ABSTRACT

Title of Dissertation: PHYLOGEOGRAPHY OF AN INTRODUCED INSECT PEST AND CONSEQUENCES OF AN INSECT INTRODUCTION

Robert G. Ahern, Doctor of Philosophy, 2007

Directed By: Professor Michael J. Raupp
Department of Entomology

Adelges cooleyi (Hemiptera: Adelgidae), a host-alternating gall-making insect pest native to the Rocky Mountains and Pacific Northwest and introduced into the eastern United States during the mid-19th century, was studied to address questions about phylogeography, to determine effects of introduction on genotypic and phenotypic variation, and to compare genetic variation associated with host use in native and introduced ranges. In Chapter One, sequence data from two mitochondrial (mtDNA) genes and amplified fragment length polymorphisms (AFLPs) were used to quantify the structure of genetic variation in the insect's native range. Several well-supported, divergent mtDNA lineages were identified. The structure of genetic variation among sampled locations is consistent with patterns shaped by glaciations. Samples from the southern edge of the insect's distribution are genetically isolated from the rest of the species, and hybridization of divergent mtDNA lineages via secondary contact was inferred from AFLP data.

Changes in genetic and phenotypic variation associated with introduction were quantified in Chapter Two. Introduced populations had decreased genetic variation relative to native populations. Variation in an ecologically important trait, host preference, was also significantly lower in introduced populations than in native populations. An association between mtDNA haplotypes and host preference was identified. *Adelges cooleyi* in the eastern US have low genetic and phenotypic variation but appear to be sufficiently adapted for persistence. My results call into question the utility of neutral genetic variation to assess the probability of persistence in new environments by introduced species.

Host-plants that *A. cooleyi* requires to complete its lifecycle are not native to the eastern US and occur together in patches that are often widely separated. In Chapter Three, analyses of mtDNA and AFLP genetic variation were conducted to determine the distribution of genetic variation within and among host plants in the native range and identify discrepancies that may be consistent with an incomplete lifecycle in the introduced range. Distribution of genetic variation within and among host-plants in the introduced range was not significantly different than that in the native range, as indicated by fixation indices, and I found no evidence for asexual populations in the introduced range.

PHYLOGEOGRAPHY OF AN INTRODUCED INSECT PEST AND CONSEQUENCES OF AN INSECT INTRODUCTION

By

Robert G. Ahern

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Michael Raupp, Chair
Associate Professor David Hawthorne
Professor Galen Dively
Adjunct Professor Sonja Scheffer
Assistant Professor Matthew Hare
Associate Professor Kerry Shaw

© Copyright by
Robert G. Ahern
2007

Acknowledgements

I would like to thank my committee members for their time, assistance, and guidance with my project. In particular, I would like to thank Kerry Shaw for inviting me into her lab group and encouraging me to pursue questions about speciation, and Dave Hawthorne for sharing his lab space, time, expertise and friendship with me. Also, the freedom that Mike Raupp gave me to design my own project is the greatest gift that an adviser can give, and I am more confident and a better scientist because of this opportunity.

I received help from many people, without whom I certainly would have failed. Joan West provided invaluable laboratory assistance and technical expertise, and members of the Shaw and Hawthorne labs helped with questions about theory and logistical problems. Members of the Raupp lab helped with collections, sample counting, and myriad other chores, and I am indebted to each of them for their contributions and friendship. My project would have taken far longer in not for Greg Hess's computer expertise. John Davidson's early support and encouragement were vital in my pursuit of a study organism. I thank Holly Menninger for her friendship and humor, and Steve Frank for being a superlative colleague and friend. My friends at 3706 Alton Place made my time at Maryland better than I could have possibly imagined, and I would like to thank Susan Duberstein for always being there for me, as a friend, an accomplice, a partner, and so much more- I am truly fortunate to have met you. Finally, I would like to thank my brother, sister, and parents for all of their interest and support, in this endeavor and all others.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1: Phylogeography of a specialist insect, *Adelges cooleyi* (Hemiptera: Adelgidae): historical and contemporary processes shape the distribution of population genetic variation

Abstract

The distribution of genetic variation of specialist insects is constrained and defined by distribution of host plants, and glacial activity has had a profound effect on the distribution of genetic variation across large geographic scales. *Adelges cooleyi* is a host-alternating gall-making insect native to the Rocky Mountains and Cascades in the western United States. The insect's primary hosts include various spruce (*Picea*) species, and its secondary host is Douglas fir, *Pseudotsuga menziesii*. In this study I created molecular phylogenies of geographically separate samples of *A. cooleyi* using sequence data from two mitochondrial (mtDNA) genes and amplified fragment length polymorphisms (AFLPs). Three divergent mtDNA lineages were identified. Analysis of mtDNA and AFLP genetic variation identified samples from southeastern Arizona that are genetically isolated from all other samples. Finally, AFLP data identified an area in the central Rockies were hybridization between individuals from divergent mtDNA lineages can be inferred. Factors that likely affected divergence within *A. coolyei* were identified by comparing my conclusions with previous phylogeographic work conducted on this specialist insect's host-plants.

Introduction

Plant diversity affects insect diversity (Futuyma and McCafferty 1990, Mitter et al. 1991), and the geographic distribution of specialized herbivorous insects is directly affected by the distribution of their plant hosts (Funk et al. 1995, Brown et al. 1996, Brown et al. 1997). Phylogeographic studies of specialist insects are fundamentally different than other types of phylogeographic studies because phylogeography of insect specialists is constrained and defined by the distribution of hosts. Good examples of congruence between insect and host phylogenies can be found in several beetle genera (Futuyma and McCafferty 1990, Becerra and Venable 1999), but little work has been done comparing specialist and host-plant phylogeographies. Phylogeographic congruence of an herbivore and its host plants may be a signature of coevolution, as proposed by Thompson (Thompson 1994, Thompson 1999), or it may be a remnant of drift associated with allopatric divergence and reduced gene flow among isolated populations (Hartl and Clark 1997, Knowles and Richards 2005). Regardless of the evolutionary processes underlying phylogenetic congruence, analysis of phylogeographic data from specialist insects and their hosts may be used to examine patterns of host distribution and differentiation that may have affected differentiation of specialist insects.

Phylogeographic analysis provides inferences into the effects of geographic structures on genetic fragmentation of populations (Avise et al. 1987). In addition to barriers such as mountains, rivers, and plains, glacial activity may have a profound effect on the distribution of genetic variation across large geographic scales (Hewitt

1996).Variation in mitochondrial DNA (mtDNA) has been used extensively to map genetic variation of natural populations (Avise 2000), and a large number of studies have categorized the population structure of a variety of organisms, including plants (Chiang et al. 2001), salamanders (Shaffer and McKnight 1996), vertebrates (Vila et al. 1999), and insects (Brown et al. 1996).

Mitochondrial DNA is maternally inherited and does not undergo substantial recombination (Hartl and Clark 1997, Sunnucks 2000) so it is difficult or impossible to use mtDNA data to detect important population biological processes such as sex-limited dispersal and intraspecific hybridization. In contrast, amplified fragment length polymorphisms (AFLPs) reflect a large, random sample of the nuclear genome (Mendelson and Shaw 2005) and are sensitive to contemporary forces shaping population structure (Sunnucks 2000). AFLP variation has been used to characterize genetic variation within and among closely related organisms, including plants (Elias et al. 2000), fish (Albertson et al. 1999), aquatic invertebrates (Barki et al. 2000), terrestrial vertebrates (Bensch et al. 2002), and insects (Parsons and Shaw 2001).

Reduced gene flow caused by reproductive isolation leads to population divergence due to genetic drift and selection, and these forces may differentially affect nuclear and organelle genomes. In relatively large populations, alleles may become fixed more quickly in maternally inherited organelles such as mtDNA because the genome's effective population size is smaller (< 1/4) that of the nuclear genome (Hare 2001). If populations are isolated for enough time, both mitochondrial and nuclear alleles will diverge from other populations and fixed differences will accrue. MtDNA and AFLP phylogenies that show similar topologies and levels of

support provide evidence for limited historical gene flow and isolation of populations, while incongruent phylogenies may have several interpretations. MtDNA alleles are often sorted and become fixed more quickly than nuclear alleles, and well-supported mtDNA branches that are not supported in AFLP phylogenies may indicate incomplete sorting of nuclear variation. This pattern may also indicate hybridization between divergent mtDNA lineages, and additional information about the ecology and history of organism(s) under consideration may aid in interpretation of incongruent mtDNA and AFLP phylogenies. Finally, AFLP phylogenies that identify groups that are not supported by mtDNA phylogenies may indicate very recent, ecologically driven divergence.

Analyses of mtDNA and AFLP variation in population genetic and phylogeographic studies are widespread and researchers have confidence in their utility. However, combination of data from both markers to study intraspecific phylogeography is still uncommon (Creer et al. 2004), and few researchers have tackled the difficult issue of incongruence between marker signals for identification of isolated or divergent populations (Parsons and Shaw 2001, Bensch et al. 2002, Shaw 2002, Bensch et al. 2006). Empirical studies that separate historical and contemporary gene flow by comparing mtDNA and AFLP data demonstrate the utility of this technique to make evolutionarily important observations, such as identification of hybrid zones, in natural populations. If insects at sample locations have been reproductively isolated, divergence should be evident in both mtDNA and AFLP phylogenies. Incongruence between mtDNA and AFLP phylogenies will indicate incomplete lineage sorting, intraspecific hybridization, or recent

ecologically-driven divergence. Here I show that combination of mtDNA and AFLP data enables inference into both historical and contemporary processes that shape population genetic variation in a specialist herbivore.

Study system. Gillette described *Adelges cooleyi* in 1907 (Gillette 1907, Annand 1928). The insect's primary hosts include various spruce (*Picea*) species, and its secondary host is Douglas fir, *Pseudotsuga menziesii*. This insect forms large galls at the end of terminal branches on spruce hosts and yellows and distorts needles on Douglas fir (Annand 1928). It requires both tree species to complete a complex lifecycle over two years (fig. 1.1) (Annand 1928). Both sexual and asexual reproduction occurs on primary hosts, while all reproduction on the secondary host is asexual.

The native range of CSGA resembles an inverted 'V', extending along the Rocky Mountains into Canada and back down the West Coast into northern California. This distribution is roughly congruent with the distribution of its primary and secondary hosts (Little 1971). At least three species of spruce are found in the native range of CSGA and serve as primary hosts: Colorado blue spruce, *Picea pungens*; Engelmann spruce, *Picea engelmannii*; and Sitka spruce, *Picea sitchensis* (Annand 1928, Little 1971). Variation among spruce hosts is compounded with substantial variation in *P. menziesii*. Previous studies on phenotypic variation (Rehfeldt 1977) and allozymes (Li and Adams 1989) have confirmed the existence of two varieties of Douglas fir, a coastal variety (var. *menziesii*) and an interior variety (var. *glauca*), that have been distinct since the Miocene,13 million years ago

(Critchfield 1984, Aagaard et al. 1998). The interior variety is further separated into 2 major races (northern and southern) distinguished by significant divergence in allozyme variation (Li and Adams 1989).

In this study I created molecular phylogenies of geographically separate samples of *A. cooleyi* using data from mitochondrial and nuclear genomes. I conducted analyses of mtDNA variation to identify related samples, and I used nested clade analysis to test the effect of geography on the distribution of mitochondrial haplotypes. I also conducted analyses of AFLP variation to identify related samples and samples from isolated locations. Results generated by analysis of data from non-recombining mitochondrial regions and recombining nuclear regions were considered together in an effort to determine whether both markers generated comparable conclusions about the distribution of genetic variation throughout the native range of *A. cooleyi*. Factors that likely affected divergence within *A. coolyei* were identified by comparing my conclusions with previous phylogeographic work conducted on this specialist insect's host-plants.

Methods

Material used. CSGA was collected throughout its native range on *Picea pungens*, *P. engelmannii*, and *P. sitchensis*. Insects were removed from galls in mid-summer (June-July) and stored in 100% EtOH until DNA extraction. With few exceptions, molecular analyses were conducted on at least 4 insects from separate galls from 24 locations distributed throughout the range (table 1.1). All locations were separated by at least 25km.

Laboratory procedures. Whole genomic DNA was extracted using DNeasy Tissue Kits (Qiagen) following a slightly modified protocol B for insects. Polymerase chain reaction (PCR) was used to amplify two fragments of the mitochondrial genome, cytochrome oxidase I (COI) and cytochrome oxidase II (COII). The target region of COI was amplified in a 20ul reaction using 1ul each of 5uM primers "Ron" (5'-GGA TCA CCT GAT ATA GCA TTC CC-3') and "Nancy" (5'-CCC GGT AAA ATT AAA ATA TAA ACT TC-3')(Simon et al. 1994) with 1ul DNA template under the following conditions: initial denaturing for 5:00 at 95 ºC, followed by 10 cycles of a touchdown-type reaction of denaturing for 0:10, annealing at 39.8 -0.7 ºC per cycle for 1:00, extension at 72 ºC for one minute. The touchdown reaction was followed by 30 cycles of denaturing for 0:10, annealing at 32.8 ºC for 1:00, and extension at 72 ºC for 1:00 + 0:01 per cycle; the reaction concluded with a final 4:00 extension at 72 ºC.

The target region of COII was amplified in a 20ul reaction using 1ul each of 5uM primers AdelCOIIF1 (5'-GCA GAA ACC AAT GCA ATG AAC-3': N. Havill, personal communication) and AdelCOIIR1 (5'-CGT CCR GGA ATT GCA TCT ATT-3': N. Havill, personal communication) with 1ul DNA template under the following conditions: initial denaturing at 94 ºC for 0:30, followed by 30 cycles of denaturing at 94 ºC for 0:10, annealing at 45 ºC for 1:00, and extension at 72 ºC for 1:00.

Unincorporated primers were removed from all PCR products used for sequencing with exonuclease in shrimp phosphatase. Cleaned PCR products were prepared for sequencing via standard protocols for BigDye Terminator version 3.1

(Applied Biosystems). Sequencing reactions were conducted using both forward and reverse primers to increase precision and reduce ambiguity of resulting sequences. Sequences were generated at the University of Maryland CORE facility using ABI 3100 and ABI 3730 automated sequencers (Applied Biosystems). ABI traces were aligned and manually checked using Sequencher (Gene Codes Corporation).

Amplified fragment length polymorphisms (AFLPs) were generated using a slightly modified version of the procedure originally proposed by Vos (1995). Digests were conducted with EcoRI and PstI enzymes. For selective amplification, EcoRI primers were labeled with florescent 6-FAM (6-carboxyl flourescein) to enable automated analysis with an ABI 3730 (Applied Biosystems). Of those screened, three EcoRI/PstI primer pairs (AG.CA, AG.GA, AG.CC) were used for selective amplifications. The screening process entailed a 2X comparison of 16 randomly chosen samples (N= 32 for each primer pair) starting from the construct step. Amplified products showed a high degree of concordance with products obtained during screening, indicating a high degree of repeatability between runs (<95%). A 1000bp internal size standard (ROX 1000, Genescan) was run with samples, and data were collected and analyzed using GeneMapper software (Applied Biosystems). To reduce subjectivity in scoring, a locus was scored if peak height was at least 100 reflectance units (rfu). After visual inspection, ambiguous peaks (50-100 rfu) were scored '?', indicating missing data.

Phylogenetic inferences. Parsimony and likelihood analyses were performed on mtDNA sequence data using PAUP* version 4.0 (Swofford 1999). Modeltest version

3.7 (Posada and Crandall 1998) was used to determine the most appropriate model of DNA substitution for the data set, and this model was used for likelihood analyses. Bootstrapping was used to determine the robustness and support of topologies obtained for both parsimony and likelihood trees. A statistical parsimony network was constructed using the program TCS (Clement et al. 2000). Nested clade analysis (NCA) (Templeton and Sing 1993, Templeton et al. 1995, Templeton 1998) was performed using GeoDis software (Posada et al. 2000), and this analysis was used to test the null hypothesis of no geographical association of mtDNA haplotypes. The inference key provided with GeoDis version 2.5 was used to analyze and interpret results. Interpretations were made with caution and considered within the context of all data collected during this project to account for concerns raised about the validity of inferences provided by Templeton's key (Knowles and Maddison 2002).

Genetic distances are frequently used as measures of genetic similarity within populations or between two or more populations with the expectation that smaller average genetic distances indicate greater similarity and gene flow either within or among populations (Muluvi et al. 1999, Parsons and Shaw 2001, Elderkin et al. 2004). In contrast, large average genetic distances may be interpreted as providing evidence for isolation or divergence within and among populations (Parsons and Shaw 2001). Reduced haplotype diversity may also provide evidence for isolation or divergence among populations (Barton and Slatkin 1986).

AFLP data were used to construct parsimony and distance based neighbor-joining trees. Uncorrected pairwise and Nei-Li distances (Nei and Li 1979), based on band sharing, were calculated and used to infer phylogeny. Bootstrapping was used to

determine the robustness and support of topologies obtained by parsimony and distance analysis. Analysis of molecular variance (AMOVA) was used to compare population genetic variation within and between populations (Muluvi et al. 1999, Parsons and Shaw 2001, Elderkin et al. 2004). Mantel tests (Mantel 1967) were conducted using Arlequin version 3.1 (Excoffier et al. 2005) to determine whether the average number of within-population pairwise genetic differences were correlated with geographic distances between collection locations. Structure version 2.1, a software package that uses Bayesian algorithms to determine the likelihood of multiple hypotheses (Pritchard et al. 2000), was used to elucidate the putative population structure throughout the native range.

Results

Phylogenetic inference- mtDNA. Seventy sequences from COI and COII (table 1.1) were combined into a 959bp alignment with 28 parsimony informative characters, 10 variable uninformative characters, and 921 constant characters. Twenty seven unique haplotypes were identified. Modeltest identified TIM+I as the best model for these data and likelihood analyses were conducted using the following parameters: frequency A = 0.3938, frequency C = 0.1444, frequency G = 0.0701, frequency T = 0.3871; substitution A-C = 1.0000, substitution A-G = 17.2917, substitution A-T = 0.0509, substitution C-G = 0.0509, substitution C-T = 7.7867, substitution G-T = 1.0000; proportion of invariable sites = 0.9160, equal rates for all sites. Uncorrected character differences varied from 0.001 to 0.018, and distances based on likelihood estimates varied from 0.001 to 0.032. Parsimony and likelihood analyses of mtDNA

data identified three divergent mtDNA lineages and several shallow groupings (fig. 1.2). Samples collected from the eastern face of the Rocky Mountains and the Pacific Northwest form one clade, and *A. cooleyi* from the western slope of the Rockies and the Great Basin form another. In addition, individuals from two southeastern Arizona populations are highly differentiated from the Western Rockies clade (fig. 1.2). Although unsupported by bootstrap analysis, unique groups of individuals within the Pacific Northwest (OR, WA), the northern Rockies (ID, MT), and the southern Rockies (CO, WY) were consistently identified by both parsimony and likelihood analyses (fig. 1.2).

Nested clade analysis. Templeton (1995) limits inference to those clades that show significant geographical association. Five nested groups showed significant geographical associations in nested contingency tests (table 1.2, figure 1.3). Clade distance (Dc) is a measure of the average distance that an individual bearing a haplotype lies from the geographical center of all individuals bearing haplotypes from the same clade, representing how geographically widespread individuals with a particular haplotype are (Templeton et al. 1995). Haplotypes from several lower level clades in the central Rockies (1-2, 2-1, 2-5) and from one higher level clades (3-3) in southern Arizona were significantly more widespread than expected with no effect of geography on haplotype distribution, while Dc of a clade found in southern Arizona (1-17) was significantly smaller (table 1.3). Nested distance (Dn) measures average distance that an individual bearing a haplotype lies from the geographical center of all individuals bearing haplotypes from the nested higher level clade that contains the

clade of interest, representing how far evolutionarily similar individuals are from one another (Templeton et al. 1995). The geographic distribution of nested haplotypes in lower level clades found in the eastern and central Rockies and southern Arizona (1-2, 1-17, 2-1, 2-5) and higher level clades found in southern Arizona (3-3) was greater than expected with no effect of geography on haplotype distribution (table 1.3). Contrasts between interior and tip (I-T) nodes represent comparisons of the geographic distribution of relatively young tip clades with older, related interior clades, and may be useful for inferring potential causes of geographical structuring of variation (Templeton et al. 1995). Dc for an higher level I-T clade in southeastern Arizona (3-3) was larger than expected and Dn for I-T clades from the eastern Rockies (2-1, 3-1) was significantly smaller than expected with no effect of geography on the distribution of mtDNA genetic variation (table 1.3). Although significant observations were made at the total cladogram level, inferences could not be made because tip and interior states could not be determined with confidence.

Phylogenetic inference- AFLP. Three primer combinations generated 445 polymorphic bands in 94 samples (table 1.1) that were used for data analyses. Fifteen sites were constant, 399 sites were variable and parsimony informative while 31 were uninformative. Total character differences varied from 39 to 142, mean character differences varied from 0.09 to 0.34, and Nei-Li distances varied from 0.05 to 0.18. Distance analysis of mean pairwise and Nei-Li distances produced trees with several well-supported clades, based on 1000 bootstrap replicates (fig. 1.4). Individuals from locations in southern Arizona comprise one clade, while individuals from central

Idaho, Washington, and Oregon comprise another clade (fig. 1.4). In contrast to mtDNA analysis, no differentiation was detected between samples collected in Colorado, Utah, Wyoming, and northern Arizona. These results were corroborated by analysis using the program Structure (Pritchard et al. 2000). The most probable hypothesis inferred by Bayesian analysis is that genetic data arose from four populations (table 1.4, fig. 1.5). Of these four populations, one is comprised of individuals from the Pacific Northwest; one is comprised of individuals from southern Arizona; and the remaining two populations are highly mixed and poorly differentiated (fig. 1.5). A Mantel test that compared corrected pair-wise genetic distances with geographic distances for samples collected from Boulder (CO), Colorado Springs (CO), Ft. Collins (CO), Denver (CO), Durango (CO), Steamboat Springs (CO), Laramie (WY), Heber (UT), Logan (UT), Ogden (UT), Driggs (ID), Pocatello (ID), Missoula (MT), Coeur d'Alene (ID), Puyallup (WA), and Corvallis (OR) confirmed isolation by distance (IBD) of populations from the Pacific Northwest from those is the Rocky Mountains (correlation coefficient (r) = 0.7498; $p < 0.0001$).

Discussion

Geographic structure of genetic variation. Population genetic structure as inferred by sequence variation in mtDNA is useful for examining questions dealing with population structuring events, such as glaciations, that have occurred within the last several million years (Hewitt 2001, 2004). Remington (1968) proposed that, in addition to other locations, the Rocky Mountains may be a “suture zone” where

effects of glacial events may be pronounced. This is largely due to the geology of the area, where large mountains may have served as refuges during periods of glacial retreat. The geographic distribution of *A. cooleyi* is relatively contiguous, and the insect is found throughout the Rocky Mountains and into the Cascade Range. However, my data show clear patterns of historically reduced gene flow between locations where *A. cooleyi* occurs. Nested clade analysis identified several portions of cladogram where the distribution of genetic variation was either significantly larger or smaller (table 1.2) than would be expected if the distribution of haplotypes were unaffected by geographic distance separating locations where samples were collected. Several possible conclusions can be drawn from my data using the inference key provided in GeoDis version 2.5 (2005). Well-supported branches observed in the mtDNA tree are likely due to differentiation that occurred as a result of allopatric fragmentation. The high degree of differentiation observed in samples from locations in southeastern Arizona is likely the result of past fragmentation and suggests that this group has been isolated for a relatively long period of time, and my AFLP data corroborate this conclusion. The pattern of genetic variation observed in Utah and Arizona is consistent with population structure shaped by gradual range expansion and population fragmentation that occurred during warming periods after glacial maxima. I also infer that genetic variation in the central and northern Rockies and the Pacific Northwest is categorized by restricted gene flow and isolation by distance; however, some long-distance dispersal likely occurs. Again, my AFLP data support these conclusions by showing phylogenetic concordance and by identifying isolation by distance of samples from the Pacific Northwest.

My data suggest the existence of at least three well-differentiated lineages within *A. coolyei*. Given the current distribution of *A. coolyei*, it is likely that insects currently found in the western Rockies (Utah, Southern Idaho, and Arizona) and southern Arizona were isolated after glacial periods (Templeton et al. 1995, Templeton 1998, Avise 2000, Hewitt 2001, 2004). These lineages have persisted and, in the case of the western Rockies populations, spread from refuge areas. MtDNA lineages found in samples collected in the eastern Rockies are comparably young, as indicated by the relatively small number of mtDNA haplotypes found there (Templeton et al. 1995, Templeton 1998), and are most closely related to samples found in the Pacific Northwest. High levels of mtDNA differentiation between samples that are relatively close in proximity, such as those found throughout Colorado and those found in neighboring Utah (fig. 1.2), reflects limited female dispersal (Hartl and Clark 1997). *Adelges cooleyi* lineages were separated from one another for a period of time sufficient for significant divergence before coming back into close proximity with one another. Similar to observations made by Ungerer et al. (1998) in sunflowers and Bensch et al. (2002) in warblers (Ungerer et al. 1998, Bensch et al. 2002), data from the nuclear genome provide evidence that divergent *A. cooleyi* lineages hybridize in the central Rockies. However, contemporary gene flow between locations in the eastern and western Rockies is highly limited. This may be due to combined effects of differential mating associated with the presence of different primary hosts in the two regions (*P. sitchensis* in the Pacific Northwest and *P. pungens* and *P. engelmannii* in the Rockies) (fig. 1.2) and isolation by distance.

Secondary contact and hybridization. I did not find complete congruence between mtDNA and AFLP phylogenies. Lack of AFLP structure between well-supported mtDNA lineages can occur due to incomplete nuclear lineage sorting or intraspecific hybridization. Given the depth and support of branches in my mtDNA phylogeny and the prevalence of glacial activity in the insect's native range, it is unlikely that my AFLP data represent incompletely sorted nuclear lineages. A more likely interpretation of my data is that well-differentiated mtDNA lineages of *A. cooleyi* share nuclear material in the central Rockies due to intraspecific hybridization.

Hybridization is the most rapid and significant source of shared genetic variation in large populations (Lewontin and Birch 1966). In addition to creating novel genotypes, hybridization between differentiated populations may lead to the development of novel phenotypes (Ungerer et al. 1998, Albertson et al. 1999). These novel phenotypes may be innocuous and inconsequential or they may affect important ecological traits related to growth and habitat use (Johnston et al. 2004), jaw or beak shape (Albertson et al. 1999, Grant and Grant 2002, respectively), or colonization of new environments (Seehausen 2004). While no direct consequences of hybridization were observed in this study, the creation of novel genotypes and phenotypes is especially relevant in this species because CSGA is a pest in the eastern United States (Johnson and Lyon 1988). Adelgids found in introduced populations currently possess a very limited amount of genetic variation (R. Ahern, personal observation), and subsequent introductions of hybrid individuals may exacerbate control efforts by providing large amounts genetic and phenotypic variation to currently bottlenecked populations (Ellstrand and Schierenbeck 2000, Seehausen 2004).

Combined mtDNA and AFLP data sets. Speciation histories based on mtDNA can be misleading (Shaw 2002), and it is important to use multiple neutral genetic markers and loci to infer population genetic structure because individual gene genealogies may offer incomplete or inaccurate estimates of overall population genetic variation (Zhang and Hewitt 2003). DNA from uniparentally-inherited organelles such as mitochondria is thought to experience little or no recombination, making assessment of interbreeding difficult or impossible in organisms with sex-limited dispersal (Hartl and Clark 1997). Recent studies have challenged long-held beliefs of no recombination and neutrality of mtDNA. Relatively high levels of recombination (Saville et al. 1998, Rokas et al. 2003) and combination with portions of the nuclear genome (Bensasson et al. 2001) have been reported, and evidence exists for fairly stringent selective regimes within the mitochondrial genome (Rand 2001). While these findings do not discount the significance of previous work done on mtDNA variation, they encourage caution in the interpretation of results and underscore the importance of analyzing data from multiple genomes.

In contrast to organelle DNA, amplified fragment length polymorphisms (AFLPs) are a random sample of the nuclear genome (Vos et al. 1995, Mendelson and Shaw 2005). Analysis of population genetic variation as inferred by shared AFLP phenotypes may provide insight into relatively contemporary events (< 10,000 years) that affect population structure (Hewitt 2004), and may be useful for identifying intraspecific hybrid zones (Ungerer et al. 1998). Consideration of mtDNA and AFLP data together provides more robust estimates of phylogeography and the degree to

which populations are isolated from one another, and also provides insights into patterns of contemporary gene flow, including intraspecific hybridization. Using multiple markers, I was able to corroborate isolation of samples from southern Arizona. Using a genealogical species concept (Baum and Shaw 1995) and current barcoding thresholds (Kerr et al. 2007), these samples are evolutionarily distinct from the rest of the species and should be considered as such. Also, I was able to detect limited nuclear gene flow in the Pacific Northwest and infer hybridization in the central Rockies because I used multiple markers. While studies that rely on single genes obviously represent important research, I agree with Hare (2001) and argue that a more complete understanding of patterns affecting the geographic distribution of genetic variation is arrived at through application of multiple genetic markers.

Host variation. In addition to the effects of glaciations on the distribution of *A. cooleyi*, current population structure is likely mediated, at least in part, by the distribution of host plants (Futuyma and McCafferty 1990, Via 1991, Funk et al. 1995, Ruiz-Montoya et al. 2003). Glacial activity affects the distribution of plant species (Hewitt 1996, 2004) and it is possible that effects of large-scale climate changes might be especially profound for long-lived woody plants, such as conifers found in the Pacific Northwest and Rocky Mountains. The distribution of specialized herbivorous insects is directly related to the distribution of their host-plants (Futuyma and Peterson 1985, Futuyma and McCafferty 1990, Funk et al. 1995), and it is important to consider the effects of climate change on the historic distribution of host

plants when trying to understand current patterns of genetic variation observed in specialist insect herbivores.

The distribution of organisms affected by glacial activity is often characterized by movement into lower latitudes during cold periods and retraction to higher elevations and higher latitudes when climate warms (Hewitt 1996, 2004). Colorado spruce is currently found in the central Rockies, while the distribution of Engelmann spruce is far larger and encompasses an area that stretches from Arizona into Canada (Little 1971) (figs. 1.2, 1.4). Sitka spruce is very common in the Pacific Northwest (Little 1971) (figs. 1.2, 1.4). Recent molecular work confirms significant divergence between spruce species (*P. engelmannii*, *P. pungens*, and *P. sitchensis*) found in the Pacific Northwest and Rocky Mountains (Ran et al. 2006). Coastal (var. *menziesii*) and interior varieties (var. *glauca*) of Douglas fir have been confirmed by phenotypic (Rehfeldt 1977), allozyme (Li and Adams 1989), and RAPD (Aagaard et al. 1998) variation. The coastal variety is common to relatively low-lying areas in California, Oregon, Washington, and British Columbia (Little 1971) (figs. 1.2, 1.4), while the interior variety has a far larger distribution and is found throughout the Cascade and Rocky Mountain ranges from Mexico into northern Canada (Little 1971) (figs. 1.2, 1.4). The two varieties are thought to have been distinct since the Miocene (13 million years ago)(Critchfield 1984, Aagaard et al. 1998), and data collected from seed middens identify the Pacific Northwest as the pre-glacial range of *Pseudotsuga menziesii* in the United States. Evidence collected from woodrat middens also shows that *P. pungens* and *P. menziesii* grew on the south slope of the Dutch John Mountains in northeastern Utah during the last glacial period (10,000 – 15,000 YBP)

(Jackson et al. 2005). Data collected by Jackson et al. (2005) suggests that a broad band of confiers, including *P. pungens* and *P. menziesii*, extended from the Grand Canyon to southern Wyoming during the last glacial period, indicating an extensive shift in geographic distribution of both species in response to glacial activity.

The process of isolation of *A. cooleyi* lineages followed by range expansion and secondary contact was likely mediated by changes in the distribution of spruce species and Douglas fir in the central and southern Rockies in response to glacial activity (Futuyma and Peterson 1985, Futuyma and McCafferty 1990, Hewitt 1996, 2004). Multiple spruce species and multiple races of Douglas fir are present in the central Rockies (figs. 1.2, 1.4). Populations of *A. cooleyi* found in the central Rockies are comprised of individuals from two divergent clades, and this pattern may be explained by range expansion from southern refuges into northern latitudes by one or both hosts after the last glacial period. Significant mitochondrial and nuclear differentiation of southeastern Arizona populations indicates long term isolation from both of the other two well-supported lineages of *A. cooleyi* (Hartl and Clark 1997). Southern interior Douglas fir and Engelmann spruce thrive in the White Mountains where samples were collected (Lipscomb 1993) (figs. 1.2, 1.4), and these host species live at relatively high elevation in an environment that is discontinuous or nonexistent at lower elevations (Little 1971, Lipscomb 1993). This habitat discontinuity likely acts as an isolating factor for *A. cooleyi* because it impedes the distribution of host plants.

In addition to the deep, well-supported clades found in the Rockies, several haplotype groups occur in the Pacific Northwest that may represent locally adapted

genotypes that have not spread throughout the range of the insect (figs. 1.2, 1.4). This may be due to limited dispersal or actual adaptation to either the coastal variety of Douglas fir; a spruce host found only in the Pacific Northwest (*P. sitchensis*); or both of these host-plants (figs. 1.2, 1.4). The structure of population genetic variation in this area is characterized by shallow branches and clade infidelity (see Corvallis sample grouped with northern Rockies), indicating relatively ineffective isolating mechanisms, incomplete lineage sorting, or both (Coyne and Orr 2004). I did not directly test for differentiation among trees were insects were collected but previous studies have shown genetic differentiation between both primary and secondary hosts found in the Pacific Northwest and the Colorado Rockies (Li and Adams 1989, Aagaard et al. 1998, Ran et al. 2006). *Adelges cooleyi* spend several months developing inside modified plant parts on *Picea* hosts and complete the majority of one life stage within the tumor-like structures that they create (Annand 1928). In general, galling insects show a high degree of host specificity (Price 1997), and it is likely that inter- and intraspecific variation in spruce hosts as well as variation within and among races of Douglas fir affect population genetic structure of *A. cooleyi.*

Combined phylogeographic studies. Intraspecific phylogeography as originally described by Avise provides a framework for analyzing geographical distribution of genealogical lineages (Avise et al. 1987). The relationship between specialist insects and their plant host is often tightly coupled, and small changes in one may result in changes in the other in a process commonly referred to as coevolution (Ehrlich and Raven 1964, Thompson 1994). Researchers have considered phylogenetic congruence

of hosts and pathogens (Roy 2001), hosts and bacteria (Moran 2001), and hosts and herbivores (Brown et al. 1996). However, few studies have used multiple molecular markers to separate historic and contemporary gene flow within the context of introduced insects (Tsutsui et al. 2001, Suarez et al. 2004). Further, few previous studies on specialist herbivores have interpreted the geographic distribution of genetic variation within the context of geographic distribution of host variation. Here I present evidence for the importance of understanding interspecific interactions that affect the geographic distribution of genetic variation in a specialist insect.

Table 1.1. Collection information for samples used in molecular analyses. N1 and N2 columns show the number of samples from each location used to construct mitochondrial and AFLP phylogenies, respectively. All collections were made by R. Ahern except those from Arizona (A. Lynch), Vernon BC (B. Lord), and Vancouver BC (M. Hall).

Location	Date	Host species	Map Code	N1	N 2
Boulder, CO	5-Jul-05	*P. pungens*	A	3	7
Coeur d'Alene, ID	2-Jul-04	*P. pungens*	B	4	4
Colorado Springs, CO	28-Jun-05	*P. pungens*	C	4	6
Corvallis, OR	7-Jul-04	*P. pungens*	D	4	5
Denver, CO	27-Jun-05	*P. pungens*	E	3	7
Driggs, ID	1-Jul-04	*P. pungens*	F	4	2
Durango, CO	29-Jun-05	*P. pungens*	G	3	3
Ft. Collins, CO	28-Jun-05	*P. pungens*	H	4	8
Heber, UT	2-Jul-05	*P. pungens*	I	4	4
Laramie, WY	3-Jul-05	*P. pungens*	J	4	3
Logan, UT	1-Jul-05	*P. pungens*	K	4	6
Ogden, UT	1-Jul-05	*P. pungens*	L	4	5
Missoula, MT	2-Jul-04	*P. pungens*	M	3	4
Pocatello, ID	1-Jul-05	*P. pungens*	N	3	8
Puyallup, WA	6-Jul-04	*P. sitchensis*	O	4	4
Steamboat Springs, CO	30-Jun-04	*P. pungens*	P	1	3
Vancouver, BC	21-Jul-03	*P. engelmannii*	Q	2	2
Vernon, BC	22-Jul-03	*Picea sp.*	U	0	2
White Mt. Reservoir, AZ	20-Jul-04	*P. engelmannii*	R	4	3
Paradise Butte, AZ	20-Jul-04	*Picea sp.*	S	4	4
Piñaleno Mts., AZ	20-Jul-04	*P. engelmannii*	T	4	4
			Total	70	94

Table 1.2. Clades showing significant geographic associations as determined by nested clade analysis. Asterisk represents clade level for which inferences could not be made because tip and interior states could not be determined.

Clade	chi-square statistic	Probability
2-1	50.51	<0.001
2-7	7.00	0.024
3-1	23.37	0.04
3-2	62.00	<0.001
Entire Cladogram*	121.18	<0.001

Table 1.3. Results from permutation tests on clade (D_C) and nesting (D_N) distances as determined by nested clade analysis. Asterisks denote level of significance (* $p < 0.05$, ** $p < 0.001$) based on 1000 re-samples.

	Postion	*Dc*	*P*	*Dn*	*P*
Clade 2-1					
Clade 1-1	tip	172.76	0.03*	830.52	1.000
Clade 1-2	interior	114.86	<0.001**	348.45	<0.001**
Clade 1-6	tip	281.05	0.37	353.15	0.18
I-T clades	─	-87.43	0.22	-351.88	0.001**
Clade 2-7					
Clade 1-16	interior	0.00	0.148	1.37	1.000
Clade 1-17	tip	0.00	0.024**	1.03	0.024*
I-T clades	─	0.00	0.654	0.34	1.000
Clade 3-1					
Clade 2-1	interior	479.03	0.001**	517.04	0.001**
Clade 2-2	tip	141.38	0.16	1105.4	1.000
I-T clades	─	337.65	0.78	-588.36	0.001**
Clade 3-2					
Clade 2-3	tip	594.7	0.69	978.56	1.000
Clade 2-4	interior	0.00	0.075	704.19	0.73
Clade 2-5	tip	286.88	<0.001**	315.59	<0.001**
I-T clades	─	-371.80	0.32	205.71	0.74
Entire Cladogram					
Clade 3-1	tip	565.62	0.27	603.04	0.50
Clade 3-2	tip	537.72	0.15	515.69	0.049*
Clade 3-3	tip	1.21	<0.001**	966.29	0.99
No I-T clades in this group					

Table 1.4. Results of Structure analysis. Log likelihoods for hypotheses that AFLP haplotypes originate from different numbers of parent populations (K) as inferred by Bayesian analysis. K=4 has the highest posterior probability.

Population number (K)	Log Likelihood
2	-11323.6
3	-10996.7
4	-10717.2
5	-10897.5
6	-14043.1
7	-13106.3
8	-13820.4

Figure 1.1. Cooley spruce gall adelgid lifecycle. Seasons and lifecycle progress clockwise. Notice both migratory and sessile insects are present during the spring on Douglas fir.

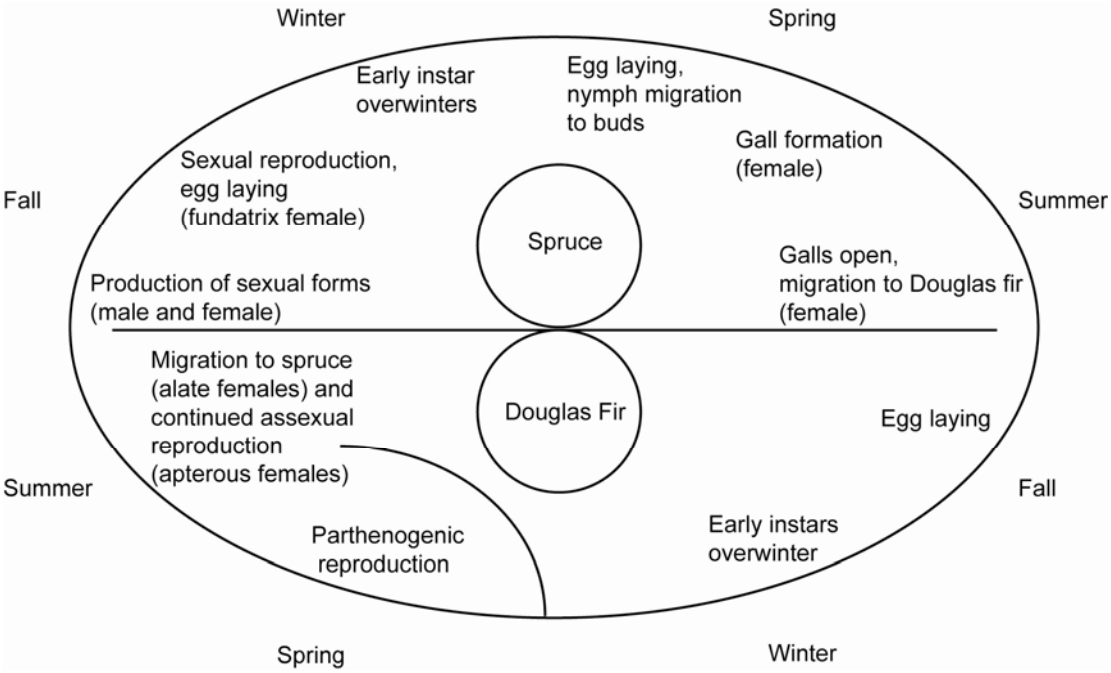


Figure 1.2. Midpoint rooted neighbor-joining tree based on maximum likelihood distances calculated from mtDNA data. Numbers above nodes indicate bootstrap support after 1000 replications. Color of sample names corresponds to predominant spruce species and Douglas fir races found where samples were collected; samples with names in black originated from the central Rockies where two spruce species and two races of Douglas fir may be found.

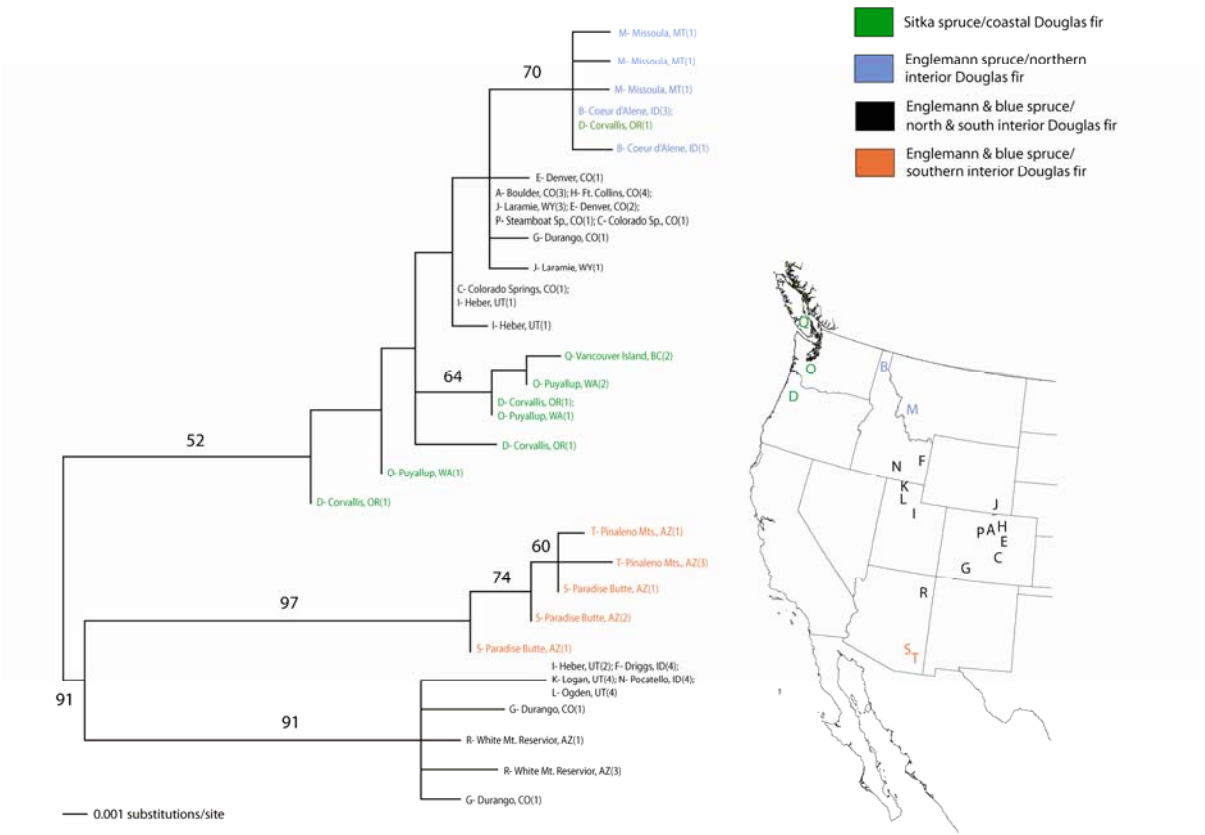


Figure 1.3. Nested clade haplotype network for 27 haplotypes observed from 20 localities. Each line represents a single mutational change and open circles represent inferred haplotypes that were not observed in this study. Haplotype names and distribution are listed in table A. Haplotype distribution was as follows: bo2 – Boulder (3), Ft. Collins (4), Laramie (3), Denver (2), Steamboat Springs (1), Colorado Springs (1); cd1 – Coeur d'Alene (3), Corvallis (1); cd7 – Coeur d'Alene (1); cv15 – Corvalis (1), Puyallup (1); cv110 – Corvallis (1); cv112 – Corvallis (1); la4 – Laramie (1); mt12 – Missoula (1); mt27 – Missoula (1); mt22 – Missoula (1); sm1 – Denver (1); py15 – Puyallup (2); py11 – Puyallup (1); du1 – Durango (1); du3 – Durango (1), Colorado Springs (2); du4 – Durango (1); cs5 – Colorado Springs (1), Heber (1); he2 – Heber (1); he3 – Heber (2), Driggs (4), Logan (4), Pocatello (3), Ogden (4); vi2 – Vancouver Island (2); az34 – White Mountain reservoir (3); az32 – White Mountain reservoir (1); az42 – Paradise Butte (1); az45 – Paradise Butte (2); az42 – Paradise Butte (1); az52 – Pinaleno Mountains (3); az59 – Pinaleno Mountains

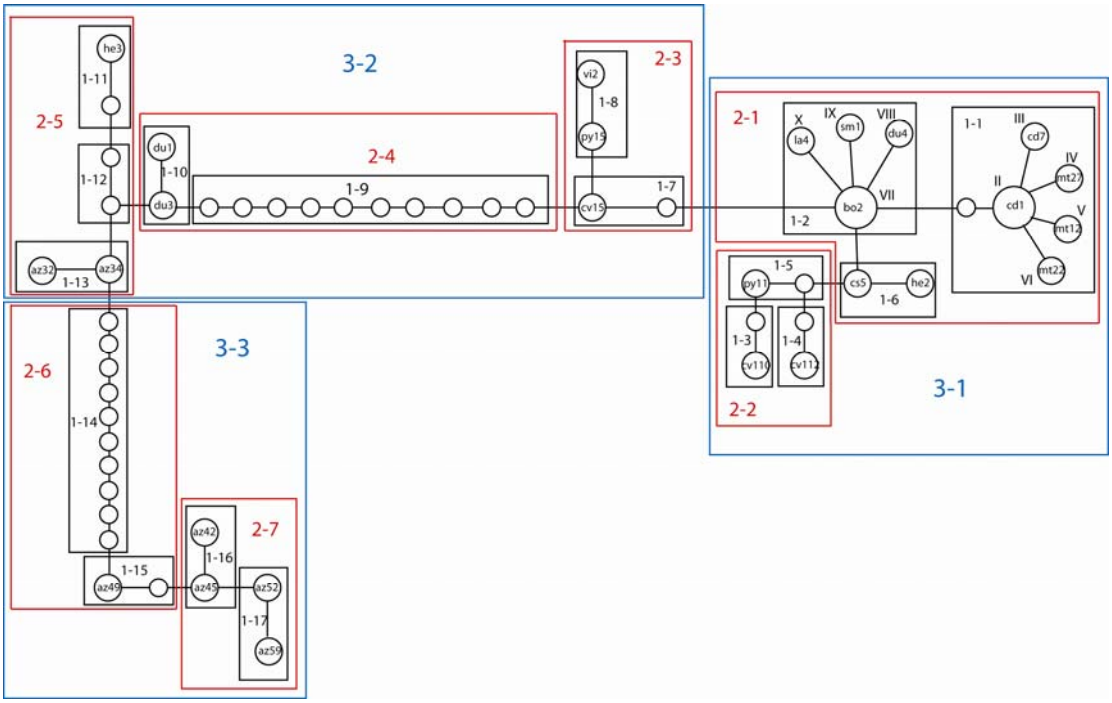
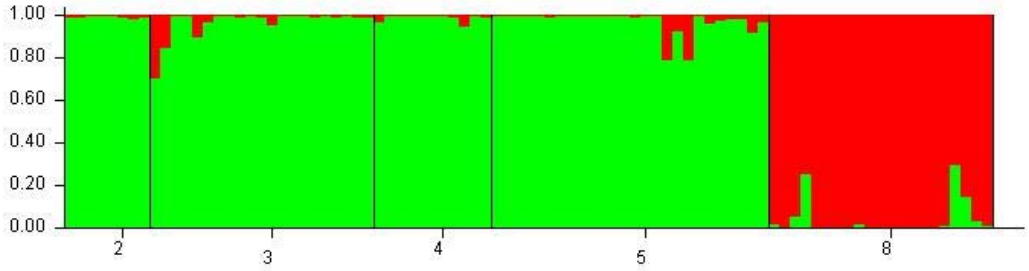


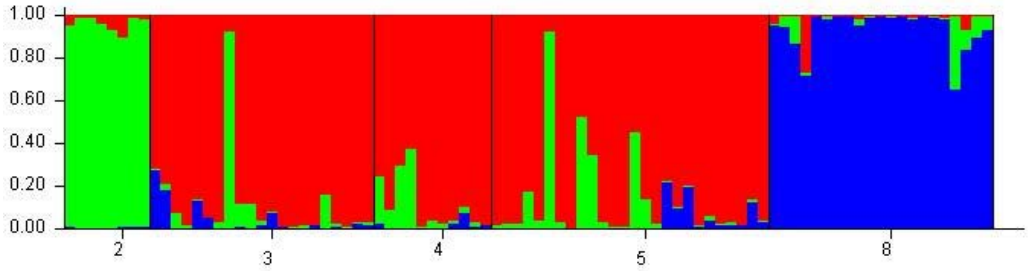
Figure 1.4. Midpoint-rooted neighbor-joining tree based on Nei-Li distances generated by AFLPs. Numbers above nodes indicate bootstrap support (distance / parsimony) after 1000 replications. Color of sample names corresponds to predominant spruce species and Douglas fir races found where samples were collected; samples with names in black originated from the central Rockies where two spruce species and two races of Douglas fir may be found.

Figure 1.5. Results of Structure analysis. Panels A-D represent K=2 through k=5, respectively. Colors are arbitrarily assigned and represent putative populations. Values on the x-axis represent locations from which samples originated. Scale on the y-axis represents the proportion of an individual's genetic variation assigned to each population. Populations originating from the Pacific Northwest (8) and Southern Arizona (2) and are well-supported, while those originating from the western, southern, and eastern Rocky Mountains (3, 4, and 5) are not clearly differentiated from one another.

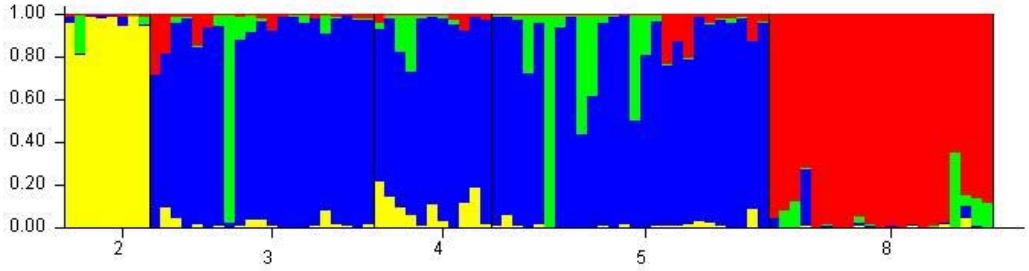
A. K=2



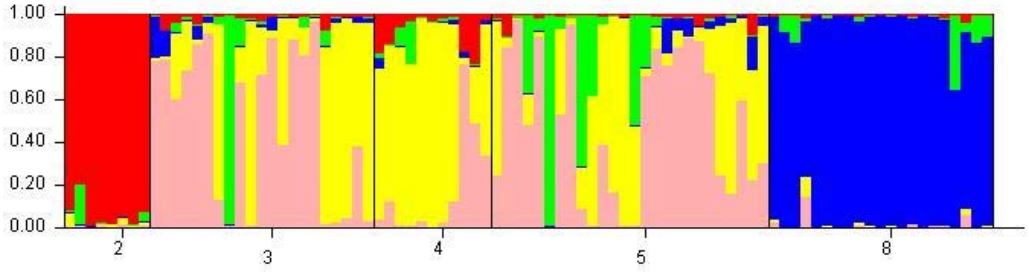
B. K=3



C. K=4



D. K=5



Chapter 2: Founder effects and phenotypic variation in Cooley spruce gall adelgid, *Adelges cooleyi* (Hemiptera: Adelgidae), an insect pest introduced to the eastern United States

Abstract

Introduced organisms experience founder effects and may undergo genetic bottlenecks that result in significant reductions in genetic variation. Genetic bottlenecks due to founder effects may constrain the evolution of phenotypic traits that facilitate success in novel habitats. I examined the effect of introduction into novel environments on genetic diversity using an insect pest, *Adelges cooleyi*, which was introduced into the eastern United States during the mid 19th century. I measured variation in mitochondrial and nuclear genomes in native and introduced samples to determine the effect of introduction on genetic variation experienced by this insect. I also measured an important ecological phenotype, variation in host preference, in both native and introduced samples to determine whether genetic variation is correlated with phenotypic variation. *Adelges cooleyi* in the eastern US has low genetic and phenotypic variation but appears to be sufficiently adapted to persist in this novel environment. My results call into question the utility of neutral genetic variation to assess the probability of persistence in new environments by introduced species.

Introduction

The introduction of organisms into non-native habitats is an important type of human-mediated environmental change (Vitousek et al. 1997). Specifically, introduction of organisms into areas in which they are not native constitutes a threat to biodiversity and stability of a ecosystems (Vitousek et al. 1997). Introduced species account for an estimated $137 billion in damage and losses in the United States each year (Pimentel et al. 2000), and insect damage to forest resources in native, cultivated, and urban settings accounts for a significant proportion of these losses (Pimentel et al. 2000). Introduced phytophagous insects persist in novel habitats due to release from natural enemies, insufficient defenses in host plants, and superior competitive ability of introduced organisms (reviewed in Price 1997). Despite these potential advantages, adaptation to conditions in novel environments may be required, or may facilitate success of introductions (Hedrick et al. 1976, Mackauer 1976). In particular, selective forces present in introduced environments (natural enemies, plant defenses, and competition) may differ from those in native environments and may favor previously unsuccessful genotypes and phenotypes (Mack et al. 2000, Mooney and Cleland 2001, Hastings et al. 2005).

Variation necessary for successful establishment can occur in introduced organisms in several ways. Baker (1965) described general-purpose genotypes that facilitate establishment of introduced populations and enable a high degree of phenotypic plasticity, which may be adaptive and lead to success in novel

environments (Agrawal 2001). Depending on the mode of introduction, genetic variation in the colonist population(s) can be enhanced or substantially reduced. For example, colonization by individuals from previously isolated populations may lead to enhanced genetic diversity and hybridization (Kolbe et al. 2004, Hastings et al. 2005), resulting in creation genotypes and phenotypes that are unlike those found in native populations. In addition, some introduced organisms may hybridize with closely related species, resulting in the creation of novel genotypes (Rhymer and Simberloff 1996, Mooney and Cleland 2001, Hastings et al. 2005). Novel genotypes can be important for success and spread of introduced organisms (Ellstrand and Schierenbeck 2000), and genetic variation created by hybridization may play important roles in successful establishment and persistence after introduction (Ellstrand and Schierenbeck 2000). Alternatively, colonization by few individuals from a common population represents a founder effect and may lead to a genetic bottleneck that results in very low genetic diversity in introduced populations (Hartl and Clark 1997). Introduced organisms with low genetic variation can persist, or even thrive, if genotypes that are introduced are well-selected, general-purpose, or both (Baker 1965, Mackauer 1976, Hawthorne 1997, Tsutsui et al. 2000).

Knowledge about the amount of genetic variation present in introduced pests may be useful to pest managers, as pests with reduced genetic variation may be more easily controlled than those with more genetic diversity (Mackauer 1976, Burdon and Thrall 2004). For example, pathogenic biological control agents are less effective against populations with multiple different resistant phenotypes compared with populations lacking such diversity (Burdon and Thrall 2004). Genetic variation

affects host use (Futuyma and Peterson 1985) and, subsequently, may affect range expansion following colonization. Selection of oviposition sites by females is a critical component of host use (Thompson 1988) and is controlled, in part, by genetic traits (Hawthorne and Via 2001).

Some introductions result in decreased genetic variation (Tsutsui et al. 2000), while others results in increased genetic variation (Kolbe et al. 2004). Likewise, a broad suite of phenotypes are hypothesized to affect introduction success (Price 1997, Kolar and Lodge 2001). Genotypic and phenotypic comparisons of phytophagous insects from native and introduced ranges may provide insight into processes that facilitate persistence after introduction more completely than consideration of either component separately. An understanding of how genotypic and phenotypic variation affects introduction success will improve control strategies because pest managers can target control efforts at susceptible aspects of a pest's biology such host use (Thompson 1988, Via 1999), dispersal (Paterson et al. 1995, Ingram and Gordon 2003), natural enemy resistance (Hufbauer and Via 1999), and insecticide resistance (Roush and McKenzie 1987).

Study system. Gillette described *Adelges cooleyi* in 1907 (Gillette 1907, Annand 1928). The insect's primary hosts include various spruce species (*Picea pungens*, *P. engelmannii*, and *P. sitchensis*), and its secondary host is Douglas fir, *Pseudotsuga menziesii* (Gillette 1907, Annand 1928). *Adelges cooleyi* forms large galls at the end of terminal branches on spruce hosts and yellows and distorts needles on Douglas fir (Annand 1928, Johnson and Lyon 1988). It requires both tree species to complete a

complex lifecycle over two years (Annand 1928). Both sexual and asexual reproduction occurs on primary hosts, while all reproduction on the secondary host is asexual. The native range of CSGA in the United States resembles an inverted 'V', extending along the Rocky Mountains into Canada and back down the West Coast into northern California. This distribution is roughly congruent with the distribution of its primary and secondary hosts (Little 1971).

Ornamental plant materials were exported to Europe and the eastern US in the early to mid-nineteenth century, leading to removal and distribution of many plants and their associated insect faunas from native ranges to new locations (Reveal 1992). Members of the Lewis and Clark expedition (1803-1806) were likely the first people of European descent to come into contact with spruce and fir species indigenous to the United States (Reveal 1992). Samples collected by the expedition and shipped back to the east thrived in eastern American gardens by 1820 (Reveal 1992). Despite being described in 1907, *A. cooleyi* was first collected and misidentified by Cooley in Massachusetts around 1897 (Gillette 1907), suggesting that the insect was introduced into the eastern United States through an unknown number of introductions during the late 19th century.

In this study I collected molecular data from mitochondrial and nuclear genomes, and phenotypic data related to host preference. I compared the number of mitochondrial (mtDNA) haplotypes found in the native range with that found in the introduced range in an attempt to quantify a change in mtDNA variation due to the introduction. I analyzed amplified fragment length polymorphisms (AFLPs) to assess relative levels of nuclear genetic variation in native and introduced samples. I

expected to find a relatively small number of mtDNA haplotypes, consistent with few introductions, introductions from the same location within the native range, or both. Introductions often result in founder effects that may lead to significant genetic bottlenecks, and I expected to find less AFLP variation in the introduced range than in the native range. Analysis of data from non-recombining mitochondrial regions and recombining nuclear regions were considered together to determine whether both sets of markers supported the same conclusions regarding the changes in genetic variation experienced by *A. cooleyi* as a result of its introduction.

In addition to neutral nuclear variation, I measured an ecologically important phenotype, host preference. Host preference, like all phenotypes, is mediated by genetic variation (Futuyma and Peterson 1985). Genetic bottlenecks may affect variation in these traits (Hartl and Clark 1997), and I predicted lower phenotypic variation in samples from the introduced range compared with samples from the native range. Results from genetic analyses were used to direct interpretation of results from my phenotypic assay. Factors, such as the ability to utilize genetically diverse host plants, which likely affect persistence of introduced *A. cooleyi*, were identified.

Methods

Collection of specimens. CSGA was collected on *Pseudotsuga menziesii* and *Picea pungens* in its native and introduced ranges (table 2.1). Insects collected on *Pseudotsuga menziesii* in early spring (April-May) were removed directly from foliage and stored in 100% EtOH until DNA was extracted, while those collected on

P. pungens were removed from galls in mid-summer (June-July) and stored in 100% EtOH until DNA was extracted. With few exceptions, molecular analyses were conducted on approximately 8 insects from separate galls and 8 insects from different needles from different trees from 7 native and 7 introduced populations (table 2.1). All sampled locations are separated by at least 25km.

Laboratory procedures. Whole genomic DNA was extracted using DNeasy Tissue Kits (Qiagen) following a slightly modified protocol B for insects. All samples for sequencing were processed at Canadian Centre for Barcoding, Biodiversity Institute of Ontario, University of Guelph. Extractions were done as described in Ivanova et al. (2006). Primers were modified for Lepidoptera: LepF (ATTCAACCAATCATAAAGATATTGG) and LepR (TAAACTTCTGGATGTCCAAAAAATCA) from Hajibabaei et al. (2006). Sequencing reactions were conducted with ABI BigDye v. 3.1 and fragments were visualized on an ABI 3730 (Eric Maw, personal communication). Sequences were aligned with BioEdit (Ibis Therapeutics).

AFLPs were generated using a slightly modified version of the procedure originally proposed by Vos (1995). Digests were conducted with EcoRI and PstI enzymes. For selective amplification, EcoRI primers were labeled with florescent 6-FAM (6-carboxyl flourescein) and analyzed with an ABI 3730 capillary electrophoresis device (Applied Biosystems). Three EcoRI/PstI primer pairs (AG.CA, AG.GA, AG.CC) were chosen for selective amplifications after screening. The screening process entailed a 2X comparison of 16 randomly chosen samples (N= 32

for each primer pair) starting from the construct step. Amplified products showed a high degree of concordance with products obtained during screening, indicating a high degree of repeatability between runs (<95%). A 1000bp internal size standard (ROX 1000, Bioventures) was run with samples, and data were collected and analyzed using GeneMapper software (Applied Biosystems). To reduce subjectivity in scoring, a locus was scored if peak height was at least 100 reflectance units (rfu). After visual inspection, ambiguous peaks, those between 50-100 rfu, were scored '?', indicating missing data.

Genetic data analysis. Analysis of mtDNA variation was conducted on 650bp of the COI gene. Relatedness of mtDNA haplotypes was inferred by distance-based clustering analyses using PAUP* version 4.0 (Swofford 1999). Collapse (version 2.1) software was used to identify identical mtDNA haplotypes. A t-test was used to compare the number of mtDNA haplotypes found in the native range with the number of haplotypes present in the introduced range. The average number of pairwise mtDNA differences within sample locations was calculated using Arliquin (version 3.1) (Excoffier et al. 1992, Excoffier et al. 2005). Mean number of pairwise differences within native and introduced sample locations was compared by a one-tailed t-test. A one-tailed t-test was performed because of an *a priori* assumption of decreased genetic variation in samples collected from the introduced range (Zar 1998). Analysis of molecular variance (AMOVA) was used to determine the distribution of genetic variation within and among sample locations in native and introduced ranges (Excoffier et al. 2005).

AFLP data were used to determine the average number of pairwise differences within native and introduced sample locations (Excoffier et al. 1992, Excoffier et al. 2005). A one-tailed t-test was used to compare the average number of pairwise differences within native locations with that of introduced locations after it was confirmed that data met assumptions of normality (Zar 1998). AMOVA was conducted to compare distribution of genetic variation within and among locations sampled in native and introduced ranges (Muluvi et al. 1999, Parsons and Shaw 2001, Elderkin et al. 2004). Measures of genetic distance, including Nei Li distances (Nei and Li 1979) distances, were computed and distance-based clustering methods in PAUP* version 4.0 (Swofford 1999) were used to infer relatedness of AFLP phenotypes. Structure version 2.1, a software package that uses Bayesian algorithms to determine the likelihood of multiple hypotheses (Pritchard et al. 2000), was used to determine whether introduced samples were significantly differentiated from native samples.

Phenotype assay. To measure variation in an ecologically important trait, I measured variation in preference of insects from native and introduced samples for two Douglas fir varieties. The two varieties were: coastal Douglas fir, *Pseudotsuga menziesii* var. *menziesii*; and interior Douglas fir, *Pseudotsuga menziesii* var. *glauca*. The assay consisted of choice tests in which one gall was placed inside a mesh cage with one 30 – 50cm tall (2-3 year-old) coastal Douglas fir host (Brooks Tree Farm, OR) and one comparably sized 2-year-old interior Douglas fir (Musser Forests, PA), representing 2 levels of host plant origin treatment. Coastal Douglas fir is native to

the Pacific Northwest and interior Douglas fir used for assays was grown from seed collected in the Lincoln National Forest in New Mexico. Thus, the two varieties represent plant genotypes from geographic extremes of the host range (Little 1971, Lipscomb 1993). Because both host plant varieties originate from locations outside of areas where insect samples were collected, treatments represent variation in samples' use of secondary hosts and may mimic the ability to adapt to genetically diverse host plants.

Collection of insects for the phenotype assay was confined to *Picea pungens* to eliminate the confounding effect of different host plant species on secondary host preference. Seven locations in the native range and four locations in the introduced range were sampled (n=11 [7 native, 4 introduced]). One gall per tree was sampled from 5-10 trees per location. Samples were collected from native spruce stands in the native range to increase the likelihood of accurately sampling the pool of genetic variation from which insects in the east originated. Samples in the introduced range were collected from locations where primary and secondary host plants were both present to reduce confounding effects associated with isolation from secondary hosts. Samples in the native range (N = 27) were collected from Colorado (CO), Utah (UT), Wyoming (WY), and Idaho (ID), and samples in the introduced range (N = 22) were collected from New York (NY), Michigan (MI), and Pennsylvania (PA) (table 2.2).

Galls were placed into arenas and adelgids were allowed to emerge and establish on the variety of their choice for a minimum of 2 weeks. The number of insects established on each host plant was counted. Dead individuals, representing those that failed to establish, were also counted but were left out of my analysis

because the biological significance of these individuals could not be determined. It was unclear whether they were averse to both hosts or could not establish for some other reason.

Phenotypic data analysis. A preference variable was calculated using the formula $(N_1-N_2) / (N_1+N_2)$ where N_1 and N_2 are the number of adelgids found on coastal and interior varieties, respectively (Cock 1978). Preference data were analyzed using 1-way ANOVA to determine the main effect of insect population origin (Zar 1998). Preference was standardized for analysis such that data from galls from which more adelgids established were weighted greater than those with low establishment. Standardization values were calculated by determining mean establishment across all samples, and data were multiplied by a standardizing value that represented higher or lower establishment than the overall mean. This accounted for potentially important differences in emergence while protecting against biases caused by very low levels of emergence. I hypothesized that variation in secondary host preference would be greater for native populations than for introduced populations due to expectations about populations that experience founder effects. A variance ratio test with a one-tailed distribution was used to test this hypothesis (Zar 1998).

Results

Genetic variation- mtDNA. Two hundred and ten (210) individuals were sampled (106 in the native range; 104 in the introduced range) (table 2.1). There were approximately half as many mtDNA haplotypes found in introduced samples (N = 2)

than in native samples (N = 4.2). However, these data overestimate the mitochondrial genetic variation found in the introduced range. Only 4 haplotypes were found in the introduced range compared with 16 found in the native range, and the average number of pairwise differences in mtDNA haplotypes within introduced locations (0.224) was significantly smaller (P = 0.002; df = 12) than within native locations (4.1) (table 2.3). Furthermore, 98% of individuals sampled in the introduced range had either haplotype 1 (N = 92) or haplotype 2 (N = 10) (table 2.4). In contrast, only 25% of individuals in the native range had haplotype 1. The most common haplotype in the native range, haplotype 8, was found in 37% of individuals and was absent from samples from the introduced range (table 2.4). Two unique haplotypes were found in the introduced range (table 2.4), possibly indicating the presence of insects introduced from a portion of the native range, likely the Pacific Northwest, which was not sampled for this project. The majority (98%) of introduced individuals had haplotypes also found in individuals from locations in the eastern Rocky Mountains.

Results from AMOVA support my observation of reduced genetic variation in the introduced range (table 2.5), as evidenced by far smaller sums of squares and variance components in introduced samples compared to native samples. A large amount of mtDNA variation occurs both among and within sample locations in the native range. In contrast, almost all mtDNA variation observed in the introduced range occurs within locations. Further evidence of significant between-location genetic structure in the native range is provided by a large fixation index value (0.58825), compared with a value of 0.04461 in the introduced range.

Genetic variation- AFLP. Three primer combinations generated 445 polymorphic AFLP bands that were used to analyze nuclear DNA variation. Total character differences between individuals varied from 30 to 162, average number of pairwise differences within sample locations varied 44.2 to 86.0, average number of pairwise differences among sample locations varied from 52.5 to 89.2, and corrected average pairwise differences among sample locations varied from 0.8 to 14.9. The average number of pairwise differences within native sample locations was significantly greater than that of introduced sample locations ($P = 0.0259$; $df = 12$) (fig. 2.1). Genetic variation within locations in the introduced range was then compared with within location genetic variation from areas where the introduction likely originated (Boulder, Denver, Ft. Collins, and Colorado Springs), and the average number of pairwise differences was significantly smaller in introduced locations ($P = 0.023$; $df = 9$). AMOVA conducted on samples from native and introduced locations separately showed that genetic variation within and among locations were partitioned similarly in the two ranges (table 2.5). Within-location genetic variation was the greatest source of genetic variation (table 2.5) in both native and introduced ranges.

Bayesian analysis supported a hypothesis of 6 putative populations when data from native and introduced ranges were compared together (table 2.6), although interpretation of these data is difficult. Genetic structure in both ranges was characterized by variation within locations, and no pattern for the distribution of AFLP haplotypes in the native range could be ascertained (fig. 2.4). No solid evidence of separate native and introduced populations was observed, although two large groups were identified in the introduced range (fig. 2.4) and appeared somewhat

differentiated from native samples. These data indicate that groups of samples in the introduced range are more similar to each other than to samples from any specific location in the native range.

Phenotypic variation. Data were collected for adelgids that emerged from 49 galls (N = 27 for native locations, N = 22 for introduced locations) (table 2.2). Establishment varied from 4 to 380 (mean = 143.87), and preference varied from 0.03 to 0.85. Overall mean establishment for native and introduced samples did not differ. No difference in preference was observed between native and introduced samples ($F_{1, 47} = 0.41$; $P = 0.5246$) (fig. 2.2). However, results from a variance ratio test indicate significantly greater variance in preference by native samples than by introduced samples ($F_{22, 27} = 2.18$; $P < 0.05$).

Because I found distinct mitochondrial divergence between samples from the Rocky Mountains and those from the Great Basin (R. Ahern, personal observation), I tested the hypothesis that samples with similar mitochondrial haplotypes have similar host preference. All samples from both ranges were grouped by haplotype (table 2.1) and the main effect of host preference was tested by ANOVA. I observed a significant effect of haplotype on host plant preference ($F_{1, 40} = 4.55$; $P = 0.0391$). When only native populations were compared, this main effect was nearly significant ($F_{1, 25} = 3.73$; $P = 0.065$), while comparison of haplotype 'A' samples from the introduced range with haplotype 'B' samples from the native range produced a highly significant result ($F_{1, 30} = 5.64$; $P = 0.024$), supporting the previous observation of decreased variation in preference in introduced samples. Very little preference for

either variety was observed for introduced samples and samples from the Rocky Mountains (haplotype group A), while samples from the Great Basin preferred the coastal variety of Douglas fir (haplotype group B) (Fig. 2.3).

Discussion

Founder effects. Here, I sought to measure the genetic footprint of an introduction by analysis of genetic diversity in native and introduced ranges. Many researchers have considered genetic variation in native and introduced populations (Scheffer and Grissell 2003, Iline and Phillips 2004, Taylor et al. 2004, Grapputo et al. 2005, Stadler et al. 2005). However, very few studies include comprehensive sampling of both native and introduced populations of insects (although see Tsutsui et al. 2000, 2001, Suarez et al. 2004) and my data provide a unique picture of the size of founder effects experienced by introduced insects. The strength of the work presented here lies in my ability to sample both native and introduced ranges, providing important context for my findings.

As was expected, genetic variation in introduced samples of *A. cooleyi* is substantially reduced in terms of the number of mtDNA haplotypes that occur in the introduced range and the average number of pairwise genetic differences within locations, as measured by both mtDNA and AFLPs. Data from mitochondrial and nuclear genomes indicate that *A. cooleyi* introduced into the eastern United States have undergone a significant founder effect and persist with greatly reduced genetic variation relative to insects found in the native range.

Reduced phenotypic variation. Variation in host plant preference for secondary host varieties is greater in the native range than in the introduced range. Increased variation in secondary host preference may reflect adaptation to locally specialized host plants (Futuyma and McCafferty 1990, Via 1991), such as those found in geographically and climatically heterogeneous areas, or it may reflect variation that is neutral because these insects do not encounter the tested variation in nature. The environment within the native range of *A. cooleyi*, including the Rocky Mountains and the Great Basin, is more heterogeneous than the introduced range of Michigan and the eastern Appalachians and Piedmont, and patterns observed in this study may reflect both the age of this insect-host assemblages (Moran 1988) and local adaptation by adelgids and host plants in the native range (Futuyma and McCafferty 1990, Via 1991). Alternatively, low phenotypic variation in introduced adelgid populations may be due to cultivation of phenotypically homogeneous hosts in the eastern US, a genetic bottleneck, or both.

Genetic variation corresponds with phenotypic variation. Mitochondrial haplotypes are associated with secondary host preference. However, this is not to say that mitochondrial genes I sampled play any role in host selection or preference. The correlation between mtDNA haplotype and host preference is likely due to historical host plant relationships that have been structured by glaciations, isolation of populations, and host plant differentiation within the native range (Hewitt 1996, Avise 2000, Hewitt 2001). Importantly, divergent lineages of *A. cooleyi* hybridize in the wild when previously isolated mtDNA lineages experience secondary contact (R.

Ahern, personal observation). The result of hybridization between divergent *A. cooleyi* lineages with different host preferences is unknown. The management implications of this hybridization may be profound if hybridization leads to development of more pestiferous (for example, generalist, plastic, or resistant) phenotypes (Ellstrand and Schierenbeck 2000).

My data show that there is very little genetic variation found in introduced *A. cooleyi*. MtDNA haplotypes from the eastern US are also found in the Rocky Mountains, strongly suggesting that introduced populations originated from within the Colorado Rockies. Therefore, the similarity in host variety preference by samples from these two regions is not surprising. However, significant variation in host plant preference within the native range is notable given the close geographic proximity of collection sites and the continuous distribution of host plants within the range. One possible explanation for greater levels of genetic and phenotypic variation in samples from the native range is that these insects have been subjected to hosts and climatic regimes that are more variable than those found in the introduced range. Plant genotypes may affect adelgid survival and performance (Stephan 1987), and Rehfeldt (1978) found a high degree of inbreeding in stands of Douglas fir, indicating potentially profound genetic differences in cultivars from widely separated locations. Needle color, which is determined in part by the presence epicuticular waxes (Eigenbrode 1995), is one of the primary distinctions between coastal and interior varieties of Douglas fir (Lipscomb 1993). Needles of coastal varieties have relatively little wax and appear green, while needles of interior varieties have more wax and appear blue (Lipscomb 1993). Waxy crop varieties are often less susceptible to insect

herbivores than non-waxy varieties (Eigenbrode 1995), and, in addition to affecting feeding behavior, cuticular waxes may affect host use by reducing the ability of insects to adhere to host plants, as has been shown with psyllids (Brennan and Weinbaum 2001). Regarding *A. cooleyi*, Cranshaw (1989) found that *P. pugens* showing blue coloration tended to have fewer galls than greener trees but, because some blue trees were heavily galled, the author suggests that tree color alone does not confer resistance.

Given the correlation between elevation, temperature, and UV intensity, Douglas fir in the Rocky Mountains may be adapted to a colder, higher-UV environment and produce protective wax (Rehfeldt 1977, Sullivan and Teramura 1988, Li and Adams 1989, Apple et al. 2000) while Douglas fir in the lower elevation Great Basin do not. Consequently, if preference for the most common host plant phenotypes has evolved in *A. cooleyi*, when adelgids from the Great Basin are given a choice they prefer less-waxy secondary hosts, or are unable to use waxy hosts.

Overall implications. It appears as though the introduction of *A. cooleyi* into the eastern United States led to a founder effect, as would be expected for the introduction of relatively few individuals (Hartl and Clark 1997). This is notable because this introduction has occurred on a continuous land area (North America), suggesting that there may not be ongoing gene flow between native and introduced ranges. At least two introduction scenarios are possible given the data that I collected. First, a limited number of introductions to the eastern half of the United States may have occurred from a relatively small part of the native range. Alternatively, it is

possible that many introductions have taken place and only individuals from a relatively small portion of the native range have persisted due to selection against unfavorable genotypes. At present it is impossible to distinguish between these two scenarios.

There are no historical records that map the spread of *A. cooleyi* in the eastern US. Consequently, it is not possible to determine whether a lag period occurred. If a lag period occurred, I do not know its duration. None of *A. cooleyi*'s hosts are native to the eastern US (Little 1971), and population growth was likely mediated by the importation of plant material for landscapes and for establishment of Christmas tree farms in regions where Douglas fir and spruce species can thrive (R. Ahern, personal observation). My data provide no evidence for unique genotypes or hybridization of divergent lineages found in the introduced range, and I found no evidence that extensive selection on rare or novel genotypes facilitated this introduction. *A. cooleyi* in the eastern US have limited genetic variation but, apparently, are sufficiently adapted to have persisted in this novel environment for over 100 years (Gillette 1907). Given this finding, my results call into question the utility of neutral genetic variation to assess the probability of persistence in new environments by introduced species.

In contrast to neutral genetic variation, information about genetic variation at loci directly related to ecologically important phenotypes may provide a powerful predictive tool about the likelihood of establishment success (Lee 2002). Unfortunately, this information is not available for the overwhelming majority of organisms, and collection of these data is time-consuming and expensive (Sunnucks

2000). When information about quantitative trait loci is not available, information about variation in ecologically important phenotypic traits may provide indirect information about variation at these loci (Hawthorne and Via 2001), and may be a better predictor of establishment success than neutral genetic variation alone (Kolar and Lodge 2001, Roderick and Navajas 2003).

Table 2.1. Collection information for samples used for molecular analyses. Collection information, with the exception of N, is the same for phenotype assays. For range column, 'I' indicates samples from the introduced range and 'N' indicates samples from the native range. For 'N-mtDNA' and 'N-AFLP' columns, the first value is the number of samples analyzed from *Pseudotsuga menziesii* and the second value is the number of samples analyzed from *Picae pungens*. 'Group' column indicates haplotypic identity for genotype by phenotype analysis. Group assignment in based on geography, not sequencing.

Range	Location	N-mtDNA	N-AFLP	Collector	Group
I	Bedford, PA	8, 5	7, 6	R. Ahern	A
	Cadillac, MI	8, 8	8, 8	J. O'Donnell	A
	Petersburg, WV	8, 7	8, 7	R. Ahern	A
	Montrose, PA	8, 8	7, 7	R. Ahern	A
	Elma, NY	8, 7	7, 7	R. Ahern	A
	Lehighton, PA	8, 6	7, 8	R. Ahern	A
	Newport, VA	7, 8	8, 8	R. Ahern	A
	Total	104	103		
N	Ft. Collins, CO	7, 7	7, 8	R. Ahern	A
	Pocatello, ID	8, 8	7, 8	R. Ahern	B
	Boulder, CO	8, 7	5, 7	R. Ahern	A
	Colorado Springs, CO	7, 7	6, 6	R. Ahern	A
	Ogden, UT	8, 8	7, 5	R. Ahern	B
	Logan, UT	8, 8	8, 6	R. Ahern	B
	Denver, CO	7, 8	8, 8	R. Ahern	A
	Total	106	96		

Table 2.2. Collection information for samples used to determine phenotypic variation in secondary host preference. Samples from Michigan were collected by J. O'Donnell; all other samples were collected by R. Ahern

Range	Location	N
Introduced	Lehighton, PA	6
	Cadillac, MI	6
	Montrose, PA	5
	Elma, NY	5
	Total	22
Native	Ft. Collins, CO	5
	Pocatello, ID	3
	Boulder, CO	3
	Colorado Springs, CO	3
	Ogden, UT	4
	Logan, UT	3
	Laramie, WY	6
	Total	27

Table 2.3. Average number of pairwise mtDNA differences within introduced and native populations. The average number of pairwise differences within introduced populations was significantly smaller (P = 0.002; df = 12) than the average number of pairwise differences within native populations. Differences were computed by Arlequin.

		Differences
Introduced	Bedford, PA	0.28
	Cadillac, MI	0.59
	Petersburg, WV	0.13
	Montrose, PA	0
	Elma, NY	0.13
	Lehighton, PA	0.44
	Newport, VA	0
	Mean	0.22
Native	Ft. Collins, CO	6
	Pocatello, ID	4.25
	Boulder, CO	0.93
	Colorado Springs, CO	8.1
	Ogden, UT	0
	Logan, UT	5.71
	Denver, CO	3.69
	Mean	4.1

Table 2.4. Haplotype frequencies in native and introduced populations. Asterisks represent haplotypes found in the introduced range that likely originated from a portion of the native range that was not sampled due to the absence of *Picea pungens*, the primary host where samples were collected for this analysis.

		Haplotype																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Introduced	Bedford, PA	11	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cadillac, MI	12	2	1*	1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Petersburg, WV	14	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Montrose, PA	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Elma, NY	14	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Lehighton, PA	10	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Newport, VA	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	92	10	1*	1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Native	Ft. Collins, CO	6	2	0	0	1	2	3	0	0	0	0	0	0	0	0	0	0	0
	Pocatello, ID	0	1	0	0	0	0	0	12	1	1	1	0	0	0	0	0	0	0
	Boulder, CO	5	1	0	0	0	0	7	0	0	0	0	1	1	0	0	0	0	0
	Colorado Springs, CO	6	1	0	0	5	1	0	0	0	0	0	0	0	1	0	0	0	0
	Ogden, UT	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0
	Logan, UT	0	1	0	0	0	0	0	12	0	0	0	0	0	0	1	1	1	0
	Denver, CO	10	0	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	1
	Total	27	6	0	0	8	3	12	40	1	1	1	1	1	1	1	1	1	1

Table 2.5. AMOVA for mtDNA and AFLP data from native and introduced populations. All values are significant by permutation test (1000 iterations). All values were calculated by Arlequin (version 3.1)

Marker	Range	Source of Variation	df	Sums of squares	Variance components	Percentage of variation
mtDNA						
	Native					
		Among locations	6	272.566	2.869	58.83
		Within locations	99	198.793	2.008	41.17
		Total	105	471.358	4.877	
		Fixation index (FST)	0.588			
	Introduced					
		Among locations	6	1.137	0.005	4.46
		Within locations	97	10.854	0.112	95.54
		Total	103	11.99	0.117	
		Fixation index (FST)	0.045			
AFLP						
	Native					
		Among locations	6	401.967	2.356	6.35
		Within locations	89	3093.44	34.758	93.65
		Total	95	3495.41	37.114	
		Fixation index (FST)	0.063			
	Introduced					
		Among locations	6	434.452	2.946	9.07
		Within locations	95	2805.80	29.535	90.93
		Total	101	3240.23	32.480	
		Fixation index (FST)	0.091			

Table 2.6. Results of combined Structure analysis of samples from native and introduced ranges. Log likelihoods for hypotheses that AFLP haplotypes originate from different numbers of parent populations (K), as inferred by Bayesian analysis. K = 6 has the highest posterior probability.

Population number (K)	Log likelihood
3	-24255.4
4	-23827.2
5	-23755.1
6	-23434.4
7	-23951.6
8	-24108.4
9	-31473.1

Figure 2.1. Mean number of pairwise differences within populations as measured by shared AFLP bands in the native (mean = 70.13; standard error = 3.87) and introduced (mean = 59.15; standard error = 3.29) ranges. Means with different letters are significantly different at alpha = 0.05 by one tailed t-test (p = 0.026; df = 12).

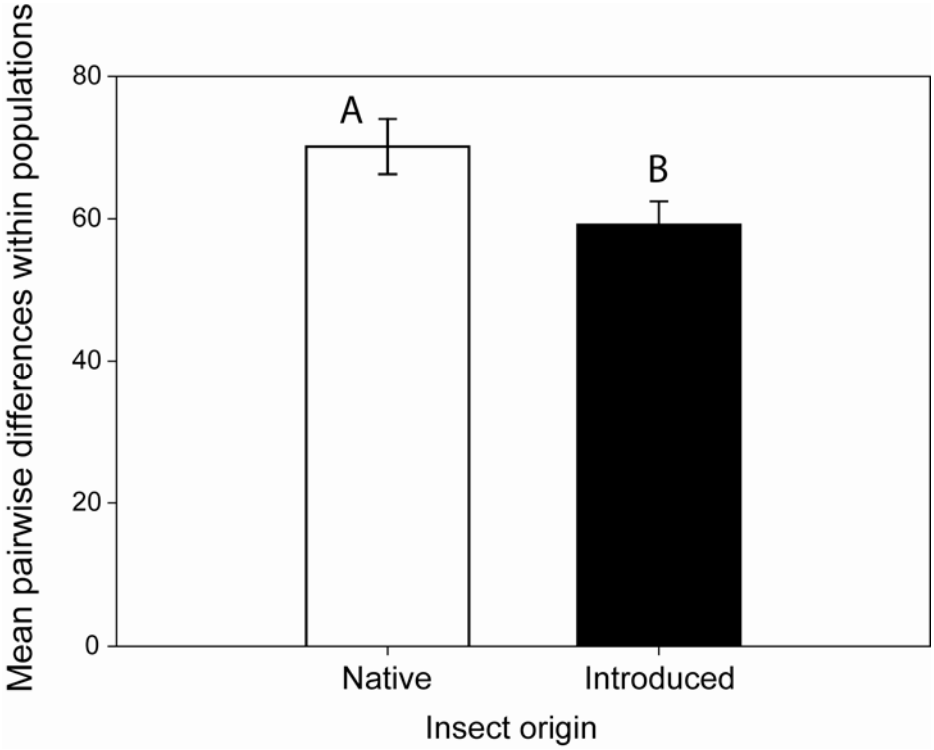


Figure 2.2. Mean establishment of adelgids from native and introduced populations. Bars indicate one standard error. No significant difference was observed in establishment between adelgids from the two ranges.

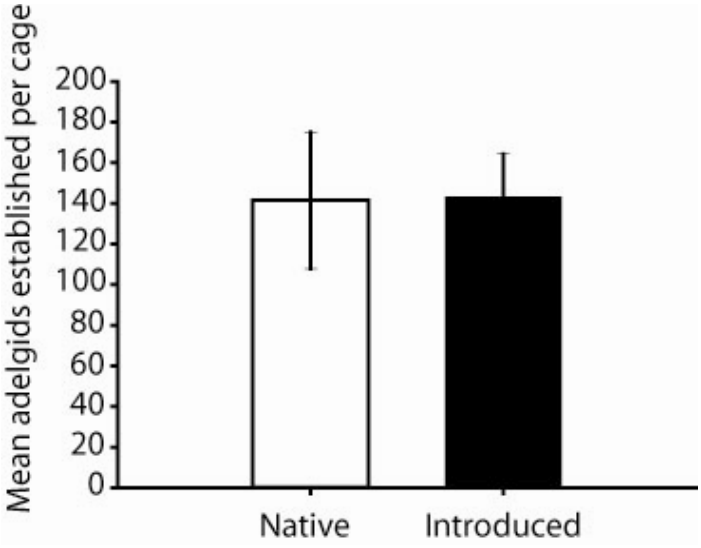


Figure 2.3. Variation in preference phenotype by haplotype group for all samples. Error bars represent one standard error. A significant effect of haplotype group was observed ($F_{1, 47} = 4.66$; P 0.0356). Haplotype groups are presented in table 2.1.

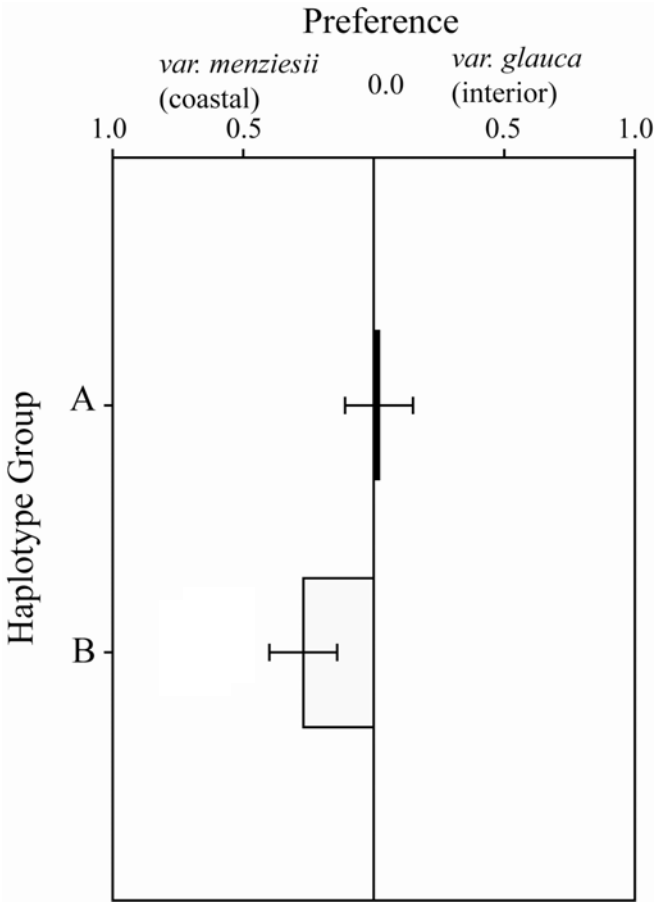
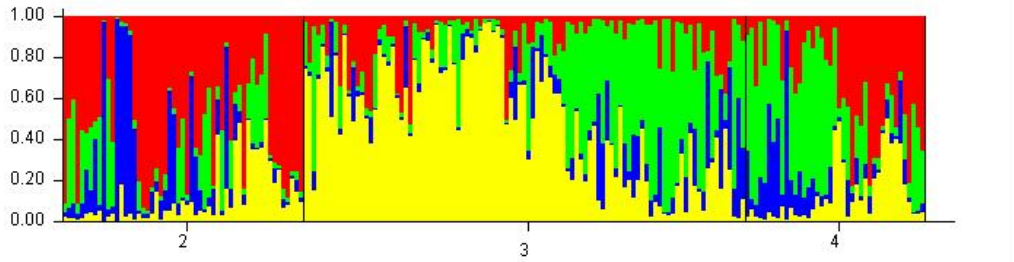
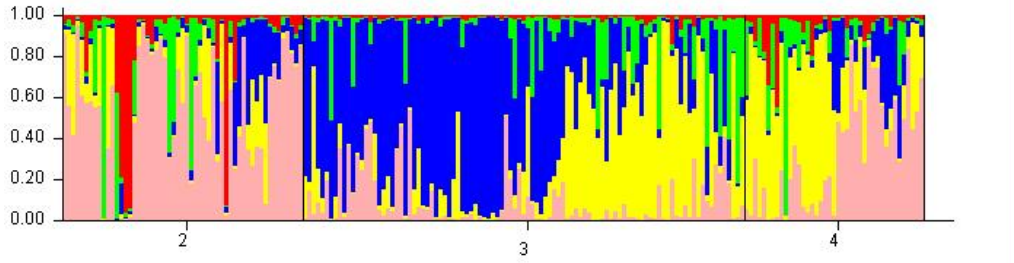


Figure 2.4. Structure analysis of native and introduced populations. Panels A-D represent hypotheses K = 4 through k = 7, respectively. Values on the x-axis represent ranges were samples originated: 2 = eastern Rockies, 3 = eastern US (introduced), and 4 = central Rockies. Scale on the y-axis represents the proportion of an individual's genetic variation assigned to each putative population. There is no pattern for the distribution of AFLP haplotypes in the native range and no evidence of separate native and introduced populations, although two large groups are apparent in the introduced range.

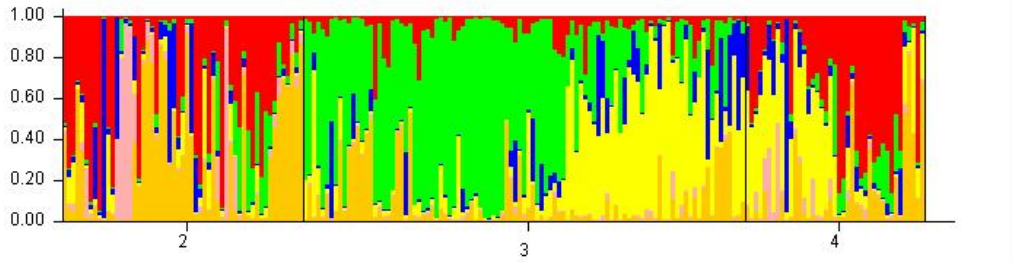
K=4



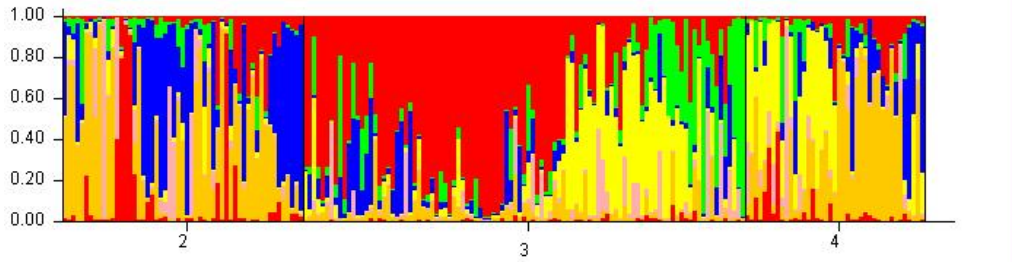
K = 5



K=6



K=7



Chapter 3: Distribution of genetic variation in a host-alternating insect pest, *Adelges cooleyi* (Hemiptera: Adelgidae), within and between primary and secondary hosts

Abstract

Adelges cooleyi, a host-alternating gall-making insect native to the Rocky Mountains and Cascades in the western United States, was introduced into the eastern United States during the mid-nineteenth century. Its host plants, *Pseudotsuga menziesii* and *Picea* species, are not native to the eastern US and occur together in patches that are often widely separated. Isolated, asexually reproducing populations are thought to exist on *P. menziesii*, although the existence of these populations has never been confirmed empirically and, if such populations are common, the degree to which they are genetically differentiated from holocyclic populations is not known. Analyses of mitochondrial and nuclear genetic variation were conducted to examine the distribution of genetic variation within and among host plants in the native range. These data were then used to determine expected values against which data from the introduced range were tested to identify discrepancies that may be consistent with an incomplete, or asexual, lifecycle. Although there is greater mitochondrial (mtDNA) haplotypic diversity in the native range, comparable numbers of mtDNA haplotypes were found on each host in native and introduced ranges. Distribution of nuclear genetic variation, as measured by amplified fragment length polymorphisms (AFLPs), within and among host plants in the

introduced range was not significantly different than in the native range. I found no evidence for asexual populations in the introduced range, although further analyses with co-dominant markers are required to confirm this observation. Factors that likely affect persistence of introduced *A. cooleyi* were identified.

Introduction

Parthenogenesis and specialization. The efficiency of selection is mediated by two lifecycle attributes: the number of genomes transmitted to an offspring and the extent of genetic exchange between different, independently selected organisms (Kondrashov 1997). Parthenogenesis is a form of asexual mitotic reproduction in which female insects give rise to clonal, genetically identical offspring (Kondrashov 1997, Forneck et al. 2001, although see Lushai and Loxdale 2002, Loxdale and Lushai 2003 for critique of the clone concept). Parthenogenic reproduction is common within the Aphidoidea (Dixon 1985, Blackman and Eastop 1994), with either facultative or obligate asexual reproductive periods. Asexual organisms experience adaptive and evolutionary tradeoffs compared to sexually reproducing organisms (Williams 1975). Sexually reproducing populations should show more genetic variation than asexually reproducing populations (Williams 1975, Wohrmann and Hales 1989, Hales et al. 1997), and genetic variation should decrease as asexual reproduction increases (Williams 1975, Dixon 1985, Hales et al. 1997). However, specialization, perhaps in the form of increased performance, may be more likely in asexual populations because advantageous genotypes are not fractured by recombination during sexual reproduction (Hales et al. 1997). Observations by Sunnucks et al. (1998) of host preference by grain aphids support this hypothesis, while Vorburger

(2004) found no difference in adaptation to cold temperatures between cyclical and obligate parthenogens.

Specialist insects, such as those that make galls on or inside host plants, have an intimate association with their hosts (Price 1997), often spending several months developing inside modified plant parts and occasionally completing their entire lives within tumor-like structures that they create (Wool 2004). Most galling insects show a high degree of host specificity, indicating a closely coupled relationship between plant and insect (Price 1997). This specificity may be due to the particular way in which galling insects interact with host plants, as aphid galls are often induced by hormonal changes in the plant host caused by female feeding (Wool 2004). Some gall-making insects, including members of the Adelgidae, require stimulation by both maternal and offspring feeding to induce galls (Ozaki 1994).

Host alternation. Host alternation in the Aphidoidea often involves seasonal migration between two unrelated hosts, typically a tree and an herbaceous host (Dixon 1985, Blackman and Eastop 1994). Moran (1988) developed several hypotheses to explain the development of complex aphid life cycles within the context of host alternation. Aphid species may be historically constrained and "fundatrix specialization", or extreme adaptation to the primary (original) host, may impede a complete switch to an alternate host. This constraint is detrimental because the alternate host may provide better and more consistent nutrient resources (Moran 1988). Hence, aphids are trapped in a sub-optimal situation by historical host relationships. A second hypothesis deals with the optimal use of host resources. Deciduous trees have an obvious phenology, with a flush

of activity during the spring and summer and senescence during the fall and winter. Migration to herbaceous hosts during the period of decreased arboreal activity provides aphids with access to otherwise scarce resources (Moran 1988). Given the seasonal consistency of gymnosperms relative to angiosperms, arguments based on fluctuating resource availability are likely less applicable in situations where both the primary and secondary hosts are conifers, as in the case of the Adelgidae.

Moran also provided some insight into how a host can be excluded from an aphid life cycle. Reduced migration to alternate hosts by some aphid clones may lead to the reduction of a complex life cycle and result in increased performance by these more specialized lineages (Moran 1988). Work by Via and Hawthorne (2001) and Via (1999) showed that increased aphid performance on one host plant may lead to decreased performance on another due to genetic correlations between performance on the two hosts. Unless loci that affect success on both hosts are linked, it is unlikely that specialization in the form of increased performance on one host can occur in species that require multiple hosts to complete their lifecycle (Hartl and Clark 1997, Hawthorne and Via 2001).

In holocyclic populations, genetic variation should be approximately equal on both host plants because insects frequently migrate from hosts where sexual reproduction occurs to those where asexual reproduction occurs (Williams 1975, Hales et al. 1997). In contrast, genetic variation in populations that have persisted for many generations with only parthenogenic reproduction is expected to be lower than that of holocyclic populations (Williams 1975, Hartl and Clark 1997). Linkage disequilibrium is expected to develop in persistent asexual populations (Chen and McDonald 1996, Hartl and Clark

1997, Sharbel et al. 2000), and a positive signal of asexually reproducing populations is a high frequency of genetically identical individuals, a high degree of linkage disequilibrium, or both. Molecular genetic tools can help to determine whether populations complete their lifecycle or are confined to the secondary host in several ways. First, fixation indices for populations known to reproduce sexually can be calculated and compared with putatively asexual populations. Similar signals in fixation indices may be achieved by long distance colonization, colonization by several founders, or both; however, similarity of indices supports completion of the holocycle. Second, molecular tools, especially co-dominant markers, can be used to measure linkage disequilibrium and test the distribution of alleles against null models of no linkage to infer a persistent asexual lifecycle (Chen and McDonald 1996, Sharbel et al. 2000).

Aphids serve as an excellent model for a closely related group of important insect pests, the Adelgidae. Fewer than two-dozen species within the genus *Adelges* have been identified worldwide (Blackman and Eastop 1994, Havill and Foottit 2007) and, given the paucity of empirical work on the reproductive biology of adelgids, empirical studies of aphid relatives may provide useful insights into the biology of this group. In particular, extensive work on host alternation and parthenogenesis in aphids provides a theoretical framework within which patterns of host associated genetic variation in adelgids may be better understood.

Adelgid biology. Gillette described *Adelges cooleyi* in 1907 (Gillette 1907, Annand 1928). The insect's primary hosts include various spruce species (*Picea pungens*, *P. engelmannii*, and *P. sitchensis*), and its secondary host is Douglas fir, *Pseudotsuga*

menziesii. This insect forms large galls at the end of terminal branches on spruce hosts and yellows and distorts needles on Douglas fir (Annand 1928). It requires both tree species to complete a complex lifecycle over two years (Annand 1928). Both sexual and asexual reproduction occurs on primary hosts, while all reproduction on the secondary host is asexual. The native range of CSGA in the United States resembles an inverted 'V', extending along the Rocky Mountains into Canada and back down the West Coast into northern California. This distribution is roughly congruent with the distribution of its primary and secondary hosts (Little 1971).

In his original description of *A. cooleyi*, Gillette (1907) proposed 2 species, distinguishing those found on the secondary host, *Pseudotsuga menziesii*, as *Chermes* (*Adelges*) *cooleyi* var. *coweni*. Annand (1928) also described two distinct groups, referring to those found with both primary and secondary hosts (*Picea* spp. and *P. menziesii*, respectively) as dieceous and those found with only the secondary host (*P. menziesii*) as monecious. Although Cameron (1936) later synonymised var. *coweni* with *A. cooleyi*, speculation persists about the existence of isolated, asexually reproducing populations on *P. menziesii* (Havill and Foottit 2007). The existence of isolated asexual populations has never been confirmed empirically in this species and, if such populations are common, the degree to which they are genetically differentiated from holocyclic populations is not known.

Adelges cooleyi was introduced into the eastern United States through an unknown number of introductions during the mid 19th century (Gillette 1907), likely when novel ornamental plant materials were exported to the eastern US (Reveal 1992). *Pseudotsuga menziesii* and *Picea* species used by *A. cooleyi* are not native to the eastern

US (Little 1971) and occur together in patches that are often widely separated in the introduced range of the insect (R. Ahern, personal observation). Habitat discontinuity may impede migration between populations and promote local adaptation (Via 1991). Asexual populations, thought to be present in the native range, may also be present in the introduced range, but no empirical data supports their existence. Understanding host use in the introduced range is also important from a management perspective (Granett et al. 2001). Removal of one host from proximity of the other may break the lifecycle of *A. cooleyi*, providing an important type of cultural control in areas where it is an economic pest (Hawthorne and Dennehy 1991, Granett et al. 2001).

The degree to which sexual reproduction contributes to maintenance and production of genetic variation in *A. cooleyi* is unknown and very little is known about sex ratios in native populations (Annand 1928, Havill and Foottit 2007), although the majority of the lifecycle is comprised of parthenogenic females (Annand 1928). In this study I collected molecular data from mitochondrial and nuclear genomes from samples collected on spruce and fir hosts in native and introduced ranges. Analysis of mitochondrial (mtDNA) sequence data from the cytochrome oxidase I (COI) gene can provide information about population structure and identify divergent lineages (Avise 1992, Templeton et al. 1995), and amplified fragment length polymorphisms (AFLPs) can be used to sample the nuclear genome and measure genetic structure and contemporary gene flow (Mueller and Wolfenbarger 1999, Parsons and Shaw 2001, Mendelson and Shaw 2005). The numbers of mtDNA haplotypes found on each host in the native range were compared to determine whether haplotypic diversity was partitioned among hosts. The same count was made in the introduced range to determine

whether the distribution of mtDNA haplotypes among hosts was consistent in the two ranges. I used AFLPs to assess relative levels of genetic diversity on both hosts within and among locations in native and introduced ranges. Fixation indices (Excoffier et al. 2005) were computed for the native range and compared with indices from the introduced range to determine whether the structure of genetic variation in the two ranges was significantly different. Results generated by analysis of data from non-recombining mitochondrial regions and recombining nuclear regions were then considered together to determine whether both sets of markers generated the same conclusions regarding the structure of genetic variation within and between hosts in native and introduced samples of *A. cooleyi*.

Methods

Material used. Adelges cooleyi was collected on Douglas fir (*Pseudotsuga menziesii)* and Colorado blue spruce (*Picea pungens)* throughout its native and introduced ranges (table 3.1). Insects collected on *Pseudotsuga menziesii* in early spring (April-May) were removed directly from foliage and stored in 100% EtOH until DNA was extracted, while those collected on *P. pungens* were removed from galls in mid-summer (June-July) and stored in 100% EtOH until DNA was extracted. With few exceptions, molecular analyses were conducted on approximately 8 insects from separate galls and 8 insects from different needles from different trees from 7 native and 7 introduced populations. All populations are separated by at least 25km.

MtDNA and AFLP data were analyzed to enable inference into both historical and contemporary processes that shape genetic variation. Corrie and Hoffmann (2004) used microsatellites to identify parthenogenic populations of grape phylloxera on roots in Australian vineyards, and it is difficult to quantify the contribution of meiotic processes to overall population genetic variation using dominant markers (Corrie et al. 2002). However, Jensen et al. (2002) used AFLPs to unambiguously separate sexual and asexual lineages, and studies of genetic structure using dominant markers may provide some insights into the likelihood that populations reproduce asexually (Forneck et al. 2001).

Laboratory procedures. Whole genomic DNA was extracted using DNeasy Tissue Kits (Qiagen) following a slightly modified protocol B for insects. All samples for sequencing were processed at Canadian Centre for Barcoding, Biodiversity Institute of Ontario, University of Guelph. Extractions were done as described in Ivanova et al. (2006). Primers were modified for Lepidoptera: LepF (ATTCAACCAATCATAAAGATATTGG) and LepR (TAAACTTCTGGATGTCCAAAAAATCA) from Hajibabaei et al. (2006). Sequencing reactions were conducted with ABI BigDye v. 3.1 and fragments were visualized on an ABI 3730 (Eric Maw, personal communication). Sequences were aligned with BioEdit (Ibis Therapeutics).

AFLPs were generated using a slightly modified version of the procedure originally proposed by Vos (1995). Digests were conducted with EcoRI and PstI enzymes. For selective amplification, EcoRI primers were labeled with florescent 6-FAM (6-carboxyl flourescein) to enable automated analysis with an ABI 3730 (Applied

Biosystems) capillary electrophoresis instrument. Of those screened, three EcoRI/PstI primer pairs (AG.CA, AG.GA, AG.CC) were used for selective amplifications. The screening process entailed a 2X comparison of 16 randomly chosen samples (N= 32 for each primer pair) starting from the construct step. Amplified products showed a high degree of concordance with products obtained during screening, indicating a high degree of repeatability between runs (<95%). A 1000bp internal size standard (ROX 1000, Bioventures) was run with samples, and data were collected and analyzed using GeneMapper software (Applied Biosystems). A locus was scored if peak height was at least 100 reflectance units (rfu) to reduce subjectivity in scoring. After visual inspection, ambiguous peaks (50-100 rfu) were scored '?', indicating missing data.

Molecular data analysis. Relatedness of mtDNA haplotypes was inferred by distance-based clustering analyses using PAUP* version 4.0 (Swofford 1999). Collapse software (version 1.2) was used to identify identical mtDNA haplotypes. Mean number of pairwise mtDNA differences was calculated for insects originating on each host from locations in native and introduced ranges (Excoffier et al. 1992, Excoffier et al. 2005). Mean pairwise differences between hosts in native and introduced samples were compared by paired t-tests to determine whether genetic variation was significantly low in samples from either host. Paired t-tests were performed because samples collected from both hosts were collected at the same sites and data were not independent (Zar 1998). The total number of mtDNA haplotypes observed in samples from native and introduced ranges was determined. Mean number of haplotypes occurring on each host in each range was calculated and a t-test was used to compare the number of mtDNA haplotypes found on

each host in the native range and the number of haplotypes present on each host in the introduced range. Data collected from locations in the native range were used to establish an expectation about the structure of mtDNA haplotypic diversity among hosts, and data collected from introduced locations were tested against this expectation.

Comparison of the genetic structure populations on different hosts may be used to identify patterns of host-related specialization and divergence (Haack et al. 2000), and analysis of variation between populations on both hosts may reveal differential host use, reduced migration from the secondary host, and overall decreased sexual reproduction (Corrie and Hoffmann 2004). AFLP data were used to generate average pairwise genetic differences (Excoffier et al. 1992, Excoffier et al. 2005) for insects on both hosts from native and introduced locations. Mean pairwise differences between hosts in native and introduced locations were compared by paired t-tests to determine whether genetic variation was significantly low in samples from either host. Paired t-tests were performed because samples collected from both hosts were collected at the same sites and data were not independent (Zar 1998).

Data collected from locations in the native range were used to establish expectations about the structure of genetic variation among hosts, among locations, and within locations. Data collected from locations in the introduced range were tested against these expectations. AMOVA was conducted to compare genetic variation within and among spruce and Douglas fir samples in native and introduced ranges (as in Muluvi et al. 1999, Parsons and Shaw 2001, Elderkin et al. 2004), and fixation indices were calculated from variance components (Excoffier et al. 2005). Significance tests (approximately 1000 reps) were run to determine probability of calculated values. Chi-

squared tests were used to determine whether fixation indices in the native range were significantly different (alpha = 0.05) than those in the introduced range. Measures of genetic distances, including Nei Li distances (Nei and Li 1979), were computed and distance-based clustering methods in PAUP* version 4.0 (Swofford 1999) were used to infer relatedness of AFLP banding phenotypes. Structure version 2.1, a software package that uses Bayesian algorithms to determine the likelihood of multiple hypotheses (Pritchard et al. 2000), was used to determine whether spruce samples were significantly differentiated from Douglas fir samples.

Results

Genetic variation- mtDNA. Analysis of mtDNA variation was conducted on 650bp of the COI gene. Two hundred and ten (210) individuals were sampled (106 in the native range; 104 in the introduced range) (table 3.1). For AFLP analysis, three primer combinations generated 445 polymorphic AFLP bands that were used to analyze nuclear DNA variation. Total character differences between individuals varied from 30 to 162. Average number of pairwise differences within native locations varied from 96.4 - 55.5 and 71.1 - 53.8 for Douglas fir and spruce, respectively. Average number of pairwise differences within introduced locations varied from 72.4 - 49.6 and 73.9 - 36.9 for Douglas fir and spruce, respectively. Corrected average number of pairwise differences among native locations varied from 0.4 - 15.9 and 1.1 - 12.1 for Douglas fir and spruce, respectively. Corrected average number of pairwise differences among introduced locations varied from 0.2 - 8.7 and 2.9 - 13.8 for Douglas fir and spruce, respectively.

Genetic variation- AFLPs. No samples with identical AFLP haplotypes were identified. Average numbers of pairwise genetic differences, as measured by mtDNA haplotypes and shared AFLP bands, were not significantly different by host in either native or introduced samples (alpha = 0.5) (fig. 3.1). However, a far greater number of mtDNA haplotypes were observed in the native range than in the introduced range (fig. 3.1, table 3.2), and subsequent AMOVA of mtDNA data was uninformative because the data were not comparable. The mean number of mtDNA haplotypes observed on each type of host plant was not significantly different in either the native or introduced range (table 3.3), indicating similar distribution of mtDNA genetic variation among hosts in each range.

The majority of AFLP genetic variation observed in native and introduced ranges occurs within locations (table 3.4) on both hosts. The following comparisons yielded no significant differences (Chi-squared test, alpha = 0.05) between fixation indices (table 3.5): native spruce and introduced spruce; native Douglas fir and introduced Douglas fir. AMOVA of AFLP data in which both hosts were compared together (representing the group level) in each range bore similar results, with no significant differences detected when fixation indices were compared (table 3.5) (Chi-squared test, alpha = 0.05). Variation among primary and secondary hosts accounted for approximately 0-2% variation (table 3.4) in these analyses, implying little restriction in gene flow between samples occurring on either host.

Bayesian analysis supported a hypothesis of four (K = 4) putative populations in native and introduced ranges (table 3.6, figs. 3.2, 3.3). Genetic structure in both ranges was characterized by variation within locations, and no evidence of reduced gene flow between samples on Douglas fir and spruce was observed. Distribution of AFLP

haplotypes in the native range was nearly uniform, (fig. 3.2) while two large groups with somewhat different haplotypic composition (fig. 3.3) were identified in the introduced range.

Discussion

Evidence for isolated asexual populations. Parthenogenetic populations are thought to persist on Douglas fir without migration to spruce (Gillette 1907) but, to date, no empirical evidence has been provided to substantiate this claim. Patterns of genetic variation inferred by dominant markers are not ideal for assessment of paternity and inbreeding (Sunnucks 2000); however, my sample of the nuclear genome provides a means by which the gross contribution of insects on each host makes to overall genetic variation may be analyzed. Here I provide evidence against the existence of persistent asexual populations in areas where I sampled. One signature of persistent isolated asexual populations is no or low within-population genetic variation (Hartl and Clark 1997). I found no samples with significantly fewer AFLP haplotypes. Genetic variation, which is expected to be small in inbreeding and asexual populations (Kondrashov 1997), was not significantly different on spruce and fir, and may be somewhat higher on *P. menziesii*, providing no support for reduced genetic variation in samples collected on the secondary host. One explanation for higher genetic variation on Douglas fir is that *A. cooleyi* migrate to secondary hosts from several primary host species, primarily *P. pungens* and *P. engelmannii* in the area that was sampled. As a result, adelgid populations may be more heterogeneous on Douglas fir than on any individual spruce species.

The pattern of little host-mediated genetic variation is consistent between native and introduced ranges, although the distribution of host plants in the introduced range is highly patchy and discontinuous. My data provide evidence that introduced insects find and effectively use both primary and secondary hosts. In general, there is very little genetic variation in introduced samples, and previous work (R. Ahern, personal observation) shows very little structure in mitochondrial and nuclear variation in introduced samples. Lack of genetic structure, taken with results presented here, indicates dispersal between locations throughout the introduced range. These results may have implications for pest management programs aimed at suppressing other adelgid species, as long distance movement by adelgids has been hypothesized (McClure 1990) but not rigorously tested. Previous studies dealing with adelgid movement suggest that birds and other mammals play an important role in dispersal (McClure 1990). While these organisms may contribute to limited long-distance movement of a small number of adelgids, it is likely that adelgid dispersal plays an important role in maintaining the low degree of among-location structure observed in introduced populations.

Implications for specialization. Increased performance on one host plant may lead to decreased performance on another due to genetic correlations between performance on the two hosts (Via 1999, Hawthorne and Via 2001). In insect species that require multiple hosts to complete their lifecycle, it is unlikely that increased performance on one host can occur unless loci that affect success on both hosts are linked (Hawthorne and Via 2001). Increased performance and specialization may be more likely in asexual populations because advantageous genotypes are not fractured by recombination during sexual

reproduction (reviewed in Hales et al. 1997). In theory, highly specialized parthenogenic lineages may persist for many years but these lineages are evolutionarily constrained by the genetic variation present in the founding female (Williams 1975). Sexually reproducing populations should have more genetic variation than asexually reproducing populations (Williams 1975, Wohrmann and Hales 1989), providing additional raw materials for adaptive evolution and specialization. Douglas fir and spruce species often co-occur in the western US (Little 1971), providing an ideal environment for completion of the lifecycle. In contrast, host plants are not continuous in the introduced range and the proximity of host species to one another is largely governed by choices made by tree growers. Chemical, cultural, and biological control efforts almost certainly exert some type of selective pressure on pest populations. While specialization may be a favorable strategy for short-term resource exploitation, phenotypic variability may be more important for long-term persistence (Williams 1975). In instances when multiple hosts are required for sexual reproduction, an effective, basic control method may be the spatial separation of hosts (Hawthorne and Dennehy 1991, Granett et al. 2001). Pest insects may persist on one or both hosts but it is likely that populations will have decreased genetic variation and may be more sensitive to control efforts (Roush and McKenzie 1987).

Testing hypotheses about isolated populations. In environments where selection is variable, a lifecycle that includes asexual reproduction during periods of low environmental stress and sexual reproduction during periods of heavy selection provides a good strategy for survival (Kondrashov 1997). If selective regimes decrease, either due to biotic or abiotic conditions, a shift to an asexual lifecycle may be advantageous

(Kondrashov 1997). Annand (1928) proposed the existence of two groups within the *A. cooleyi*: a diecious group comprised of insects found with both hosts, and a monecious group comprised of insects found with only Douglas fir. While no direct empirical evidence was collected to support or refute this proposal, molecular tools now exist that can test the hypothesis that insects found on the secondary host are genetically isolated.

Genetic variation in holocyclic populations should be approximately equal on both host plants because insects frequently migrate from hosts where sexual reproduction occurs to those where asexual reproduction occurs (Wohrmann and Hales 1989, Hales et al. 1997). In contrast, genetic variation is expected to be lower in populations that have persisted for many generations with only parthenogenic reproduction (Dixon 1985, Hales et al. 1997). Linkage disequilibrium is expected to develop in persistent asexual populations, and a positive signal of asexually reproducing populations is a high frequency of genetically identical individuals, a high degree of linkage disequilibrium, or both (Chen and McDonald 1996, Hartl and Clark 1997, Sharbel et al. 2000). My collections in the native range were made in areas where both hosts were present and my data provide a robust estimate of expected levels of genetic variation when insects are completing their life cycle. Because I found no evidence of isolated locations or differential host use in the introduced range, my data may be used as a baseline to test putatively isolated samples for signs of reduced gene flow consistent with an incomplete lifecycle.

Historical observations of putatively parthenogenic populations were made in California at locations where spruces were not present (Annand 1928). It is common for Douglas fir to grow at lower elevations than spruce species endemic to the Rocky

Mountains and Cascades (Little 1971). It is possible that asexual adelgid populations exist and represent lineages that persist without sexual reproduction, perhaps due to relative environmental stability in areas where they occur (Annand 1928), in contrast to harsh conditions found in the Rocky Mountains and Cascade Range. An alternative hypothesis is that colonizing adelgids may have originated from populations that were out of sight, giving the appearance that populations persist on Douglas fir when populations were actually renewed each year by long-distance migration. Molecular genetic approaches provide tools to address these hypotheses, and my data provide a starting point for assumptions about expected levels of genetic variation in holocyclic populations. In areas where population genetic variation is different than what I observed, these two hypotheses may be resolved using co-dominant markers to study linkage disequilibrium.

Table 3.1. Collection information for samples used for molecular analyses. In the 'Range' column, 'I' refers to populations in the introduced range and 'N' refers to populations in the native range. In the 'Number' column, the first value indicates the number of mtDNA samples analyzed and the second value indicates the number of AFLP samples analyzed. Samples from Cadillac, MI were collected by J. O'Donnell; all samples were collected by R. Ahern.

Range	Location	Number	Date	Host
I	Bedford, PA	8, 7	11-V-06	*P. menziesii*
	Bedford, PA	5, 6	13-VI-05	*P. pungens*
	Cadillac, MI	8, 8	17-V-06	*P. menziesii*
	Cadillac, MI	8, 8	28-VII-05	*P. pungens*
	Petersburg, WV	8, 8	2-V-06	*P. menziesii*
	Petersburg, WV	7, 7	11-VII-05	*P. pungens*
	Montrose, PA	8, 7	18-V-05	*P. menziesii*
	Montrose, PA	8, 7	14-VII-05	*P. pungens*
	Elma, NY	8, 7	18-V-06	*P. menziesii*
	Elma, NY	7, 7	14-VII-05	*P. pungens*
	Lehighton, PA	8, 7	11-V-06	*P. menziesii*
	Lehighton, PA	6, 8	13-VII-05	*P. pungens*
	Newport, VA	7, 8	2-V-06	*P. menziesii*
	Newport, VA	8, 8	9-VI-06	*P. pungens*
	Total	104, 103		
N	Ft. Collins, CO	7, 7	8-V-06	*P. menziesii*
	Ft. Collins, CO	7, 8	28-VI-05	*P. pungens*
	Pocatello, ID	8, 7	6-V-06	*P. menziesii*
	Pocatello, ID	8, 8	1-VII-05	*P. pungens*
	Boulder, CO	8, 5	8-V-06	*P. menziesii*
	Boulder, CO	7, 7	5-VII-05	*P. pungens*
	Colorado Springs, CO	7, 6	8-V-06	*P. menziesii*
	Colorado Springs, CO	7, 6	28-VI-05	*P. pungens*
	Ogden, UT	8, 7	5-V-06	*P. menziesii*
	Ogden, UT	8, 5	1-VII-05	*P. pungens*
	Logan, UT	8, 8	5-V-06	*P. menziesii*
	Logan, UT	8, 6	30-VI-05	*P. pungens*
	Denver, CO	7, 8	7-V-06	*P. menziesii*
	Denver, CO	8, 8	27-VI-05	*P. pungens*
	Total	106, 96		

Table 3.2a. Haplotype frequencies in the native range from populations collected on spruce and Douglas fir. Asterisks represent haplotypes found in the introduced range that likely originated from a portion of the native range that was not sampled due to the absence of *Picea pungens*, the primary host where samples were collected for this analysis.

		Haplotype																	
Host		1	2	3*	4*	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Spruce	Ft. Collins, CO	3	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
	Pocatello, ID	0	0	0	0	0	0	0	6	0	1	1	0	0	0	0	0	0	0
	Boulder, CO	3	0	0	0	0	0	3	0	0	0	0	0	1	0	0	0	0	0
	Colorado Springs, CO	4	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
	Ogden, UT	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0
	Logan, UT	0	1	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0
	Denver, CO	4	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	1
	Total	14	3	0	0	4	0	7	21	0	1	1	0	1	0	0	0	0	1
Fir	Ft. Collins, CO	3	1	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0
	Pocatello, ID	0	1	0	0	0	0	0	6	1	0	0	0	0	0	0	0	0	0
	Boulder, CO	2	1	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	0
	Colorado Springs, CO	2	0	0	0	3	1	0	0	0	0	0	0	0	1	0	0	0	0
	Ogden, UT	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0
	Logan, UT	0	0	0	0	0	0	0	5	0	0	0	0	0	0	1	1	1	0
	Denver, CO	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	Total	13	3	0	0	4	3	5	19	1	0	0	1	0	1	1	1	1	0

Table 3.2b. Haplotype frequencies in the introduced range from populations collected on spruce and Douglas fir. Asterisks represent haplotypes found in the introduced range that likely originated from a portion of the native range that was not sampled due to the absence of *Picea pungens*, the primary host where samples were collected for this analysis.

		Haplotype																
Host		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Spruce	Bedford, PA	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cadillac, MI	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Petersburg, WV	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Montrose, PA	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Elma, NY	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Lehighton, PA	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Newport, VA	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	46	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fir	Bedford, PA	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cadillac, MI	5	1	1*	1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Petersburg, WV	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Montrose, PA	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Elma, NY	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Lehighton, PA	5	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Newport, VA	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	46	7	1*	1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3.3. Number of mtDNA haplotypes found on spruce and Douglas fir hosts in native and introduced ranges. Mean haplotypes in each range are not significantly different by paired t-test ($p = 0.407$ and $p = 0.103$ for native and introduced ranges, respectively).

Location	Host	
	Spruce	Douglas fir
Ft. Collins, CO	3	4
Pocatello, ID	3	3
Boulder, CO	3	4
Colorado Springs, CO	3	4
Ogden, UT	1	1
Logan, UT	2	4
Denver, CO	4	2
Mean	2.71	3.14
Bedford, PA	2	2
Cadillac, MI	2	4
Petersburg, WV	1	2
Montrose, PA	1	1
Elma, NY	1	2
Lehighton, PA	2	2
Newport, VA	1	1
Mean	1.43	2

Figure 3.1. Average number of pairwise genetic differences as measured by mtDNA haplotypes and shared AFLP bands. Black bars represent data collected from adelgid populations on spruce, and white bars represent data collect from adelgid populations on Douglas fir. Error bars represent one standard error. Bars that share a letter are not significantly different (alpha = 0.05) by paired t-tests. All comparisons were made within ranges, not between.

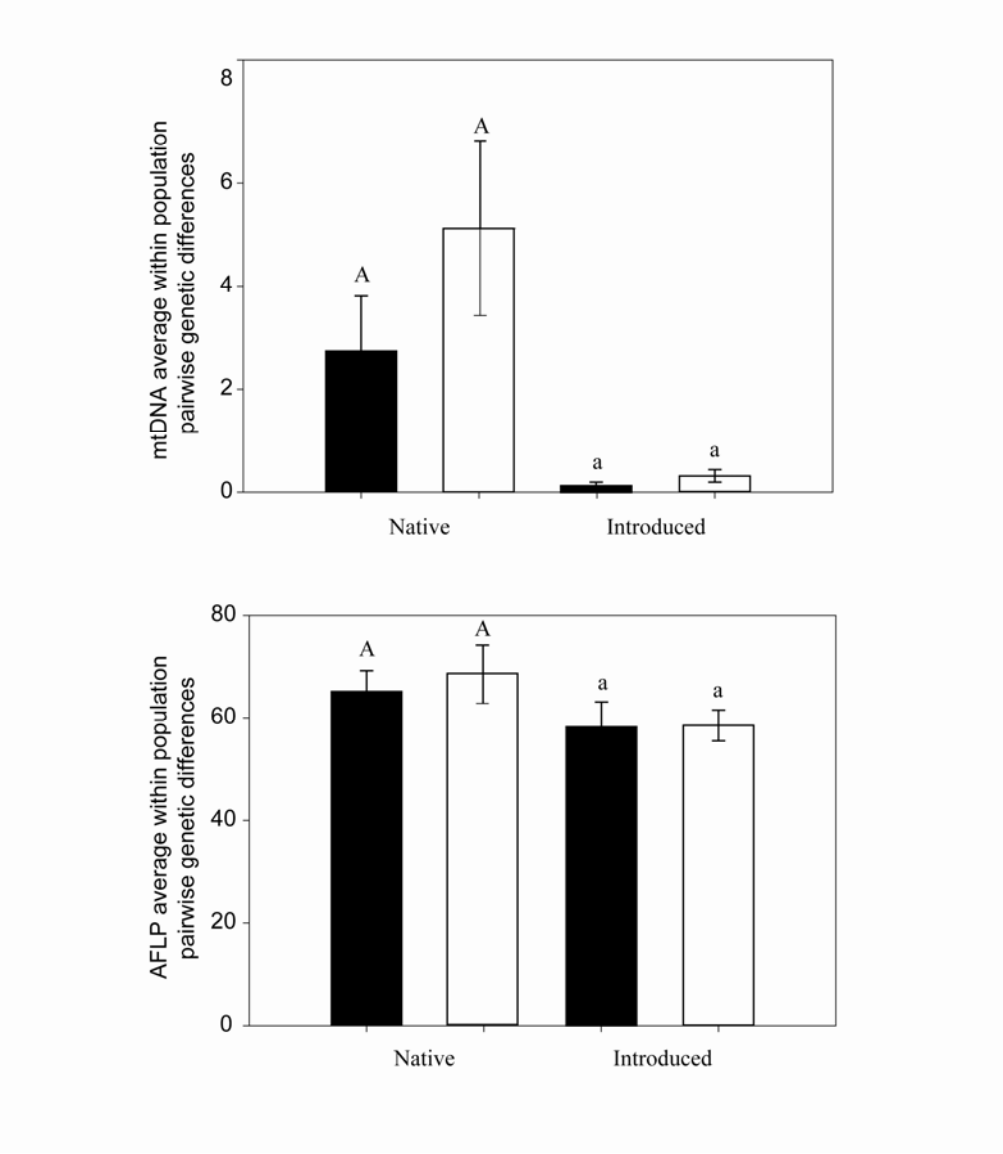


Table 3.4. AMOVA table of AFLP data. Combined analyses refer to those were data from both hosts were analyzed together. Unless otherwise noted (*), all variance components are significantly larger than 0. Note similarity in the partitioning of population genetic variation within populations and among hosts in native and introduced ranges.

Range	Host	Source of Variation	df	Sums of squares	Variance components	Percentage of variation
Native						
	Spruce					
		Among locations	6	338.404	3.459	9.54
		Within locations	41	1344.305	32.788	90.46
		Total	47	1682.708	36.247	
	Douglas fir					
		Among locations	6	325.592	2.943	7.93
		Within locations	41	1400.429	34.157	92.07
		Total	47	1726.021	37.020	
	Combined					
		Among groups	1	86.677	0.640	1.72
		Among locations	12	663.996	3.201	8.58
		Within locations	82	2744.733	33.472	89.71
		Total	95	3495.406	37.313	
Introduced						
	Spruce					
		Among locations	6	337.939	3.751	11.44
		Within locations	44	1277.512	29.034	88.56
		Total	50	1615.451	32.786	
	Douglas fir					
		Among locations	6	296.047	2.710	8.49
		Within locations	45	1315.107	29.225	91.51
		Total	51	1611.154	31.935	
	Combined					
		Among groups	1	36.201	-0.327*	-1.02
		Among locations	12	633.986	3.225	10.07
		Within locations	89	2592.619	29.131	90.95
		Total	102	3262.806	32.029	

Table 3.5. Fixation indices calculated from AFLP data following the procedure of (Excoffier et al. 2005). Unless otherwise noted (*), all values are significantly larger than a random value (1000 permutations). There are no significant differences (alpha = 0.05) between comparable indices in native and introduced ranges as determine by Chi-squared tests.

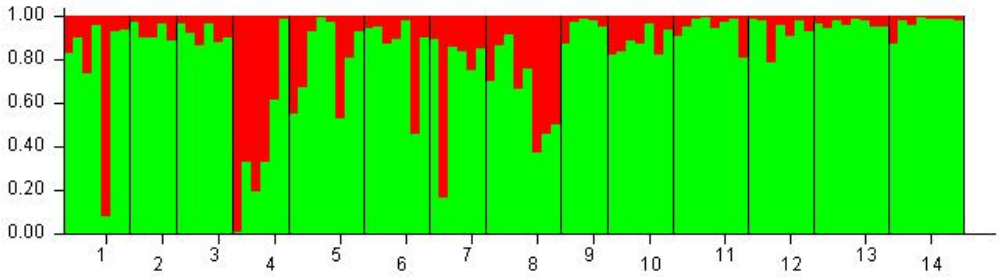
Range	Source of variation	Fst	Fsc	Fct
Native				
	Spruce	0.09543		
	Douglas fir	0.07932		
	Combined	0.10294	0.08728	0.01716
Introduced				
	Spruce	0.11442		
	Douglas fir	0.08486		
	Combined	0.0905	0.09968	-0.0102*

Table 3.6. Results of Structure analysis. Log likelihoods of hypotheses that AFLP haplotypes originate from different numbers of parent populations (K), as inferred by Bayesian analysis. K = 4 has the highest posterior probability for samples from native and introduced ranges.

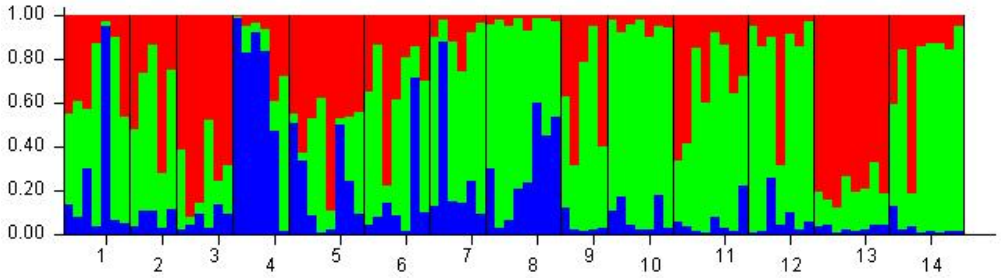
Population number (K)	Log likelihood	
	Native	Introduced
2	-12651.66	-11362.8
3	-12420.18	-11321.64
4	-12178.9	-10972.98
5	-15708.7	-11112.77
6	-13501.34	-11586.84
7	-16909.52	-12841.97
8	-13386.74	-12227.48

Figure 3.2. Structure analysis of native populations. Panels A-D represent hypotheses K = 2 through k = 5, respectively. Colors are arbitrarily assigned and represent putative populations. Values on the x-axis represent sample locations; odd numbered samples originated on spruce and even numbered samples originated on Douglas fir. Scale on the y-axis represents the proportion of an individual's genetic variation assigned to each population. No differentiation by host is observed.

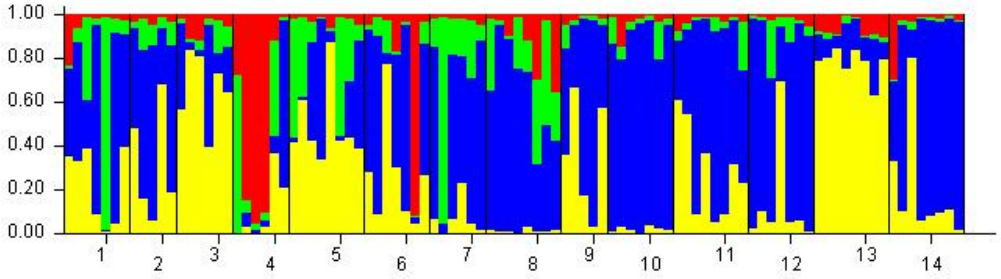
A. K = 2



B. K = 3



C. K = 4



D. K= 5

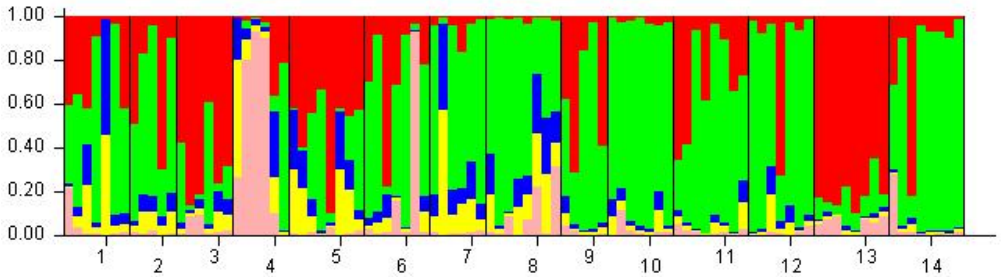
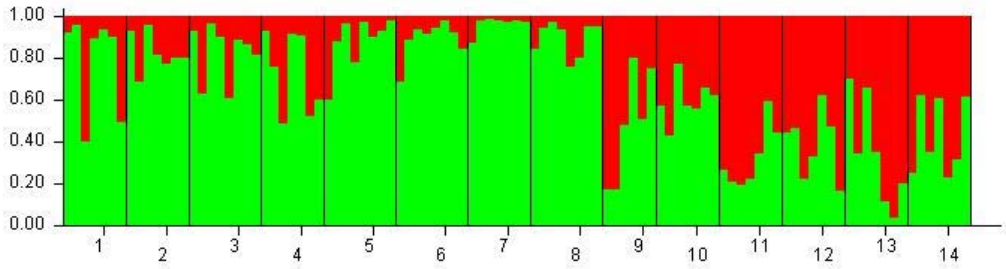
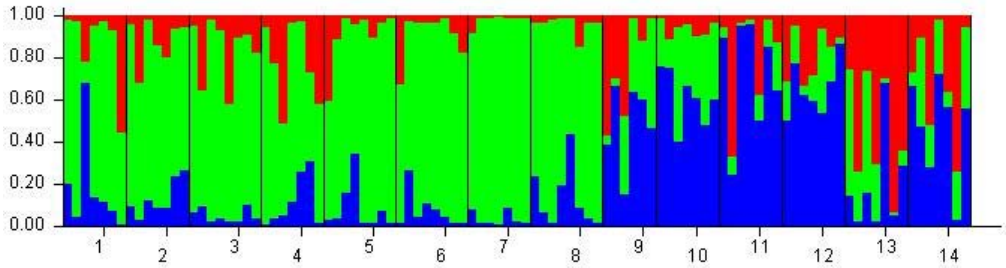


Figure 3.3. Structure analysis of introduced populations. Panels A-D represent hypotheses K = 2 through k = 5, respectively. Values on the x-axis represent sample locations; odd numbered samples originated on spruce and even numbered samples originated on Douglas fir. Scale on the y-axis represents the proportion of an individual's genetic variation assigned to each putative population. No differentiation by host is observed.

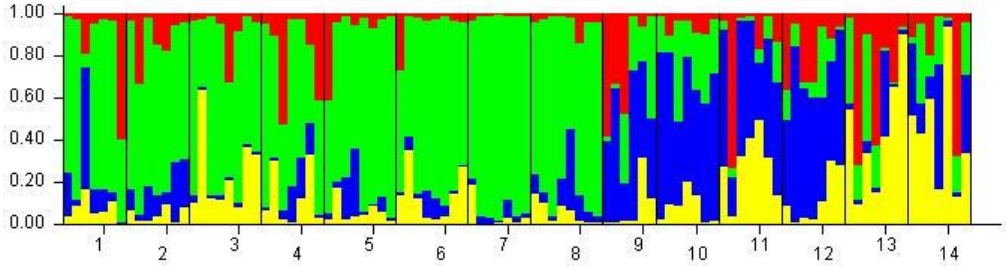
A. K = 2



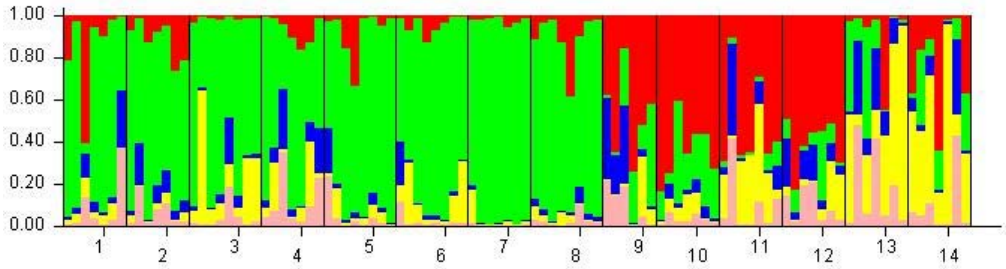
B. K = 3



C. K = 4



C. K = 5



Literature Cited

Aagaard, J. E., K. Krutovskii, and S. H. Strauss. 1998. RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. Heredity 81: 69-78.

Agrawal, A. A. 2001. Phenotypic plasticity in the interactions and evolution of species. Science 294: 321-326.

Albertson, R. C., J. A. Market, P. D. Danley, and T. D. Kocher. 1999. Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. PNAS 96: 5107-5110.

Annand, P. N. 1928. A contribution toward a monograph of the Adelginae (Phylloxeridae) of North America. Biological Sciences 6: 2-146.

Apple, M. E., D. M. Olszyk, D. P. Ormrod, J. Lewis, D. Southworth, and D. T. Tingey. 2000. Morphology and stomatal function of Douglas fir needles exposed to climate change: elevated CO_2 and temperature. International Journal of Plant Science 161: 127-132.

Avise, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna:a case history with lessons for conservation biology. Oikos 63: 62-76.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and S. N. 1987. Intraspecific phylogeography: The mitochondrial

DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489-522.

Baker, H. G. 1965. Characteristics and modes of origin of weeds, pp. 147-172. *In* H. G. Baker and J. A. Drake [eds.], The genetics of colonizing species. Academic Press, New York.

Barki, Y., J. Douek, D. Graur, D. Gateno, and B. Rinkevich. 2000. Polymorphism in soft coral larvae revealed by amplified fragment-length polymorphism (AFLP) markers. Marine Biology 136: 37-41.

Barton, N. H., and M. Slatkin. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity 56: 409-415.

Baum, D. A., and K. L. Shaw. 1995. Genealogical perspectives on the species problem, pp. 289-303. *In* P. C. Hoch and A. G. Stephenson [eds.], Molecular and experimental approaches to plant systematics. Missouri Botanical Garden, St. Louis.

Becerra, J. X., and D. L. Venable. 1999. Macroevolution of insect-host associations: The relevance of host biogeography to host affiliations. PNAS 96: 12626-12631.

Bensasson, D., D.-X. Zhang, D. L. Hartl, and G. M. Hewitt. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology and Evolution 16: 314-321.

Bensch, S., A. J. Helbig, M. Salomon, and I. Seibold. 2002. Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers. Molecular Ecology 11: 473-481.

Bensch, S., D. E. Irwin, J. H. Irwin, L. Kvist, and S. Akesson. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. Molecular Ecology 15: 161-171.

Blackman, R. L., and V. F. Eastop. 1994. Aphids on the World's Trees. CAB International, London.

Brennan, E. B., and S. A. Weinbaum. 2001. Effect of epicuticular wax on adhesion of psyllids to glaucous juvenile and glossy adult leaves of *Eucalyptus globulus* Labillardiere. Australian Journal of Entomology 40: 270-277.

Brown, J. M., W. G. Abrahamson, and P. A. Way. 1996. Mitochondrial DNA phylogeography of host races of the goldenrod ball gallmaker, *Eurosta solidaginis* (Diptera: Tephrididae). Evolution 50: 777-786.

Brown, J. M., J.-H. Leebens-Mack, J. N. Thompson, O. Pellmyr, and R. G. Harrison. 1997. Phylogeography and host association in a pollinating seed parasite *Grya politella* (Lepidoptera: Prodoxidae). Molecular Ecology 6: 215-224.

Burdon, J. J., and P. H. Thrall. 2004. Genetic structure of natural plant and pathogen populations, pp. 1-15. *In* L. E. Ehler, R. Sforza and T. Mateille [eds.], Genetics, Evolution and Biological Control. CABI Publishing, Cambridge.

Cameron, A. E. 1936. Adelges cooleyi Gillette (Hemiptera: Adelgidae) of the Douglas fir in Britain: completion of its life cycle. Annals of Applied Biology 23: 585-605.

Chen, R. S., and B. A. McDonald. 1996. Sexual reproduction plays a major role in the genetic structure of populations of the fungus *Mycosphaerella graminicola*. Genetics 142: 1119-1127.

Chiang, T. Y., Y. C. Chiang, Y. J. Chen, C. H. Chou, S. Havanond, T. N. Hong, and S. Huang. 2001. Phylogeography of *Kandelia candel* in East Asiatic mangroves based on nucleotide variation of chloroplast and mitochondrial DNAs. Molecular Ecology 10: 2697-2710.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1659.

Cock, M. J. W. 1978. The assessment of preference. The Journal of Animal Ecology 47: 805-816.

Corrie, A. M., and A. A. Hoffmann. 2004. Fine-scale genetic structure of grape phylloxera from the roots and leaves of *Vitis*. Heredity 92: 118-127.

Corrie, A. M., R. H. Crozier, R. Van Heeswijck, and A. A. Hoffmann. 2002. Clonal reproduction and population genetic structure of grape phylloxera, Daktulosphaira vitifoliae, in Australia. Heredity 88: 203-211.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer, Sunderland, MA.

Cranshaw, W. S. 1989. Patterns of gall formation by the Cooley spruce gall adelgid on Colorado blue spruce. Journal of Arboriculture 15: 277- 280.

Creer, S., R. S. Thorpe, A. Malhotra, W.-H. Chou, and A. G. Stenson. 2004. The utility of AFLPs for supporting mitochondria DNA phylogeographical analyses in the Taiwanese bamboo viper, *Trimeresurus stejnegeri*. Journal of Evolutionary Biology 17: 100-107.

Critchfield, W. B. 1984. Impact of the Pleistocene on the genetic structure of North American conifers, pp. 70-118. *In* R. M. Lanner [ed.], 8th North American Forest Biology Workshop, Utah State University, Logan UT.

Dixon, A. F. G. 1985. Aphid Ecology. Champman and Hall, New York, NY.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and plants: a study in coevolution. Evolution 18: 586-608.

Eigenbrode, S. D. 1995. Effects of plant epicuticular lipids on insect herbivores. Annual Review of Entomology 40: 171-194.

Elderkin, C. L., E. J. Perkins, P. L. Leberg, P. L. Klerks, and R. F. Lance. 2004. Amplified fragment length polymorphism (AFLP) analysis of the genetic structure of the zebra mussel, *Dreissena polymorpha*, in the Mississippi River. Freshwater Biology 49: 1487-1494.

Elias, M., O. Panaud, and T. Robert. 2000. Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. Heredity 85: 219-230.

Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? PNAS 97: 7043-7050.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47-50.

Forneck, A., M. A. Walker, and R. Blaich. 2001. An *in vitro* assessment of phylloxera (*Daktulosphaira vitifoliae* Fitch) (Hom., Phylloxeridae) life cycle. Journal of Applied Entomology 125: 443-447.

Funk, D. J., D. J. Futuyma, G. Orti, and A. Meyer. 1995. A history of host associations and evolutionary diversification of *Ophraella* (Coleoptera: Chrysomelidae): new evidence from mitochondrial DNA. Evolution 49: 1008-1017.

Futuyma, D. J., and S. C. Peterson. 1985. Genetic variation in the use of resources by insects. Annual Review of Entomology 30: 217-238.

Futuyma, D. J., and McCafferty. 1990. Phylogeny and the evolution of host plant associations in the leaf beetle *Ophraella* (Coleoptera, Chrysomelidae). Evolution 44: 1885-1913.

Gillette, C. P. 1907. Chermes of Colorado conifers. Proceedings of the Academy of Natural Sciences of Philadelphia 59: 3-22.

Granett, J., M. A. Walker, L. Kocsis, and A. D. Omer. 2001. Biology and management of grape phylloxera. Annual Review of Entomology 46: 387-412.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296: 707-711.

Grapputo, A., S. Boman, L. Lindstrom, A. Lyytinen, and J. Mappes. 2005. The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. Molecular Ecology 14: 4207-19.

Haack, L., J.-C. Simon, J.-P. Gauthier, M. Plantegenest, and C.-A. Dedryver. 2000. Evidence for predominant clones in a cyclically parthenogenic organism by combined demographic and genetic analyses. Molecular Ecology 9: 2055-2066.

Hajibabaei, M., D. H. Janzen, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences of the United States of America 103: 968-971.

Hales, D. F., J. Tomiuk, K. Wohrmann, and P. Sunnucks. 1997. Evolutionary and genetic aspects of aphid biology: a review. European Journal of Entomology 94: 1-55.

Hare, M. P. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16: 700-706.

Hartl, D. L., and A. G. Clark. 1997. Principles of Population Genetics. Sinauer Associates, Inc., Sunderland.

Hastings, A., K. Cuddington, K. F. Davies, C. F. Dugaw, S. Elmendorf, A. Freestone, S. Harrison, M. Holland, J. Lambrinos, U. Malvadkar, B. A. melbourne, K. Morre, C. Taylor, and D. Thomson. 2005. The spatial spread of invasions: new developments in theory and evidence. Ecology Letters 1: 91-101.

Havill, N. P., and R. G. Foottit. 2007. Biology and evolution of the Adelgidae. Annual Review of Entomology 52: 325-49.

Hawthorne, D. J. 1997. Ecological hostory and evolution in a novel environment: habitat heterogeneity and insect adaptation to a new host plant. Evolution 51: 153-162.

Hawthorne, D. J., and T. J. Dennehy. 1991. Reciprocal movement of grape phylloxera (Homoptera: Phylloxeridae) alates and crawlers between two differentially phylloxera-resistant grape cultivars. Journal of Economic Entomology 84: 230-236.

Hawthorne, D. J., and S. Via. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412: 904-907.

Hedrick, P. W., M. E. Ginevan, and E. P. Ewing. 1976. Genetic polymorphism in heterogeneous environments. Annual Review of Ecology and Systematics 7: 1-32.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society 58: 247-276.

Hewitt, G. M. 2001. Speciation, hybrid zones, and phylogeography - or seeing genes in space and time. Molecular Ecology 10: 537-549.

Hewitt, G. M. 2004. Genetic consequences of climate oscillations in the Quaternary. Philosophical transactions of the Royal Society B: Biological Sciences 359: 183-195.

Hufbauer, R. A., and S. Via. 1999. Evolution of an aphid-parasitoid interaction: variation in resistance to parasitism among populations specialized on different plants. Evolution 53: 1435-45.

Iline, I. I., and C. B. Phillips. 2004. Allozyme markers to help define the South American origins of *Microctonus hyperodae* (Hymenoptera: Braconidae) established in New Zealand for biological control of Argentine stem weevil. Bulletin of Entomological Research 94: 229-34.

Ingram, K. K., and D. M. Gordon. 2003. Genetic analysis of dispersal dynamics in an invading population of Argentine ants. Ecology 84: 2832-2842.

Ivanova, N. V., J. R. deWaard, and P. D. N. Hebert. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6: 998-1002.

Jackson, S. T., J. L. Betancourt, M. E. Lyford, S. T. Gray, and K. A. Rylander. 2005. A 40,000-year woodrat-midden record of vegetational and biogeographical dynamics in north-eastern Utah, USA. Journal of Biogeography 32: 1085-1106.

Jensen, L. H., H. Enghoff, J. Frydenberg, and E. D. Parker Jr. 2002. Genetic diversity and the phylogeography of parthenogenesis: comparing bisexual and thelytokous populations of *Nemasoma varicorne* (Diplopoda: Nemasomatidae) in Denmark. Hereditas 136: 184-194.

Johnson, W. T., and H. H. Lyon. 1988. Insects That Feed on Trees and Shrubs. Comstock Publishing Association, Ithaca, N.Y.

Johnston, J. A., L. A. Donovan, and M. L. Arnold. 2004. Novel phenotypes among early generation hybrids of two Louisiana iris species: flooding experiments. Journal of Ecology 92: 967-976.

Kerr, K. C. R., M. Y. Stoeckle, C. J. Dove, L. A. Weight, C. M. Francis, and P. D. N. Hebert. 2007. Comprehensive DNA barcode coverage of North American birds. Molecular Ecology In press.

Knowles, L. L., and W. P. Maddison. 2002. Statistical phylogeography. Molecular Ecology 11: 2623-2635.

Knowles, L. L., and C. L. Richards. 2005. Importance of genetic drift during Pleisttocene divergence as revealed by analyses of genomic variation. Molecular Ecology 14: 4023-4032.

Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution 16: 199-204.

Kolbe, J. J., R. E. Glor, L. R. Schettino, A. C. Lara, A. Larson, and J. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431: 177-181.

Kondrashov, A. S. 1997. Evolutionary genetics of life cycles. Annual Review of Ecology and Systematics 28: 391-435.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution 17: 386-391.

Lewontin, R. C., and L. C. Birch. 1966. Hybridization as a source of variation for adaptation to new environments. Evolution 20: 315-336.

Li, P., and W. T. Adams. 1989. Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). Canadian Journal of Forest Research 19: 149-161.

Lipscomb, B. 1993. Flora of North America North of Mexico. Oxford Univerity Press, New York and Oxford.

Little, E. L., Jr. 1971. Atlas of the United States trees, volume 1, conifers and important hardwoods, pp. 9, 200 maps. U.S. Department of Agriculture.

Loxdale, H. D., and G. Lushai. 2003. Rapid changes in clonal lines: the death of a 'sacred cow'. Biological Journal of the Linnean Society 70: 3-16.

Lushai, G., and H. D. Loxdale. 2002. The biological improbability of a clone. Genetical Research, Cambridge 79: 1-9.

Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: causes, epidemiology, global consequences, and control. Ecological Applications 10: 689-710.

Mackauer, M. 1976. Genetic problems in the production of biological control agents. Annual Review of Entomology 21: 369-385.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209-220.

McClure, M. S. 1990. Role of wind, birds, deer, and humans in the dispersal of hemlock woolly adelgid (Homoptera: Adelgidae). Environmental Entomology 19: 36-43.

Mendelson, T. C., and K. L. Shaw. 2005. Use of AFLP markers in surveys of arthropod diversity. Methods in Enzymology 395: 161-177.

Mitter, C. M., B. Farrell, and D. J. Futuyma. 1991. Phylogenetic studies of insect-plant interactions: Insights into the genesis of diversity. Trends in Ecology and Evolution 6: 290-293.

Mooney, H. A., and E. E. Cleland. 2001. The evolutionary impact of invasive species. PNAS 98: 5446-5451.

Moran, N. A. 1988. The evolution of host-plant alternation in aphids: evidence for specialization as a dead end. The American Naturalist 132: 681-706.

Moran, N. A. 2001. The coevolution of bacterial endosymbionts and phloem-feeding insect. Annals of the Missouri Botanical Garden 88: 35-44.

Mueller, U. G., and L. L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. Trends in Ecology and Evolution 14: 389-394.

Muluvi, G. M., J. I. Sprent, N. Soranzo, J. Provan, D. Odee, G. Folkard, J. W. McNicol, and W. Powell. 1999. Amplified fragment length polymorphism (AFLP) analysis of genetic variation in *Moringa oleifera* Lam. Molecular Ecology 8: 463-470.

Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy Science 76: 5269-5273.

Ozaki, K. 1994. Role of fundatrix and gallicola in the gall formation in Adelges japonicus (Monzen) (Hom., Adelgidae). Journal of Applied Entomology 118: 151-157.

Parsons, Y. M., and K. L. Shaw. 2001. Species boundaries and genetic diversity among Hawaiian crickets from the genus *Laupala* identified using amplified fragment length polymorphism. Molecular Ecology 10: 1765-1772.

Paterson, A. H., K. F. Schertz, Y.-R. Lin, S.-C. Liu, and Y.-L. Chang. 1995. The weediness of wild plants: molecular analysis of genes influencing dispersal

and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. PNAS 92: 6127-6131.

Pimentel, D., L. Lach, R. Zuniga, and D. Morrison. 2000. Environmental and economic costs associated with non-indigenous species in the United States. BioScience 39: 53-65.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Posada, D., K. A. Crandall, and A. R. Templeton. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Molecular Ecology 9: 487-488.

Price, P. W. 1997. Insect Ecology. John Wiley & Sons, Inc., New York.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Ran, J.-H., X.-X. Wei, and X.-Q. Wang. 2006. Molecular phylogeny and biogepgraphy of *Picea* (Pinaceae): Implications for phylogeographical studies using cytoplasmic haplotypes. Molecular Phylogenetics and Evolution 41: 405-419.

Rand, D. M. 2001. The units of selection on mitochondrial DNA. Annual Review of Ecology and Systematics 32: 415-448.

Rehfeldt, G. E. 1977. Growth and cold hardiness of intervarietal hybrids of Douglas-fir. Theoretical and Applied Genetics 50: 3-15.

Rehfeldt, G. E. 1978. The genetic structure of a population of Douglas-fir (*Pseudotruga menziesii* var. glauca) as reflected by its wind-pollinated progenies. Silvae Genetica 27: 49-52.

Remington, C. L. 1968. Suture-zones of hybrid interaction between recently joined biotas. Evolutionary Biology 2: 321-428.

Reveal, J. 1992. Gentle Conquest: The Botanical Discovery of North America with Illustrations from The Library of Congress. Starwood Publishing, Washington, DC.

Rhymer, J. M., and D. Simberloff. 1996. Extinction by hybridization and introgression. Annual Review of Ecology and Systematics 27: 83-109.

Roderick, G. K., and M. Navajas. 2003. Genetics in new environments: genetics and evolution in biological control. Nature Reviews Genetics 4: 889-899.

Rokas, A., E. Ladoukakis, and E. Zouros. 2003. Animal mitochondrial DNA recombination revisited. Trends in Ecology and Evolution 18: 411-417.

Roush, R. T., and J. A. McKenzie. 1987. Ecological genetics of insecticide and acaricide resistance. Annual Review of Entomology 32: 361-380.

Roy, B. A. 2001. Patterns of association between crucifers and their flower-mimic pathogens: host jumps are more common than coevolution or cospeciation. Evolution 55: 41-53.

Ruiz-Montoya, L. R., J. Nunez-Farfan, and J. Vargas. 2003. Host-associated genetic structure of Mexican poppulations of the cabbage aphid *Brevicoryne brassicae* L. (Homoptera: Aphididae). Heredity 91: 415-421.

Saville, B. J., Y. Kohli, and J. B. Anderson. 1998. mtDNA recombination in natural populations. Proceedings of the National Academy Science 95: 1331-1335.

Scheffer, S. J., and E. E. Grissell. 2003. Tracing the geographical origin of *Megastigmus transvaalensis* (Hymenoptera: Torymidae): an African wasp feeding on a South American plant in North America. Molecular Ecology 12: 415-421.

Seehausen, O. 2004. Hybridization and adaptation radiation. Trends in Ecology and Evolution 19: 198-207.

Shaffer, H. B., and M. L. McKnight. 1996. The polytypic species revisited: genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. Evolution 50: 417-433.

Sharbel, T. F., B. Haubold, and T. Mitchell-Olds. 2000. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Molecular Ecology 9: 2109-2118.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. PNAS 99: 16122-16127.

Simon, C., F. Frati, A. Backenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and the phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 78: 651-701.

Stadler, T., M. Frye, M. Neiman, and C. M. Lively. 2005. Mitochondrial haplotypes and the New Zealand origin of clonal European *Potamopyrgus*, an invasive aquatic snail. Molecular Ecology 14: 2465-73.

Stephan, B. R. 1987. Differences in the resistance if Douglas fir provenances to the woolly aphid *Gilletteella cooleyi*. Silvae Genetica 36.

Suarez, A. V., N. D. Tsutsui, D. A. Holway, and T. J. Case. 2004. Behavioral and genetic differentiation between native and introduced population of the Argentine ant. Biological Invasions 1: 43-53.

Sullivan, J. H., and A. H. Teramura. 1988. Effects of ultra-violet-B irradiation on seed growth in the Pinaceae. American Journal of Botany 75: 225-230.

Sunnucks, P. 2000. Efficient genetic markers for population biology. Trends in Ecology and Evolution 15: 199-203.

Sunnucks, P., D. Chisholm, E. Turak, and D. F. Hales. 1998. Evolution of an ecological trait in parthenogenic *Sitobion* aphids. Heredity 81: 638-647.

Swofford, D. L. 1999. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer Associates, Sunderland, MA.

Taylor, A. C., P. E. Cowan, B. L. Fricke, S. Geddes, B. D. Hansen, M. Lam, and D. W. Cooper. 2004. High microsatellite diversity and differential structuring among populations of the introduced common brushtail possum, *Trichosurus vulpecula*, in New Zealand. Genetical Research 83: 101-11.

Templeton, A. R. 1998. Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7: 381-397.

Templeton, A. R., and C. F. Sing. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping: IV. Nested analysis with cladogram uncertainty and recombination. Genetics 134: 659-669.

Templeton, A. R., E. Routman, and C. A. Phillips. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics 140: 767-782.

Thompson, J. N. 1988. Evolutionary ecology of the relationship between oviposition preference and performance of offspring in phytophagous insects. Entomologia Experimentalis et Applicata 47: 3-14.

Thompson, J. N. 1994. The coevolutionary process. Chicago Press, Chicago.

Thompson, J. N. 1999. Specific hypothese on the geographic mosaic of coevolution. The American Naturalist 153: S1-S14.

Tsutsui, N. D., A. V. Suarez, D. A. Holway, and T. J. Case. 2000. Reduced genetic variation and the success of an invasive species. PNAS 97: 5948-5953.

Tsutsui, N. D., A. V. Suarez, D. A. Holway, and T. J. Case. 2001. Relationships among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. Molecular Ecology 10: 2151-2161.

Ungerer, M. C., S. J. E. Baird, J. Pan, and L. H. Rieseberg. 1998. Rapid hybrid speciation in wild sunflowers. Evolution 95: 11757-11762.

Via, S. 1991. The genetic structure of host plant adaptation in a spatial patchwork: demographic variability among reciprocally transplanted pea aphid clones. Evolution 45: 827-852.

Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. Evolution 53: 1446-1457.

Vila, C., I. R. Amorim, J. A. Leonard, D. Posada, J. Castroviejo, F. Petrucci-Fonseca, K. A. Crandall, H. Ellegren, and R. K. Wayne. 1999. Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. Molecular Ecology 8: 2089-2103.

Vitousek, P. M., C. M. D'Antonio, L. L. Loope, M. Rejamek, and R. Westbrooks. 1997. Introduced species: a significant component of human-caused global change. New Zealand Journal of Ecology 21: 1-16.

Vorburger, C. 2004. Cold tolerance in obligate and cyclical parthenogens of the peach-potato aphid, *Myzus persicae*. Ecological Entomology 29: 498-505.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.

Williams, G. C. 1975. Sex and Evolution. Princeton University Press, Princeton, NJ.

Wohrmann, K., and D. F. Hales. 1989. Life cycle components and genetic variability in aphids. Journal of Applied Entomology 107: 71-77.

Wool, D. 2004. Galling aphids: specialization, biolgical complexity, and variation. Annual Review of Entomology 49: 175-192.

Zar, J. H. 1998. Biostatistical Analysis. Prentice Hall, Upper Saddle River, New Jersey.

Zhang, D.-X., and G. M. Hewitt. 2003. Nuclear DNA analysis in genetic studies of populations: practice, problems, and prospects. Molecular Ecology 12: 563-584.