns the individual to the population where its genotype is most likely (Paetkau et al., 1995). This value then is corrected to account for differences in diversity across populations (AIc), where a positive index suggests that the individual's genotype is likely to appear within the population from which it was sampled (resident), while a more negative value indicates that the genotype is rare in that population (immigrant) (Favre et al., 1997). Consequently, mAIc estimates are expected to be lower for the less philopatric sex, as the genotypes of immigrant individuals are less likely to be assigned to the population from which they were sampled. Moreover, estimates of vAIc are expected to be higher in the less philopatric sex, reflecting how assignment probabilities will show a large variance if there are immigrant snakes detected. This test also was repeated within

den groups containing two or more den sites with at least 2 individuals of each sex ($n=8$). Given that Western Rattlesnakes are polygynous and, along with other Crotaline species, are known to exhibit male mate searching, we predicted that we would find a male bias in dispersal for this species, if any bias existed at all (Clark et al., 2008; King & Duvall, 1990).

3.3 Results

3.3.1 GT-seq Genotyping

A total of 758 samples were genotyped with an average of 363,834 raw reads per locus (range: 34,362-693,305). After quality filtering for 30 percent missing data and minor allele frequency >0.05 , 636 samples and 308 loci remained in our dataset. Mean genotyping error across all repeated samples where both replicates amplified was 0.00384. After repeated samples that amplified were removed ($n=21$), a total of 615 individuals remained with an average read depth of 305 and 3.87 percent missing data. After those individuals with an inferred parent in the dataset ($n=27$) were removed, a total of 588 individuals were used for subsequent population genetic analysis (Appendix B; Table S3.1).

3.3.2 Inferred Regional Population Structure and Connectivity

STRUCTURE analyses inferred an optimal value of $K=3$ for all samples range wide ($n=588$), (Appendix B; Figure S3.2). Each unique cluster represented Thompson-Nicola (KAM), Vernon (VER), and the rest of the geographic regions across the range, respectively. The same pattern of structure was found with the dataset of 461 individuals that could be assigned back to a den site (Figure 3.2). A second iteration of STRUCTURE revealed genetic substructure aligning with the SOK, RC, GF, and WA sampling regions ($K=3$). The SOK, RC, and WA regions each were composed of distinct genetic clusters, while the GF region exhibited admixture between the SOK and RC groups, and the WA-S region displayed admixture between the SOK and WA-N regions.

Results from an AMOVA did not find significant variation among the three genetic clusters identified with STRUCTURE (4.76%, $df=2$); however, significant variation ($p<0.0001$) was detected among geographic regions within each genetic cluster (10.53%, $df=3$), as well as within sampling regions (84.72%, $df=944$) (Appendix B; Table S3.2). Estimates of pairwise differentiation (F_{ST}) between regions all were significant ($p<0.001$) and aligned with the pattern

identified with multiple iterations of STRUCTURE (Table 3.1). Among geographic regions, there was no evidence for a significant pattern of isolation by distance ($r = -0.186$, $p = 0.66$). There also was no detected pattern of isolation-by-distance among den groups range-wide ($r = -0.038$, $p = 0.50$) (Figure 3.3).

3.3.3 Within-region Population Structure and Connectivity

Within the three genetically distinct groups inferred range-wide, hierarchical STRUCTURE analyses and AMOVAs identified genetic substructure within the KAM, VER, SOK, and WA regions. A total of eleven discrete den groups were identified (Figure 3.1). Three distinct den groups ($K=3$), were detected in the KAM region (KAM-W, KAM-C, KAM-E) (Figure 3.2), while two discrete den groups ($K=2$) were found within each of VER and SOK regions, respectively (VER-KL, VER-CB; WL, OS). Within the KAM region, results of an AMOVA identified significant structure among the KAM-W, KAM-C, and KAM-E den groups (20.69%, $df=2$, $p<0.05$), as well as among dens within each group (5.29%, $df=3$, $p<0.0001$) and within individual dens (74.02%, $df=212$, $p<0.0001$) (Appendix B; Table S3.2). In the VER region, AMOVA results detected structure among the VER-CB and VER-KL den groups (10.66%, $df=1$); however, significant variation ($p<0.0001$) was only found among dens within the den groups (2.18%, $df=4$) and within dens (87.16%, $df=168$) (Appendix B; Table S3.2). For the SOK region low, but significant, variation was found among the OS and WL den groups (4.17%, $df=1$, $p<0.05$) as well as within den groups (1.17%, $df=7$) and among individuals within dens (94.66%, $df=299$) (Appendix B; Table S3.2). Within the WA region, significant genetic variation ($p < 0.05$) was detected among the WA-N and WA-S den groups (4.76%, $df=1$), as well as among dens within den groups (8.5%, $df=9$, $p<0.001$) and within individual dens (86.75%, $df=103$, $p<0.001$) (Appendix B; Table S3.2). Across all identified den groups range-wide, estimates of pairwise genetic differentiation (F_{ST}) ranged from 0.0459 to 0.526 and were all significant ($p<0.001$) (Table 3.2). Results of an AMOVA across all 11 den groups found significant ($p<0.001$) genetic variation among den groups (17.02%, $df=10$), among dens within den groups (3.20%, $df=25$), and within individual dens (79.77%, $df=886$) (Appendix B; Table S3.2).

Within regions with >1 den site, a significant pattern of isolation-by-distance was detected among dens in VER ($r = 0.873$, $p = 0.03$, SOK ($r = 0.811$, $p < 0.001$), and WA ($r = 0.57$, $p < 0.001$). Estimates of recent migration found significant gene flow from the OS den group to

WA-S and from OS to WL. All other estimates of gene flow were low ($m < 0.1$) and not significant (Table 3.3). Average pairwise relatedness was higher between snakes within a den group than among den groups. The most related den group was the KAM-E den group ($r = 0.781$) and the least related was the RC den group ($r = -0.047$) (Table 3.4). For both VER den groups, pairwise relatedness estimates within each den site were larger than those among dens. Mean relatedness within VER-CB was 0.184, ranging from 0.191-0.236 within dens, while in VER-KL the mean pairwise relatedness was 0.0666 among dens and ranged from 0.0758-0.250 within dens. For the KAM-W den group, mean pairwise relatedness was 0.143, while the mean relatedness within dens ranged from 0.132-0.147. Within the KAM-C group, the mean pairwise relatedness was 0.268 and ranged from 0.219- 0.409, while the GF den group, had a mean pairwise relatedness of -0.0330 and within den relatedness ranged from -0.0467 to 0.0445 (Table 3.4).

3.3.4 Den Level Genetic Diversity

Observed values of heterozygosity varied slightly across the range, with the highest value in the VER-CB den group ($H_o = 0.379$) and the lowest in the KAM-C group ($H_o = 0.290$) (Table 3.5). A significant heterozygote deficit was found in the WA-N, WA-S, WL, KAM-C, and KAM-W den groups, all accompanied by inbreeding coefficients significantly above zero (range: 0.0219-0.102). The VER-KL den group also exhibited a significant level of inbreeding ($F_{IS} = 0.017$) (Table 3.5), while a significant heterozygote excess was found in the VER-CB group. Across den groups the proportion of polymorphic sites within each group varied widely. The largest proportion of polymorphism was found in the OS, WL, and WA den groups ($P = 0.994-0.961$), while the least polymorphism was found in the KAM-E den group ($P = 0.302$) (Table 3.5). There was no significant negative correlation detected between latitude and proportion of polymorphism ($r^2 = 0.196$, $p = 0.10$); however, gene diversity (N_g) was significantly and negatively correlated with latitude ($r^2 = 0.365$, $p = 0.03$) (Figure 3.4).

At the individual den level, estimates of heterozygosity tended to be higher than those calculated at the den group level (Table 3.5). Values of observed heterozygosity ranged from 0.305 in KAM-C02 to 0.603 in WA-S01. Evidence for a significant heterozygote deficit was found in the WA-N03, WA-N04, WL-05, and KAM-C02 den sites, all of which also exhibited significant levels of inbreeding (Table 3.5). In contrast, the WA-N06, WA-S01, VER-KL03, VER-CB01, and VER-CB02 den sites all displayed an excess of heterozygotes, which was not

observed in WA-N, WA-S, or VER-KL at the den group level. Estimates of gene diversity and the proportion of polymorphism among individual dens showed similar patterns to those determined at the den group scale. A significant decrease in gene diversity (N_g) was found with increasing latitude ($r^2=0.311$, $p<0.001$), while there was no correlation between latitude and the proportion of polymorphism (P) found in a den site ($r^2=0.026$, $p= 0.75$) (Figure 3.4).

3.3.5 Sex Bias in Dispersal

Overall, no significant pattern of sex bias in dispersal was detected among den groups of Western Rattlesnakes. Within den groups, the only detectable bias was in the VER-KL den group where the estimate of $vAIC$ was significantly higher for females than males ($p<0.05$) (Table 3.6, Appendix B; Figure S3.2). All other comparisons showed no difference between estimates of $mAIC$ and $vAIC$ across the sexes.

3.4 Discussion

3.4.1 Population Structure and Connectivity

In this study, we found evidence for significant genetic structure and low connectivity among populations of Western Rattlesnakes sampled across Southern British Columbia and Washington State (USA). At the regional scale, each of the five main geographic regions across the distribution of rattlesnakes in BC as defined by COSEWIC were determined to be genetically distinct from one another. While the Thompson-Nicola (KAM) and the Vernon (VER) regions each exhibited a greater degree of genetic differentiation from the other regions as evidenced by the three main genetic clusters determined range-wide, the SOK, GF and RC regions also exhibited significant genetic separation from one another (Figure 3.2, Table 3.1). Additionally, no relationship between genetic and geographic distance was detected across regions, indicating the presence of barriers to gene flow across the landscape. The broad scale pattern of genetic differentiation identified across the distribution aligns with the geographic disjunction between regions due to the presence of unsuitable habitat (COSEWIC, 2015). This is especially apparent between the KAM and VER regions, where there is a large and significant level of genetic differentiation (Table 3.1) that corresponds to a large disconnect in the species range. It is thought that this separation has existed since the Holocene Climate Optimum (Hobbs, 2013),

after which temperatures began to cool roughly 6,000-5,000 years before present, the start of a period known as the “Mesothermic Interval” in BC (Walker, 2004) (Figure 1.1).

Within geographic regions, we found varying evidence for barriers to gene flow among groups of dens, similar to the pattern observed among regions (Figure 3.2). No pattern of isolation-by-distance was detected among the eleven den groups defined across the entire distribution, with low and generally insignificant estimates of migration. Particularly, there was support for the presence of barriers to gene flow among den groups identified in the KAM region. Large values of significant pairwise differentiation among den groups (Table 3.2), as well as no significant pattern of isolation-by-distance among den sites within this region are indicative of reduced population connectivity due to factors besides geographic distance. Previous studies of other snake species have found roads to not only alter snake movement and behaviour (Andrews & Gibbons, 2005; Shepard et al., 2008), but also to serve as barriers to gene flow at multiple spatial scales. For example, roads have been found to limit gene flow and connectivity among hibernacula in Timber Rattlesnakes (*C. horridus*) (Clark et al., 2010), where roads of varying traffic volumes had significantly impacted the genetic structure of populations. Roads have also been shown to impact connectivity within and among regional populations of Eastern Massasauga Rattlesnakes in Ontario, Canada (DiLeo et al., 2013), where bodies of water also were found to affect patterns of population structure. Within the KAM region, the KAM-W, and KAM-C den groups are separated by major highways such as Highway 97 and Highway 1, as well as the Thompson Rivers and Kamloops Lake, while the KAM-C and KAM-E den groups are separated by Highway 5, the North Thompson River, and extensive human development. Therefore, our findings of significant genetic structure and limited population connectivity in this region may be a result of the roads and rivers present amongst the den groups that were sampled, in addition to the geographic distances between these groups. While it is unclear as to exactly which factors are specifically responsible for the observed genetic separation in KAM, further, more spatially-explicit analyses using resistance-surface modelling could help parse apart the influence of roads, water bodies, and other landscape features on patterns of genetic differentiation in this area.

In contrast to the KAM region, our data provided weak evidence for distinct geographical or anthropogenic barriers to gene flow within both the SOK and WA regions. Despite significant estimates of pairwise differentiation between den groups within these regions (Table 3.2), all

values were low and fall within the level of ‘inbreeding connectivity’ as defined by Lowe & Allendorf (2010). This distinction suggests that gene flow may be occurring between den groups in each of these regions at a level that can counteract the negative effects of inbreeding (Lowe & Allendorf, 2010), and indicates some level of population connectivity. Within SOK, significant migration was detected between the OS and WL den groups, consistent with low estimates of differentiation, demonstrating some connectivity within this region. Significant patterns of isolation-by-distance also were detected among den sites in both SOK and WA, suggesting that geographic distance may be responsible for limiting population connectivity among dens in these regions. Interestingly, no significant genetic structure was detected among den sites within the WL den group (data not shown). In this area, the sampled den sites are situated along a road with average traffic of 350 vehicles per day, which is responsible for an average of 124 snake deaths per year (Winton et al., 2018). Despite the known effects of roads on snake behaviour and population structure (Andrews & Gibbons, 2005; Clark et al., 2010; DiLeo et al., 2013; Shepard et al., 2008), and the impact of this particular roadway on snake mortality in WL, it does not seem to be impeding genetic connectivity among den sites on opposite sides of the road as was observed in the KAM den groups. This may be due to the lower traffic volume on this road in WL as compared to the major highways in KAM, or potentially due to larger population sizes of dens in this area, that may be able to counteract the detrimental effects of road mortality on genetic connectivity. Generally, a minimum of one migrant per generation has been shown to be sufficient for maintaining genetic connectivity among subpopulations of a species (Mills & Allendorf, 1996; Wright, 1931). Given the proximity of the dens sampled in WL, it is likely that adequate movement of snakes is occurring among these populations to facilitate genetic admixture.

Consistent with previous predictions of connectivity between *C. oregonus* in BC and populations further south in the USA (Stebbins, 2003), a significant estimate of gene flow was detected among the SOK and WA regions, specifically between the OS and WA-S den groups. Given that the OS den group is closer in proximity to the WA-N den group, we expected to find that more migration would be occurring between these populations. While further investigation found a significant pattern of isolation-by distance among dens across the OS and WA-N groups ($r=0.518$, $p=0.002$), the magnitude of this relationship was low, suggesting that other factors are contributing to the genetic differentiation between OS and WA-N. In addition to the geographic

distance between these populations of rattlesnakes, there are several high-elevation mountain peaks within the Pasayten Wilderness that are situated between the OS and WA-N den groups that may be inhibiting snake movement. In order for gene flow to occur between these groups, snakes would have to successfully traverse over these mountains at elevations over 2000m, well above the elevation at which this snake species has been shown to occur in studies across its range in North America (Ashton, 2001; Gomez et al., 2015; Loughran et al., 2015; Macartney & Gregory, 1988; Maida et al., 2018; Winton et al., 2018). In contrast, elevation remains relatively low (<834m) across the landscape from OS to WA-S, better facilitating connectivity between these two areas.

Overall, estimates of migration were low to absent between most den groups (Table 3.3). While male Western Rattlesnakes have been found to move considerable distances during seasonal migrations, limits to migration behavior and den site fidelity may be limiting overall population connectivity between groups of dens across the landscape. Adult males of this species in BC have been shown to travel a maximum of ~4 kilometers from their den of origin prior to retuning to hibernate, although migration behaviour is not necessarily consistent among individual snakes or populations within the same geographical region (Harvey, 2015). Differences in migration patterns could potentially be a result of differences in prey availability, thermoregulatory opportunity, and habitat suitability, all of which are key factors for snake persistence (Lomas et al., 2015). In the present study, the straight-line distances between den groups defined across the landscape ranged from roughly 15 to 413 kilometers (Table 3.2), distances well beyond that undertaken by an individual snake in a single season. Therefore, it could take several years for snakes to make contact with those snakes in other den groups. In addition to large straight-line distances between den groups and other barriers to gene flow, the potential environmental variability among sampled regions of the range, may be decreasing the likelihood that individuals will disperse far enough to find mates from dens outside of a given den group. Based upon migration estimates, this seems probable; however, as we were unable to sample from all known den sites within each geographical region, the den groups that we defined may be excluding intermediary populations that potentially act as sources of connectivity.

Patterns of relatedness among den groups also suggest that the behavior of snakes may influence levels of migration and connectivity within geographic regions. Snakes sampled within den groups were found to be more related to each other than to other snakes across a given

region, while individuals within den sites also were found to be as related or more related to each other than to individuals from other dens. Higher within-den group relatedness suggests that snakes are exhibiting fidelity to den sites within a given area, leading to the majority of snakes mating among den sites within den groups.

3.4.2 Genetic Diversity

Across both den groups and individual dens, identified patterns of genetic variation aligned with predictions for populations that exist at the periphery of a species distribution. According to the central-marginal hypothesis, populations of a species at the boundary of their range are expected to have lower genetic diversity (Lesica & Allendorf, 1995), and exhibit a greater extent of genetic differentiation than populations towards the core (Eckert et al., 2008). Peripheral populations are thought to be smaller and susceptible to environmental pressures that differ from those found at the core of the species range. Here, we saw reduced levels of genetic variation within the northernmost region of the snake distribution, as evidenced by a significant negative relationship detected between latitude and within-population estimates of gene diversity (N_g). The proportion of polymorphism detected within den groups was also qualitatively lower at the northern extent of the species range in KAM and higher in the southern populations in SOK (Table 3.5). Based upon inferred patterns of migration and genetic differentiation, it is likely that diversity measures are higher among den groups within the SOK and WA regions due to inferred connectivity between these groups, which may extend further south into other populations of *C. oregonus* closer to the core of the range in North America.

Generally, observed levels of heterozygosity were similar to those detected for Eastern Massasauga Rattlesnakes using SNPs (Sovic et al., 2019). Although not pervasive throughout all sampled populations of rattlesnakes in this study, almost all instances of significant heterozygote deficiency within den groups and individual dens also were associated with significant levels of inbreeding. However, within the KAM-W and WA-S den groups, den sites were not found to deviate from Hardy-Weinberg expectations of heterozygosity or show evidence for significant levels of inbreeding. For the KAM-W group, the reduced heterozygosity at the den group level may be explained by the Wahlund effect (Wahlund, 1928), which results from the pooling of individuals from genetically structured groups, or in this case, dens. A low but significant ($F_{ST}=0.0763$; $p<0.001$) pairwise differentiation was found between the KAM-W01 and KAM-W02 den sites, suggesting that underlying substructure may be responsible for the reduced

heterozygosity of the den group as a whole. In the WA-S den group, the only significant ($p < 0.05$) pairwise differentiation detected was between the WA-S02 and WA-S03 den sites ($F_{ST} = 0.0813$). Rather than genetic substructure influencing heterozygosity at the den group level in WA-S, relatedness among sampled individuals may be leading to a ‘Family Wahlund Effect’ (Castric et al., 2002). In this den group, mean pairwise relatedness was higher within than among den sites, suggesting these dens may be comprised of separate family groups. As a result, the probability of having sampled sets of closely related individuals across this entire den group is higher than expected if individuals were randomly mating across den sites, thereby decreasing observed levels of heterozygosity.

A slight yet significant excess of heterozygosity was detected within the VER-CB den group, as well as within both the VER-CB01 and VER-CB02 den sites. Heterozygote excess has been shown to occur in finite populations where there may be a heterozygote advantage (Cornuet & Luikart, 1996), or the number of effective breeding individuals (N_{EB}) is small, as allele frequencies can differ among males and females due to chance (Kirby, 1975; Luikart et al., 2003; Pudovkin et al., 1996). Given negative estimates of inbreeding coefficients (Table 3.5), low genetic differentiation ($F_{ST} = 0.025$), and a very short straight-line distance between den sites in VER-CB01 and VER-CB02 (260 m), it is quite likely that there is substantial gene flow occurring between these den sites. Additionally, mean pairwise relatedness also was similar within dens and among dens (Table 3.4), indicating that individuals in this den group are strongly related. Therefore, the excess of heterozygosity found within VER-CB may be due to a small population of breeders among these two den sites, potentially the result of a population bottleneck (Cornuet & Luikart, 1996). However, given that the departure from expected heterozygosity in this den group was slight, further analysis is required to examine the cause of this observed heterozygosity excess.

3.4.3 Sex Bias in Dispersal

Contrary to initial predictions, our findings show no evidence for an overall sex bias in dispersal among adults across the sampled distribution. The only instance of a significant bias was detected from the comparison of vA_{IC} between males and females within the VER-KL den group, suggesting a female bias in dispersal among dens in this area. The lack of an observable difference in dispersal between sexes found in this system may potentially be due to the low rates of migration detected among den groups. When limited movement is occurring between

populations, the presence of rare genotypes are unlikely to affect the overall variation in population assignment between males and females, resulting in no detectable difference in dispersal (Goudet et al., 2002). In a biological context, it is also possible that there is no sex bias in dispersal for this species as mature female Western Rattlesnakes are known to adjust their behavior according to their reproductive status (Macartney & Gregory, 1988). Gravid females in BC have been found to remain within 300m of dens at rookery sites that provide both rock cover for safety as well as basking ground for adequate thermoregulation. While there is not extensive empirical data on the movement of non-gravid female snakes for this species in BC, they have been found to move comparable distances to male snakes (Bertram et al., 2001; Macartney, 1985). Therefore, non-gravid female snakes may be dispersing as much as males in a given seasonal migration both in terms of distance and frequency, resulting in no detectable sex bias in dispersal.

Additionally, the lack of a sex bias in dispersal detected in this study may be due to the inclusion of subadult snakes in our analyses. As not all sampled snakes were explicitly assigned an age class upon sampling, SVL length was used as a proxy for age. The SVL length threshold used to distinguish adults from other age classes was based upon data derived from a single area of the range (Osoyoos), and may not be representative of the body size variation that exists spatially or temporally across the range of Western Rattlesnakes in BC and Washington (Macartney et al., 1988; Winton 2018). The potential inclusion of subadult snakes may have skewed our results, in that the *hierfstat* analysis is intended to estimate population assignment for snakes post-dispersal, where unique genotypes have been brought into the population by immigrants and can be passed along to the next generation (Goudet et al., 2002). By including individuals that have not yet dispersed (i.e. neonates, juveniles, subadults) the proportion of immigrant and resident individuals detected for each sex may be biased towards residents, as subadults would likely be assigned to their population of origin. In turn, this can influence mean population assignment index values between the sexes, impacting the overall ability to detect a bias in dispersal. To this end, a more detailed study combining both empirical and genetic data from adult snakes could better resolve patterns of dispersal across the sexes. Especially with the inclusion of empirical movement data, comparisons of travel distance and frequency among males and non-gravid females could provide insight into the lack of a sex bias in dispersal determined in this study.

3.4.4 Implications for the Conservation of Snakes

Taken together, our results provide valuable genetic information that can help inform the conservation management of this snake in BC, and denning populations of snakes elsewhere. Evidence of significant population differentiation and limited population connectivity within and among the main geographic regions of rattlesnake occurrence suggest that management of this species as a single conservation unit may not be adequate for maintaining genetic diversity and population connectivity range wide. In particular, the magnitude of genetic differentiation between the KAM region, VER region, and the other sampled populations across the range may warrant the delineation of additional designatable units (DUs), although our results only provide evidence for population discreteness. Further evaluation of the genetic variation and local adaptation found within and among populations of this species is necessary to assess the evolutionary significance of populations for defining DUs.

In addition to informing conservation unit designation, the results of this study can help guide the implementation of targeted conservation management strategies for populations of this species. For most populations that occur in landscapes that become fragmented or heavily developed, re-establishing connectivity often is challenging. The main management options that exist in most situations include protecting remaining suitable habitat between areas, translocating individuals, and reintroducing captive-bred juveniles or neonates. Translocation is a common conservation management tool intended to increase the genetic diversity of populations but is known to be less effective for reptile and amphibian species (Germano & Bishop, 2009). Generally, the translocation of herpetofauna is not advised as it has been shown to alter snake movement patterns, increase rates of overwintering mortality, and has the potential to lead to outbreeding depression; however, if biologically informed, this method could prove to be successful (Reinert, 1991; Reinert & Rupert, 1999). Combined with ecological data, the genetic information about population connectivity and diversity inferred from this study can help develop translocation strategies for those populations where genetic diversity is greatly impacted by barriers to gene flow. Alternatively, captive rearing or breeding programs could be used in circumstances where translocations of adult snakes are ineffective. The introduction or translocation of captively-raised neonate snakes into wild populations has been shown to be moderately successful for Timber Rattlesnakes in Eastern Texas, with three of nine neonates surviving and maturing into the breeding population five years post-release (Conner et al., 2003).

Although the snakes in that study were not found to participate in communal denning, translocated individuals were able to locate hibernacula and survive below freezing winter temperatures. Over time, the growth of a breeding population through neonate translocation can increase genetic diversity and bolster population size, effectively providing greater resilience to stochasticity and ongoing anthropogenic development.

Overall, the continued genetic monitoring of rattlesnakes in this system can allow conservation managers to track changes in genetic diversity and population connectivity over time, as well as evaluate the outcomes of management effort. The genotyping method employed in this study (GT-seq, Campbell et al., 2015) is a reliable, quick, and relatively inexpensive method for simultaneously genotyping hundreds to thousands of individuals at hundreds of SNPs, and can be successfully applied to minimally-invasive DNA samples. As the environment is continually modified due to anthropogenic development and changing climates, leading-edge genomic tools can allow wildlife and conservation managers to survey snake populations over time and adapt conservation strategies as necessary to encourage the persistence of this species in BC.

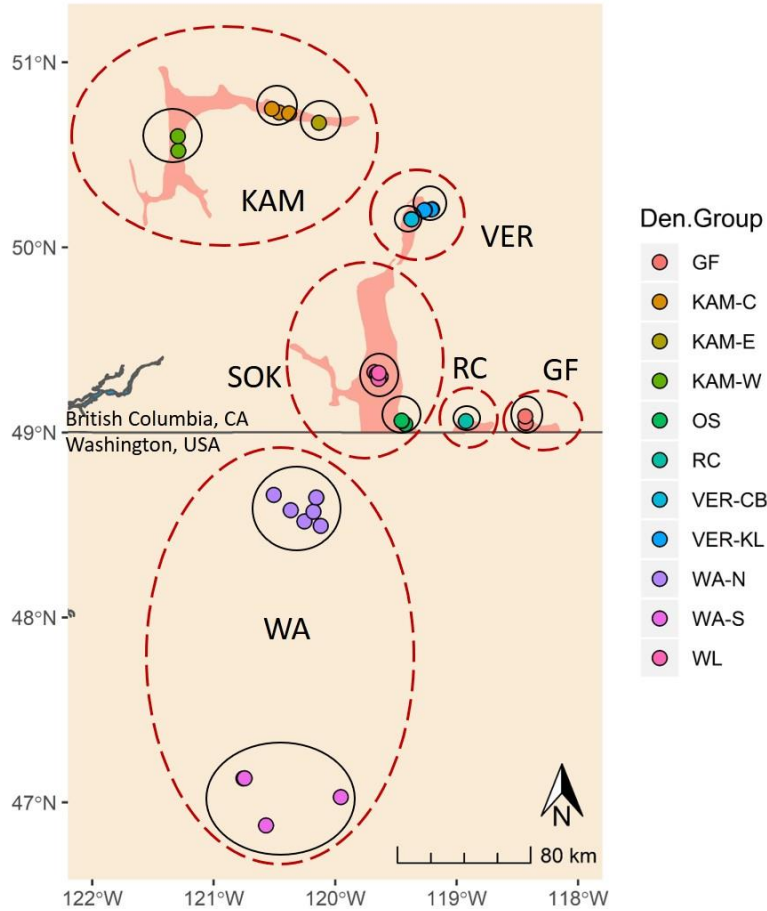


Figure 3.1 The distribution of den site sampling locations of Western Rattlesnakes in BC across the five main geographic regions: Thompson-Nicola (KAM), Vernon (VER), Okanagan-Similkameen (SOK), Midway (RC), Grand Forks (GF) and in Washington, USA (WA). Red dotted circles encompass each defined geographic region, while solid black circles designate each of the 11 den groups defined within this study. Coloured points indicate the 36 den sites sampled across den groups: KAM-W (n=2), KAM-C (n=3), KAM-E (n=1), VER-CB (n=2), VER-KL (n=4), WL (n=6), OS (n=3), RC (n=1), GF (n=3), WA-N (n=7), WA-S (n=4).

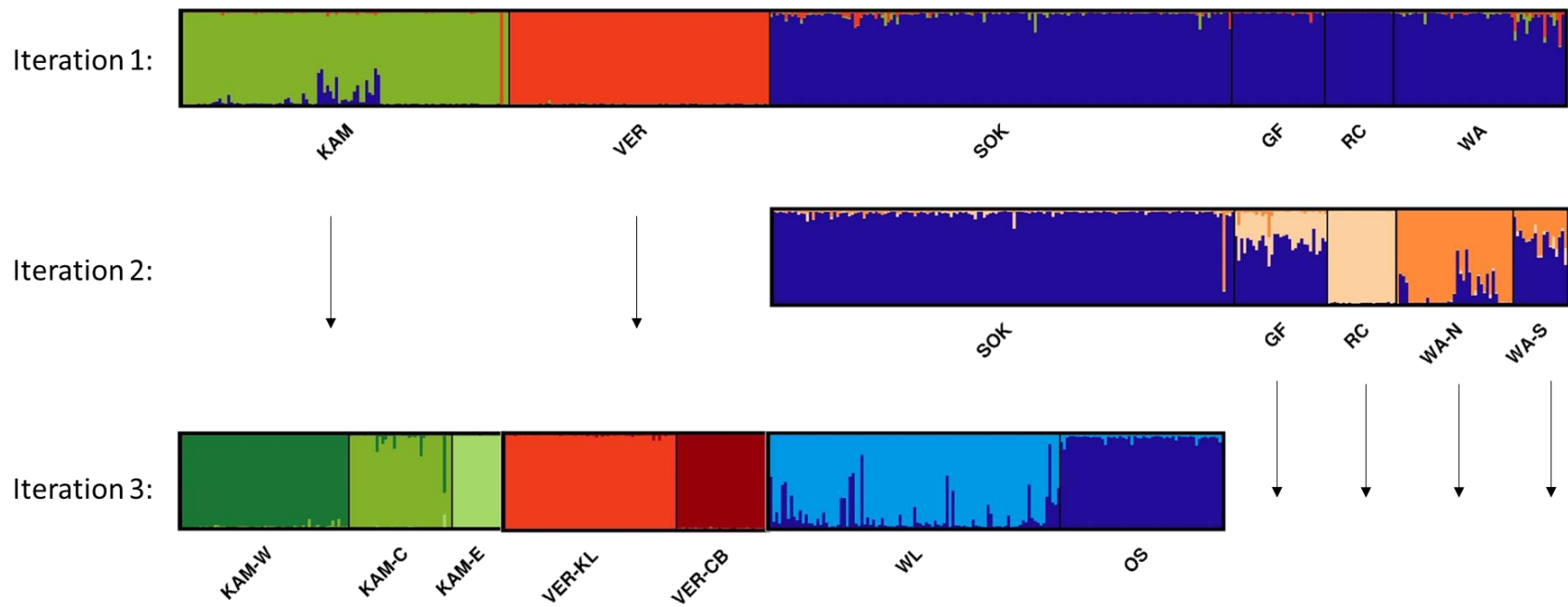


Figure 3.2 STRUCTURE bar plot displaying inferred clustering and ancestry estimation for individuals assigned to dens ($n=461$), showing $K=3$ for iteration 1, $K=3$ for iteration 2, and $K=3, 2$, and 2 , respectively for populations across iteration 3 (population acronyms as defined in Figure 3.1). Each color represents a unique genetic cluster and each vertical bar represents the proportion of ancestry to each cluster for each individual.

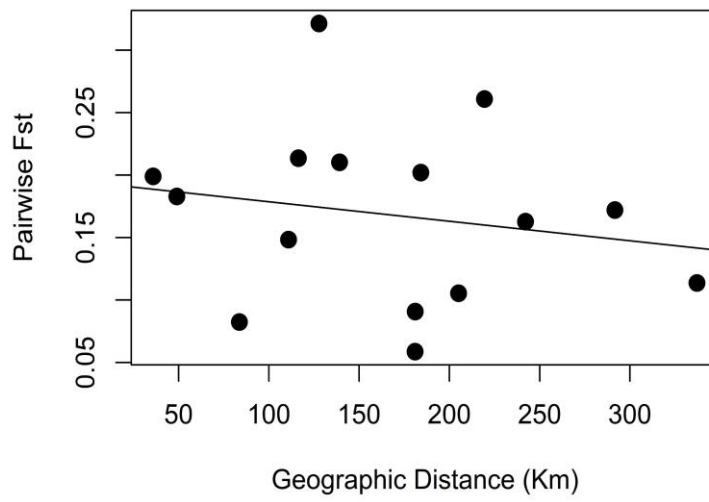
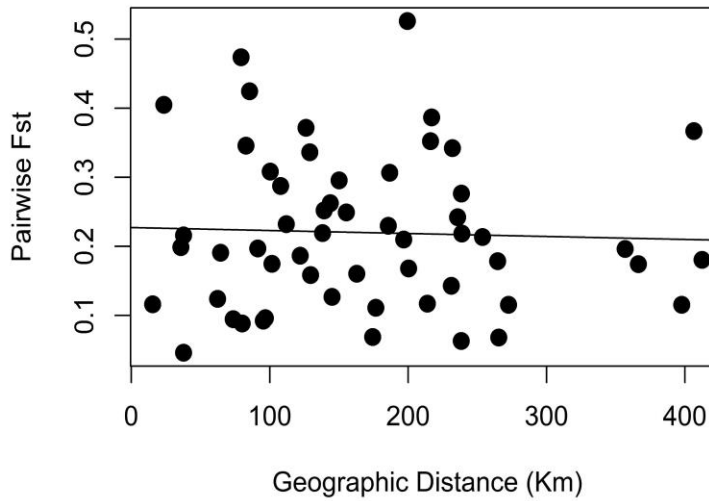
A**B**

Figure 3.3 Mantel's correlation between straight-line distance (km) and genetic distance (F_{ST}) for all 6 geographic regions ($r = -0.1863916$, $p = 0.6588$) (A) and all 11 den groups (B) ($r = -0.03812$, $p = 0.4989$).

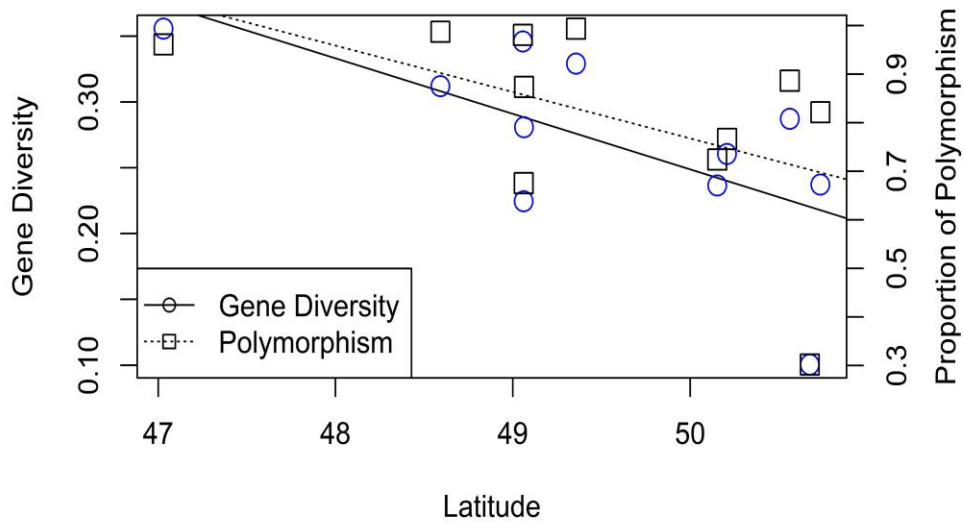
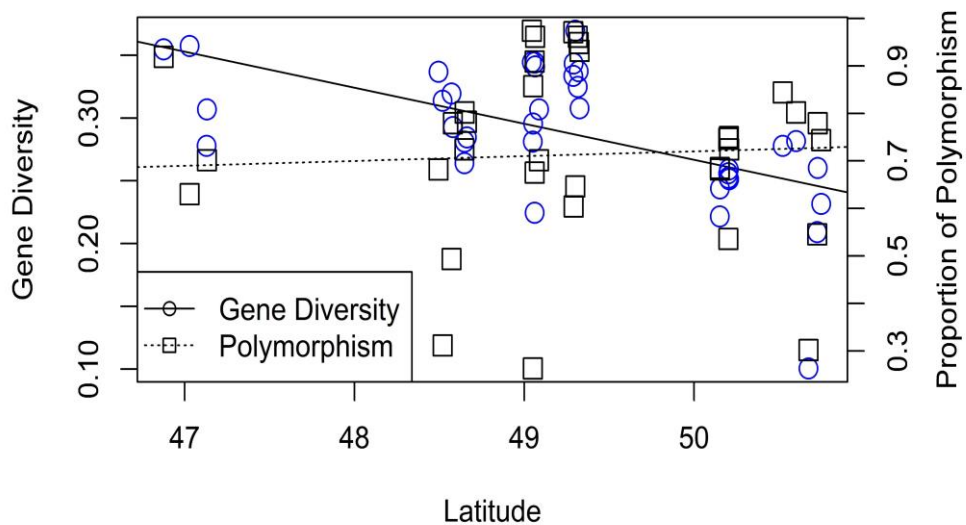
A**B**

Figure 3.4 A) Correlation between latitude and gene diversity (N_g) ($r^2= 0.365$, $p= 0.03$), and latitude and proportion of polymorphism (P) ($r^2= 0.196$, $p= 0.1$) across den groups. B) Correlation between latitude and gene diversity (N_g) ($r^2= 0.311$, $p<0.001$), and latitude and proportion of polymorphism (P) ($r^2= 0.026$, $p= 0.75$) across individual dens.

Table 3.1 Estimates of pairwise genetic distance (F_{ST}) between all 6 geographic regions across the distribution. All comparisons were significant ($p < 0.001$) with 1000 replicates.

	SOK	WA	KAM	VER	GF	RC
SOK	-----	0.059	0.091	0.148	0.082	0.183
WA		-----	0.114	0.172	0.105	0.202
KAM			-----	0.213	0.163	0.261
VER				-----	0.210	0.321
GF					-----	0.199
RC						-----

Table 3.2 Estimates of pairwise F_{ST} and straight-line distance (km) between the 11 defined den groups. Values above the diagonal represent F_{ST} and those below are straight-line distances. All pairwise comparisons of F_{ST} were significant $p < 0.001$ with 1000 replicates.

	OS	WA-N	WA-S	WL	KAM-C	KAM-W	KAM-E	VER-KL	VER-CB	GF	RC
OS	---	0.088	0.063	0.046	0.168	0.117	0.307	0.159	0.186	0.094	0.216
WA-N	80.1	---	0.069	0.092	0.219	0.143	0.342	0.210	0.230	0.127	0.232
WA-S	238.3	174.5	---	0.068	0.180	0.115	0.367	0.174	0.196	0.115	0.213
WL	37.8	95.4	265.5	---	0.160	0.111	0.296	0.175	0.197	0.096	0.191
KAM-C	200.3	238.9	412.4	162.9	---	0.124	0.405	0.287	0.308	0.242	0.352
KAM-W	213.9	231.3	397.7	176.7	62.3	---	0.346	0.249	0.263	0.178	0.276
KAM-E	186.7	232.1	406.4	150.1	23.5	82.9	---	0.424	0.474	0.387	0.526
VER-KL	129.6	196.8	366.4	101.7	107.9	155.4	85.5	---	0.116	0.219	0.336
VER-CB	122.0	185.6	356.9	91.4	100.4	143.8	79.3	15.4	---	0.252	0.372
GF	73.7	145.0	272.6	96.9	235.8	264.7	216.9	138.2	138.9	---	0.199
RC	37.8	112.0	253.8	64.5	216.2	238.5	199.5	129.0	125.7	35.9	---

Table 3.3 Estimates of contemporary migration among den groups as calculated in BAYESASS. Numbers represent the mean proportion of migrants moving from one group to another per generation averaged across 5 iterations. Italicized values along the diagonal indicate the mean proportion of non-migrants within a den group. Bolded values indicate significant estimates of migration where calculated 95% credible sets did not include zero.

		From										
		OS	WA-N	WA-S	WL	KAM-C	KAM-W	KAM-E	VER-KL	VER-CB	GF	RC
To	OS	<i>0.949</i>	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
	WA-N	0.042	<i>0.888</i>	0.013	0.010	0.007	0.007	0.007	0.007	0.007	0.007	0.007
	WA-S	0.216	0.012	<i>0.678</i>	0.014	0.012	0.012	0.011	0.011	0.012	0.011	0.012
	WL	0.065	0.006	0.003	<i>0.904</i>	0.003	0.003	0.003	0.003	0.003	0.003	0.003
	KAM-C	0.007	0.007	0.007	0.007	<i>0.917</i>	0.011	0.007	0.007	0.015	0.007	0.007
	KAM-W	0.005	0.005	0.005	0.005	0.005	<i>0.951</i>	0.005	0.005	0.005	0.005	0.005
	KAM-E	0.012	0.012	0.012	0.012	0.012	0.012	<i>0.881</i>	0.012	0.012	0.012	0.012
	VER-KL	0.005	0.005	0.005	0.005	0.005	0.005	0.005	<i>0.951</i>	0.005	0.005	0.005
	VER-CB	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.009	<i>0.918</i>	0.008	0.008
	GF	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	<i>0.921</i>	0.008
	RC	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	<i>0.902</i>

Table 3.4 Estimates of Queller and Goodnight's pairwise relatedness (rQG) between pairs of individuals within each den group and within each den.

Den Group	Den	n	rQG
KAM-W		59	0.143
	KAMW-01	36	0.147
	KAMW-02	23	0.133
KAM-C		36	0.268
	KAM-C01	9	0.409
	KAM-C02	18	0.247
	KAM-C03	9	0.219
KAM-E	KAM-E01	17	0.781
VER-CB		38	0.184
	VER-CB01	13	0.236
	VER-CB02	25	0.191
VER-KL		63	0.067
	VER-KL01	21	0.121
	VER-KL02	24	0.087
	VER-KL03	3	0.250
	VER-KL04	15	0.076
OS		57	0.054
	OS-01	30	0.068
	OS-02	8	0.069
	OS-03	19	0.056
WL		105	0.028
	WL-01	28	0.023
	WL-02	26	0.065
	WL-03	22	0.089
	WL-04	2	-0.077
	WL-05	25	0.037
	WL-06	2	0.078
WA-N		40	0.050
	WA-N01	3	0.077
	WA-N02	1	
	WA-N03	10	0.235
	WA-N04	9	0.202
	WA-N05	3	0.119
	WA-N06	5	0.135
	WA-N07	9	0.228
WA-S		18	0.030
	WA-S01	2	0.142
	WA-S02	11	0.055
	WA-S03	4	0.253
	WA-S04	1	

Den Group	Den	n	rQG
GF		32	-0.033
	GF-01	12	0.044
	GF-02	19	-0.047
	GF-03	1	
RC	RC-01	23	-0.047

Note: n refers to the number of individuals within a specified den group or den site.

Table 3.5 Estimates of mean observed and expected heterozygosity, proportion of polymorphism (P), nucleotide diversity (N_g), and inbreeding coefficients (F_{IS}) for each den group and dens within each den group. * indicates $p < 0.05$, ** indicates $p < 0.001$. Bolded values indicate a significant excess of heterozygosity.

Den Group	Den	n	P	Mean H_o	Mean H_e	F_{IS}	N_g
OS		55	0.981	0.362	0.360	-0.005	0.346
	OS-01	28	0.974	0.365	0.360	-0.014	0.344
	OS-02	8	0.909	0.378	0.382	0.011	0.344
	OS-03	19	0.961	0.376	0.371	-0.012	0.341
WA-N		39	0.987	0.312*	0.346	0.102**	0.312
	WA-N01	3	0.682	0.452	0.468	0.054	0.337
	WA-N02	1	0.312	1.000	1.000	0.000	0.314
	WA-N03	9	0.724	0.352*	0.375	0.084*	0.264
	WA-N04	9	0.802	0.353**	0.368	0.043*	0.281
	WA-N05	3	0.494	0.454	0.454	-0.025	0.319
	WA-N06	5	0.779	0.442*	0.418	-0.071	0.293
	WA-N07	9	0.782	0.369	0.380	0.027	0.285
WA-S		18	0.961	0.363*	0.375	0.031*	0.356
	WA-S01	2	0.630	0.603*	0.564	-0.109	0.357
	WA-S02	11	0.919	0.392	0.390	-0.005	0.355
	WA-S03	4	0.701	0.456	0.441	-0.028	0.307
	WA-S04	1	0.276	1.000	1.000	0.000	0.278
WL		99	0.994	0.349*	0.357	0.022**	0.329
	WL-01	28	0.971	0.357	0.362	0.012	0.334
	WL-02	26	0.961	0.360	0.357	-0.007	0.325
	WL-03	21	0.932	0.364	0.361	-0.002	0.308
	WL-04	2	0.646	0.550	0.555	0.013	0.370
	WL-05	20	0.948	0.354**	0.370	0.046*	0.337
	WL-06	2	0.604	0.554	0.552	-0.005	0.343
KAM-W		57	0.886	0.331*	0.341	0.026**	0.287
	KAM-W01	36	0.844	0.340	0.337	-0.010	0.278
	KAM-W02	21	0.802	0.367	0.361	-0.014	0.282
KAM-C		35	0.821	0.290**	0.307	0.062**	0.237
	KAM-C01	9	0.545	0.365	0.359	0.004	0.209
	KAM-C02	9	0.779	0.305**	0.340	0.111*	0.260
	KAM-C03	17	0.744	0.344	0.344	0.003	0.232
KAM-E	KAM-E01	17	0.302	0.363	0.353	-0.033	0.100
VER-KL		57	0.766	0.338	0.343	0.017*	0.260
	VER-KL01	17	0.724	0.349	0.351	0.003	0.251
	VER-KL02	23	0.750	0.348	0.343	-0.013	0.259
	VER-KL03	3	0.536	0.499*	0.465	-0.078	0.252
	VER-KL04	14	0.747	0.360	0.367	0.018	0.256

Den Group	Den	n	P	Mean Ho	Mean He	FIS	Ng
VER-CB		30	0.724	0.379*	0.353	-0.070	0.237
	VER-CB01	17	0.685	0.401**	0.376	-0.063	0.222
	VER-CB02	13	0.682	0.390**	0.351	-0.109	0.244
GF		31	0.873	0.335*	0.341	0.021	0.281
	GF-01	12	0.776	0.361	0.356	-0.020	0.257
	GF-02	18	0.857	0.346	0.355	0.024	0.296
	GF-03	1	0.263	1.000	1.000	0.000	0.281
RC	RC-01	23	0.675	0.349	0.350	0.002	0.225

Note: n refers to the number of individuals within a specified den group or den site.

Table 3.6 Comparison of mAIC and vAIC for adult male and female snakes among all den groups and within den groups containing at least 2 den sites with at least 2 males and females at each. * denotes $p < 0.05$ with 10,000 replicates.

Among den groups (n=11)	Sex	n	mAIC	vAIC
	M	171	0.549	151.546
	F	166	-0.559	147.920
Within den groups				
GF	M	18	-1.006	77.628
	F	9	2.011	58.310
KAM-C	M	8	-6.671	166.884
	F	20	2.668	96.085
KAM-W	M	20	1.394	71.188
	F	21	-1.328	65.362
OS	M	22	1.132	54.611
	F	10	-2.490	30.144
VER-CB	M	10	-2.439	66.587
	F	10	2.439	107.778
VER-KL	M	20	0.390	40.655
	F	23	-0.337	109.276*
WA-N	M	12	4.008	100.116
	F	5	-9.620	665.815
WL	M	36	-0.863	165.582
	F	44	0.706	130.672

Note: n refers to the number of individuals per sex within a specified den group.

Chapter 4: Conclusions

4.1 Findings and Significance

Understanding the population dynamics of species found across fragmented landscapes provides information useful for defining conservation units, which are critical for the effective management of species at-risk. In cases where species of conservation concern are rare, elusive, or difficult to monitor empirically, genetic assessments can be used to make inferences about population ecology to inform management. In this study, we conducted the first genetic investigation of populations of the Western Rattlesnake (*Crotalus oreganus*), a species of elevated conservation status in British Columbia, Canada. Through evaluating patterns of genetic diversity and structure within and among populations of this species across BC, we discovered distinct genetic groups range-wide, and across multiple spatial scales that experience limited connectivity and gene flow.

At the broad scale, patterns of genetic diversity and structure detected among Western Rattlesnakes in BC can help inform conservation efforts for this species. The population structure found among the Thompson-Nicola, Vernon, Okanagan-Similkameen, Midway, and Grand Forks geographic regions can inform future species status assessments for delineating Designatable Units (DUs), the conservation units adopted by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC, 2015). Specifically, we provide evidence of population discreteness for multiple geographic areas across the range, one of two main criteria used to define DUs. Although more information is required to identify separate populations that may warrant specific conservation attention, the potential designation of additional conservation units for the Western Rattlesnake in BC can help guide the preservation of diversity among populations found across the species distribution.

As a whole, this study further demonstrates the utility of combining next generation sequencing techniques with minimally invasive DNA samples (MIS). While more traditional microsatellite markers have successfully been genotyped in MIS, the ability to genotype SNP markers from MIS both increases the quality of data collected, as well as provides information spanning a broader range of the genome (Morin et al., 2004). In addition, the multiplexed approach used in this study allows for the simultaneous genotyping of hundreds to thousands of samples from a single sequencing run, making this method time and cost effective. The

efficiency of this method can enable the continuous monitoring of populations of species at-risk over time, allowing managers to adjust conservation strategies as necessary to maintain appropriate genetic variation.

4.1 Future Directions

In Canada, COSEWIC status assessments of species to help define designatable units (DUs) for conservation require evidence for both population discreteness and evolutionary significance (Green, 2005). Despite the indications of discreteness detected in this study, our ability to inform the designation of conservation units for this species is limited by our lack of evidence for evolutionary significance. Future genome-wide studies of Western Rattlesnakes in BC can help fill this knowledge gap by examining populations for evidence of local adaptation, as well as historical range disjunctions that could provide further support for additional DUs.

In particular, genome-wide approaches that involve the detection of ‘outlier’ loci, or those SNPs under selection, allow for the quantification of putative genetic variation that may reflect local adaptation in sampled populations (Allendorf et al., 2010). Coupled with the population distinctiveness detected in this study, characterizing patterns of adaptive divergence range wide can help identify evolutionarily significant populations that may warrant designatable unit designation (Funk et al., 2012). While we did not examine patterns of local adaptation using the genomic data collected from blood samples with RADseq (see Chapter 2), further analysis to detect outliers in this dataset using high quality DNA samples could provide initial targets for future investigation of adaptive variation.

Moreover, our results highlight how barriers to gene flow and their impacts across a landscape are context dependent. The levels of genetic differentiation and patterns of population connectivity identified in our study varied across geographic areas and were found to coincide with the presence of roads, waterways, elevation, unsuitable habitat, and geographic distance among populations. However, with the population genetic analyses employed here, we were unable to identify the influence of these factors to the maintenance of gene flow among populations of Western Rattlesnakes. A more comprehensive landscape genetics study could help evaluate the presence and severity of various barriers to gene flow and how they impact the genetic diversity of rattlesnakes in BC. Compared to the tests for patterns of isolation-by-distance utilized here, patterns of isolation-by-resistance (McRae, 2006) may be more informative in discerning the causes of genetic differentiation across the distribution, as diverse

landscape features will have differing impacts on the ability for snakes to disperse. For instance, Clark et al. (2008) found a greater correlation between genetic and geographic distance along the ‘path of least resistance’ between populations of *C. horridus*, a pattern that may also exist among populations of *C. oreganus* in BC. However, as other aspects of rattlesnake biology can confound patterns of dispersal and genetic connectivity, further investigation into the spatial and dispersal ecology of this species may be warranted to complement results of landscape genetic analyses.

Along with evaluating the severity of barriers to gene flow across the range of rattlesnakes in BC, further examination genome wide can also assess the contemporary and historical factors and processes that may shape genetic variation among populations. As the analyses used to infer patterns of population structure and connectivity throughout this thesis only provide estimates of current genetic differences among populations, they offer no information regarding the evolution of this variation (Sirén et al., 2011). Along with standard population genetics, the integration of methods such as Approximate Bayesian Coalescent (ABC) modeling (Beaumont et al., 2002) and Ecological Niche Modeling (ENM) can be employed to test hypotheses regarding how different demographic scenarios, environmental changes, and shifts in species distributions may have led to observed patterns of population genetic structure (Csilléry et al., 2010; He et al., 2013; Reid et al., 2019). ABC modeling also allows for the estimation of effective population sizes (Tallmon et al., 2004) and admixture among populations (Excoffier et al., 2005), as well as the inference of demographic history (Segelbacher et al., 2010), all of which can provide critical insight for the conservation of species at risk (Frankham et al., 2002). For the Western Rattlesnake, this type of modeling can help distinguish between the impacts of natural and anthropogenic processes on the genetic differentiation of populations to help define conservation units and prioritize management action.

Notably, the minimally-invasive buccal swab samples included in this study were successfully amplified and genotyped with GT-seq. A previous assessment conducted by Ford et al. (2017) found buccal swab samples collected from Western Rattlesnakes to be sub-optimal for collecting genotypes using microsatellites, producing the highest amount of genotyping error of all sample types tested. In this study, however, buccal samples showed similar amplification success compared to all other samples genotyped with GT-seq (data not shown). As buccal swabs were obtained from juvenile and neonate snakes that were too small for cloacal swab

sampling, the success of applying GT-seq to these samples allows for snakes of all age classes to be included in future genetic analyses. The inclusion of various age classes can enable further study into the demographics of sampled populations to test hypotheses regarding natal recruitment and dispersal at den sites, as well as other aspects of population ecology. More broadly, genotyping individuals using buccal swabs also has implications for a host of other species where this type of sampling is common, including birds, amphibians, and other reptiles (Broquet et al., 2007; Handel et al., 2006; Miller, 2006).

In summary, this study presents the first genetic assessment of the Western Rattlesnake in the northwest portion of the species range. Here we detected patterns of population structure within and among regions of *C. oreganus* occurrence in BC and Washington, as well as found evidence for limited population connectivity among groups of den sites. Together, these results help to fulfill the population discreteness criteria for informing conservation unit delineation in Canada and provide a basis for future genetic studies of this species to assess evolutionary significance. This study also highlights the utility of applying GT-seq, a targeted amplicon sequencing approach, to minimally invasive samples to support studies in conservation and molecular ecology.

References

- Adams JR, Kelly BT, Waits LP (2003) Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). *Molecular Ecology*, 12, 2175–2186.
- Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11, 697–709.
- Ando A, Camm J, Polasky S, Solow A (1998) Species Distributions, Land Values, and Efficient Conservation. *Science*, 279, 2126–2128.
- Andrews KM, Gibbons JW (2005) How Do Highways Influence Snake Movement? Behavioral Responses to Roads and Vehicles. *Copeia*, 2005, 772–782.
- Andrews KR, Barba MD, Russello MA, Waits LP (2018) Advances in using non-invasive, archival, and environmental samples for population genomic studies. In P. Hohenlohe & O. Rajora (Eds.), *Population Genomics: Wildlife* (pp. 1–37) Springer
- Andrews S (2010) FastQC: A quality control tool for high throughput sequence data.
- Ashton KG (2001) Body Size Variation Among Mainland Populations of the Western Rattlesnake (*Crotalus viridis*). *Evolution*, 55, 2523–2533.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, ... Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, 3, 1–7.
- Barba MD, Waits LP, Genovesi P, Randi E, Chirichella R, Cetto E (2010) Comparing opportunistic and systematic sampling methods for non-invasive genetic monitoring of a small translocated brown bear population. *Journal of Applied Ecology*, 47, 172–181.
- Bayerl H, Kraus RHS, Nowak C, Foerster DW, Fickel J, Kuehn R (2018) Fast and cost-effective single nucleotide polymorphism (SNP) detection in the absence of a reference genome using semideep next-generation Random Amplicon Sequencing (RAMseq). *Molecular Ecology Resources*, 18, 107–117.
- Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian Computation in Population Genetics. *Genetics*, 162, 2025–2035.
- Beebee TJC (2013) Effects of Road Mortality and Mitigation Measures on Amphibian Populations: Amphibians and Roads. *Conservation Biology*, 27, 657–668.

- Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources*, 9, 1279–1301.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, Windows TM software for population genetics. Laboratoire Génome, Populations, Interactions, CNRS UMR
- Bellemain E, Swenson JE, Tallmon D, Brunberg S, Taberlet P (2005) Estimating Population Size of Elusive Animals with DNA from Hunter-Collected Feces: Four Methods for Brown Bears. *Conservation Biology*, 19, 150–161.
- Bertram N, Larsen KW, Surgenor J (2001) Identification of critical habitats and conservation issues for the Western Rattlesnake and Great Basin Gopher Snake within the Thompson-Nicola region of British Columbia. Prepared for the Ministry of Water, Land and Air Protection, Kamloops, BC and the Habitat Conservation Trust Fund of BC, Victoria, BC
- Bivand R, Rundel C (2014) Rgeos: Interface to Geometry Engine–Open Source (GEOS). R Package Version 0.3–6. Vienna: Comprehensive R Archive Network. Software
- Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, ... Bruyn MD (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, 29, 358–367.
- Brooks RJ, Brown GP, Galbraith DA (1991) Effects of a sudden increase in natural mortality of adults on a population of the common snapping turtle (*Chelydra serpentina*). *Canadian Journal of Zoology*, 69, 1314–1320.
- Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics*, 8, 509–511.
- Brown GW, Bennett AF, Potts JM (2008) Regional faunal decline – reptile occurrence in fragmented rural landscapes of south-eastern Australia. *Wildlife Research*, 35, 8.
- Campbell NR, Harmon SA, Narum SR (2015) Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, 15, 855–867.
- Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, Nickerson DA (2003) Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nature Genetics*, 33, 518–521.

- Carroll EL, Bruford MW, DeWoody JA, Leroy G, Strand A, Waits L, Wang J (2018) Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary Applications*
- Castric V, Bernatchez L, Belkhir K, Bonhomme F (2002) Heterozygote deficiencies in small lacustrine populations of Brook Charr *Salvelinus fontinalis* Mitchell (Pisces, Salmonidae): A test of alternative hypotheses. *Heredity*, 89, 27–35.
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140.
- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM (2015) Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1, e1400253.
- Chiou KL, Bergey CM (2018) Methylation-based enrichment facilitates low-cost, noninvasive genomic scale sequencing of populations from feces. *Scientific Reports*, 8, 1975.
- Clark RW, Brown WS, Stechert R, Zamudio KR (2008) Integrating individual behaviour and landscape genetics: The population structure of Timber Rattlesnake hibernacula. *Molecular Ecology*, 17, 719–730.
- Clark RW, Brown WS, Stechert R, Zamudio KR (2010) Roads, Interrupted Dispersal, and Genetic Diversity in Timber Rattlesnakes: Roads and Population Genetics. *Conservation Biology*, 24, 1059–1069.
- Conner RN, Rudolph DC, Saenz D, Schaefer RR, Burgdorf SJ (2003) Growth rates and post-release survival of captive neonate Timber Rattlesnakes, *Crotalus horridus*. *Herpetological Review*, 34, 314.
- Cornuet JM, Luikart G (1996) Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. *Genetics*, 144, 2001–2014.
- COSEWIC (2015) Western rattlesnake (*Crotalus oreganus*). Ottawa: Environment Canada
- Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, 15, 290–295.
- Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, 25, 410–418.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, ... Durbin R (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.

- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12, 499–510.
- Díaz S, Fargione J, Iii FSC, Tilman D (2006) Biodiversity Loss Threatens Human Well-Being. *PLOS Biology*, 4, e277.
- DiLeo MF, Rouse JD, Dávila JA, Lougheed SC (2013) The influence of landscape on gene flow in the eastern massasauga rattlesnake (*Sistrurus c. Catenatus*): Insight from computer simulations. *Molecular Ecology*, 22, 4483–4498.
- Dixon JD, Oli MK, Wooten MC, Eason TH, McCown JW, Cunningham MW (2007) Genetic consequences of habitat fragmentation and loss: The case of the Florida Black Bear (*Ursus americanus floridanus*). *Conservation Genetics*, 8, 455–464.
- Dray S, Dufour A-B (2007) The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, 22.
- Driscoll DA (2004) Extinction and Outbreaks Accompany Fragmentation of a Reptile Community. *Ecological Applications*, 14, 220–240.
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361.
- Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: The central–marginal hypothesis and beyond. *Molecular Ecology*, 17, 1170–1188.
- Eklom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107, 1–15.
- Environment and Climate Change Canada (2009, September 1) Species at Risk Act: Description. [acts]
- Environment and Climate Change Canada (2017, June 7) COSEWIC guidelines for recognizing designatable units. [guidance]
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14, 2611–2620.
- Excoffier L, Estoup A, Cornuet J-M (2005) Bayesian Analysis of an Admixture Model With Mutations and Arbitrarily Linked Markers. *Genetics*, 169, 1727–1738.

- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.
- Fahrig L, Merriam G (1985) Habitat patch connectivity and population survival: Ecological archives e066-008. *Ecology*, 66, 1762–1768.
- Fahrig L, Merriam G (1994) Conservation of Fragmented Populations. *Conservation Biology*, 8, 50–59.
- Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocicidura russula*: Evidence from field data and microsatellite patterns. *Proceedings of the Royal Society B: Biological Sciences*, 264, 127–132.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.
- Ford B, Govindarajulu P, Larsen K, Russello M (2017) Evaluating the efficacy of non-invasive genetic sampling of the Northern Pacific rattlesnake with implications for other venomous squamates. *Conservation Genetics Resources*, 9, 13–15.
- Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies*, 326, 22–29.
- Frankham R, Ballou J, Briscoe D (2002) *Introduction to Conservation Genetics*. (Vol. 617) Cambridge University Press
- Funk WC, Mckay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, 27, 489–496.
- Gerlach G, Musolf K (2000) Fragmentation of Landscape as a Cause for Genetic Subdivision in Bank Voles. *Conservation Biology*, 14, 1066–1074.
- Germano JM, Bishop PJ (2009) Suitability of Amphibians and Reptiles for Translocation. *Conservation Biology*, 23, 7–15.
- Gibbons JW, Scott DE, Ryan TJ, Buhlmann KA, Tuberville TD, Metts BS, ... Winne CT (2000) The Global Decline of Reptiles, Déjà Vu Amphibians Reptile species are declining on a global scale. *BioScience*, 50, 653–666.
- Gienger CM, Beck DD (2011) Northern Pacific Rattlesnakes (*Crotalus oreganus*) use thermal and structural cues to choose overwintering hibernacula. *Canadian Journal of Zoology*, 89, 1084–1090.

- Gomez (2008) Habitat use and movement patterns of the Northern Pacific rattlesnake (*Crotalus o. oregonus*) in British Columbia. Library and Archives Canada = Bibliothèque et Archives Canada, Ottawa
- Gomez L, Larsen K, Gregory P (2015) Contrasting Patterns of Migration and Habitat Use in Neighboring Rattlesnake Populations. *Journal of Herpetology*, 49, 371–376.
- Goudet J (2005) Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184–186.
- Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology*, 11, 1103–1114.
- Graham CF, Glenn TC, McArthur AG, Boreham DR, Kieran T, Lance S, ... Somers CM (2015) Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, 15, 1304–1315.
- Green DM (2005) Designatable Units for Status Assessment of Endangered Species. *Conservation Biology*, 19, 1813–1820.
- Haddad NM, Brudvig LA, Clobert J, Davies KF, Gonzalez A, Holt RD, ... Townshend JR (2015) Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, 1, e1500052.
- Handel CM, Pajot LM, Talbot SL, Sage GK (2006) Use of Buccal Swabs for Sampling DNA from Nestling and Adult Birds. *Wildlife Society Bulletin*, 34, 1094–1100.
- Hansson L (1991) Dispersal and connectivity in metapopulations. *Biological Journal of the Linnean Society*, 42, 89–103.
- Hardy OJ, Vekemans X (2002) SPAGeDI: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, 2, 618–620.
- Harvey JA (2015) Thermal Influences on Summer Habitat Use by Western Rattlesnakes (*Crotalus oregonus*) in British Columbia.
- He Q, Edwards DL, Knowles LL (2013) Integrative Testing of How Environments from the Past to the Present Shape Genetic Structure Across Landscapes. *Evolution*, 67, 3386–3402.
- Hess JE, Campbell NR, Docker MF, Baker C, Jackson A, Lampman R, ... Narum SR (2015) Use of genotyping by sequencing data to develop a high-throughput and multifunctional SNP

- panel for conservation applications in Pacific lamprey. *Molecular Ecology Resources*, 15, 187–202.
- Hobbs J (2013) Den survey and population assessment of the Northern Pacific Rattlesnake in BC: Final Report. BC Ministry of Forests, Lands and Natural Resource Operations, Victoria, British Columbia
- Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLOS Genetics*, 6, e1000862.
- Hunter ME, Hoban S, Bruford MW, Segelbacher G, Bernatchez L (2018) Next Generation Conservation Genetics and Biodiversity Monitoring. *Evolutionary Applications*, 0–2.
- IUCN (2019) The IUCN Red List of Threatened Species.
- Jakobsson M, Rosenberg NA (2007) CLUMPP : a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. , 23, 1801–1806.
- Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL (2017) The K = 2 conundrum. *Molecular Ecology*, 26, 3594–3602.
- Jones R, Cable J, Bruford MW (2008) An evaluation of non-invasive sampling for genetic analysis in northern European reptiles. *The Herpetological Journal*, 18, 32–39.
- Kalinowski ST, Wagner AP, Taper ML (2006) ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, 6, 576–579.
- Kersey DC, Dehnhard M (2014) The use of noninvasive and minimally invasive methods in endocrinology for threatened mammalian species conservation. *General and Comparative Endocrinology*, 203, 296–306.
- King MB, Duvall D (1990) Prairie rattlesnake seasonal migrations: Episodes of movement, vernal foraging and sex differences. *Animal Behaviour*, 39, 924–935.
- Kirby GC (1975) Heterozygote frequencies in small subpopulations. *Theoretical Population Biology*, 8, 31–48.
- Kraus RHS, vonHoldt B, Cocchiara B, Harms V, Bayerl H, Kühn R, ... Nowak C (2015) A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf

- monitoring in Europe based on noninvasively collected samples. *Molecular Ecology Resources*, 15, 295–305.
- Lanci AKJ, Roden SE, Bowman A, LaCasella EL, Frey A, Dutton PH (2012) Evaluating Buccal and Cloacal Swabs for Ease of Collection and Use in Genetic Analyses of Marine Turtles. *Chelonian Conservation and Biology*, 11, 144–148.
- Latch EK, Boarman WI, Walde A, Fleischer RC (2011) Fine-Scale Analysis Reveals Cryptic Landscape Genetic Structure in Desert Tortoises. *PLoS ONE*, 6, e27794.
- Lehmkuhl JF, Ruggiero LF (1991) Forest fragmentation in the Pacific Northwest and its potential effects on wildlife. *Wildlife and Vegetation of Unmanaged Douglas-Fir Forests*, 35–46.
- Lemay MA, Russello MA (2015) Genetic evidence for ecological divergence in kokanee salmon. *Molecular Ecology*, 24, 798–811.
- Lesbarrères D, Ashpole SL, Bishop CA, Blouin-Demers G, Brooks RJ, Echaubard P, ... Lougheed SC (2014) Conservation of herpetofauna in northern landscapes: Threats and challenges from a Canadian perspective. *Biological Conservation*, 170, 48–55.
- Lesica P, Allendorf FW (1995) When Are Peripheral Populations Valuable for Conservation? *Conservation Biology*, 9, 753–760.
- Lew RM, Finger AJ, Baerwald MR, Goodbla A, May B, Meek MH (2015) Using Next-Generation Sequencing to Assist a Conservation Hatchery: A Single-Nucleotide Polymorphism Panel for the Genetic Management of Endangered Delta Smelt. *Transactions of the American Fisheries Society*, 144, 767–779.
- Lomas E, Larsen KW, Bishop CA (2015) Persistence of Northern Pacific Rattlesnakes masks the impact of human disturbance on weight and body condition: Effects of human disturbance on rattlesnakes. *Animal Conservation*, 18, 548–556.
- Longmire JL, Maltbie M, Baker RJ (1997) Use of “lysis buffer” in DNA isolation and its implication for museum collections. *Occasional Papers, Museum of Texas Tech University*, 163, 1–3.
- Loughran CL, Beck DD, Weaver RE (2015) Use of Communal Shedding Sites by the Northern Pacific Rattlesnake (*Crotalus Oreganus Oreganus*) in Central Washington State. *Northwestern Naturalist*, 96, 156–160.

- Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, 19, 3038–3051.
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, 4, 981–994.
- Luiselli L, Capizzi D (1997) Influences of area, isolation and habitat features on distribution of snakes in Mediterranean fragmented woodlands. *Biodiversity & Conservation*, 6, 1339–1351.
- Lukacs PM, Burnham KP (2005) Review of capture–recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, 14, 3909–3919.
- Macartney (1985) The ecology of the Northern Pacific Rattlesnake, *Crotalus viridis oregonus*, in British Columbia. M.Sc. Thesis, University of Victoria, Victoria, B. C .
- Macartney JM (1989) Diet of the Northern Pacific Rattlesnake, *Crotalus viridis oregonus*, in British Columbia. *Herpetologica*, 45, 299–304.
- Macartney JM, Gregory PT (1988) Reproductive Biology of Female Rattlesnakes (*Crotalus viridis*) in British Columbia. *Copeia*, 1988, 47–57.
- Macartney JM, Gregory PT, Charland MB (1990) Growth and Sexual Maturity of the Western Rattlesnake, *Crotalus viridis*, in British Columbia. *Copeia*, 1990, 528.
- Macartney JM, Gregory PT, Larsen KW (1988) A Tabular Survey of Data on Movements and Home Ranges of Snakes. *Journal of Herpetology*, 22, 61.
- Magin CD, Johnson TH, Groombridge B, Jenkins M, Smith H (1994) Species extinctions, endangerment and captive breeding. In P. J. S. Olney, G. M. Mace, & A. T. C. Feistner (Eds.), *Creative Conservation: Interactive management of wild and captive animals* (pp. 3–31) Dordrecht: Springer Netherlands
- Maida JR, Kirk DA, McKibbin O, Row JR, Larsen KW, Stringam C, Bishop CA (2018) Population Estimate, Survivorship, and Generation Time of the Northern Pacific Rattlesnake (*Crotalus o. Oregonus*) at its Northern-most Range Limits. , 12.
- Mantel N (1967) The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, 27, 209–220.

- McAllister JM, Maida JR, Dyer O, Larsen KW (2016) Diet of Roadkilled Western Rattlesnakes (*Crotalus oreganus*) and Gophersnakes (*Pituophis catenifer*) in Southern British Columbia. *Northwestern Naturalist*, 97, 181–189.
- McRae BH (2006) Isolation by Resistance. *Evolution*, 60, 1551–1561.
- Meek MH, Larson WA (2019) The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, 19, 795–803.
- Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4, 792–794.
- Miller HC (2006) Cloacal and buccal swabs are a reliable source of DNA for microsatellite genotyping of reptiles. *Conservation Genetics*, 7, 1001–1003.
- Mills LS, Allendorf FW (1996, December 1) The One-Migrant-per-Generation Rule in Conservation and Management.
- Mittermeier RA, Carr JL, Swingland IR, Werner TB, Mast RB (1992) Conservation of amphibians and reptiles. *Herpetology: Current Research on the Biology of Amphibians and Reptiles*. K. Adler (Ed.). Society for the Study of Amphibians and Reptiles Publication, Missouri, 59–80.
- Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19, 208–216.
- Moritz C (1994) Defining “Evolutionarily Significant Units” for conservation. *Trends in Ecology & Evolution*, 9, 373–375.
- Murphy JB (1971) A method for immobilising venomous snakes at Dallas Zoo. *International Zoo Yearbook*, 11, 233–233.
- NatureServe and IUCN (International Union for Conservation of Nature) 2007. *Crotalus oreganus*. The IUCN Red List of Threatened Species. Version 2019-2. <http://www.iucnredlist.org>. Downloaded on 06 August 2019.
- Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12, 443–451.
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 4, 347–354.

- Paquette SR, Behncke SM, O'Brien SH, Brenneman RA, Louis EE, Lapointe F-J (2007) Riverbeds demarcate distinct conservation units of the radiated tortoise (*Geochelone radiata*) in southern Madagascar. *Conservation Genetics*, 8, 797–807.
- Pease CM, Lande R, Bull JJ (1989) A Model of Population Growth, Dispersal and Evolution in a Changing Environment. *Ecology*, 70, 1657–1664.
- Perry GH, Marioni JC, Melsted P, Gilad Y (2010) Genomic-scale capture and sequencing of endogenous DNA from feces. *Molecular Ecology*, 19, 5332–5344.
- Poinar HN, Höss M, Bada JL, Pääbo S (1996) Amino Acid Racemization and the Preservation of Ancient DNA. *Science*, 272, 864–866.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155, 945–959.
- Pudovkin AI, Zaykin DV, Hedgecock D (1996) On the Potential for Estimating the Effective Number of Breeders From Heterozygote-Excess in Progeny. *Genetics*, 144, 383–387.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, ... Daly MJ (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575.
- Purvis A, Gittleman JL, Cowlishaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences*, 267, 1947–1952.
- Queller DC, Goodnight KF (1989) Estimating Relatedness Using Genetic Markers. *Evolution*, 43, 258–275.
- R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. URL <http://www.R-project.org/>.
- Reid BN, Kass JM, Wollney S, Jensen EL, Russello MA, Viola EM, ... Naro-Maciel E (2019) Disentangling the genetic effects of refugial isolation and range expansion in a transcontinentally distributed species. *Heredity*, 122, 441.
- Reinert HK (1991) Translocation as a Conservation Strategy for Amphibians and Reptiles: Some Comments, Concerns, and Observations. *Herpetologica*, 47, 357–363.
- Reinert HK, Rupert RR (1999) Impacts of Translocation on Behavior and Survival of Timber Rattlesnakes, *Crotalus horridus*. *Journal of Herpetology*, 33, 45–61.

- Ribeiro R, Santos X, Sillero N, Carretero MA, Llorente GA (2009) Biodiversity and land uses at a regional scale: Is agriculture the biggest threat for reptile assemblages? *Acta Oecologica*, 35, 327–334.
- Rosenberg NA (2004) distruct: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137–138.
- Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8, 103–106.
- Row JR, Blouin-Demers G, Weatherhead PJ (2007) Demographic effects of road mortality in black ratsnakes (*Elaphe obsoleta*). *Biological Conservation*, 137, 117–124.
- Russello MA, Kirk SL, Frazer KK, Askey PJ (2012) Detection of outlier loci and their utility for fisheries management. *Evolutionary Applications*, 5, 39–52.
- Russello MA, Waterhouse MD, Etter PD, Johnson EA (2015) From promise to practice: Pairing non-invasive sampling with genomics in conservation. *PeerJ*, 3, e1106.
- Ryder OA (1986) Species conservation and systematics: The dilemma of subspecies. *Trends in Ecology & Evolution*, 1, 9–10.
- Sarre SD (1998) Demographics and Population Persistence of *Gehyra variegata* (Gekkonidae) Following Habitat Fragmentation. *Journal of Herpetology*, 32, 153.
- Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14, 807–820.
- Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy OJ, ... Manel S (2010) Applications of landscape genetics in conservation biology: Concepts and challenges. *Conservation Genetics*, 11, 375–385.
- Senn H, Ogden R, Cezard T, Gharbi K, Iqbal Z, Johnson E, ... McEwing R (2013) Reference-free SNP discovery for the Eurasian beaver from restriction site-associated DNA paired-end data. *Molecular Ecology*, 22, 3141–3150.
- Shepard DB, Kuhns AR, Dreslik MJ, Phillips CA (2008) Roads as barriers to animal movement in fragmented landscapes. *Animal Conservation*, 11, 288–296.
- Shine R, Lemaster M, Wall M, Langkilde T, Mason R (2004) Why Did the Snake Cross the Road? Effects of Roads on Movement and Location of Mates by Garter Snakes (*Thamnophis sirtalis parietalis*). *Ecology and Society*, 9.

- Sirén J, Marttinen P, Corander J (2011) Reconstructing Population Histories from Single Nucleotide Polymorphism Data. *Molecular Biology and Evolution*, 28, 673–683.
- Southern Interior Reptile Recovery Team (2016) Recovery strategy for the Western Rattlesnake (*Crotalus oreganus*) in British Columbia. , 21.
- Sovic M, Fries A, Martin SA, Gibbs HL (2019) Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, *Sistrurus catenatus*. *Evolutionary Applications*, 12, 664–678.
- Stebbins RC (2003) Western reptiles and amphibians. Houghton Mifflin Company Boston, MA
- Stephens SH, White VC, Bremer JRA (2010) Assessment of Genetic Tissue Sampling Methods for the Critically Endangered Kemp’s Ridley Sea Turtle (*Lepidochelys kempii*). *Marine Turtle Newsletter*, 19.
- Swingland IR (2001) Biodiversity, Definition of. In *Encyclopedia of Biodiversity* (pp. 377–391) Elsevier
- Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: Look before you leap. *Trends Ecology Evolution*, 14, 323–327.
- Tallmon DA, Luikart G, Beaumont MA (2004) Comparative Evaluation of a New Effective Population Size Estimator Based on Approximate Bayesian Computation. *Genetics*, 167, 977–988.
- Todd B, Willson J, Gibbons J (2010) The Global Status of Reptiles and Causes of Their Decline. In *Ecotoxicology of Amphibians and Reptiles* (pp. 47–67)
- Topp CM, Winker K (2008) Genetic Patterns of Differentiation Among Five Landbird Species from the Queen Charlotte Islands, British Columbia. *The Auk: Ornithological Advances*, 125, 461–472.
- Vähä J-P, Erkinaro J, Niemelä E, Primmer CR (2007) Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology*, 16, 2638–2654.
- Wahlund S (1928) Zusammensetzung Von Populationen Und Korrelationserscheinungen Vom Standpunkt Der Vererbungslehre Aus Betrachtet. *Hereditas*, 11, 65–106.
- Walker IR (2004) Climate Change, The Last Fifteen Thousand Years in the Okanagan. In M. A. Roed & J. D. Greenough (Eds.), *Okanagan Geology* (pp. 55–62) Kelowna: Kelowna Geology Committee

- Waples RS (1991) Pacific Salmon, *Oncorhynchus* spp., and the Definition of “Species” Under the Endangered Species Act. *Marine Fisheries Review*, 12.
- Watling JI, Donnelly MA (2007) Multivariate correlates of extinction proneness in a naturally fragmented landscape. *Diversity and Distributions*, 13, 372–378.
- Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38, 1358–1370.
- Wilson GA, Rannala B (2003) Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. *Genetics*, 163, 1177–1191.
- Winton SA, Taylor R, Bishop CA, Larsen KW (2018) Estimating actual versus detected road mortality rates for a northern viper. *Global Ecology and Conservation*, 16, e00476.
- Wright S (1931) Evolution in Mendelian populations. *Genetics*, 16, 97.
- Wright S (1943) Isolation by Distance. *Genetics*, 28, 114–138.

Appendices

Appendix A

Chapter 2 Supplementary Material

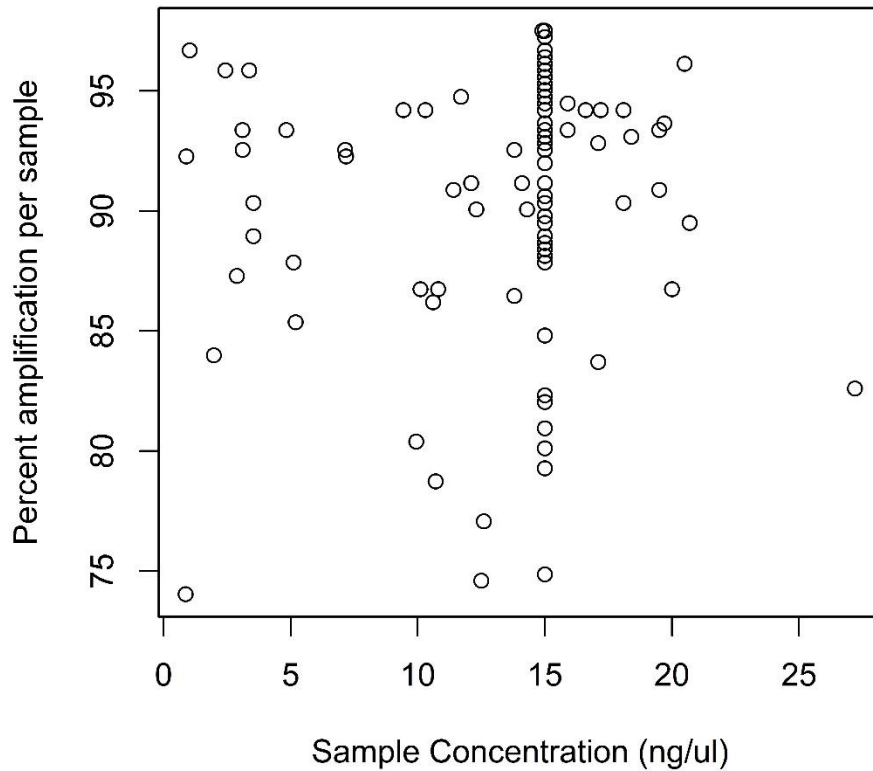


Figure S2.1. Scatterplot displaying the relationship between percent amplification across 311 loci for samples remaining after quality filtering (n=103) and initial DNA sample concentration in ng/ul. Samples with concentrations larger than 30 ng/uL were diluted to 15 ng/ul to avoid preferential amplification of samples with relatively high concentrations.

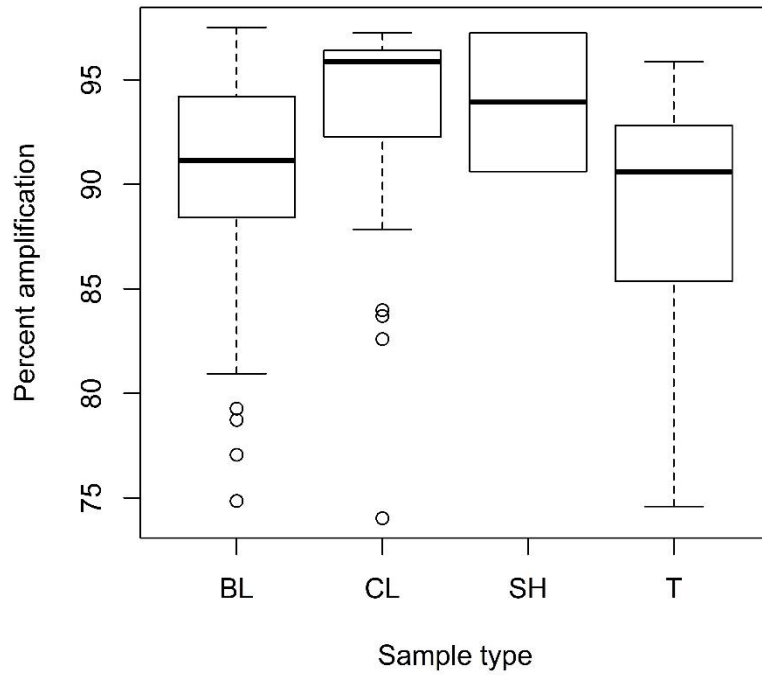


Figure S2.2. Boxplot displaying the mean and range of percent amplification for all samples remaining after quality filtering (n=103) across 311 loci for all sample types including: blood (BL; n=61), cloacal swab (CL; n=26), shed (SH; n=2), and roadkill tissue (T; n=14).

Table S2.1. The number of individual samples that were collected within each region by sample type (region acronyms as in Figure 1).

Sampling Region	Tissue Type				Total
	Cloacal	Blood	Roadkill Tissue	Shed	
KL	6	8	0	0	14
SOK	6	17	17	0	40
WSH	1	15	0	0	16
VN	5	25	0	0	30
GF	0	0	0	4	4
Total	18	65	17	4	104

Table S2.2. The average, minimum, and maximum DNA concentrations per sample type employed in this study across all sampled regions.

Tissue Type	n	Avg concentration (ng/ul)	Min	Max
Cloacal swab	18	24.0	0.31	178.0
Blood sample	65	29.1	4.82	138.0
Roadkill Tissue	17	13.3	2.43	52.1
Shed	4	34.4	19.2	59.2

Table S2.3. STRUCTURE HARVESTER output for all K values tested ($K=1-7$) for the RADseq_9568, GT-seq_311, and RADseq_311 datasets. The posterior log-probability for each K (Mean $\text{LnP}(K)$), standard deviation in $\text{LnP}(K)$ (Stdev $\text{LnP}(K)$), as well as the first order rate of change $\text{Ln}'(K)$ and absolute value of the second order rate of change $|\text{Ln}''(K)|$ used to estimate ΔK are presented. Bolded rows represent the optimal K value selected for each dataset ($K=3$) where a plateau in the log probability across all K values was identified, variation was low, and where ΔK was maximized.

Dataset	K	Reps	Mean $\text{LnP}(K)$	Stdev $\text{LnP}(K)$	$\text{Ln}'(K)$	$ \text{Ln}''(K) $	$\square K$
RADseq_9568	1	10	-392492.09	17.6302	NA	NA	NA
	2	10	-364757.12	15.8772	27734.97	8381.43	527.892487
	3	10	-345403.58	11.1799	19353.54	11738.23	1049.938407
	4	10	-337788.27	25.9509	7615.31	3593.83	138.485549
	5	10	-333766.79	41.1361	4021.48	1980.32	48.140715
	6	10	-331725.63	13.6966	2041.16	3156.29	230.442685
	7	10	-332840.76	100.2142	-1115.13	NA	NA
GT-seq_311	1	10	-13321.93	1.0023	NA	NA	NA
	2	10	-12452.16	3.3457	869.77	235.03	70.248161
	3	10	-11817.42	1.2882	634.74	390.76	303.329188
	4	10	-11573.44	1.8763	243.98	94.85	50.552029
	5	10	-11424.31	3.6528	149.13	474.49	129.896277
	6	10	-11749.67	338.7073	-325.36	263.32	0.777426
	7	10	-11811.71	505.4022	-62.04	NA	NA
RADseq_311	1	10	-13302.69	0.941	NA	NA	NA
	2	10	-12367.9	1.7114	934.79	217.86	127.29931
	3	10	-11650.97	1.149	716.93	487.35	424.165876
	4	10	-11421.39	0.9803	229.58	90.97	92.797548
	5	10	-11282.78	5.1426	138.61	153.93	29.932394
	6	10	-11298.10	291.8692	-15.32	10.79	0.036969
	7	10	-11324.21	77.6416	-26.11	NA	NA

Appendix B

Chapter 3 Supplementary Material

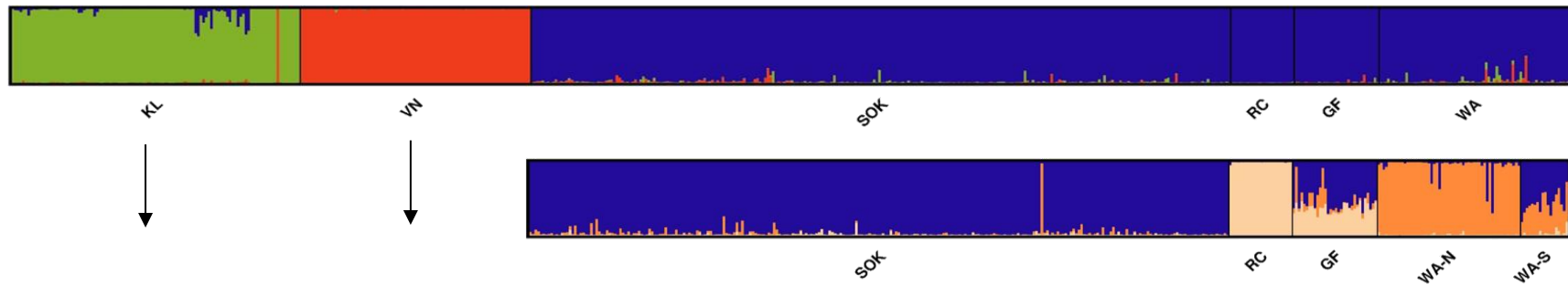


Figure S3.1 Barplot of STRUCTURE results for K=3 for both iterations of all samples (n=588) regardless of assignment to den.

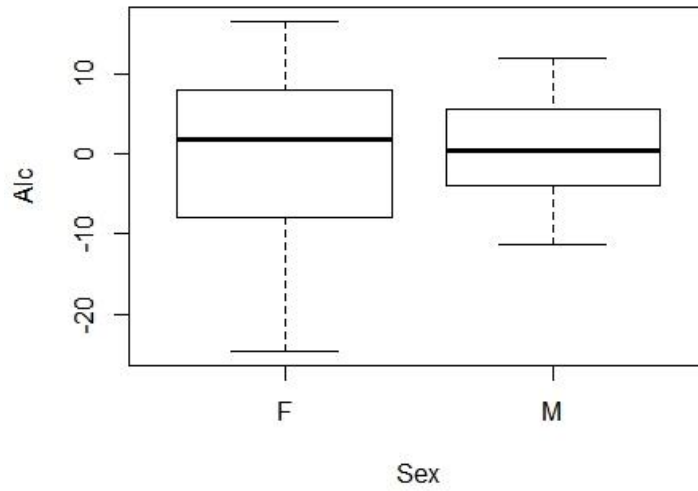


Figure S3.2 Boxplots displaying corrected assignment probabilities for Male and Female adult snakes in the VER-KL den group. The variance of assignment probabilities (vAIC) was significantly higher ($p < 0.05$) for females than males as evidenced by the range of values for each sex in this plot.

Table S3.1 Counts of individual samples used in broad scale population genetic analyses (n=588) divided by region and sample type. Sample types include cloacal swab samples (CL), blood samples (BL), roadkill tissue samples (T), buccal swab samples (BU), and shed skins (SH).

Region	Sample Type				
	CL	BL	T	BU	SH
GF	31	0	1	0	2
KAM	103	3	0	1	0
VN	64	22	0	1	0
SOK	154	13	97	0	0
RC	23	0	1	0	0
WA	57	15	0	0	0

Table S3.2 Results of AMOVAs across various groupings. * indicates $p < 0.05$, while ** denotes $p < 0.01$ after 10000 permutations.

Source of variation	df	Percent Variation	F - statistic
3 Genetic cluster			
Among clusters	2	4.76	0.048
Among regions within clusters	3	10.53	0.110*
Within regions	944	84.72	0.153*
11 den groups			
Among den groups	10	17.02	0.170*
Among dens within groups	25	3.2	0.039*
Within dens	886	79.77	0.202*
KAM			
Among den groups	2	20.69	0.207*
Among dens within groups	3	5.29	0.067**
Within dens	212	74.02	0.260**
VER			
Among den groups	1	10.66	0.107
Among dens within groups	4	2.18	0.024**
Within dens	168	87.16	0.107**
SOK			
Among den groups	1	4.17	0.042*
Among dens within groups	7	1.17	0.012**
Within dens	299	94.66	0.053**
WA			
Among den groups	1	4.76	0.048*
Among dens within groups	9	8.5	0.090**
Within dens	103	86.75	0.133**