

Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

Insular Biogeography and Population Genetics of Dwarf Mistletoe
(*Arceuthobium americanum*)

by
Roy Hill

A thesis
submitted in partial fulfillment
of the requirements for the degree of
Master of Science in the Department of Biological Sciences
Idaho State University
Spring 2016

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Roy Hill find it satisfactory and recommend that it be accepted.

Dr. Ken Aho,
Major Advisor

Dr. F. Charles Williams,
Committee Member

Dr. Michael McCurry,
Graduate Faculty Representative

TABLE OF CONTENTS

LIST OF FIGURES.....	vi
LIST OF TABLES.....	vii
Abstract.....	viii
Chapter I: Development and description of microsatellite primers for <i>Arceuthobium americanum</i>	1
Abstract	1
Introduction.....	1
Methods and Results.....	3
Conclusions.....	6
Acknowledgements	6
Literature cited	9
Chapter II: Population genetics of <i>Arcethobium americanum</i> in a biogeographic context	11
Abstract	11
Introduction.....	12
<i>Arceuthobium americanum</i>	12
Insular biogeography and population genetics	15
Methods	18
Results	21
Genetic dispersal	21
F_{ST} , F_{IS} , F_{IT} , and allelic richness.....	22
Biogeographic assessments	23
Cluster diagram and analyses	24
Bottlenecks	25
Discussion	25
Genetic dispersal and F -statistics	26
Insular biogeography	28
Ontogeny of <i>A. americanum</i> islands.....	29
Figures and Tables	32
Literature cited	40

Appendix	49
Molecular approaches	49
<i>F</i> -statistics	49
Cluster diagram and cluster analysis	51

LIST OF FIGURES

- Figure 1. Map of Southeast and Central Idaho. Hatched areas denote coarse hulls of contiguous *Pinus contorta* cover (U.S. Geological Survey, 1999), green areas define bounds for sampling, and black dots indicate individual sample locations. **Error! Bookmark not defined.**
- Figure 2. Cluster diagram for *Arceuthobium americanum* displaying two clusters from the sampled populations. Pairwise F_{ST} 's were calculated using Weir (1996) pairwise F_{ST} with the ENA correction procedure (Chapuis and Estoup 2007). **Error! Bookmark not defined.**
- Figure 3. Scatterplot of allelic richness as a function of gross estimate of contiguous lodgepole pine cover (km²). Solid line depicts the reduced major axis (RMA) regression (slope= 9.3×10^{-4} , $R^2=0.41$, upper tail $p=0.093$), while the dashed line depicting ordinary least squares (OLS) regression (slope= 9.0×10^{-4} , $R^2=0.41$, upper tail $p=0.093$). Approximate bootstrapped (9999 iterations) standard errors for average allelic richness are also shown. **Error! Bookmark not defined.**
- Figure 4. Scatterplot of allelic richness as a function of gross estimate of contiguous lodgepole pine cover (km²). Solid line depicts the reduced major axis (RMA) regression (slope= 9.3×10^{-4} , $R^2=0.41$, lower tail $p=0.093$), while the dashed line depicting ordinary least squares (OLS) regression (slope= 9.0×10^{-4} , $R^2=0.41$, lower tail $p=0.093$). Approximate bootstrapped (9999 iterations) standard errors for average allelic richness are also shown. **Error! Bookmark not defined.** 35
- Figure 5. Scatterplot of allelic richness as a function of Euclidean distance from nearest mainland (km). Solid line depicts the reduced major axis (RMA) regression (slope= -2.8×10^{-2} , $R^2=0.68$, $p=0.080$), whereas dashed line depicting ordinary least squares (OLS) regression (slope= -3.5×10^{-2} , $R^2=0.68$, $p=0.080$). Bootstrapped (9999 iterations) 95% confidence intervals for average allelic richness are also shown. 36

LIST OF TABLES

Chapter I

Table 1. Description of 7 microsatellite primer pairs from <i>Arceuthobium americanum</i> , using 241 individuals from 6 populations, with GenBank accession numbers.	Error! Bookmark not defined.
Table 2. Results for Hardy-Weinberg equilibrium tests of 7 loci within 6 populations of <i>Arceuthobium americanum</i> .	Error! Bookmark not defined.

Chapter II

Table 1. Table 1. Names of collection sites for <i>Arceuthobium americanum</i> with population designation, and other pertinent information. Climate information were obtained from the National Climatic Data Center (NCDC) which is a department of the National Oceanic and Atmospheric Administration (NOAA, Asheville, North Carolina, USA). Gross estimates of <i>P. contorta</i> cover were based on topographical and satellite maps for the study area.	37
Table 2. Euclidean distance (in km) (lower triangle) and Weir (1996) ENA corrected pairwise F_{ST} (upper triangle) matrix for the sampled populations.	37
Table 3. Bayesian proportional assignment of <i>Arceuthobium americanum</i> sites into two clusters.	38
Table 4. F_{IS} , allelic richness, and presence (+) or absence (-) of recent bottleneck for <i>Arceuthobium americanum</i> populations using the 4L dataset.	38
Table 5. Overall F_{ST} or G_{ST} and F_{IS} values for <i>Arceuthobium americanum</i> and comparisons to other species in the literature.	39

Abstract

Lodgepole pine dwarf mistletoe (*Arceuthobium americanum*) is an obligate parasite of lodgepole pine (*Pinus contorta*) in Western North America, and strongly influences both fire severity and host timber harvest yield. Seven polymorphic microsatellite markers were developed for use in population genetic analysis of *A. americanum*. Six populations in Idaho were examined for population differences. Genetic evidence suggests species is dispersal limited. Genetic differentiation, when accounted for geographic area, was found to be comparable to other species with similar life history traits. Allelic richness was found to be correlated with geographic size of the population. Cluster diagram, and Bayesian cluster analysis indicate two groupings of populations based around two mainland populations, with exception of two populations which are believed to differ due to the effect of neighboring populations outside of study area and population bottlenecks.

Chapter I:

Development and description of microsatellite primers for *Arceuthobium americanum*

Abstract

Lodgepole pine dwarf mistletoe (*Arceuthobium americanum*) is an obligate parasite of lodgepole pine (*Pinus contorta*), and strongly influences fire severity and timber harvest yield in montane regions of Western North America. Seven microsatellite loci primer pairs were developed to allow examination of genetic variability in populations of *A. americanum*. All loci were unlinked and polymorphic with a range of 4 to 24 alleles per locus. The seven loci were applied to 241 *A. Americanum* individuals from six geographically distinct populations in the Northern Rocky Mountains, USA. The populations often demonstrated significant deviations from Hardy-Weinberg equilibrium at multiple loci, suggesting the presence of null alleles. As a result I recommend that primer users check, and if necessary apply corrective measures for null alleles when quantifying microsatellite reads.

Introduction

Dwarf mistletoes (Santalaceae: *Arceuthobium* sp.) comprise 42 phloem-limited plant species that parasitize conifers in both the old and new world (Hawksworth & Wiens 1996, Nickrent et al. 2004). Natural populations of Rocky Mountain lodgepole pine (*Pinus contorta* Dougl. var. *latifolia* Engelm.), a species characteristic of montane forests in Western North America, are often infected by the obligate parasite lodgepole pine

dwarf mistletoe (*Arcethobium americanum* Nutt ex Engelm.). Parasitism from *A. americanum* is of great concern to foresters, because *A. americanum*, like all dwarf mistletoes, captures nutrients and decreases starch content in the needles of infected trees, decreasing host lifespan and fitness (Broshot et al. 1986). Infections also induce “witches broom” asymmetries in host plant growth that increase fire severity and decrease the value of timber harvests (Hawksworth & Wiens 1996).

Dwarf mistletoes may experience low rates of gene flow both within and among populations because of putatively limited seed and pollen dispersal systems and distributions constrained to host “islands.” Dwarf mistletoes use hydrostatic pressure and thermogenesis to forcibly propel seeds from the parent plant at high speeds but at distances ≤ 16 m (Hinds et al. 1963 *qtd.* in Hawksworth & Wiens 1996, deBruyn et al. 2015). Pollen dispersal is also believed to be limited, as primary pollinators are wingless insects (Penfield et al. 1976).

Previous genetic research on dwarf mistletoes has focused on genus-level phylogenies using nuclear ribosomal internal transcribed spacer (ITS) sequences (e.g. Nickrent et al. 2004). However, population-level genetic processes of species within *Arceuthobium* have been ignored, largely because of the absence of whole-genome microsatellite markers that allow, (because of their high mutation rates and potential for genetic variability) detection of fine-scale genetic variation among and within populations (Nickrent *qtd.* In Hawksworth & Wiens 1996). As a response to this deficiency I sought to develop and provide microsatellite tools to allow genetic studies of dwarf mistletoes, particularly *A. americanum*, at the population scale.

Methods and Results

Tissues of fifteen *A. americanum* individuals were randomly selected from a much larger set of samples ($n = 241$; see below) collected from Rocky Mountain lodgepole pines in the Targhee, Caribou, and Sawtooth National Forests in Southern and Central Idaho, USA. The tissues were supplied to Ecogenics GmbH (Schützenstrasse 15, 9436 Balgach, Switzerland) for the identification of microsatellite markers and the development of associated primers. Briefly, enriched, size-selected fragments were obtained from genomic DNA for simple sequence repeat (SSR) content using magnetic streptavidin beads and biotin-labeled GATA, GTAT, AAAC and AAAG repeat oligonucleotides. The SSR-enriched library was then analyzed for VNTR motifs on a Roche 454 platform (© Roche Diagnostics, Indianapolis, Indiana) using 454 GS FLX Titanium reagents (© Roche Diagnostics, Indianapolis, Indiana). A total of 8993 reads were completed, with an average read length of 422 base pairs. Of the reads, 2394 contained a microsatellite insert with a tetra- or a trinucleotide of at least 6 repeat units or a dinucleotide of at least 10 repeat units. Suitable primer design was possible in 266 reads. Of these, seven reads were selected for primer development based on primer exclusivity and stability and length of the SSR. The loci were then analyzed for allelic trends and frequencies (Table 1).

Allelic frequencies of the seven markers were determined for 241 *A. americanum* individuals randomly sampled from the six geographically distinct populations within the locations described above. DNA was extracted from the stem tissues of individual plants using PowerPlant Pro DNA Isolation Kit (© Mo Bio, Carlsbad, California, USA). Labeled

primers were obtained from Applied Biosystems (Foster City, California, USA).

Fluorescent dyes associated with primers belonged to the DS-33 dye set. PCR amplification was performed using HotStarTaq DNA Polymerase (© Qiagen, Valencia, California, USA). 10 µL reactions were used with the following final concentrations: 1× of 10× PCR buffer, 200 µM dNTP mix (Phenix, Candler, North Carolina, USA), 0.16 µM F primer, 0.16 µM R primer, 0.5 units of HotStarTaq, and 2-10ng Genomic DNA. Thermal cycling protocol was as follows: 1 cycle at 95°C for 15 min, 30 cycles at 94°C for 30 s, 56°C for 45 s, 72°C for 45 s, another 8 cycles at 95°C for 30 s, 53°C for 45 s, 72°C for 45 s, with a final extension of 1 cycle at 72°C for 30 min.

After amplification, samples were submitted to the Molecular Research Core Facility (MRCF) at Idaho State University for fragment analysis. Samples were analyzed using a 3130XL Genetic Analyzer (© Applied Biosystems, Carlsbad, California, USA). Peaks were scored by hand using GeneMapper Software 4.0 (© Applied Biosystems, Carlsbad, California, USA) with close attention paid to potential stuttering effects, and binned using Flexibin v2 (Amos et al. 2007).

Linkage disequilibrium was quantified using GENEPOP software (Rousset 2008). To account for $(7^2 - 7)/2 = 21$ marker pairwise comparisons, the Bonferroni family-wise type I error rate was defined to be $\alpha = 0.05/21 = 0.0024$ in statistical tests. From this perspective, none of the marker pairs were found to deviate significantly from the null hypothesis of loci independence. I note, however, that one locus pair (Arcame_001198

and Arcame_005927) in one population approached the adjusted significance level ($P = 0.0067$).

The number of alleles per locus per population was estimated using GENEPOP 4.4.3 software (Rousset 2008). Observed and expected heterozygosity, and tests of Hardy-Weinberg equilibrium (HWE) were computed using Arlequin 3.5.2.2 software (Excoffier & Lischer 2010; Table 2). In these analyses, 64% of loci deviated significantly from HWE (Bonferroni adjusted P -value $P < 0.05/7 = 0.0071$; Table 2). These deviations were due to higher than expected levels of homozygosity, and prompted assessments of the loci for null alleles (alleles that fail to amplify due to non-ideal PCR conditions or mutations in the primer binding region; Selkoe and Toonen 2006).

Diagnostic tests for null alleles were applied using MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004) which calculates Bonferroni-adjusted permutation confidence intervals for true heterozygosity under HWE and compares this to observed heterozygosity. Under this approach, evidence of null alleles were found in all populations for loci Arcame_006627 and Arcame_006790. The loci Arcame_001198, Arcame_002494, Arcame_005390, and Arcame_005927 showed evidence of null alleles in 50-83% of the populations. Arcame_005674 showed evidence of null alleles in only one population. Given the strong evidence of null alleles I recommend that primer users perform diagnostics for null alleles and if necessary, employ statistical software that corrects for the presence of null alleles, e.g., FREENA (Chapuis & Estoup 2007), and INEST (Chybicki & Burczyk 2009).

Of note, in an exploratory project I tested all seven primers on limber pine dwarf mistletoe (*A. cyanocarpum*). Explicit microsatellite analyses were not performed. Electrophoresis gels, however, demonstrated the presence of fragments at appropriate lengths for four of the seven loci (Arcame_001198, Arcame_002494, Arcame_005927, and Arcame_006627), suggesting the applicability of our markers for close relatives of *A. americanum*.

Conclusions

Development of short sequence repeat microsatellite primers for *A. americanum* resulted in seven loci, displaying a useful range of allelic variability for population analyses. There was no evidence of linkage disequilibrium among the loci after controlling for family-wise type I error. Presence of null alleles are suggested to play a factor in the significant deviations from Hardy-Weinberg equilibrium.

Acknowledgements

I thank Rick Williams (Idaho State University) and Andre Buser (Ecogenics GmbH) for their comments on this manuscript.

Table 1. Description of 7 microsatellite primer pairs from *Arceuthobium americanum*, with respect to 241 individuals from 6 populations, with GenBank accession numbers.

Locus	Primer Sequences (5'-3')	Repeat type	Size bp	# of alleles	GenBank Accession #
Arcame_001198	F: AGTGCGATTTGAGCGACATC R: GTGTGTATGTCTGTGCGTGC	(ACAT) ₁₆	114 - 166	14	Pr032816302
Arcame_002494	F: CGAAACCTGCGAACGAAGTG R: ACACTCGCTCTAAACCCCTC	(CA) ₁₁	90 - 110	11	Pr032816303
Arcame_005390	F: CACTCCTGACTCCTGAGAGC R: CACTAACAAAATCACTTGAAAATCCG	(TTCT) ₉	234 - 394	14	Pr032816304
Arcame_005674	F: TGCTGATCTACTGTATCACCCCTG R: ACTTGGGCTCATTATTACAACGG	(GAT) ₇	195 - 204	4	Pr032816305
Arcame_005927	F: ATTTGGGGGATGCTACCGAG R: AGGTGAAGACGAGGTTGTCC	(GA) ₁₂	166 - 218	21	Pr032816306
Arcame_006627	F: TAGGTCGCCCTCTTCTCTTC R: TAGGTGGTTCCGACAAGGTG	(CT) ₁₂	193 - 277	24	Pr032816307
Arcame_006790	F: CGCACTCAGCTCCTCAAAAC R: AGACTATGAGGCAGTTCGTCC	(ATA) ₇ G(TAA) ₁₅	178 - 220	14	Pr032816308

Table 2. Results for Hardy-Weinberg equilibrium tests of 7 loci within 6 populations of *Arceuthobium americanum*.

Locus	Central Idaho (n=75)				Pike Mountain (n=25)				Pomerelle (n=15)				Island Park (n=85)				George Town Canyon (n=19)				Skinner Canyon (n=25)			
	A	H_o	H_e	HWE	A	H_o	H_e	HWE	A	H_o	H_e	HWE	A	H_o	H_e	HWE	A	H_o	H_e	HWE	A	H_o	H_e	HWE
Arcame_001198	13	0.500	0.874	0.000	10	0.500	0.872	0.000	9	0.733	0.874	0.443	13	0.500	0.909	0.000	9	0.571	0.862	0.039	9	0.524	0.835	0.001
Arcame_002494	9	0.027	0.611	0.000	5	0.211	0.627	0.000	1	0.000			5	0.042	0.273	0.000	8	0.308	0.618	0.000	1	0.000		
Arcame_005390	11	0.476	0.789	0.000	6	0.556	0.843	0.000	6	0.778	0.843	0.095	11	0.577	0.827	0.000	7	0.750	0.833	0.063	10	0.429	0.858	0.000
Arcame_005674	3	0.080	0.102	0.180	2	0.045	0.045	1.000	1	0.000			3	0.065	0.064	1.000	3	0.056	0.160	0.030	2	0.080	0.078	1.000
Arcame_005927	15	0.703	0.836	0.005	5	0.545	0.736	0.345	7	0.600	0.786	0.052	17	0.652	0.835	0.003	10	0.533	0.867	0.002	8	0.625	0.758	0.185
Arcame_006627	19	0.246	0.915	0.000	3	0.333	0.733	0.200	5	0.000	0.758	0.000	16	0.174	0.896	0.000	6	0.231	0.846	0.000	10	0.143	0.894	0.000
Arcame_006790	12	0.419	0.885	0.000	3	0.500	0.833	0.332	4	0.000	0.788	0.001	12	0.300	0.902	0.000	5	0.143	0.824	0.000	6	0.100	0.784	0.000

Note: A = number of alleles; H_o = observed heterozygosity; H_e = expected heterozygosity (under Hardy Weinberg equilibrium); HWE = P values concerning a null hypotheses of Hardy-Weinberg equilibria.

Literature cited

- Amos, W., J. I. Hoffman, A. Frodsham, L. Zhang, S. Best, and A. V. S. Hill. 2007. Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes* 7:10–14.
- Broshot, N., L. Larsen, and R. ; Tinnin. 1986. Effects of *Arceuthobium americanum* on twig growth of *Pinus contorts*. Research Notes, Pacific Northwest Research Station, US Department of Agriculture, Forest Service (USA).
- Chapuis, M.-P., and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–631.
- Chen, C., E. Durand, F. Forbes, and O. François. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* 7:747–756.
- Chybicki, I. J., and J. Burczyk. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100:106–113.
- deBruyn, R. A. J., M. Paetkau, K. A. Ross, D. V. Godfrey, and C. R. Friedman. 2015. Thermogenesis-triggered seed dispersal in dwarf mistletoe. *Nature Communications* 6:6262.
- Excoffier, L., and H. E. L. Lischer. 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.
- Hawksworth, F. G., and D. Wiens. 1996. Dwarf Mistletoes: Biology, Pathology, and Systematics. United States Department of Agriculture Forest Service.

- Morgante, M., M. Hanafey, and W. Powell. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* 30:194–200.
- Nickrent, D. L., M. A. García, M. P. Martín, and R. L. Mathiasen. 2004. A phylogeny of all species of *Arceuthobium* (Viscaceae) using nuclear and chloroplast DNA sequences. *American Journal of Botany* 91:125–138.
- Penfield, F. B., R. E. Stevens, and F. G. Hawksworth. 1976. Pollination ecology of three Rocky Mountain dwarf mistletoes. *Forest Science* 22:473–484.
- Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8:103–106.
- Schlötterer, C. 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109:365–371.
- Selkoe, K. A., and R. J. Toonen. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9:615–629.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4:535–538.

Chapter II:

Population genetics of *Arcethobium americanum* in a biogeographic context

Abstract

This study considers the population genetics of lodgepole pine dwarf mistletoe (*Arceuthobium americanum*), an aggressive parasite of lodgepole pine (*Pinus contorta*), that is distributionally limited to host montane “islands” in Western North America. To quantify insular biogeographic effects on *A. Americanum* population genetics, I applied seven novel microsatellite markers to six populations (four island and two mainland locations) at the western edge of the Central Rocky Mountains. This work had three facets: 1) I assessed the effect of purportedly limited pollen and seed dispersal mechanisms of *A. americanum* using fixation indices and meta analyses, 2) I tested the predictions for allelic diversity and gene flow in *A. americanum* suggested by classic theoretical island models from population genetics and ecology, and 3) I generated models for predicting the origin and relatedness of island and mainland populations in the study area. I found that measures of gene fixation and inbreeding in *A. americanum* were similar to those reported for dispersal limited species, with similar life history characteristics. As predicted by biogeographic theory, allelic richness was correlated with geographic size of the population, and distance from source (mainland) populations. UPGMA dendrograms and Bayesian cluster analysis identified two groupings of populations with respect to the two mainland populations. Notably, two of

the island populations may be affected by both gene flow from populations outside of the study area and population bottlenecks.

Introduction

Forest ecologists and land managers often require detailed species information, including genetic data to effectively understand and maintain the natural diversity, stability, and resilience of forest ecosystems (Loo et al. 2014).

Particularly relevant are considerations of the population genetics of forest species in montane environments, as these systems are expected to become increasingly fragmented due to global climate change and other anthropogenic impacts (Romme and Turner 1991). Such pressures are likely to increase vulnerability of montane populations to extinction by altering patterns of genetic differentiation/fixity and gene flow.

Arceuthobium americanum.

A. americanum Nutt ex Engelm. is an obligate parasite of lodgepole pine (*Pinus contorta* Dougl.), a species limited to montane ecosystems in Western N. America. The dwarf mistletoe genus *Arcethobium* (Santalaceae), includes 42 species, 39 of which are endemic to North America. All dwarf mistletoes are obligate parasites of conifers, and require a host tree to be able to obtain water and nutrients (Hawksworth and Wiens 1996). Dwarf mistletoes damage host conifers by depleting nutrients, decreasing

respiration (Wanner and Tinnin 1986) and by causing a tumor-like growth called a “witches broom” in most host species (Nicholls et al. 1987, Hawksworth and Wiens 1996). Witches brooms tend to be more flammable than normal growth and thus increase fire severity and coincident mortality of conifers in forest fires (Harrington and Hawksworth 1990). On the other hand, dwarf mistletoes can positively affect forest ecosystems by providing food and shelter for animal species, and increasing understory light, allowing higher forest plant diversity (Nicholls et al. 1984, Ostry and Nicholls 1998, Hawksworth and Weins 1996, Nicholls 2014).

Pollen dispersal. Evidence suggest that *A. americanum* has poor dispersal mechanisms. *A. americanum* plants are dioecious, with a subequal sex ratios, and the potential for hundreds of flowers per inflorescence (Penfield et al. 1976, Hawksworth and Wiens 1996). *A. americanum* produces approximately 11,000 pollen grains per ovule; the pollen grains, however, are in bundles and nectaries are open and flat, resulting in a morphology more suited to generalist insect dispersal (Penfield et al. 1976). On average, wind dispersal of pollen occurs over short distances; ≈ 1.6 m (Penfield et al. 1976), with a maximum reported range of 512 meters (Coppola 1989). Because of these properties, entomophily has been suggested as the main pollen dispersal mechanism, over anemophily. However, the relative contributions of wind and insect pollination to sexual reproduction in *A. americanum* remain unclear (Penfield et al. 1976, Coppola 1989).

The major animal pollinators of *A. americanum* are insects, including the hymenopteran *Formica fusca* Linnaeus (Formicidae), and dipterans, particularly *Philygria debilis* Loew (Ephydriidae). Only unwinged workers of the primary pollinator, *F. fusca*, a common circumboreal forest ant species (Mackay and Mackay 2002), visit inflorescences for nectar (Penfield et al. 1976). The secondary pollinator, *P. debilis*, however, obtains floral resources as a winged adult (Penfield et al. 1976, Mathis and Mathis 2008). Other, smaller insects (e.g. *Protophormia terraenovae* Robineau-Desvoidy, *Hylemya ceralis* Gillette, *H. cinerella* Fallen, and *H. platura* Meigen [Anthomyiidae]) provide pollination services for *A. Americanum* to a lesser degree (Penfield et al. 1976).

Seed dispersal. Seed dispersal of *A. americanum* (and other dwarf mistletoes) also appears constrained. Under the primary method of seed dispersal, hydrostatic pressure is created in the seed pod and thermogenesis – using the alternative oxidase (AOX) pathway - is used to trigger ejection of the seed from the parent plant at velocities up to 27 meters per second (Hinds et al. 1963 *qtd.* in Hawksworth and Wiens 1996, deBruyn et al. 2015). This mechanism propels seeds up to 16 meters, however seeds usually travel 2 to 4 meters before encountering and affixing to a host as the result of a viscous coating (Hawksworth and Wiens 1996). Approximately 40% of propelled seeds encounter viable hosts; although this percentage may vary with stand density (Hawksworth and Wiens 1996).

Animals have been implicated as secondary seed dispersal agents of *A. americanum*. Specifically, while ingested seeds are non-viable after passing through the digestive system, external animal dispersal may occur via small mammals, and birds (Hudler et al. 1979, Nicholls and Hawksworth 1983, Nicholls et al. 1984a, Nicholls et al. 1984b, Hawksworth et al. 1987, Nicholls 2014). This mechanism, particularly dispersal via migratory birds – whose fall migratory times coincide with *A. americanum* seed release – may allow the establishment of remote new populations, and/or facilitate meta population gene flow over long distances.

Dispersal and genetics. The strong physical limitations to pollen and seed dispersal given above suggest that populations of *A. americanum* should have extremely low levels of gene flow, low levels of genetic diversity, and may thus be prone to speciation given geographical isolation. Jerome and Ford (2002), however, argued that most populations of *A. americanum* will not be negatively affected by genetic drift because of large population sizes and dioecious gender expression. This striking conflict prompts further consideration of the dispersal characteristics of *A. americanum*, and its genetic consequences.

Insular biogeography and population genetics

The ecology of *A. americanum* requires its distribution to be highly insular. Specifically, *A. americanum* is obligate to host “islands” of lodgepole pine, whose Central Rocky

Mountain distribution is limited to montane sites with elevations ranging between 1670-2380m (Barrett and Arno 1991).

Two notable theoretical models have been used to predict genetic patterns for island populations. First, the infinite island model, (Wright 1931) posits that: 1) gene flow will occur primarily from mainland/large islands to smaller islands, and 2) islands closer to a mainland (and to each other) will have higher rates of inter-population gene flow. Second, the theory of island biogeography (MacArthur and Wilson (1963, 1967) proposed that, given sufficient time, species diversity on islands will result in an equilibrium point resulting from processes of immigration and extinction. Specifically, the theory predicts that higher species diversity will occur when island size is large (due to lower extinction rates) and distance from mainland is short (due to higher immigration rates). Although originally intended for oceanic islands, the theory of island biogeography has been applied to terrestrial montane “islands” as well. Research from this perspective includes studies of mammal populations in the Great Basin (Brown 1971), the mountain chains of the American Southwest (Lomolino et al. 1989), and the Olympic plateau (Lomolino and Perault 2001). Genetic predictions of the theory, noted above, have also been tested in the context of terrestrial islands (Wheeler and Guries 1982, Floyd et al. 2005, Marlowe and Hufford 2008).

Jaenike (1973) suggested that the infinite island model and theory of island biogeography implied similar mechanisms for gene flow, and could be conflated to explain or predict genetic diversity of species’ populations. This view has been

substantiated by Frankham (1996, 1997) and Vellend (2003) whose meta-analyses found that small isolated islands had the lowest levels of genetic diversity. This decrease was attributed to a number of factors including inbreeding and genetic drift, i.e., random changes in allele frequencies, not attributable to natural selection.

I considered the predictions of the infinite island model and the theory of island biogeography for *A. americanum* by studying insular populations along the Western edge of the Central Rocky Mountains near the Northern edge of the Great Basin (Fig. 1). In this region *P. contorta* host distributions are limited to a series of small, relatively isolated montane landscapes with potential metapopulation connections to each other and to extensive “mainland” populations to the north and northeast.

Hypotheses

Given the background provided above, I have three hypotheses:

- 1) Given putatively limited capacities for pollen and seed dispersal, F -statistic (see Appendix) values for *A. americanum* will be comparable to other dispersal limited species.
- 2) Based on the predictions of Wright’s Island Model (1931) and the Theory of Island Biogeography (MacArthur and Wilson 1963, 1967) I posit that large contiguous populations of *A. americanum* closest to a figurative “mainland” will have the highest levels of genetic diversity.

3) I will test two hypotheses concerned with the geographical source of island populations of *A. americanum* at the northern edge of the Great Basin.

3a) Eastern island populations are the result of colonization from the Island Park/Yellowstone “mainland” area to the northeast.

3b) Western island populations are the result of colonization from a Central Idaho “mainland” area to the north.

Methods

Stems from 241 *A. americanum* individuals were collected from Rocky Mountain lodgepole pines (*Pinus contorta* Dougl. var. *latifolia* Engelm.) in the Targhee, Caribou, and Sawtooth National Forests within montane regions constituting the western edge of the Central Rocky Mountains. The locations of sampling (Fig. 1) corresponded to two “mainland” areas: Island Park (IP), Central Idaho (CI), and four outlying “island” areas of varying size and degree of isolation. These were designated Skinner Canyon (SC), Georgetown Canyon (GTC), Pomerelle (POM), and Pike Mountain (PM). Sample numbers per site, as well as physical characteristics are given in Table 1.

Tissue samples were acquired using a tessellated stratified random process (Stevens et al. 2004). Specifically, for each island and mainland population, inferential expanses of *P. contorta* forest were designated and grids with one square mile area cells were

overlaid using mapping software. Cells were randomly selected from overlays, with the number of selected cells reflecting size island/mainland size. Cell numbers were: IP = 17, CI = 15, SC = 5, GTC = 4, POM = 3, and PM = 5. From within selected cells five *P. contorta* hosts were then randomly chosen using mapping software for extraction of *A. americanum* tissue.

To consider *A. americanum* population genetics, seven microsatellite markers and primer pairs were developed by this project specifically for the study organism, (Chapter I) and applied to tissue samples. DNA extraction from *A. americanum* tissues was performed using PowerPlant Pro DNA Isolation Kit (© Mo Bio, Carlsbad, California, USA). Primers labeled with the DS-33 dye set were obtained from Applied Biosystems (Foster City, California, USA). PCR amplification was performed using HotStarTaq DNA Polymerase (© Qiagen, Valencia, California, USA). 10 µL reactions were used with the following final concentrations: 1× of 10× PCR buffer, 200 µM dNTP mix (Phenix, Candler, North Carolina, USA), 0.16 µM F primer, 0.16 µM R primer, 0.5 units of HotStarTaq, and 2-10ng Genomic DNA. Thermal cycling protocol was as follows: 1 cycle at 95°C for 15 min, 30 cycles at 94°C for 30 s, 56°C for 45 s, 72°C for 45 s, another 8 cycles at 95°C for 30 s, 53°C for 45 s, 72°C for 45 s, with a final extension of 1 cycle at 72°C for 30 min. Fragment analysis was performed by the Molecular Research Core Facility (MRCF) at Idaho State University (Pocatello, Idaho, USA) using a 3130XL Genetic Analyzer (© Applied Biosystems, Carlsbad, California, USA). GeneMapper Software 4.0 (© Applied Biosystems, Carlsbad, California, USA) was used to score the peaks, and Flexibin v2 (Amos et al. 2007) was used to bin the “called” peaks. Primer sequences and

characteristics, number of alleles per loci, and Hardy Weinberg equilibrium tests are detailed in Ch. 1.

There is strong evidence that the datasets used for this study contain null alleles (Chapter I). As such, I used two different datasets for analyses, our full dataset with null alleles included (NAI), and a trimmed dataset using the four loci that contained the least amount of evidence of null alleles (4L). Statistics corrected for null alleles using the NAI dataset will be abbreviated CNA.

Euclidean spatial distance of sample locations was calculated using GenAlEx 6.502 (Peakall and Smouse 2006, 2012). Spatial autocorrelation was performed on the 4L dataset using GenAlEx 6.502 (Peakall and Smouse 2006, 2012). Using diveRsity 1.9.89 (Keenan et al. 2013) Weir and Cockerham's (1984) F_{ST} and Wright's (1951) F_{IS} were calculated for the NAI and 4L datasets; Wright's (1951) subpopulation F_{IS} , and average allelic richness (AR) per subpopulation were calculated using the 4L dataset.

Bootstrapped (10000 iterations) 95% confidence intervals were calculated for Weir and Cockerham's (1984) F_{ST} , Wright's (1951) F_{IS} , Wright's (1951) subpopulation F_{IS} , and average allelic richness (AR) per subpopulation using the bias corrected and accelerated (BCa) bootstrap method (Efron 1987) in diveRsity 1.9.89 (Keenan et al. 2013). FREENA (Chapuis & Estoup 2007) was used with the NAI dataset to correct for the presence of null alleles and estimate Weir (1996) F_{ST} and pairwise F_{ST} using the ENA method with bootstrapped (10000 iterations) 95% confidence intervals for F_{ST} and pairwise F_{ST} . INEST (Chybicki & Burczyk 2009) was used with the NAI dataset to correct for the presence of

null alleles and estimate F_{IS} and test for significance of F_{IS} . Linear regression analyses were performed using R 3.1.3 (R Core Team 2015). Hierarchical agglomerative cluster analyses of CNA pairwise F_{STs} were performed using average group linkage (UPGMA; Sokal and Michener 1958). Dendrograms were generated using the R package cluster (Maechler et al. 2015). Pruning solutions of dendrogram hierarchies were assessed using average silhouette width (Kaufman and Rousseeuw 2009, Aho et al. 2008).

Bayesian cluster analyses were performed with 50000 iterations, following a 5000 iteration burn-in using TESS 2.3 (Chen et al. 2007). The maximum number of clusters (K_{MAX}) in the dataset was verified using GENELAND 4.0.5 (Guillot et al. 2005). The software BOTTLENECK 1.2.02 (Cornuet and Luikart 1996) was used with the NAI and 4L datasets to identify recent bottlenecks in the populations using the NA and 4L datasets.

Results

Genetic dispersal

An autocorrelation diagram (Figure 3) displays a clear pattern of spatial dependence in genetic similarity for up to 50 meters. Note that at distances greater than or equal to 50 meters all bootstrapped (9999 iterations) 95% confidence intervals for the true correlation coefficient (ρ) include 0. This result suggests that spatial genetic structure resulting from limited pollen and seed dispersal is fairly pronounced.

F_{ST} , F_{IS} , F_{IT} , and allelic richness

Bootstrapping across all loci was used to obtain 95% confidence intervals for fixation indices, F_{ST} and F_{IS} , using three different datasets (Table 5). For F_{ST} I found that, the NAI dataset $F_{ST} = 0.035$ (0.020, 0.054), whereas for the 4L and CNA datasets: $F_{ST} = 0.045$ (0.029, 0.063), and $F_{ST} = 0.044$ (0.010, 0.112), respectively. For F_{IS} I found that for the NAI and 4L datasets $F_{IS} = 0.516$ (0.459, 0.558), $F_{IS} = 0.396$ (0.349, 0.442), respectively. For the CNA data I used two different Bayesian models to assess/describe F_{IS} . In the first model F_{IS} was defined as > 0 while the second model $F_{IS} = 0$. A Deviance information criterion (DIC , Spiegelhalter et. al 2002) comparison of the two models indicated greater parsimony for the $F_{IS} > 0$ perspective ($\Delta DIC = 51$). The Bayesian posterior distribution null allele corrected median $F_{IS} = 0.104$ (average $F_{IS} = 0.107$) with 95% credible interval (0.057, 0.167) (Table 5).

The overall F_{IT} calculated using Eq. 3 (Appendix). For the NAI dataset $F_{IT} = 0.533$, 4L dataset $F_{IT} = 0.423$, and using the CNA F_{ST} and F_{IS} then $F_{IT} = 0.146$.

Bootstrapped (across all loci) confidence intervals were calculated for the CNA pairwise F_{ST} values to make inferences concerning the true pairwise F_{ST} values. The resulting CIs were assessed with respect to a null value describing no genetic differentiation between the populations (i.e. pairwise $F_{ST} = 0$). Under this null, 12 statistically significant F_{ST} values were identified (Table 2). The only pairwise comparisons lacking significance were: CI & POM, CI & GTC, and POM & SC.

Analogously, bootstrapped (across all loci) confidence intervals were calculated for the 4L dataset for F_{IS} values of both island and mainland locations. I note that F_{IS} values for island and mainland location could not be calculated using the CNA dataset due to program limitations of INEST (Chybicki & Burczyk 2009). In this case CIs were considered with respect to H_0 : true $F_{IS} = 0$. All sites except Pomerelle had significant F_{IS} values (Table 4). Central Idaho had the highest $F_{IS} = 0.481$ (0.416, 0.537), followed by Island Park $F_{IS} = 0.450$ (0.346, 0.536), Pike Mountain $F_{IS} = 0.399$ (0.129, 0.562), Skinner Canyon $F_{IS} = 0.335$ (0.156, 0.467), Georgetown Canyon $F_{IS} = 0.305$ (0.148, 0.384), and Pomerelle $F_{IS} = 0.122$ (-0.096, 0.239).

Allelic richness (AR) of the 4L dataset varied dramatically across sites (Table 4).

Bootstrapping (9999 iterations) was used to calculate location estimates (means) and 95% confidence intervals for true site richness. *A. americanum* individuals in Central Idaho ($n=75$) had the highest average number of alleles per *A. americanum* individual = 6.86 (5.50, 8.25), followed by Georgetown Canyon ($n=18$) mean AR = 6.33 (5.00, 7.50), Island Park ($n=83$) mean AR = 5.86 (4.50, 7.25), Skinner Canyon ($n=25$) mean AR = 5.37 (4.25, 6.50), Pomerelle ($n=15$) mean AR = 5.10 (4.25, 5.75), and Pike Mountain ($n=25$) mean AR = 5.05 (4.00, 5.75).

Biogeographic assessments

Estimates of island and mainland contiguous *P. contorta* cover (estimated island size) and allelic richness were significantly positively associated at $\alpha = 0.1$ using both reduced

major axis (RMA) and ordinary least squares (OLS) regression (Figure 4, RMA slope= 9.3×10^{-4} , $R^2=0.41$, $n=6$, upper tailed $p=0.093$) Additionally, island allelic richness was significantly negatively associated with Euclidean distance from nearest mainland (Figure 5, RMA slope= -2.8×10^{-2} , $R^2=0.68$, $n=4$, lower tail $p=0.080$).

Cluster diagram and analyses

An UPGMA dendrogram (Figure 2; see Appendix) based on CNA pairwise F_{ST} values and subsequent average silhouette width analysis suggested two clusters for our populations. The first included CI, GTC, IP, POM, and SC. The second cluster contained only PM.

In concordance with UPGMA results evidence of only two Bayesian clusters were found ($K_{MAX}=2$). All *A. americanum* individuals at POM were assigned membership to cluster one, the larger of the two clusters. Other important contributors to cluster one were SC (98% of individuals), and IP (89% of individuals). Lesser contributors were GTC (62%) and CI (59%) (Table 3). Cluster two had strong contribution from PM (84%) with very little input (<50%) from all other sites (Table 3). Notably, there was no difference in contributions to the clusters when using the CNA and 4L datasets.

Bottlenecks

Patterns of allelic diversities and heterozygosity indicated recent bottlenecks in some populations (Table 4). In particular, when using the NAI dataset, Pike Mountain and Pomerelle were both identified as having a recent bottleneck. Using the 4L dataset, however, only Pike Mountain was identified as having a recent bottleneck. Detection of a bottleneck indicates a recent reduction of individuals in the population, or that the population is the result of a recent colonization.

Discussion

This study used customized microsatellite tools to describe the population genetics of lodgepole pine dwarf mistletoe on two montane “mainlands” and four montane “islands” across a large region (139648 km²) in the Central Rocky Mountain. I had three research foci. First, I predicted that *A. americanum* would have limited dispersal capacities resulting in F_{ST} and F_{IS} values similar to other dispersal-limited species. Second, based on Wright’s Island Model and the genetic application of the theory of insular biogeography I posited that small islands far from mainlands would have relatively low levels of genetic diversity, whereas large islands near mainlands would have relatively high diversity. Third, in addressing the ontogeny of isolated islands of *A. americanum* at the northern edge of the Great Basin, I hypothesized that the SC and GTC populations were transplants from the Island Park mainland whereas POM and PM were transplants from the Central Idaho mainland

Genetic dispersal and F -statistics

Genetic data analyses suggest that *A. americanum* is a dispersal limited species. In particular, a depiction of spatial autocorrelation (Figure 3) indicated significant correlations for distances less than 30 meters.

The overall F_{IT} (CNA F_{IT} =0.146, 4L F_{IT} =0.423, and NAI F_{IT} =0.533) indicates deviation from HWE within the complete metapopulation. Despite variability among F_{ST} , all values indicate that the main source of variation in the total population is from the subpopulation deviations from HWE (F_{IS}), not from genetic differentiation of our subpopulations (F_{ST}). Our assumption is that the CNA F -statistic values are the most accurate because they considered more alleles while accounting for the effect of null alleles. Although, we note that the 4L dataset gave similar results for F_{ST} .

The overall F_{IS} was particularly high for a mandatory outcrossing species given extant work (CNA F_{IS} =0.107, 4L F_{IS} =0.396, NAI F_{IS} =0.516, Table 4). In comparison, lodgepole pine, the *A. americanum* host, is typically outcrossing and has much lower value of F_{IS} in most populations (F_{IS} =0.06, Table 5; Yeh et al. 1985). Notably, some of the collection sites were populated with naturally and anthropogenically reforested lodgepole pine that have since been naturally infected. Thomas et al. (1999) observed that reforested lodgepole pine had increased F_{IS} values compared to old growth areas (Table 5). Our methods, however, likely control for this effect because sites were randomly sampled, with no preference for old stands versus reforested stands. Small insects, which are less

vagile, increase inbreeding in insect pollinated plants (Loveless and Hamrick 1984), thus increasing F_{IS} values. Given that entomophily is favored over anemophily in *A.*

americanum, our data supports the current movement-limited pollinators proposed by Penfield et al. (1976).

The genetic differentiation among populations (F_{ST}) was low (CNA F_{ST} = 0.044, 4L F_{ST} = 0.045, NAI F_{ST} = 0.035, Table 5). This value is much lower than Jerome and Ford's (2002) values (G_{ST} = 0.286, Table 5). I do note, however, that our sampled geographical area (220 km latitudinal by 280 km longitudinal area) is much smaller than Jerome and Ford's (2800 km latitudinal by 2400 km longitudinal area). I note that our pairwise F_{ST} values were mostly less than 0.2, in accordance with Jerome and Ford's (2002) F_{ST} values for data subsets with similar geographic areas and distances. Further, while Jerome and Ford's (2002) total populations G_{ST} value is likely a bit large for comparison, the authors also have a G_{ST} value limited to *A. americanum* distributed in the United States of America (USA) on *P. contorta* var. *latifolia* hosts. This USA G_{ST} value (G_{ST} = 0.142) lines up between the outcrossing (G_{ST} = 0.091), explosive dispersed seeds (G_{ST} = 0.092), and animal pollination (G_{ST} = 0.178) plants as would be expected.

I would expect fixation index values for *A. americanum* to be similar to those reported for other plant species with similar life-history traits— e.g., those with outcrossing (G_{ST} = 0.091), explosively dispersed seeds (G_{ST} = 0.092), and animal pollination (G_{ST} = 0.178) (Table 5). Our F_{ST} value (CNA F_{ST} = 0.044), however, is much smaller than all of these,

likely a consequence, at least in part, of our relatively small study area. Of interest, I found that our CNA F_{ST} value was 57% larger than the reforested *P. contorta* F_{ST} value ($F_{ST} = 0.028$, Table 5), in a study with a geographic extent similar to ours. *P. contorta*'s seed dispersal capabilities – up to 60 m (Alexander 1986) – is expected to be 375% greater than *A. americanum*'s seed dispersal – up to 16 m seed dispersal (Hawksworth and Wiens 1996) – which may explain this difference. Pollen dispersal cannot be properly compared due to a lack of separation of seed and pollen limitations to *A. americanum* dispersal in our work. However, previous studies have shown that seed dispersal affects genetic differentiation and gene flow more than pollen dispersal due to seeds carrying twice as much genetic information than pollen in diploid plants (Loveless and Hamrick 1984, Williams and Guries 1994).

Insular biogeography

In agreement with our second hypothesis, and previous comparisons of genetic diversity and island size (Frankham 1996), I found a positive correlation between island size and allelic richness (Figure 4), and negative correlation between distance from mainland and allelic richness (Figure 5). Thus our results support the predictions of the infinite island model and the theory of island biogeography.

As a caveat I note our associations were significant at $\alpha = 0.1$, but not $\alpha = 0.05$. This result, however, is likely due to inflation of standard errors as a result of small sample sizes ($n=6$ and $n=4$ for island size and island distance, respectively).

Ontogeny of *A. americanum* islands

For our third hypothesis, concerning the genetic origins of *A. americanum* islands; our UPGMA dendrogram and subsequent pruning analysis, and Bayesian cluster analysis were in accordance. Both analyses identified the presence of two clusters with one cluster including Central Idaho, Georgetown Canyon, Island Park, Pomerelle, and Skinner Canyon, and the other cluster consisting of Pike Mountain.

These results partially agree with our predictions. In particular, the Bayesian cluster analysis indicates that the Skinner Canyon population is more genetically similar to the Island Park mainland than Central Idaho, and that Pike Mountain is more genetically similar to Central Idaho than Island Park. Georgetown Canyon, however, appears to be a genetic mixture of the mainland populations. I do note that *A. americanum* distributional maps (Hawksworth and Weins 1996) indicate that Georgetown Canyon has the potential for additional island and mainland populations (e.g., Colorado populations) in closer proximity to GTC than other islands in this study. Skinner Canyon is spatially proximal to Georgetown Canyon (32 km), but from distribution maps (Hawksworth and Weins 1996) it may have greater genetic connectivity with the Utah populations of *A. americanum*, which could explain the observed genetic differentiation between Skinner Canyon and Georgetown Canyon. Also in contradiction to our

prediction, the Pomerelle population was more genetically similar to the Island Park mainland than Central Idaho. The Pomerelle population, however, evinced bottlenecks in the NAI dataset and had low allelic richness (Table 4). Luikart et al. (1998) demonstrated that bottlenecks may cause a reduction or loss of rare alleles, exaggerating low diversity in that population. This loss of alleles for Pomerelle could explain its apparent divergence from Central Idaho and Pike Mountain.

Pike Mountain also displayed evidence of a bottleneck in both NAI and 4L datasets and had low allelic richness (Table 4). This suggests that Pike Mountain and Pomerelle (spatially adjacent sites) were recently colonized and/or have had recent dramatic reductions in population size. Possible causes of recent population reduction may include: undocumented mistletoe management, fire, and reduction in habitat through logging and other anthropogenic impacts. Both Pike Mountain and Pomerelle are used for recreational activities (e.g. ski resort, camping), which have prompted the removal of trees and active control of *A. americanum* infection to maintain tree health (K. Fuelling, personal communication, 2016). Pike Mountain has also had a history of recent fires which likely reduced both population size and host habitat (Poppino 2008, Klass 2012).

Conclusions

I found evidence, using genetic assays, that *A. americanum* is a dispersal limited species and that may have similar genetic differentiation and gene flow properties to outcrossing, explosive dispersed seed, and animal pollinated plants. Additionally, this study suggests that *A. americanum* islands are affected by their insular biogeography.

These effects include higher levels of genetic diversity on large island sizes with close proximity to a mainland. Finally, my work suggests ontogenic patterns for *A.*

americanum populations at the Northern edge of the Great Basin. In particular, the Pike Mountain population appears to be colonized from Central Idaho and the Skinner Canyon population appears to be colonized from Island Park mainland population. It is my hope that this information will help foresters and ecologists to better manage and understand forest ecosystems in the study site region and locations like it in Western North America.

Figures and Tables

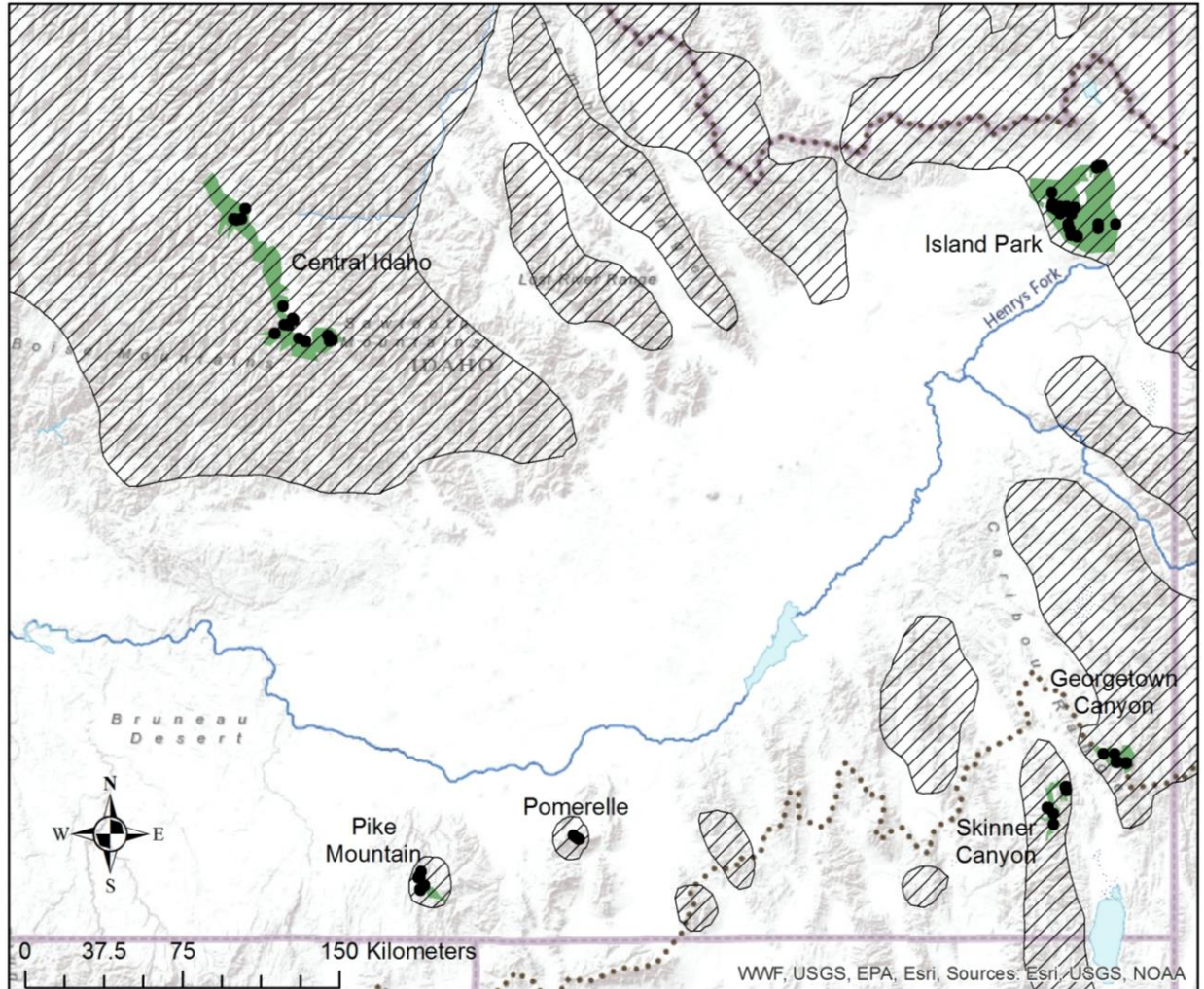


Figure 1. Map of Southeast and Central Idaho. Hatched areas denote coarse hulls of contiguous *Pinus contorta* cover (U.S. Geological Survey, 1999), green areas define bounds for sampling, and black dots indicate individual sample locations.

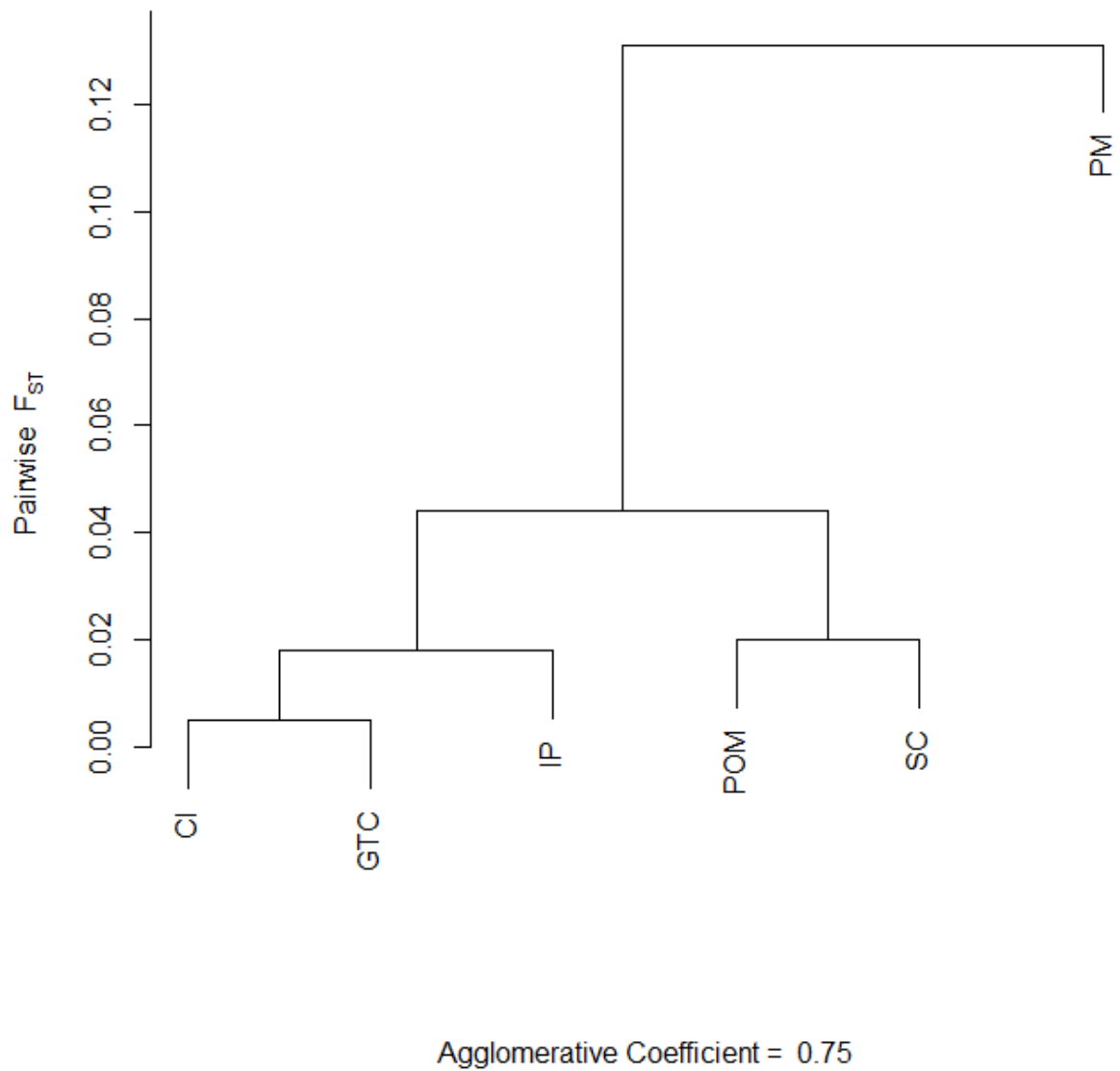


Figure 2. Cluster diagram for *Arceuthobium americanum* displaying two clusters from the sampled populations. Pairwise F_{ST} 's were calculated using Weir (1996) pairwise F_{ST} with the ENA correction procedure (Chapuis and Estoup 2007).

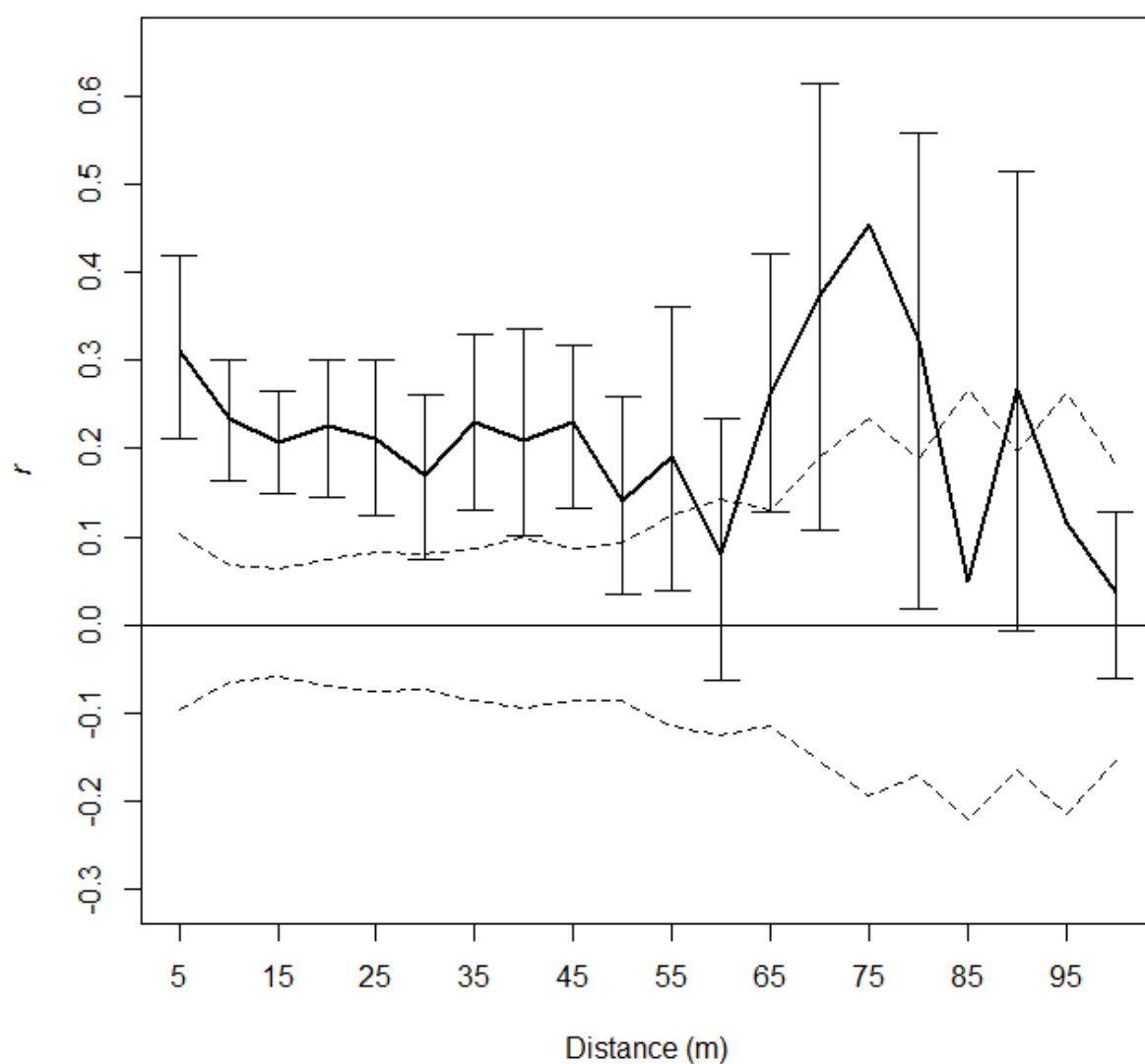


Figure 3. Spatial autocorrelation diagram for *A. americanum* given 5 meter distance classes. The thick solid line is the degree of genetic similarity (r) between plants within that distance class, with bootstrapped 95% confidence intervals for the true correlation. The dashed lines are the upper and lower 95% confidence intervals for 0.

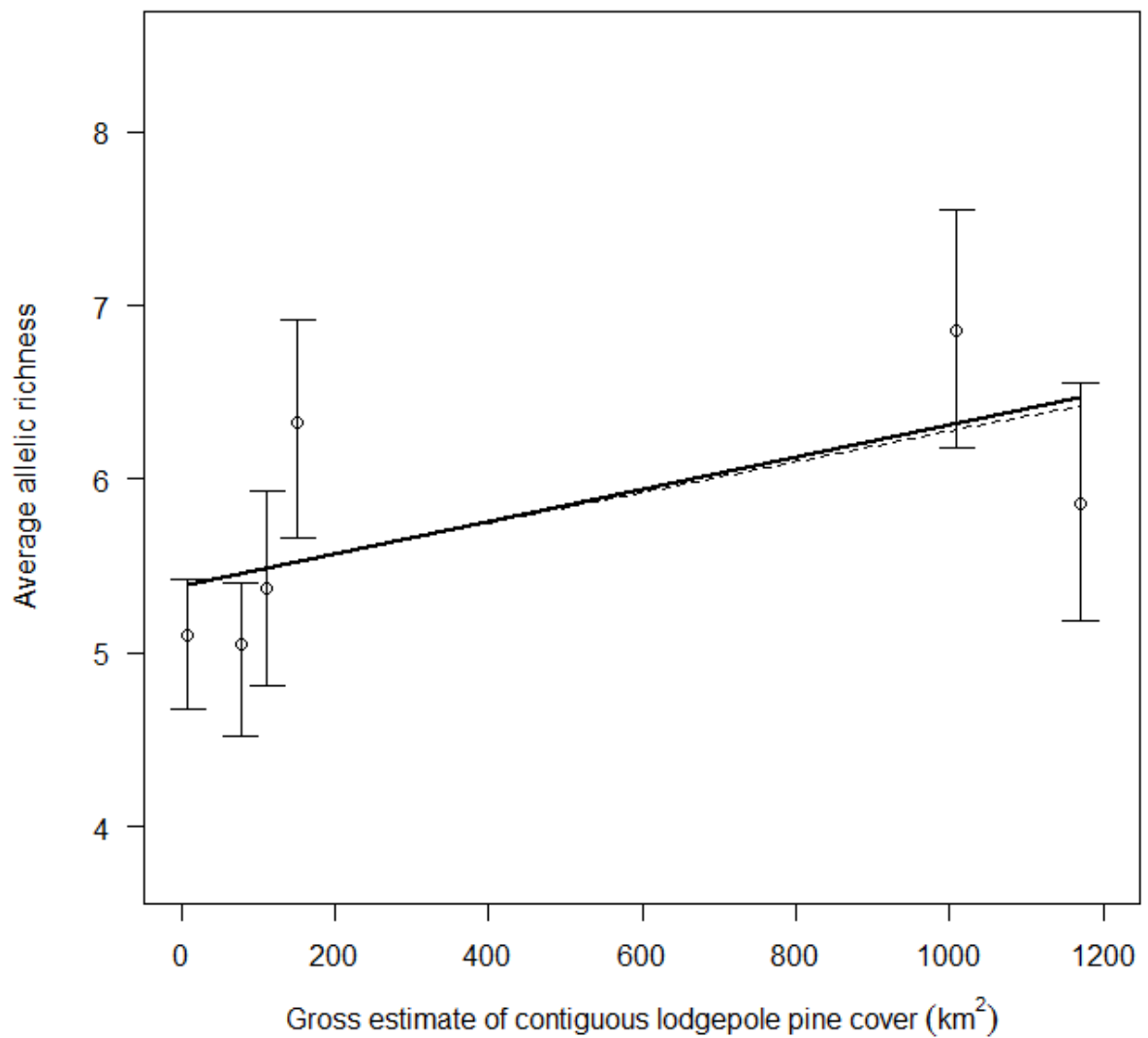


Figure 4. Scatterplot of allelic richness as a function of gross estimate of contiguous lodgepole pine cover (km²). Solid line depicts the reduced major axis (RMA) regression (slope= 9.3×10^{-4} , $R^2=0.41$, upper tail $p=0.093$), while the dashed line depicting ordinary least squares (OLS) regression (slope= 9.0×10^{-4} , $R^2=0.41$, upper tail $p=0.093$). Approximate bootstrapped (9999 iterations) standard errors for average allelic richness are also shown.

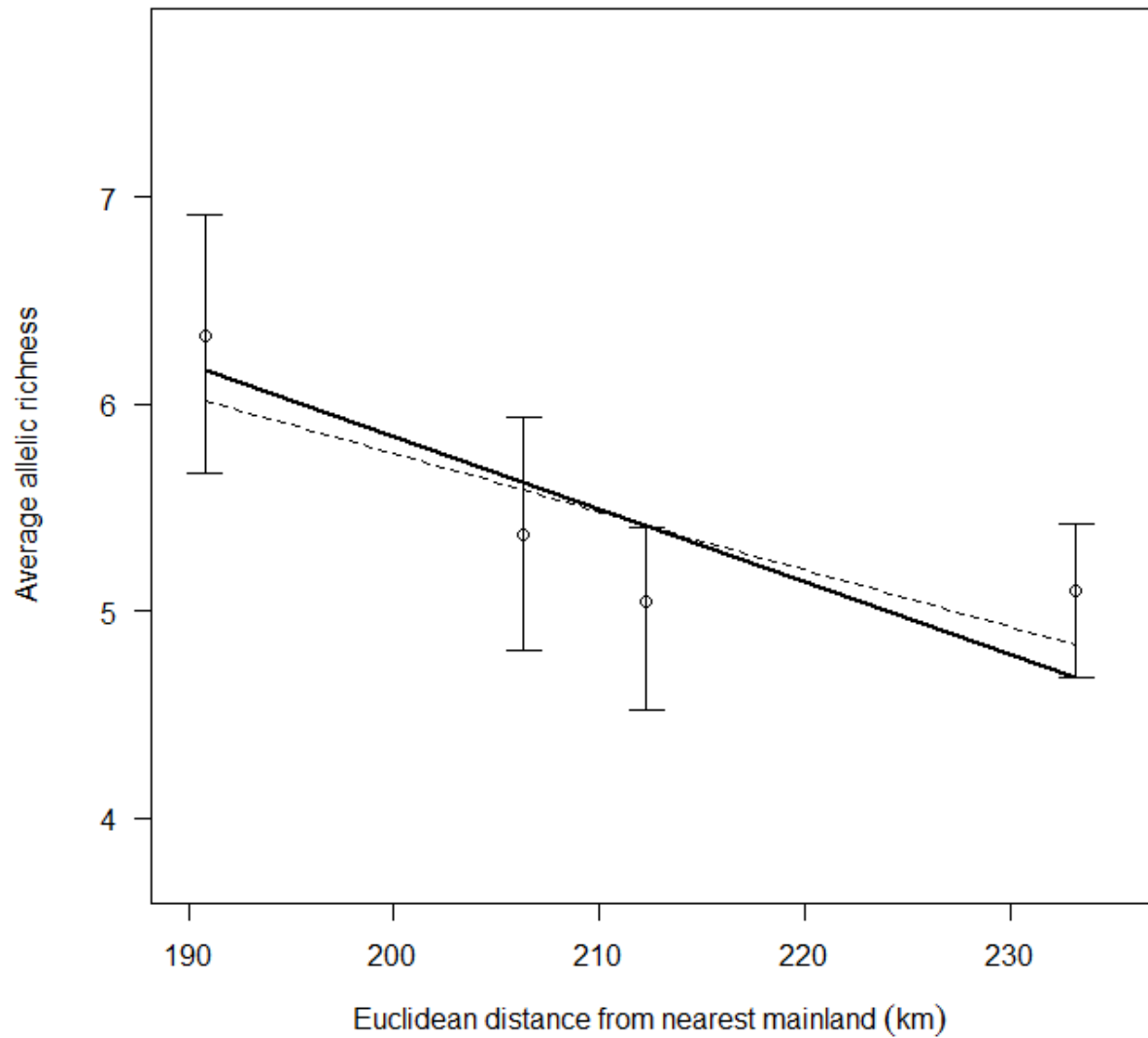


Figure 5. Scatterplot of allelic richness as a function of Euclidean distance from nearest mainland (km). Solid line depicts the reduced major axis (RMA) regression (slope= -2.8×10^{-2} , $R^2=0.68$, lower tail $p=0.080$), whereas dashed line depicting ordinary least squares (OLS) regression (slope= -3.5×10^{-2} , $R^2=0.68$, lower tail $p=0.080$). Approximate bootstrapped (9999 iterations) standard errors for average allelic richness are also shown.

Table 1. Names of collection sites for *Arceuthobium americanum* with population designation, and other pertinent information. Climate information were obtained from the National Climatic Data Center (NCDC) which is a department of the National Oceanic and Atmospheric Administration (NOAA, Asheville, North Carolina, USA). Gross estimates of *P. contorta* cover were based on topographical and satellite maps for the study area.

Population	Type of population	Number of samples collected	Average Latitude	Average Longitude	Average elevation (m)	Average yearly precipitation (mm)	Average min. Jan. temp (°C)	Average max Aug. temp (°C)	Gross estimate of contiguous <i>P. contorta</i> cover (km ²)
Central Idaho	Mainland	75	44.00	-114.86	2165	29.4	-18.3	25.8	1009
Pike Mountain	Island	25	42.19	-114.27	2102	23.4	-7.1	29.4	77
Pomerelle	Island	15	42.32	-113.61	2407	23.4	-7.1	29.4	8
Island Park	Mainland	83	44.28	-111.48	1958	60.2	-16.1	25.8	1170
Georgetown Canyon	Island	18	42.57	-111.30	2221	36.9	-13.3	28.6	151
Skinner Canyon	Island	25	42.42	-111.55	2275	36.9	-13.3	28.6	112

Table 2. Euclidean distance (in km) (lower triangle) and Weir (1996) ENA corrected pairwise F_{ST} (upper triangle) matrix for the sampled populations.

Sites	Central Idaho	Pike Mountain	Pomerelle	Island Park	Georgetown Canyon	Skinner Canyon
Central Idaho	0.0	0.0603*	0.0587	0.0204*	0.0050	0.0506*
Pike Mountain	207.6	0.0	0.1968*	0.1220*	0.0752*	0.2015*
Pomerelle	212.9	56.8	0.0	0.0227*	0.0494*	0.0201
Island Park	271.5	324.7	277.7	0.0	0.0155*	0.0344*
Georgetown Canyon	329.2	248.2	191.6	190.3	0.0	0.0489*
Skinner Canyon	320.7	225.9	169.8	206.2	26.1	0.0

* indicates significance with bootstrapped 95% confidence interval (10000 bootstraps).

Table 3. Bayesian proportional assignment of *Arceuthobium americanum* sites into two clusters.

Site	Clusters	
	1	2
Central Idaho	0.5914	0.4086
Georgetown Canyon	0.6192	0.3808
Island Park	0.8941	0.1059
Pike Mountain	0.1566	0.8434
Pomerelle	1.0000	0.0000
Skinner Canyon	0.9820	0.0180

Table 4. F_{IS} , allelic richness, and presence (+) or absence (-) of recent bottleneck for *Arceuthobium americanum* populations using the 4L dataset. $\pm$ indicates a positive bottleneck result for NAI dataset and a negative result for 4L dataset.

Site	F_{IS}	Allelic Richness	Bottleneck
Central Idaho	0.481*	6.86 (5.50, 8.25)	-
Pike Mountain	0.399*	5.05 (4.00, 5.75)	+
Pomerelle	0.122	5.10 (4.25, 5.75)	$\pm$
Island Park	0.450*	5.86 (4.50, 7.25)	-
Georgetown Canyon	0.305*	6.33 (5.00, 7.50)	-
Skinner Canyon	0.335*	5.37 (4.25, 6.50)	-

* indicates significant at $\alpha = 0.05$.

Table 5. Overall F_{ST} or G_{ST} and F_{IS} values for *Arceuthobium. americanum* and comparisons to other species in the literature.

Species	F_{IS}	F_{ST} or G_{ST}	Reference
<i>A. americanum</i> - full dataset no NA corrections	0.516	0.035	This Study
<i>A. americanum</i> - full dataset with NA corrections	0.107	0.044	This Study
<i>A. americanum</i> - with 4 best loci	0.396	0.045	This Study
<i>A. americanum</i>	N/A	0.286*	Jerome & Ford (2002)
<i>A. americanum</i> - USA on <i>P. contorta</i> var. <i>latifolia</i>	N/A	0.142*	Jerome & Ford (2002)
<i>Pinus contorta</i>	0.060	0.007	Yeh et al. (1985)
Reforested <i>Pinus contorta</i>	0.360	0.028	Thomas et al. (1999)
Non-colonizing conifers	N/A	0.002	Govindaraju (1988b)
Colonizing conifers	N/A	0.002	Govindaraju (1988b)
Ingested dispersed seed plants	N/A	0.051*	Hamrick et al. (1992)
Sexual and asexual reproductive plants	N/A	0.051*	Hamrick et al. (1992)
Animal attached dispersed seed plants	N/A	0.065*	Hamrick et al. (1992)
Wind dispersed seed plants	N/A	0.076*	Hamrick et al. (1992)
Wind pollinated plants	N/A	0.084*	Govindaraju (1988c)
Sexual reproductive plants	N/A	0.086*	Hamrick et al. (1992)
Outcrossing plants	N/A	0.091*	Govindaraju (1988a)
Explosive dispersed seed plants	N/A	0.092*	Hamrick et al. (1992)
Gravity dispersed seed plants	N/A	0.131*	Hamrick et al. (1992)
Animal pollinated plants	N/A	0.178*	Govindaraju (1988c)
Mixed mating plants	N/A	0.216*	Govindaraju (1988a)
Primarily outcrossing plants	N/A	0.277*	Govindaraju (1989)
Self pollinating plants	N/A	0.446*	Govindaraju (1988a, 1988c, 1989)

* indicates a G_{ST} value

Literature cited

- Aho, K., D. W. Roberts, and T. Weaver. 2008. Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods. *Journal of Vegetation Science* 19:549–562.
- Barrett, S. W., and S. F. Arno. 1991. Classifying fire regimes and defining their topographic controls in the Selway-Bitterroot Wilderness. *Proceedings of the Society of American Foresters National Convention (USA)*.
- Brown, J. H. 1971. Mammals on mountaintops: nonequilibrium insular biogeography. *The American Naturalist* 105:467–478.
- Chapuis, M.-P., and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–631.
- Chen, C., E. Durand, F. Forbes, and O. François. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* 7:747–756.
- Chybicki, I. J., and J. Burczyk. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100:106–113.
- Coppola, J. 1989. Wind dispersal of *Arceuthobium americanum* pollen. *The Southwestern Naturalist* 34:291.
- Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001–2014.

- deBruyn, R. A. J., M. Paetkau, K. A. Ross, D. V. Godfrey, and C. R. Friedman. 2015. Thermogenesis-triggered seed dispersal in dwarf mistletoe. *Nature Communications* 6:6262.
- Efron, B. 1987. Better bootstrap confidence intervals. *Journal of the American Statistical Association* 82:171–185.
- Escudero, A., J. M. Iriondo, and M. E. Torres. 2003. Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation* 113:351–365.
- ESRI. 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.
- Fazekas, A. J., and F. C. Yeh. 2006. Postglacial colonization and population genetic relationships in the *Pinus contorta* complex. *Botany* 84:223–234.
- Floyd, C. H., D. H. Van Vuren, and B. May. 2005. Marmots on great basin mountaintops: using genetics to test a biogeographic paradigm. *Ecology* 86:2145–2153.
- Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. *Conservation Biology* 10:1500–1508.
- Frankham, R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity* 78:311–327.
- Fuelling, K. 2016. Personal communication.
- Guillot, G., F. Mortier, and A. Estoup. 2005. Geneland: a computer package for landscape genetics. *Molecular Ecology Notes* 5:712–715.
- Harrington, M., and F. G. Hawksworth. 1990. Interactions of fire and dwarf mistletoe on mortality of southwestern ponderosa pine. United States Department of Agriculture, Forest Service General Technical Report RM-191:234–240.

- Hawksworth, F. G., T. H. Nicholls, and L. M. Merrill. 1987. Long-distance dispersal of lodgepole pine dwarf mistletoe. General technical report RM-Rocky Mountain Forest and Range Experiment Station, United States Department of Agriculture, Forest Service.
- Hawksworth, F. G., and D. Wiens. 1996. Dwarf Mistletoes: Biology, Pathology, and Systematics. United States Department of Agriculture Forest Service.
- Heywood, J. S. 1991. Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*:335–355.
- Hinds, T. E., F. G. Hawksworth, and W. J. McGinnies. 1963. Seed discharge in *Arceuthobium*: a photographic study. *Science* 140:1236–1238.
- Hudler, G. W., N. Oshima, and F. G. Hawksworth. 1979. Bird dissemination of dwarf mistletoe on ponderosa pine in Colorado. *The American Midland Naturalist* 102:273–280.
- Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal. *Heredity* 77:282–291.
- Jaenike, J. R. 1973. A steady state model of genetic polymorphism on islands. *The American Naturalist* 107:793–795.
- Jerome, C. A., and B. A. Ford. 2002. The discovery of three genetic races of the dwarf mistletoe *Arceuthobium americanum* (Viscaceae) provides insight into the evolution of parasitic angiosperms. *Molecular Ecology* 11:387–405.
- Kaufman, L., and P. J. Rousseeuw. 2005. Finding Groups in Data: An Introduction to Cluster Analysis. 1 edition. Wiley-Interscience, Hoboken, N.J.

- Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. 2013. *diveRsity*: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 4:782–788.
- Klass, K. 2012. South Hills fire grows-road still closed.
<http://newsradio1310.com/breaking-another-fire-in-south-hills/>.
- Lomolino, M. V., J. H. Brown, and R. Davis. 1989. Island biogeography of montane forest mammals in the American Southwest. *Ecology*:180–194.
- Lomolino, M. V., and D. R. Perault. 2001. Island biogeography and landscape ecology of mammals inhabiting fragmented, temperate rain forests. *Global Ecology and Biogeography* 10:113–132.
- Loo, J., O. Souvannavong, and I. K. Dawson. 2014. Seeing the trees as well as the forest: The importance of managing forest genetic resources. *Forest Ecology and Management* 333:1–8.
- Loveless, M. D., and J. L. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15:65–95.
- Luikart, G., F. W. Allendorf, J. M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89:238–247.
- Lunn, D., C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter. 2012. *The BUGS Book: A Practical Introduction to Bayesian Analysis*. 1 edition. Chapman and Hall/CRC, Boca Raton, FL.
- MacArthur, R. H., and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17:373–387.

- MacArthur, R. H., and E. O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press.
- MacKay, W. P., and E. Mackay. 2002. *The ants of New Mexico:(Hymenoptera: Formicidae)*. Edwin Mellen Press Lewiston, NY.
- Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2015. *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.1.
- Marlowe, K., and L. Hufford. 2008. Evolution of *Synthyris* sect. *Dissecta* (Plantaginaceae) on sky islands in the Northern Rocky Mountains. *American journal of botany* 95:381–392.
- Mathis, W. N., and D. Mathis. 2008. Shore flies (Insecta: Diptera: Ephydriidae) of Plummers Island, Maryland. *Bulletin of the Biological Society of Washington* 15:173–191.
- Nicholls, T., L. Egeland, F. G. Hawksworth, D. Johnson, and K. Robbins. 1987. Control of dwarf mistletoes with a plant growth regulator. United States Department of Agriculture, Forest Service General Tech Report RM-149:154–156.
- Nicholls, T. H. 2014. Evolution of a Short-Term Study of Lodgepole Pine Dwarf Mistletoe Vectors that Turned Into a Long-Term Study of the Remarkable Gray Jay on the Fraser Experimental Forest, Colorado 1982–2009. Pages 251–270 in D. C. Hayes, S. L. Stout, R. H. Crawford, and A. P. Hoover, editors. United States Department of Agriculture Forest Service Experimental Forests and Ranges. Springer New York.
- Nicholls, T. H., and F. G. Hawksworth. 1983. Animal vectors of lodgepole pine dwarf mistletoe. *Phytopathology* 73:836.

- Nicholls, T. H., F. G. Hawksworth, and L. M. Merrill. 1984a. Animal vectors of dwarf mistletoe, with special reference to *Aceuthobium americanum* on lodgepole pine. In: Hawksworth FG, Scharpf RF (eds) Biology of dwarf mistletoes: proceedings of the symposium, 1984 August 8, Fort Collins, CO. General Technical Report RM-111. US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO:102–111.
- Nicholls, T. H., F. G. Hawksworth, and L. M. Merrill. 1984b. A preliminary look at animal vectors of lodgepole pine dwarf mistletoe. In: Dubreuil SH (Comp) Proceedings of the 31st Western International Forest Disease Work Conference, 1983 August 22–26, Coeur d’Alene, ID. US. Department of Agriculture, Forest Service, Northern Region, Missoula, MT:54–55.
- Nickrent, D. L. 1996. Molecular Systematics. In Dwarf Mistletoes: Biology, Pathology, and Systematics (F.G. Hawksworth & D. Wiens, Eds.). United States Department of Agriculture Forest Service.:155–172.
- Nickrent, D. L., M. A. García, M. P. Martín, and R. L. Mathiasen. 2004. A phylogeny of all species of *Arceuthobium* (Viscaceae) using nuclear and chloroplast DNA sequences. American Journal of Botany 91:125–138.
- Ostry, M. E., and T. H. Nicholls. (n.d.). The forest beyond the trees: beneficial roles of disease in forest health. In: Proceedings of the 1997 Society of American Foresters National Convention, Meeting in the Middle, 1997 October 4–8; Memphis, TN. Society of American Foresters, Bethesda, MD:167–172.

- Parker, P. G., A. A. Snow, M. D. Schug, G. C. Booton, and P. A. Fuerst. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* 79:361–382.
- Peakall, R., and P. E. Smouse. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295.
- Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)* 28:2537–2539.
- Penfield, F. B., R. E. Stevens, and F. G. Hawksworth. 1976. Pollination ecology of three Rocky Mountain dwarf mistletoes. *Forest Science* 22:473–484.
- Poppino, N. 2008. South Hills on fire.
http://magicvalley.com/news/local/south-hills-on-fire/article_76a1490d-fd6b-5f40-929a-c806fd16e967.html.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Romme, W. H., and M. G. Turner. 1991. Implications of global climate change for biogeographic patterns in the Greater Yellowstone Ecosystem. *Conservation Biology*:373–386.
- Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787–792.
- Sokal, R. R., and C. D. Michener. 1958. A Statistical Method for Evaluating Systematic Relationships. University of Kansas.

- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64:583–639.
- Stevens Jr., D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association* 99:262–278.
- Thomas, B. R., S. E. Macdonald, M. Hicks, D. L. Adams, and R. B. Hodgetts. 1999. Effects of reforestation methods on genetic diversity of lodgepole pine: an assessment using microsatellite and randomly amplified polymorphic DNA markers. *TAG Theoretical and Applied Genetics* 98:793–801.
- U.S. Geological Survey. 1999. Digital representation of “Atlas of United States Trees” by Elbert L. Little, Jr..
- Vellend, M. 2003. Island biogeography of genes and species. *The American Naturalist* 162:358–365.
- Wanner, J., and R. O. Tinnin. 1986. Respiration in lodgepole pine parasitized by American dwarf mistletoe. *Canadian Journal of Forest Research* 16:1375–1378.
- Weir, B. S. 1996. *Genetic Data Analysis 2: Methods for Discrete Population Genetic Data*. 2 Sub edition. Sinauer Associates, Inc., Sunderland, Mass.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *evolution*:1358–1370.
- Wheeler, N. C., and R. P. Guries. 1982. Biogeography of lodgepole pine. *Canadian Journal of Botany* 60:1805–1814.

- Williams, C. F., and R. P. Guries. 1994. Genetic consequences of seed dispersal in three sympatric forest herbs. I. Hierarchical population-genetic structure. *Evolution* 48:791–805.
- Williams, C. F., J. Ruvinsky, P. E. Scott, and D. K. Hews. 2001. Pollination, breeding system, and genetic structure in two sympatric *Delphinium* (Ranunculaceae) species. *American Journal of Botany* 88:1623–1633.
- Wright, S. 1931. Evolution in mendelian populations. *Genetics* 16:97–159.
- Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics* 15:323–354.
- Yeh, F. C., W. M. Cheliak, B. P. Dancik, K. Illingworth, D. C. Trust, and B. A. Pryhitka. 1985. Population differentiation in lodgepole pine, *Pinus contorta* spp. *latifolia*: a discriminant analysis of allozyme variation. *Canadian Journal of Genetics and Cytology* 27:210–218.

Appendix

Molecular approaches

A myriad of molecular techniques exist for gathering population genetics data. Examples include allozymes (enzymes with differing tertiary or quaternary structures, coded by different alleles at the same locus), restriction fragment length polymorphisms (RFLPs) including terminal-restriction and amplified fragment length polymorphism analysis (T-RFLP, AFLP), single nucleotide polymorphisms (SNPs) which often use the above methods for identification, and microsatellite variable number tandem repeat polymorphisms (VNTRs). Microsatellites are particularly useful in studies of population genetics due to being a co-dominant marker, the ability of identification of individual loci genotype, typically having high polymorphism per locus, and having a low cost per sample after primers have been developed (Parker et al. 1998).

F-statistics

Patterns in population genetics are typically quantified with statistics designed to estimate genetic differentiation, gene flow, and heterozygosity. Wright's (1951) *F*-statistics – F_{IT} , F_{ST} , and F_{IS} – allow measurements of genetic differentiation, gene flow, and heterozygosity. Variants of Wright's *F*-statistics (Weir and Cockerham 1984, Weir 1996) have been developed largely to account for the characteristics and constraints of new molecular tools (see above). .

F_{ST} (Genetic differentiation; Eq. 1) measures the genetic dissimilarity of subpopulations, and is often used to identify plant evolutionary and distributional histories (Wheeler and Guries 1982, Williams et al. 2001, Fazekas and Yeh 2006). F_{ST} ranges from zero, indicating complete panmixis to one, indicating complete genetic differentiation among subpopulations. Formally speaking, gene flow (Nm ; i.e., the denominator in Eq. 1) is the spatial movement of genetically independent individuals (i.e. plant seeds or pollen) between populations or subpopulations, and thus is negatively correlated with F_{ST} . Gene flow tends to increase the genetic diversity of a population as new genes are moved between populations (Slatkin 1987). A more diverse gene pool, as determined by genetic differentiation and gene flow, is more likely to provide a population with a diverse array of phenotypic answers to selective pressures, thus protecting against local extinction (Frankham 1996). F_{IS} (Eq. 2) is a numerical value ranging from -1 to 1. The F_{IS} quantifies heterozygosity relative to the Hardy-Weinberg equilibrium (HWE) for a population. Negative values indicate an excess of heterozygosity, whereas positive values indicate a deficit of heterozygosity. F_{IT} (Eq. 3) measures the variation of heterozygotes from the combined individual and subpopulation heterozygosity, giving the total fixation index.

$$F_{ST} = \frac{(H_T - H_S)}{(H_T)} = \frac{1}{(4Nm + 1)} \quad (1)$$

$$F_{IS} = \frac{(H_S - H_I)}{(H_S)} \quad (2)$$

$$F_{IT} = \frac{(H_T - H_I)}{(H_T)} \quad (3)$$

H_I = observed heterozygosity, per individual within subpopulations

H_S = expected heterozygosity within a subpopulation

H_T = expected heterozygosity in the total population

Cluster diagram and cluster analysis

Distance based phylogenetic trees constitute a useful method for visually expressing the genetic relations of a group of populations. Two methods are often used: cluster diagram and Bayesian cluster analyses. The UPGMA cluster diagram method (Sokal & Michener 1958) uses a genetic differentiation measure, e.g., Weir (1996) genetic differentiation, making it more appropriate for population analyses compared to character based phylogenetic trees which are often used for species evolution questions. Bayesian cluster analysis uses Markov-Chain Monte Carlo (MCMC) computational methods to quantify membership probabilities of individuals without assuming predefined populations (Chen et al. 2007).