BEAUTIFUL BUT LACKING DIVERSITY: POPULATION GENETICS OF PACIFIC DOGWOOD (*Cornus nuttallii* Audobon ex Torr. & A. Gray)

by

Karolyn R. Keir

Hons. B.Sc., University of British Columbia, 2004

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

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Forestry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

August 2008

ABSTRACT

In the past, conifers have been the primary focus of population and conservation genetic studies in Pacific Northwest (PNW) trees. These studies have provided tremendous insight as to how genetic diversity varies across species ranges for these wind-pollinated and mostly wind-dispersed species. With this study of Pacific dogwood (*Cornus nuttallii*), a broadleaved, PNW species, which utilizes biological vectors for pollen and seed dispersal, we hope to broaden our understanding of tree evolutionary dynamics.

Marker development for *C. nuttallii* found few useful polymorphisms. Of eight microsatellite markers (SSRs) developed from a closely related species, three were monomorphic, while the other five averaged only 4.4 alleles/locus. Furthermore, only a single base pair substitution was found in the *rpl16* region of the chloroplast genome after sequencing 2,262 non-coding base pairs in 100 individuals. This lack of diversity, which was found to be ubiquitous throughout the range of *C. nuttallii*, suggests this species may have endured a prolonged bottleneck in a single glacial refugium prior to recolonization. The cpDNA phylogeographic pattern and a significant decline in both SSR allelic richness ($r^2 = 0.42$, $p<0.01$), and expected heterozygosity ($r^2 = 0.51$, $p<0.01$) support this theory. Low levels of population structure, documented in both chloroplast ($D = 0.153$) and nuclear genomes ($F_{ST} = 0.071$, $R_{ST} = 0.036$) may suggest high levels of contemporary gene flow between populations are also influencing current patterns of diversity. Despite variation being the precursor for adaptation, a comparison of Q_{ST} (0.088 for first-year height and 0.113 for bud burst timing) with a refined F_{ST} estimate

(0.053), indicated that *C. nuttallii* had either retained or recovered significant phenotypic variation for differential selection to act.

Such uniformly low diversity raises the issue of how genetic conservation efforts should proceed with this and other species sharing a similar degree of genetic depauperateness. So that signs of decline may be detected, we suggest population monitoring, especially for those populations occurring at high elevations. Furthermore, we advocate the transfer of seeds from the nearest southern source, in the event that restorative efforts are required to assist this species to cope with the rapidly changing climate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Many hands and minds touch the pages of the following thesis, and it is to their gracious owners, for which I am indebted. Firstly, I must thank my wonderfully supportive supervisor, Dr. Sally Aitken. She has taught me a great deal, both within and beyond the realm of academia. Her unfaltering patience and leadership are admirable and I feel incredibly fortunate to have her as a mentor as I learn to navigate the ups and downs of scientific research. I would also like to express my gratitude for my committee members, Dr. Jeannette Whitton, and Dr. Peter Arcese, as well as my funding source, the Forest Investment Account of B.C., through the Forest Genetics Council.

Additionally, I would like to thank the US Forest Service, specifically, Fred Krueger, Dean Davis, Jenny Haas, Fletcher Linton and Melody Lardner for easing the task of finding small populations of Pacific dogwood in foreign lands. I am also grateful to the assistance of Don Pigott and Paulus Vrigmoed, for advice regarding the storage and germination of dogwood seeds.

For their expertise, I was grateful to have the guidance of Carol Ritland, Hesther Yueh and Allison Miscampbell. For their assistance in the laboratory, as well as their engaging and entertaining conversations I thank Charlotte Whitney, Amy Tchir, Jason Buckwald, Leyla Tabanfar, Pia Smets, Jordan Bemmels, Jill Hamilton, Melissa Bodner and Amy Wilson. They have been my sunshine in a windowless room.

I must also bestow praise for my friend, and colleague, Christine Chourmouzis. Her council, humour, and courage were invaluable through all portions of this work. In the field, and in life, her determination to make the seemingly impossible, possible is awe inspiring. And finally I must thank Colin Huebert, whose love and support were integral to the completion of this thesis. As it was this thesis work which saw our paths cross, I will forever see this time of my life as an immeasurable achievement.

CHAPTER 1 – Introduction and Objectives

1.1 Introduction

1.1.1 Pacific dogwood

Pacific dogwood (*Cornus nuttallii* Audubon ex. Torr. and Gray) is a beautiful deciduous tree found in forests along the Pacific coast from the lowlands of southwestern British Columbia to the mountains of southern California (Little 1976). In addition to this continuous range, disjunct populations of *C. nuttallii* are found in northern Idaho (Figure 1.1) (Arno 1977). In Idaho, over 60 plant species and numerous other animals and fungi are currently separated from the coastal portion of their distributions by 300km of arid habitat (Brunsfeld *et al*. 2001). These disjunct populations are thought to have originated through either ancient vicariance (e.g. uprising of the Cascade Mountain range during the Pliocene epoch 2-5 million years ago (mya)) or through a more recent inland dispersal (Brunsfeld *et al.* 2001, Carstens *et al.* 2005).

Within its range, *C. nuttallii* can be found at low densities underneath tall conifers, including Douglas-fir (*Psuedotsuga menziesii*), Sitka spruce (*Picea sitchensis*), and in the southern portion of its range, coastal redwood (*Sequoia sempervirens*), and giant Sequoia (*Sequoiadendron giganteum*). It also grows abundantly in open areas, such as roadsides (Arno 1977). Pacific dogwood is easily identified in the spring because it is covered with showy white floral bracts, usually occurring in groups of 4 to 6. These bracts subtend inflorescences of small, true flowers and have evolved to attract insect pollinators. After pollination, the bracts fall off and the single seeded fruits begin to mature. These bright red drupes reach maturity in the fall and attract a wide variety of birds and mammals, which act as vectors for seed dispersal (Klinka *et al.* 2000). Relative

to gravity or wind seed dispersal mechanisms, this method has the potential to spread Pacific dogwood seeds over great distances.

Pacific dogwood's closest relative is an eastern species, flowering dogwood (*Cornus florida* L.) (a.k.a. eastern flowering dogwood, Florida dogwood). Chloroplast DNA (cpDNA) restriction site analysis and sequence data, as well as a morphological comparison, placed these two species within the same lineage, the big-bracted dogwoods (Xiang *et al*. 1996). These two species are thought to have diverged over 15 million years ago (J Xiang pers. comm. U of North Carolina 2008). Besides these phylogenetic analyses, no genetic work has been published for Pacific dogwood. Up to present, any information regarding the genetics of populations for this species has been largely anecdotal.

Threats to *C. nuttallii* include habitat loss and fragmentation from human development, as well as an introduced fungal parasite, *Discula destructiva.* This pathogen is known to cause a potentially fatal disease, dogwood anthracnose. *Cornus florida* has also been heavily impacted by this disease (Caetano-Anolles *et al*. 1996). Dogwood anthracnose has resulted in the listing of *Cornus nuttallii*, as a Priority 1 species in the State of Idaho. Pacific dogwood trees are considered to be in danger of becoming extirpated from Idaho in the future if "identifiable factors contributing to their decline continue to operate" (Idaho Rare Plant Conference 2004).

1.2 Factors influencing the genetic structure of populations

Estimates of population genetic diversity and differentiation across the native range of a species are important for understanding the genetic effects of historic

evolutionary processes (e.g. post glacial recolonization), making informed predictions regarding the future and determining the conservation value of peripheral populations (Lesica and Allendorf 1995). This information can aid in the setting of conservation priorities and the effective allocation of finite conservation resources, by revealing unique and genetically diverse populations within a species. Simply put, population genetic diversity and differentiation are dictated by genetic drift, mutation, gene flow or migration, and natural selection. The relative strength of these forces is, in turn, determined by population size and the spatial separation of neighbouring populations. However, complexities arise as not all species behave in a similar, predictable manner. Life history traits (e.g. seed dispersal in plants) have a significant effect on phylogeographic structure (Irwin 2002), as well as the distribution of genetic diversity in species' ranges (Hamrick and Godt 1996, Nybom 2004).

Historically, plant species occurring along the west coast of North America have been a popular focus for phylogeographic studies (Soltis *et al.* 1997, Brunsfeld *et al.* 2001), often focusing on geographic patterns of intraspecific genetic variation in organelle and nuclear genomes. This interest is perhaps owing to the complex paleoclimatic history of the region (Hewitt 2000), the unique and relatively restricted distributions of species (Brunsfeld *et al.* 2001), the large number of endemics and/or the economic value associated with many tree species found in this region of the world. Although this research has produced informative results regarding the genetic structuring of populations of many wind pollinated and wind dispersed trees, species employing biological vectors for pollination and seed dispersal are largely under-represented in the published body of knowledge. The inclusion of sympatric species employing such

mechanisms for gene flow may broaden our understanding of tree evolutionary dynamics and add clarity to this burgeoning body of research. Pacific dogwood presents an excellent opportunity to fill this gap.

1.2.1 Phylogeographic structure

Owing to its uniparental inheritance, lack of recombination, an effective population size one quarter that of nuclear markers, and ease of marker amplification using universal primers, chloroplast DNA (cpDNA) is the most popular molecular tool employed by those interested in investigating the phylogeographic structure of plant species. Numerous universal primers have been identified for this purpose (Taberlet *et al*. 1991, Hamilton 1999). Shaw *et al*. (2005) tested the relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analyses at low taxonomic levels in numerous taxa. This summary was further refined to regions useful for phylogeographic studies in angiosperms (Shaw *et al*. 2007).

Comparison of phylogeographic patterns of sympatric plant and animal species in the Pacific Northwest region of North America identified Pleistocene glaciation as a major force affecting the current geographic distribution of haplotypes in this region (Soltis *et al*. 1997, Brunsfeld *et al*. 2001). During this epoch, Milankovitch cycles, created by variations in the earth's orbit every 100,000 years, resulted in approximately 90,000-year glacial periods, with alternating interglacial periods lasting 10,000 years (Pielou 1991). These cycles repeatedly covered the PNW in cordilleran and alpine ice, forcing species to more hospitable climatic zones at lower latitudes and elevations (Pielou 1991). During the last glacial maximum (approximately 20,000 years before present

(ybp)), it is hypothesized that many temperate species existed only in glacia refugia. Geological data, pollen records, and fossil evidence have led to formal proposition that these refugia existed along the coast south of glaciation (Whittaker 1961, Heusser 1985, Smith and Sawyer 1988), as well as in northern unglaciated regions such as central Alaska (Elliot-Fisk 1988), northwestern Vancouver Island (Heusser 1960, Pojar 1980), the Queen Charlotte Islands (Heusser 1989) and northern Idaho (Daubenmire 1952, Detling 1968). Furthermore, after Soltis *et al.* (1997) observed a common north-south partitioning of haplotypes amongst plant species from this region, two post-glacial recolonization hypotheses were proposed. These were the "north-south" recolonization hypothesis (involving multiple refugia) and the "leading edge hypothesis" (involving only a single southern refugium). As an extension of this work, researchers have now formalized hypotheses relating to the origin of disjunct mesic forests found in northern Idaho as well (Brunsfeld *et al*. 2001, Carstens *et al*. 2005, Brunsfeld and Sullivan 2005, Brunsfeld *et al*. 2007).

Irwin (2002) has shown, using simulations, that strong phylogeographical discontinuities could be observed in the absence of a geographic barrier, when both population size and dispersal distance are relatively small. Conversely, weak phylogeographic structure would be the result of large dispersal distances and large population sizes (Irwin 2002). As Pacific dogwood seeds have the capacity to travel great distances after ingestion by birds and mammals, it is likely that this life history trait would strongly influence the phylogeographic structure of this species.

Frequent, long distance dispersal events are often offered as explanations for the observation of unusual or weak phylogeographic structure. Other Pacific Northwest plant

species that deviate from the expected north-south phylogeographic pattern include *Heuchera micrantha* (crevice alumroot), and to a lesser degree, *Polystichum munitum* (sword fern). Long distance seed or spore dispersal events were cited as a potential cause for the lack of structure in both of these examples (Soltis *et al*. 1989, Soltis *et al.* 1997). In eastern North American *Prunus* species, Shaw and Small (2005) hypothesized that historical (eg. refugial zones) and contemporary forces (eg. dispersal of plum seeds by humans over large distances) acted in concert to create the unique phylogeographic pattern observed. Furthermore, the hypothesized long-distance dispersal capabilities of the now extinct passenger pigeon were cited as a potential reason for the weak phylogeographical structure observed in eastern North America's *Quercus rubra* L. (northern red oak) (Magni *et al*. 2005).

1.2.2 Structuring of neutral nuclear genetic variation

Under the assumptions of the 'abundant centre model' (Sagarin and Gaines 2002), the size and amount of gene flow received by individual populations are predicted to vary geographically. Optimum ecological conditions and large population sizes often occur synchronously at the centre of a species' range (Hengeveld and Aaeck 1982, Brown 1984, Brussard 1984) and small, more isolated populations are frequently found at the range peripheries (Lesica and Allendorf 1995). Owing to the resulting asymmetrical gene flow and different effective population sizes, populations occurring at the centre of the range are expected to harbour the highest levels of genetic diversity and lowest levels of genetic differentiation while populations occurring at the peripheries should exhibit the opposite. A special type of peripheral population is found in disjunct populations, or

those populations that are genetically isolated from their central counterparts. Here, alternative selective pressures and reduced levels gene flow from neighbouring populations can result in genetically distinct populations, thought to be important for the process of speciation (Davis and Shaw 2001).

Empirical evidence for this theory has been found in numerous conifers including pitch pine (*Pinus rigida* Mill) (Guries and Ledig 1982), lodgepole pine (*Pinus contorta* Douglas ex Loudon (Aitken and Libby 1994), Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) (Li and Adams 1989), and Sitka spruce (*Picea sitchensis* (Bong.) Carr) (Mimura and Aitken 2007). However, this is not the case for all. Norway spruce (*Picea abies* (L.) Karst (Muona *et al*. 1990), and red alder (*Alnus rubra* Bong.) (Hamann *et al*. 1998) both display high levels of variation throughout their respective ranges, perhaps owing to high levels of gene flow from the centre to the periphery of the species range. Both western redcedar (*Thuja plicata* D. Don) (O'Connell 2003) and bigleaf maple (*Acer macrophyllum* Pursh) (Iddrisu and Ritland 2004) show low levels of genetic diversity throughout their respective ranges, potentially owing to a bottleneck event in glacial refugia prior to recolonization. Extending beyond tree species, Eckert *et al*. (2008) found that 64.2% of 134 studied plant species showed reduced within population diversity towards range periphery, and 70.3% showed increased among-population differentiation towards the margins. However, it was also noted that these differences were small (Eckert *et al*. 2008).

Life history traits are known to influence the genetic structuring of populations. In a study comparing various methods and marker systems for estimating intraspecific genetic diversity in plants, Nybom (2004) found that estimates for expected (H_E) and

observed heterozyosity (H_O), as well as among-population differentiation (F_{ST}) (estimated from microsatellite markers) were significantly correlated with variables including life form, breeding system, sucessional status and seed dispersal mechanism ($p< 0.001$ for H_E and H_O, $0.05< p < 0.01$ for F_{ST}, for all variables). In addition, Nybom (2004) also discovered that plant species employing an 'ingested' mode of seed dispersal displayed the highest levels of diversity relative to other mechanisms of seed dispersal. This corroborates the findings of an earlier meta-analysis conducted by Hamrick *et al*. (1992), which reported that species whose seeds are ingested and dispersed by animals have the highest overall genetic diversity, and within population genetic diversity, and the lowest population differentiation, relative to species that use other methods of seed dispersal. Furthermore, Peterson and Denno (1998) found the phenomenon of isolation by distance (IBD) to be affected by dispersal ability in some taxa. IBD, first coined by Wright (1943), references the increasing genetic differences and decrease in gene flow between populations separated by increasing geographic distances. This trend is commonly observed in species demonstrating a central-peripheral structuring of genetic variation. Echoing this observation, Crispo and Hendry (2005) identified dispersal ability as a factor obscuring the logical relationship between IBD and time since colonization.

In Pacific dogwood, I hypothesized that high levels of interpopulation gene flow should mitigate the effect of genetic drift, resulting in similar levels of population genetic diversity across the species range, and low within and among-population genetic structuring. However, Idaho's disjunct populations may exhibit more genetic differentiation and less genetic variation than their coastal counterparts due to the limited opportunity for gene flow across the Cascade mountain range.

1.2.3 Quantitative traits

In addition to selectively neutral nuclear genetic markers, phenotypic variation of quantitative traits, which is subject to natural selection, was also used to investigate the genetics of populations. Under the assumptions of the 'abundant centre model', high levels of gene flow from the centre of the range to its peripheries are thought to limit adaptation of peripheral populations to new environmental conditions. This, in turn, is thought to check range expansion (Garcia-Ramos and Kirkpatrick 1997, Kirkpatrick, and Barton 1997, Bridle and Vines 2007). However, disjunct, or other genetically isolated populations may escape this effect, leading to local adaptation.

Common garden experiments have led to the characterization of substantial local adaptation in numerous conifers, including; Sitka spruce (*Picea sitchensis*) (Mimura and Aitken 2007), and whitebark pine (*Pinus albicaulis* Englem.) (Bower and Aitken 2008). These studies have found that traits associated with adaptation to cold (e.g. timing of bud set and bud flush) are often well-differentiated among populations and that genetic clines follow climatic gradients (see Table 1 from Howe *et al*. 2003). However the degree to which populations exhibit local adaptation appears to vary considerably between sympatric species (Howe *et al*. 2003).

1.3 Thesis objectives

Unpublished, previously observed low genetic diversity in *C. nuttallii* for isozymes (SJ Brunsfeld, U. Idaho, pers. comm.) and a range that appears climatically restricted led to the *a priori* hypothesis that this species may possess low levels of genetic variation. As genetic studies of such depauperate species are often hampered by a lack of polymorphic markers, we decided to employ multiple molecular tools as a means to understand the historic and contemporary evolutionary forces acting on Pacific dogwood. As so little genetic work has been done in this species, it was necessary to first develop a set of useful molecular markers for this species. In chapter 2 I address the following research questions; 1) What is the utility of 8 microsatellite markers, originally developed for *C. florida*, in *C. nuttallii*? 2) How polymorphic are the non-coding regions of the chloroplast genome of Pacific dogwood?

In chapter 3 I present the results from chloroplast sequence and nuclear microsatellite analyses, and compare these to preliminary results on population differentiation and local adaptation from a common garden experiment. These data were used to address the following research questions; 1) Where was the location(s) of glacial refugia and what was the post-glacial recolonization strategy for Pacific dogwood? 2) Are patterns of diversity, for both organelle and nuclear markers, similar to those of other Pacific coastal tree species, or is there evidence that animal dispersal has played an important role in shaping current patterns of variation? 3) Can Pacific dogwood be considered a genetically depauperate species and if so, to what extent? 4) How should conservation of this species proceed?

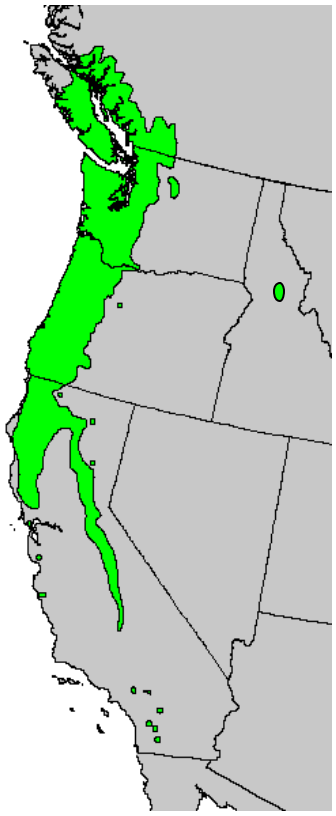

Figure 1.1. Range-wide distribution of Pacific dogwood (*Cornus nuttallii*), shown in green shading (Little 1976).

CHAPTER 2 - Chloroplast and microsatellite marker development in Pacific dogwood (*Cornus nuttallii*)

2.1 Introduction

Beginning a genetic investigation of a previously unstudied species can be a daunting task. However, there are a number of marker development options available for researchers assuming just such an endeavour.

Owing to their uniparental inheritance pattern, little or no recombination, and a different rate of evolution than nuclear DNA, haploid organelle markers are often the marker of first choice to resolve phylogenetic relationships among different taxa, reveal the dynamics of hybrid zones, and determine phylogeographical structure within species. These unique features, both of mitochondrial (mtDNA) and chloroplast (cpDNA) genomes, result in an effective population size that is ¼ that of nuclear markers (McCauley 1995). For this reason evidence of population structure should be stronger with organelle markers, providing valuable insight into the demographic history of individual species.

Although much of the gene content and order of the chloroplast genome is conserved among all land plants (Downie and Palmer 1992), phylogeographical studies of plant species exploit the higher levels of intraspecific variation found within the non-coding regions of the chloroplast genome. Universal primers have been designed to amplify theses regions (e.g. Taberlet *et al*. 1991) and several of these have been shown to be especially useful in population level studies in numerous plant taxa (Shaw *et al*. 2005).

Due to their high level of variability and codominant inheritance, microsatellites or simple sequence repeats (SSR) are useful in measuring genetic structure and diversity, and for determining relatedness among individuals in a population (Ritland 2000).

Microsatellite primers developed for a focal species can often be used to amplify such loci in closely related non-focal species, due to the presence of shared flanking regions (Zane *et al*. 2002). The practice of cross-species amplification for microsatellite loci is efficient as it renders unnecessary the laborious and costly procedures used to isolate microsatellite loci in previously unstudied species.

Pacific dogwood (*Cornus nuttallii* L.), named by John James Audubon (1780-1851) in honour of its collector, Thomas Nuttall (1786-1859), is a deciduous tree that grows along the Pacific coast of North America, from southwestern British Columbia to southern California and disjunctly in Northern Idaho (Little 1976). Currently the genetic structure and mating system of *C. nuttallii* is unknown, perhaps due to the lack of known variable genetic markers or low economic value in the wild.

Pacific dogwood is a close relative of a dogwood found in Eastern North America, flowering dogwood (*Cornus florida* L.) (a.k.a. eastern flowering dogwood, Florida dogwood). Evidence combining morphology, *matK*, ITS, *rbcL*, and 26S rDNA sequence data placed these two species in the same group, "the big-bracted dogwoods" and under the same subgenus, *Cynoxylon* (Xiang *et al*. 1996, Xiang *et al.* 2006). It is estimated that these two species diverged approximately 15 million years ago (J Xiang, North Carolina State U, pers. comm.).

Little chloroplast sequencing has been previously done in *Cornus nuttallii*. Similarly, microsatellite markers have only been developed for *Cornus florida* (Cabe and Liles 2002). For the purpose of a range-wide population level study of Pacific dogwood, we report the utility of seven non-coding chloroplast regions of *C. nuttallii* highlighted in

Shaw *et al*. (2005), as well 8 microsatellite loci, originally designed for *C. florida* (Cabe and Liles 2002).

2.2 Materials and methods

2.2.1 Sampling

Fresh foliage was collected from 595 individuals from 20 wild populations of *C. nuttallii* during April and May of 2006 and then stored at -80°C until DNA isolation. Four individuals/population from all 20 populations were randomly selected for microsatellite analysis. For chloroplast sequencing, two individuals/population from 3 populations were selected for preliminary sequencing. These individuals were chosen from populations at the geographic extremes of the species range (SB in the south, CL in the east, and PM in the north).

2.2.2 DNA extraction and amplification

Total genomic DNA from *C. nuttallii* was isolated using a modified CTAB method (Doyle and Doyle 1987). Primer sequences for microsatellite amplification were obtained from Cabes and Liles (2002). Additional microsatellite sequence data for *C. florida*, available through the GenBank database (AF387359, AF356102, AF356096), was also used in this study. For these loci, novel primers were designed by eye and also used to amplify loci in *C. nuttallii*. Primer sequences for chloroplast regions were obtained from Shaw *et al*. (2005). These loci were selected as they had been previously mapped to several different physical locations in the tobacco (*Nicotiana)* chloroplast genome (Wakasugi *et al*. 1998) and had demonstrated intraspecific variation in other asterids (Shaw *et al.* 2005).

Polymerase chain reactions (PCR) were performed to amplify microsatellite loci in *C. nuttallii.* Using a MJ Research PTC-100 thermal cycler (MJ Research, Inc.), 10 µL reactions contained 20-40 ng of total genomic DNA (Table 1), 1.0 µL of 2.0 mM dNTP, 1X *Taq* buffer (10 mM Tris, 1.5-2.5 mM $MgCl_2$ (Table 1), 50 mM KCl, pH 8.3) (Roche Inc.), 0.15 mM *Taq* DNA Polymerase (Roche Inc.), 0.5-0.8 pmol of M13 Infrared Label Primer (LiCor Inc.) (Table 1) and tailed primers (0.5 pmol each). Samples were amplified using the following PCR program: 2 min at 94°C, followed by 30 cycles of 30 s at 94°C, 45 s at the optimal annealing temperature (Table 2.1), 30 s at 72°C, and followed by an extension cycle of 5 min at 72°C. Amplification products were electrophoresed on 5% (Long Ranger™) polyacrylamide gels using a LiCor 4200 automated sequencer (LiCor Inc., Lincoln, NE). If scorable bands were not produced after 2 attempts, that particular individual was scored as 'missing data' for that locus.

For amplification of chloroplast regions, reactions contained 20 µL total volume, including 20-60 ng of total genomic DNA, 2.0 µL of 2.0 mM dNTP, 1X *Taq* buffer (10 mM Tris, 1.5-2.5 mM $MgCl_2$ (Table 1), 50 mM KCl, pH 8.3) (Roche Inc.), 0.15 mM *Taq* DNA Polymerase (Roche Inc.), and M13 tailed primers (10 pmol each). Samples were amplified using appropriate PCR programs described in Table 2.2. Amplified products were visualized on 2% agarose gels to ensure reactions were successful.

2.2.3 Sequencing

To confirm short sequence repeats in *Cornus nuttallii,* PCR products from two homozygous individuals were sequenced using SequiTherm EXCEL™ II Long-Read DNA Sequencing kits-LC (Epicentre Technologies) on a LiCor 4200 automated

sequencer (LiCor Inc., Linclon, NE). Chloroplast regions were sequenced in an identical manner.

2.3 Results

Five out of the eight published primer pairs originally developed for *C. florida* produced scorable bands in *C. nuttallii* (Cn-J7, G8, K2, N5, N10) (Table 2.1). Three additional loci were discovered as a result of unpublished sequence data for *C. florida* (Cn-N4, G13, G4). In total, 8 microsatellite loci were identified in *C. nuttallii.* Of these 8 loci, 5 demonstrated polymorphism across the range (Table 2.1).

Evidence of variation was seen in 5 of the 7 chloroplast regions (*rpS16, 5' rpS12-rpL20, psbB-psbH, rpL16,* and *rpS4R2-trnT^{UGU}(within 5'trnL^{UAA}- trnS^{UGA}*). To ensure variation was not a result of poor sequence quality, nested primers were designed with Primer 3 on the WWW (Rozen and Skaletsky 2000). These primers found only one region (*rpL16B*) to possess true variation in the form of a single base pair substitution. The region was subsequently sequenced in 300 individuals (15 individuals/population) in all 20 sampled populations at the Genome Sciences Centre (GSC) (CHAPTER 3). Sequence information for each region can be found in Appendix I.

2.4 Discussion

Despite the inclusion of individuals from the entire species range, *Cornus nuttallii* demonstrates lower levels of polymorphism at most published *Cornus florida* microsatellite markers (See Figure 2.1). During microsatellite development, loci selected are often biased towards those having the longest repeats, as these are most likely to have

high levels of polymorphism (Primmer *et al*. 1996). This bias is thought to be the reason why lower levels of polymorphism are observed at the same loci in closely related species (Zane *et al*. 2002).

During preliminary sequencing, 5547 base pairs (bp) of the Pacific dogwood's chloroplast genome were sequenced. Although no published size for this species' chloroplast genome exists to date, it is likely that between 7.6 and 10.8% of the non-coding elements were sequenced in this study. Despite our efforts, only 2 cpDNA haplotypes were identified. These haplotypes were distinguished by a single base pair mutation occurring 881bp from the 5' end of the *rpL16* region.

The results from this chapter indicate that Pacific dogwood possesses low genetic diversity at both nuclear microsatellite and chloroplast loci throughout its range. Genetic depauperateness, which is ubiquitous through a species range, is often indicative of a bottleneck suffered in glacial refugium prior to recolonization. This idea will be discussed in greater detail in CHAPTER 3.

Table 2.1 Characterization of 8 microsatellite loci in *Cornus nuttallii*.

Locus	Primer Sequence (5'- 3')	T_a (^oC)	DNA (ng)	M13 (µL)	[Mg] (mM)	N	Repeat Motif	Fragment size range (bp)	Alleles (n)	GenBank Accession no.
Cn-J7[2]	AACACTGCCCCATTGTTAGAG GAGGTGTCTCTCTCGTGGTTC[1]	55	20	0.5	1.5	80	$(TC)_{14}(AC)_{10}$	127-133	4	DQ223104
Cn-G8[2]	GCTGGTTGAATTATTTGAGG[1] GGGATTAAAAGAAAGATGACG	53	40	0.8	1.5	74	$(CA)_2(CT)_{11}$	150	1	DQ223105
Cn-K2[2]	GGGAGCGAGATCTCAAAGG[1] TTTCAGAGCATTTTGAATGAGG[1]	53	40	0.5	2.5	80	$(GA)_8$	104	1	DQ223106
Cn-N5[2]	GCAAGGATCGAACTTAGGG TGAATTATGTATCGAATGTCTGC[1]	59	20	0.5	1.5	53	$(AT)_{11}$	124-128	3	DQ223107
Cn-N10[2]	TGATTGAATAACCTTTTGATGC[1] GGTAGCTTCAAATGTCAACG	55	40	0.5	1.5	56	$(TA)_{25}(TG)_{12}$	177-211	9	DQ223108
Cn-N4[3]	TGGCCTTTGGAGAGGGAATGCA CATTCCCAGATGTTTGGTCTATC[1]	62	20	0.5	2.5	73	$(GA)_{11}$	251-255	3	DQ223110
Cn-G13[3]	CTCCGTCTATTCTTGAGC TCTAAGAGGTTTGATGGC[1]	48	20	0.8	1.5	75	$(TG)_{14}$	145	1	DQ223111
Cn-G4[3]	TCATGCCCCGCTAAACCGAC[1] CATCACTGTACTCAGGCC	49	20	0.5	1.5	80	$(TC)_{11}$	133-137	3	DQ223112

[1]Primer to which a forward or reverse tail sequence was used.
[2]From Cabe and Liles (2002)
[3]From designed based on Genbank sequences
T_a, annealing temperature; *n*, total number of individuals; H_O, observed heterozygosity; H_E, expected heterozygosity.

Table 2.2. Summary of chloroplast primers used in preliminary screening.

Region of chloroplast	Primer Pair	PCR Program	Length of fragment in *C. nuttallii* (bp)	Reference(s)
trnH^{GUG}-psbA	**trnH^{GUG}** CGC GCA TGG TGG ATT CAC AAT CC **psbA** GTT ATG CAT GAA CGT AAT GCT C	80°C, 5 min; 35×(94°C, 30s; 53 °C 30s; 72 °C, 1 min) 72°C 10 min	535	Tate and Simpson (2003) Sang et al. (1997)
rpS16	**rpS16F** AAA CGA TGT GGT ARA AAG CAA C **rpS16R** AAC ATC WAT TGC AAS GAT TCG ATA	80°C, 5 min; 35×(94°C, 30s; 53 °C 30s; 72 °C, 1 min) 72°C 5 min	925	Oxelman et al. (1997)
5' rpS12-rpL20	**5'rpS12** ATT AGA AAN RCA AGA CAG CCA AT **rpL20** CGT TAT CGA GCT ATA TAT CC	96°C, 5 min; 35×(96°C, 1 min; 53 °C 1 min; 72 °C, 1 min) 72°C 5 min	910	Hamilton (1999)
psbB-psbH	**psbB** TCC AAA AAN KKG GAG ATC CAA C **psbH** TCA AYR GT Y TGT GTA GCC AT	80°C, 5 min; 35×(94°C, 30s; 58 °C 30s; 72 °C, 1 min) 72°C 5 min	627	Hamilton (1999)
rpL16	**rpL16F71** GCT ATG CTT AGT GTG TGA CTC GTT G **rpL16R1516** CCC TTC ATT CTT CCT CTA TGT TG	80°C, 5 min; 35×(95°C, 1 min; 50 °C 1 min with a ramp of 0.3°C/s, 65 °C, 5 min) 65 °C 4 min	1030	Small et al. (1998)
5'trnL^{UAA}-trnT^{UGU}(within 5'trnL^{UAA}-trnS^{UGA})	**5'trnLR (Tab B)** TCT ACC GAT TTC GCC ATA TC **trnTF (Tab A)** CAT TAC AAA TGC GAT GCT CT	96°C, 5 min; 35×(96°C, 1 min; 55 °C 2 min; 72 °C, 2.5 min) 72°C 5 min	920	Taberlet et al. (1991)
rpS4R2-trnT^{UGU}(within 5'trnL^{UAA}-trnS^{UGA})	**rpS4R2** CTG TNA GWC CRT AAT GAA AAC G **trnTR** AGG TTA GAG CAT CGC ATT TG	96°C, 5 min; 35×(96°C, 1 min; 55 °C 2 min; 72 °C, 2.5 min) 72°C 5 min	600	Taberlet et al. (1991)

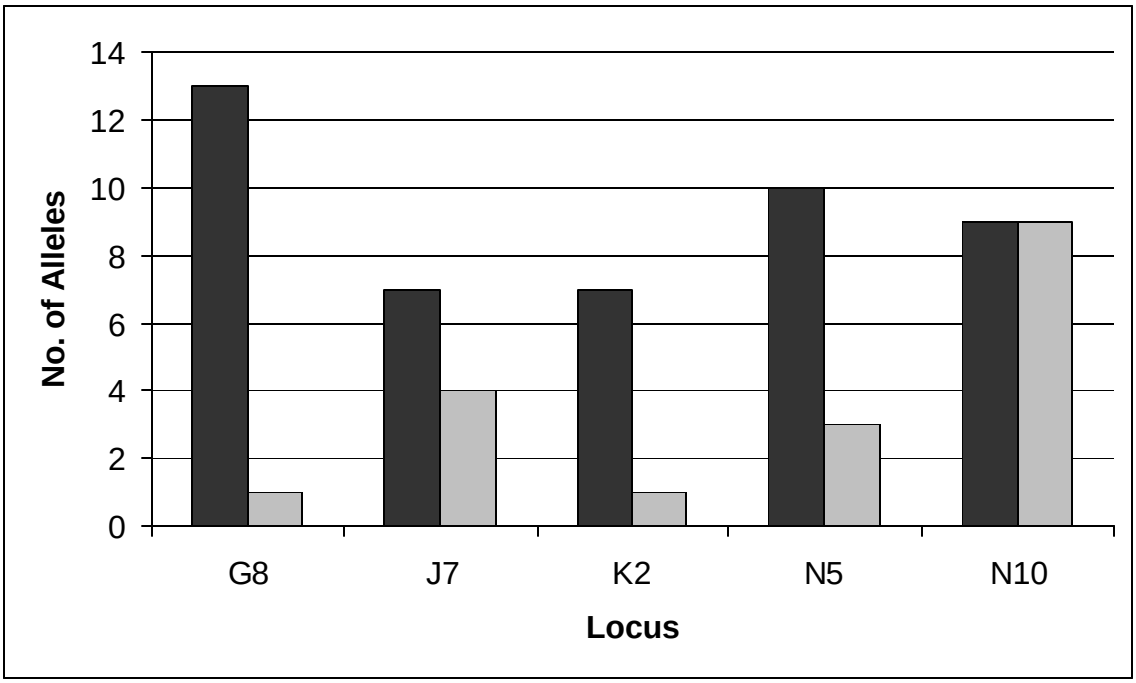

Figure 2.1. Comparison of the number of alleles discovered at identical microsatellite loci in *Cornus florida* (n = 18) (data from Cabe and Liles (2002)) (shown in dark grey), and *Cornus nuttallii* (n = 53-80) (shown in light grey).

CHAPTER 3 - Population genetics of Pacific dogwood

3.1 Introduction

Recurrent regional patterns of phylogeographic structure in sympatric plant and animal species are often observed (Europe; Taberlet *et al*. 1998, the Pacific Northwest of North America; Soltis *et al*. 1997, Brunsfeld *et al*. 2001, Eastern North America; Avise 2000, Soltis *et al*. 2006) and have led to the formal proposition of numerous glacial refugia and post-glacial migration routes. Research on western North American plant species using chloroplast DNA (cpDNA), which is most commonly maternally inherited in angiosperms, has revealed Pleistocene glaciation to be a major force affecting the population genetic structure of these species (Soltis *et al*. 1997, Brunsfeld *et al*. 2001). Numerous co-distributed plant species (including herbaceous perennials, shrubs and trees) have revealed a recurrent phylogeographic pattern, represented as a north-south partitioning of haplotypes (Soltis *et al*. 1997, Brunsfeld *et al*. 2001). The common delineation of these two clades occurs in southern to central Oregon, at a boundary referred to as 'the Soltis line' (Brunsfeld *et al*. 2007). The pattern is thought to emerge as a result of recolonization from either multiple refugia ("north-south" recolonization hypothesis) or from a single southern refugium ("leading edge hypothesis") (Soltis *et al*. 1997). Furthermore, studies of species from this region have led researchers to formalize hypotheses relating to the origin of disjunct mesic forests found in northern Idaho (Brunsfeld *et al*. 2001, Carstens *et al*. 2005, Brunsfeld and Sullivan 2005, Brunsfeld *et al*. 2007). These unique disjunct populations are thought to have originated either through ancient vicariance (e.g. uprising of the Cascade Range in the late Pliocene), or by more

recent dispersal via a northern or southern route (see Brunsfeld *et al*. 2001 for more detailed description of hypotheses).

In addition, numerous Pacific coastal tree species have supported predictions of genetic structure under the abundant centre model, including lodgepole pine (*Pinus contorta* Douglas ex Loudon) (Aitken and Libby 1994), Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) (Li and Adams 1989), and Sitka spruce (*Picea sitchensis* (Bong.) Carr) (Mimura and Aitken 2007). However, this is not the case for all. Red alder (*Alnus rubra* Bong) (Hamann *et al*. 1998) displays high levels of variation throughout its range, perhaps owing to high levels of gene flow from the centre to the periphery of the species range (Barton 2001). Both western redcedar (*Thuja plicata* D. Don) (O'Connell 2003) and bigleaf maple (*Acer macrophyllum* Pursh) (Iddrisu and Ritland 2004) showed low levels of genetic diversity throughout their respective ranges, potentially owing to a bottleneck event in glacial refugia. In a meta-analysis, Eckert *et al*. (2008) found that 64.2% of 134 studied plant species showed reduced within population diversity towards range periphery, and 70.3% showed increased among-population differentiation towards the margins. However, it was also noted that these differences were small (Eckert *et al*. 2008). Furthermore, analysis of published genetic diversity statistics in central, marginal, and disjunct tree populations found significant variation with population position (Aitken and Fady, unpubl. data). Common garden experiments have led to the discovery of local adaptation in many western North America tree species (Howe *et al*. 2003, Savolainen *et al*. 2007) including Sitka spruce (*Picea sitchensis* (Bong.)) (Mimura and Aitken 2007), and whitebark pine (*Pinus albicaulis* Engelm.) (Bower and Aitken 2008).

Despite the multitude of forest genetic studies in this region, there is a decided under-representation of tree species that employ biological vectors for both pollination and seed dispersal. The commercial value of conifers in this region, which often utilize wind for both pollen and seed dispersal, have made these trees the primary focus of forest genetic studies. The study of species employing animal pollination and seed dispersal mechanisms may broaden our understanding of the full range of tree evolutionary dynamics.

There is a growing body of literature suggesting that dispersal distance may strongly influence the genetic structure of populations. Using simulations, Irwin (2002) showed the likelihood of observing strong phylogeographic structure increases as average dispersal distance decreases. Similarly, dispersal abilities are known to influence isolation by distance (IBD) in some taxa (Peterson and Denno 1998). IBD, first coined by Wright (1943), references the increase in genetic differences and reduction in gene flow between populations separated by increasing geographic distances. This trend is commonly observed in species demonstrating a central-peripheral structuring of genetic variation, or those that show declining levels of variation from south to north. In addition, Crispo and Hendry (2005) identified dispersal ability as a factor obscuring the expected relationship between IBD and time since colonization.

In a study comparing various methods and nuclear marker systems for estimating intraspecific genetic diversity in plants, Nybom (2004) found that estimates for expected (H_E) and observed heterozyosity (H_O), as well as among-population differentiation (F_{ST}) were significantly correlated with seed dispersal mechanism (gravity, attached, wind/water, ingested). In addition, Nybom (2004) also discovered that

plant species employing an 'ingested' mode of seed dispersal displayed the highest levels of overall diversity, relative to those species utilizing other mechanisms for seed dispersal. This corroborates the findings of an earlier meta-analysis by Hamrick *et al*. (1992), which reported that species whose seeds are ingested and dispersed by animals have the highest genetic diversity both at the species level and within populations, as well as the lowest population differentiation (F_{ST}) relative to species with other methods of seed dispersal.

Pacific dogwood (*Cornus nuttallii* Audubon ex Torr. & A. Gray) represents an excellent species to add breadth to this burgeoning body of research. Revered for its beauty as an ornamental, this broadleaf tree species is endemic to western North America. Here, *C. nuttallii*'s distribution extends from the lowlands of southwestern British Columbia to the mountains of southern California (Little 1976). This species is also found disjuctly in northern Idaho (Little 1976). Other, smaller disjunct populations of this tree species can be found in the transverse ranges of southern California, including the San Bernardino Mountains. Finally, unlike many sympatric tree species studied thus far, gene flow is accomplished by way of insects, which pollinate *C. nuttallii*'s flowers, and birds and mammals that eat and disperse its seeds (Klinka 2000). Although yet to be tested explicitly, chloroplasts are assumed to be maternally inherited in *C. nuttallii*, similar to most angiosperms (Corriveau and Coleman 1988).

Eastern flowering dogwood (*Cornus florida* L.) (a.k.a. Florida dogwood) is Pacific dogwood's closest relative. These two species have been placed within the same lineage, the big-bracted dogwoods (Xiang et al. 1996; Xiang *et al*. 2006) and are estimated to have diverged approximately 15 million years ago (J Xiang, North Carolina

State U, pers. comm.). Microsatellite markers, developed for *C. florida* (Cabe and Liles 2002) have been shown to amplify polymorphic short sequence repeats (SSR) loci in *C. nuttallii* (CHAPTER 2).

Low genetic diversity previously observed in *C. nuttallii* for isozymes (SJ Brunsfeld, U. Idaho, pers. comm.) and a range that appears climatically restricted led to the hypothesis that this species may have endured a bottleneck prior to recolonization and therefore may presently possess low levels of genetic variation. As genetic studies of such depauperate species are often hampered by the lack of variation, I decided to employ both chloroplast and microsatellite markers as a means to understand the historic and contemporary evolutionary forces acting on Pacific dogwood. Furthermore, I felt such an approach was necessary as no previous results have been published on the genetics of Pacific dogwood.

Here I present the results from chloroplast sequence and nuclear microsatellite analyses, and compare those to preliminary results on population differentiation for phenotypic traits from a common garden experiment. These data were used to address the following questions: 1) Where were glacial refugia located and what was the post-glacial recolonization route of Pacific dogwood? 2) Are patterns of diversity, both organelle and nuclear, similar to those of other Pacific coastal species (e.g. showing differentiation across the Soltis line as well as central-peripheral structure), or is there evidence that animal seed dispersal has played an important role in shaping current patterns of variation? 3) Is Pacific dogwood genetically depauperate? 4) What are the implications of these results for management, restoration or conservation of this species?

3.2 Materials and methods

3.2.1 *Sampling locations and techniques*

Fresh foliage was sampled from 595 individuals in total from 20 native populations of *C. nuttallii* during April and May of 2006 (Table 3.1, Figure 3.1). These populations span the entire geographic range of the species, and include sites near all proposed glacial refugia, the Olympic Pennisula (OP), northern Vancouver Island (BT), northern Idaho (CL) and several near the southern Oregon/northern California coastline. Two disjunct populations were also sampled: one at the southern limit of the species range in the San Bernardino Mountains (SB), and the eastern species limit along the north branch of the Lochsa River in northern Idaho (CL). Populations were spaced a minimum of 75km apart; however, most were separated by much larger distances. Sites sampled in the United States were restricted to United States Forest Service National Forests. They were identified using the Pacific Northwest Forest Inventory and Analysis database (Waddell and Hiserote 2005) and through personal communication with numerous employees of the United States Forest Service. Because *C. nuttallii* grows at relatively low densities, the sampling strategy employed was opportunistic. However, 26 to 31 individual trees, spaced a minimum of 30m apart from one another, were sampled from each population. Longitude, latitude and elevation were recorded for each sampled tree (Table 3.1). After harvesting, the leaves were frozen in liquid nitrogen, and then stored at -80°C. All 595 sampled individuals were used in the microsatellite study, while a subsample of 15 individuals/population was randomly selected for chloroplast sequencing.

Seeds were collected from 164 individuals (not necessarily the same trees sampled in the spring) from 11 of the 20 populations in the fall of 2006 (Table 3.1). A minimum of 10 trees/population, spaced 30m or more apart, had their single seeded fruit picked from multiple inflorescences and stored in breathable mesh bags. Similar to spring collections, these populations were selected to encompass the entire species range and included both disjunct populations.

3.2.2 *Chloroplast sequencing*

Total genomic DNA from *C. nuttallii* leaf samples was isolated using a modified CTAB method (Doyle and Doyle 1987). Universal primers were employed to amplify seven different regions of the chloroplast genome for preliminary sequencing (CHAPTER 2). These regions were selected as they had been previously mapped to several different physical locations in the tobacco (*Nicotiana)* chloroplast genome (Wakasugi *et al*. 1998) and had demonstrated intraspecific variation in other asterids (Shaw *et al*. 2005). The chloroplast genome of most angiosperms ranges from 120 to 170 kilobase pairs in length (Downie and Palmer 1992). Among most land plants there is a relatively high degree of conservation in the size, structure, gene content and order in these genomes (Downie and Palmer 1992). However, it is estimated that 43% (10.6% introns and 32.3% intergenic spacers) of the chloroplast genome is noncoding (Wakasugi *et al.* 1998) and these regions are likely to be more polymorphic than coding regions, and useful for phylogeographical studies. During preliminary sequencing, 5547 base pairs (bp) of the Pacific dogwood's chloroplast genome were sequenced. Although no estimates of the size of this species' chloroplast genome have been published, based on the size range for angiosperms in general, it is likely that between 7.6 and 10.8% of the

noncoding elements were sequenced in this study. Nested primers, designed to amplify regions showing putative variation in preliminary sequencing, dropped this percentage of sequenced cpDNA to between and 3.1% and 4.4%.

Preliminary sequences were obtained using SequiTherm EXCEL™ II Long-Read DNA Sequencing kits-LC (Epicentre Technologies), using PCR programs outlined in CHAPTER 2, on a LiCor 4200 automated sequencer (LiCor Inc., Linclon, NE). Putative variation was detected in 5 of the 7 regions (*rpS16, 5' rpS12-rpL20, psbB-psbH, rpL16,* and *rpS4R2-trnT^{UGU}(within 5'trnL^{UAA}- trnS^{UGA}*). Nested primers were designed with the internet version of Primer 3 (Rozen and Skaletsky 2000), using data from preliminary sequences, to amplify suspected variable regions (See Table 3.2). PCR programs for nested primers were identical to those used to amplify corresponding larger regions. These regions were sequenced in 100 randomly selected trees (10 individuals/population) from 10 populations, covering the entire range of *C. nuttallii* at the Genome Sciences Centre (GSC) in Vancouver, B.C., Canada. Only one region (*rpL16B*) showed true variation in the form of a single base pair substitution. The region was subsequently sequenced in a total of 300 individuals (15 individuals/population) for all 20 sampled populations at the same facility.

Trace files were viewed and quality judged using FinchTV (Geospiza ®). Sequences were aligned using Molecular Evolutionary Genetic Analysis (MEGA) 4 (Tamura *et al.* 2007). Overall haplotype diversity (h) and unbiased Nei's pairwise population genetic distances (D) were calculated using Genetic Analysis in Excel Version 6.1 (GenAlEx) (Peakall and Smouse 2006). A Mantel test was used to test for isolation-by-distance (IBD), by estimating relationships between Nei's pairwise unbiased genetic

distances (D) and corresponding pairwise geographic distances between populations in km (estimated from longitude and latitude coordinates for populations using GenAlEx).

3.2.3 Microsatellite genotyping

Preliminary work using primers designed for *C. florida*, led to the discovery of five polymorphic microsatellite loci, which produced scorable bands in all individuals (Cn-G8, Cn-J7, Cn-N4, Cn-N5, and Cn-N10) (CHAPTER 2). For this thesis, only two polymorphic loci (Cn-G8 and Cn-J7) were genotyped for all individuals due to technical difficulties and time constraints. These data are included in this analysis.

Polymerase chain reactions (PCR) are described in CHAPTER 2, and are summarized here briefly. Each 10 µL reaction contained 20-40 ng of total genomic DNA, 1.0 µL of 2.0 mM dNTP, 1X *Taq* buffer (Roche Inc.), 0.15 mM *Taq* DNA Polymerase (Roche Inc.), 0.5-0.8 pmol of M13 Infared Label Primer (LiCor Inc.) and M13 tailed primers (0.5 pmol each). Samples were amplified using PCR programs outlined in CHAPTER 2.

Amplification products were electrophoresed on 5% (Long Ranger™) polyacrylamide gels using a LiCor 4200 automated sequencer (LiCor Inc., Lincoln, NE). Bands were scored using Gene ImagIR™ RFLP with 100bp ladders. Standard measures of genetic diversity, average number of alleles per locus (also called allelic richness, A_R), observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}) were estimated for each population using Genetic Analysis in Excel Version 6.1 (GenAlEx) software (Peakall and Smouse 2006). Measures of genetic distance, F_{ST} (Weir and Cockerham 1983) and the more microsatellite specific R_{ST} (Slatkin 1995), were also calculated using this software. Mantel tests were used to test for isolation by

distance, estimating the relationship between both F_{ST} and R_{ST} with geographic distance. Linear regressions were performed on alleles per locus, observed heterozygosity, and expected heterozygosity versus latitude. Each of the response variables was squared to meet regression assumptions.

3.2.4 Common garden

All fruits were soaked in concentrated sulphuric acid for 2 hours and then packaged in screen bags. These 'seed bags' were then immersed in running water for 48 hours and transferred to moistened peat at room temperature for 30 days. Following warm stratification, seeds were cold stratified at 4°C for 30 days. Cold stratification was shorter than recommended by Paulus Vrigmoed of Linnaea Nurseries Ltd in Langley B.C., owing to the fact that some seeds had begun to germinate. In January of 2007, 5407 seeds from 11 populations were sown two per cone in research greenhouse facilities at the University of British Columbia. Although seeds were heavily consumed by a rodent, 940 surviving seedlings, including representatives from all populations, were transferred to outside raised beds in June of 2007 in a randomized block design with 12 bocks and single-tree plots. Height and bud burst time data was collected in spring of 2008.

Results from molecular marker analyses for the 11 populations in the common garden were compared with genetic clines and Q_{ST} estimates for first year height and bud burst time from analyses by undergraduate student Jordan Bemmels. By comparing estimates of population genetic differentiation calculated with selectively neutral molecular markers (F_{ST}) and those calculated using quantitative trait variation (Q_{ST}), the extent of differential selection among populations can be evaluated (Merilä and Crnokrak

2001, McKay and Latta 2002). However, caution needs to be exercised when comparing these two values as they may have different distributions, and statistical tests explicitly testing differences are not straight forward (Whitlock 2008). Patterns of diversity observed in all marker systems (cpDNA, microsatellite, quantitative traits) were compared with geographic and climatic data to evaluate post-glacial evolutionary dynamics.

3.3 Results

3.3.1 Chloroplast sequence analysis

Despite sequencing three to five percent of the non-coding regions of the chloroplast genome of Pacific dogwood in 100 individuals, only two haplotypes were identified in this study. These haplotypes were distinguished by a single base pair (A or G) mutation 881 basepairs from the 5' end of the sequenced *rpL16* region. The more common of the two haplotypes, (1), occurs at a frequency of 0.735 in the species, while haplotype 2 has a global frequency of 0.265 across all individuals sequenced. Both haplotypes are present in all but four of the 20 populations (Figure 3.1, Table 3.3), three in California (SQ, ST, and SH) and the other in Idaho (CL). Only haplotype 1 was present in these four populations.

Total unbiased haplotype diversity (h_T) was 0.316 (se ± 0.046). Overall population genetic structure (Nei's unbiased genetic distance, *D*) averaged 0.153. A Mantel test of isolation by distance failed to show a correlation between geographic distance and Nei's unbiased genetic distance (r = 0.119, p = 0.210).

When northern populations are separated from southern populations along the hypothetical "Soltis Line" in central Oregon (Brunsfeld *et al*. 2007) (see Table 3.3 for groupings), including the Idaho population in the southern clade, some structure is observed. The southern region has frequencies of 0.88 (0.53-1.0) and 0.12 (0-0.47) for haplotype 1 and 2 respectively, with four of the 10 populations possessing only haplotype 1. The northern region shows much less variation among populations with both haplotypes present in all populations and occurring at a frequency of 0.59 (0.33-0.87) for haplotype 1 and 0.41 (0.13-0.67) for haplotype 2.

In the southern group, haplotype 2 is uncommon or absent in most populations, with the exception of the most southerly population. This small population occurs in the San Bernardino Mountains and is disjunct from the large continuous portion of the species range. Here, haplotype 1 is present at a frequency of 0.53, and haplotype 2 at 0.47.

3.3.2 Microsatellite diversity and population differentiation

A total of seven alleles were detected at the two microsatellite loci across all populations, three at Cn-G4 and four at Cn-J7. Two of the three alleles at the Cn-G4 locus were found in all populations. Only 1 allele at the Cn-J7 locus was shared among all populations (See Table 3.4 for allele frequencies). In general, southern populations were most likely to possess all alleles. Allelic richness (A_R), observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (F) over all populations and loci averaged 2.475 (se ± 0.124), 0.240 (se ± 0.022), 0.353 (se ± 0.023), and 0.308 (se ± 0.047) respectively (Table 3.5). There was a significant trend of decreasing allelic richness (A_R) ($r^2 = 0.42$, $p<0.01$), and expected heterozygosity (H_E) ($r^2 = 0.51$, $p<0.01$), from south to north (See figure 3.2, and 3.4 respectively). Although this general trend was also seen with observed heterozygosity (H_O) ($r^2 = 0.18$, $p = 0.0642$), the linear regression was not significant.

Genetic distances among populations were small, averaging 0.071 for pairwise F_{ST} estimates and 0.036 for R_{ST} pairwise estimates. No significant positive relationship was found between either estimate of pairwise genetic distance and geographical distance (Mantel test: F_{ST}; $R = 0.102$, $p = 0.14$, R_{ST}; $R = -0.201$, $p < 0.05$).

3.3.3 Common garden experiment

Significant ($p<0.05$) linear regressions of population mean quantitative traits on climatic variables for provenances included 1st year height and frost free period ($r^2 = 0.52$), as well as timing of budburst and mean coldest month temperature ($r^2 = 0.47$) and mean summer precipitation ($r^2 = 0.46$).

Q_{ST} was estimated to be 0.088 for first-year height and 0.113 for bud burst timing. Narrow sense heritability (h^2) for these traits was estimated be 0.24 and 0.38 respectively, assuming a coefficient of relatedness between siblings within open pollinated families of 1/3. Although the mating system of is currently unknown, it is suspected that open pollinated Pacific dogwood progeny are more closely related than half-sibs owing to moderate inbreeding. Both refined pairwise estimates of F_{ST} (0.053) and R_{ST} (0.048), including only those populations present in common garden experiment, were found to be lower than estimates of Q_{ST}.

3.4 Discussion

*3.4.1 Post-glacial recolonization strategy of Pacific dogwood (*Cornus nuttallii*)*

Despite the lack of diversity for cpDNA found in *C. nuttallii*, I believe the phylogeographic pattern provides support for a single southern refugium and a post-glacial migration strategy similar to the 'leading edge hypothesis'. This hypothesis was originally proposed by Cwynar and MacDonald (1987) and then applied to Pacific Northwest species by Hewitt (1993), Allen *et al*. (1996) and Soltis *et al*. (1997). This hypothesis suggests that populations persisted only south of the glacial maximum during the last ice age. Following glacial retreat, populations expanded north by long distance dispersal facilitated by individuals along the leading edge. The stochastic processes during such a range expansion event could result in a decrease of genetic variation in newly founded northern populations and potentially give rise to distinct haplotypes.

Relatively wet conditions are hypothesized to have prevailed along the coast during the last glacial maximum, creating a large regufium for mesic temperate forests south of glaciation (Brunsfeld *et al*. 2001). Further paleoecological evidence supports a refugium along the coast in the Siskiyou-Klamath mountains (Whittaker 1961, Smith and Sawyer 1988) during the last glacial maximum. However, the significant decline of allelic richness and expected heterozygosity observed south to north at microsatellite loci may indicate that the refugium was further south. Presently, three other polymorphic microsatellite loci are being genotyped in all 595 individuals for a more robust and accurate survey of nuclear variation (K Keir, in progress).

Soltis *et al*. (1997) hypothesized that a 'leading edge' recolonization process would lead to the fixation of a single haplotype in newly founded populations as a

consequence of long distance dispersal and drift. As the glaciers continued their retreat, this haplotype would subsequently spread north, filling available habitats as they became available, and restricting the spread of other haplotypes. In contrast, results from this study indicate that both Pacific dogwood haplotypes successfully, and perhaps simultaneously, migrated north out of the proposed southern refugium. This result may reflect the long distance, northern dispersal patterns of Pacific dogwood seeds at the time of recolonization. However the low levels of cpDNA haplotype diversity found in this species may also indicate that recolonization occurred relatively recently.

Although this may explain the pattern observed north of the "Soltis Line", a different pattern emerges south of the proposed refugium, with the near fixation of haplotype 1 in most populations sampled in California. As no mesic forest refugium has been identified in the Sierras or southern Cascades from fossil data (Brunsfeld *et al.* 2001) these populations likely represent post-glacial dispersal events. This result may indicate that these high elevation populations are more susceptible to the effects of genetic drift and inbreeding, as a result of small population sizes and genetic isolation from neighbouring populations. Finally, results support the hypothesis that Pacific dogwood populations in the disjunct mesic forests of northern Idaho were established via a recent inland dispersal event (Brunsfeld *et al.* 2001, Carstens *et al.* 2005). However, it is difficult to determine if this dispersal event occurred by a southern or northern route, as haplotype 1, the type for which the Idaho population is fixed, is found in all northern and southern coastal populations.

Although these results have provided some insight into the post-glacial recolonization pattern of *C. nuttallii*, further insight would be gained with increased

haplotypic diversity to analyze. Low levels of variation are a common problem in phylogeographical studies utilizing intraspecific cpDNA variation (Schaal *et al.* 1998). However, variation has recently been discovered in five regions of the chloroplast genome for *Cornus florida*, *C. nuttallii*'s closest relative, including two regions that were found to be fixed in *C. nuttallii* (*trnL-trnT* and *rpS16*) in this study (A Brooks, North Carolina State U, pers. comm.). Although our search for variation was extensive, we acknowledge the likelihood that variation exists outside of the sequenced regions. Therefore, as sequencing technology continues to advance, making whole chloroplast sequencing a viable option for phylogeographic studies, it would be advisable to continue the search for haplotypic variation in this species. This would allow for more accurate identification of suspected refugia, regions of high and low diversity, and in turn, might have implications in the realm of conservation.

3.4.2 Factors affecting contemporary patterns of diversity

Although there is no doubt that Pleistocene glaciation contributed to the contemporary pattern of variation, the phylogeographic pattern displayed by cpDNA haplotypes for Pacific dogwood indicates the role of other forces. Furthermore, although microsatellite data showed a general trend of decreasing genetic diversity from south to north, the differences between populations were small. Thus, there is only weak evidence to suggest that populations occurring at the northern range margin possess less genetic diversity than those populations occurring in the central and southern portions of Pacific dogwood's range. Finally, although the disjunct population occurring in northern Idaho displayed relatively low levels of genetic diversity as predicted, the isolated population sampled in southern California revealed the opposite trend.

Unlike many PNW species studied thus far, Pacific dogwood seeds are dispersed by many animals, including birds that have the potential to transport seeds over great distances. Therefore, it seems likely that such dispersal capabilities have influenced current patterns of genetic variation.

Frequent, long distance dispersal events are often offered as explanations for unusual or weak phylogeographic structure. Other Pacific Northwest plant species which deviate from the expected pattern include *Heuchera micrantha* (crevice alumroot), and to a lesser degree, *Polystichum munitum* (sword fern). Long distance seed or spore dispersal events were cited as a potential cause for the lack of structure in both of these species (Soltis *et al*. 1989, Soltis *et al.* 1997). In eastern North American *Prunus* species, Shaw and Small (2005) hypothesized that historical factors (e.g., refugial zones) and contemporary forces (e.g., dispersal of plum seeds by humans over large distances) acted in concert to create the unique phylogeographic pattern observed that is dissimilar to other phylogeographical patterns commonly observed in eastern North America (Soltis *et al.* 2006). Furthermore, the hypothesized long-distance dispersal capabilities of the now extinct passenger pigeon were cited as a potential reason for the weak phylogeographical structure observed in *Quercus rubra* L. (northern red oak) (Magni *et al.* 2005).

In support of these authors' conclusions, the simulation studies of Irwin (2002) showed that phylogeographical discontinuities could be observed in the absence of a geographic barrier when both population size and dispersal distance were relatively small. It is therefore understandable that some species may not show phylogeographic patterns in the same region, especially those species that have large dispersal distances or large population sizes.

Although it is unknown how far the average *C. nuttallii* seed travels, it is recognized that the red drupes, which ripen in the fall, have the potential to attract numerous migrating birds, including American robins (*Turdus migratorius*) (Willson 1994). Seeds dispersed by migrating birds are thought to result in greater than average dispersal distances (e.g., passenger pigeon, Webb 1986). Potential evidence for this type of dispersal can be observed in the cpDNA haplotype frequencies found in the most southern population sampled in the San Bernardino Mountains. Here, unlike neighbouring populations in the Sierra Nevadas, both haplotypes were found in near equal frequencies. Regions with high haplotypic diversity are often thought to have retained historically accumulated variation (i.e., while in glacial refugia) or alternatively, have more recently acquired variation through secondary admixture via dispersal (Petit *et al.* 2003, Taberlet *et al.* 1998). As there is no evidence to suggest the San Bernardino mountain range was a glacial refugium, and American robins are known to migrate from Vancouver Island and winter in the mountains of the southwestern United States (Small 1994), including the mountains in southeastern California, it is possible that over many migrations, these birds may have introduced both haplotypes to this population in a stepping stone fashion (See Figure 3.1 for an overlay of American robin distribution and chloroplast results). Furthermore, this population shows higher than average levels of allelic richness (A_R), expected (H_E) and observed heterozygosity (H_O) (See Table 3.5), the opposite of what would be expected under the abundant centre model.

The low estimates of population differentiation generated from microsatellite data provide further evidence that frequent long distance seed dispersal continues to shape the population genetic structure of Pacific dogwood. Both R_{ST} and F_{ST} estimates were low,

0.036 and 0.071 respectively, compared with other long-lived perennials (F_{ST} = 0.19) or other plants with a similar mechanism for seed dispersal (ingested, F_{ST} = 0.21) (Nybom 2004). The lack of IBD, appears to be the product of Pacific dogwood's seed dispersal ability, as has been suggested for some other taxa (Peterson and Denno 1998).

Another factor affecting the current levels of genetic diversity in populations of *C. nuttallii*, particularly those populations occurring at higher elevations in California, could be geographic isolation and founder effects. Populations sampled in California occurred at an average elevation of 1070 metres above sea level (Table 3.1) and were often found to be fixed or nearly fixed for haplotype one. Similarly, low cpSSR haplotype gene diversity was reported in Sierra Nevadan populations of whitebark pine (*Pinus albicaulis* Engelm.) when compared with more northern populations (Richardson *et al.* 2002). These populations of *P. albicaulis* were thought to represent contemporary refugia, threatened by warming climates (Richardson *et al*. 2002).

This decrease in diversity was not observed at nuclear microsatellite loci in Pacific dogwood. Instead these high elevation populations displayed high levels for both allelic richness and genetic diversity. As nuclear markers are known to have a four fold greater effective population size than that of haploid organelle markers, diminishing the effects of genetic drift and inbreeding, this result could indicate that these populations harbour more diversity than revealed with cpDNA results.

3.4.3 Low genetic diversity

Species displaying low levels of genetic diversity are often puzzling to conservation geneticists as they challenge the widely accepted dogma that genetic diversity is essential for species survival (Lehman 1998). Low levels of variation are

often considered a negative result and make statistical analyses troublesome or uninformative. I acknowledge that the above interpretations are lacking in certainty and power, owing to the decided lack of genetic diversity displayed throughout the range of Pacific dogwood, with both cpDNA sequence and microsatellite data. These results combined with earlier observations of low diversity in isozymes (SJ Brunsfeld, U. Idaho, pers. comm.) support the theory that a historic event (i.e. prolonged bottleneck prior to recolonization) has influenced population genetic diversity and structure of Pacific dogwood.

In contrast to Pacific dogwood, most tree species are found to have high levels of genetic diversity, thought to be a result of large effective population sizes, high reproductive capabilities and long life spans (Petit and Hampe 2006). The genetic depauperateness of *C. nuttallii*, documented in this study, suggests that this species may have endured a prolonged bottleneck prior to post-glacial expansion following the Last Glacial Maximum (LGM) in the Pacific Northwest.

Evidence for the duration of this bottleneck comes in the uniformity of low diversity, observed with both cpDNA haplotypes and nuclear microsatellites, throughout the range of Pacific dogwood. These homogenous low levels of variation suggest a severe species decline with the survival of a single population at some point in the species history, perhaps during the LGM. In contrast, if variation had been lost following expansion in different parts of the range, different alleles would have become fixed in different regions (e.g. *Pinus torreyana*; Ledig and Conkle 1983). Furthermore, Pacific dogwood appears to have retained low levels of diversity long after the glaciers began

their retreat. This suggests *C. nuttallii* endured a prolonged bottleneck in a glacial refugium and experienced a relatively recent and rapid range expansion.

Low levels of diversity were recently discovered at cpDNA microsatellite loci in a Mediterranean pine species, *Pinus pinea* L. (Vendramin *et al*. 2008). Similar to Pacific dogwood, this pine relies on animals, such as birds to disperse its seeds. Vendramin *et al*. (2008) believed that a scarcity of suitable dispersers at critical points in *P. pinea*'s history would lead to such low levels of diversity, and would compromise this pine's ability to colonize new territory.

In addition, despite the occupation of a large range and perceived ecological success, Grivet and Petit (2003) discovered an absence of cpDNA diversity in expanding populations of two hornbeam species in Europe, *Carpinus betula* and *C. orientalis*. This was thought to be the result of historical events occurring during the last glacial maximum in Europe (e.g. a bottleneck at the outset of colonization), biological features of these species, or the influence of humans. Furthermore, Grivet and Petit (2003) speculated as to whether certain woody species were more prone to losses of genetic diversity during glacial periods and if such a trend illustrated a route towards extinction of certain tree taxa through successive bottlenecks.

Low genetic diversity has been documented throughout the ranges of red pine (*Pinus resinosa* Ait; Fowler and Morris 1977; Walter and Epperson 2001), eastern white pine (*Pinus strobus* L.; Rajora *et al.* 1998), western redcedar (*Thuja plicata* D. Don; O'Connell 2003), and bigleaf maple (*Acer macrophyllum* Pursh; Iddrisu and Ritland 2004). Similar to inferences made regarding low haplotype diversity in chloroplast

genomes, this genetic depauperateness is thought to be the result of species experiencing bottlenecks in glacial refugia prior to recolonization.

Despite this hypothesized bottleneck for Pacific dogwood, we did find some preliminary evidence for local adaptation in populations of *C. nuttallii* with estimates of Q_{ST} being slightly larger than estimates of F_{ST} and much larger than those of R_{ST}. Despite the assumption that genetic variation is a precursor for adaptation to new environments, these findings are not all that surprising. Reed and Frankham (2001) have shown that heritability of traits and estimates of genetic diversity from molecular data are often poorly correlated. Furthermore, balancing or frequency-dependent selection is hypothesized to better preserve quantitative genetic variation than molecular diversity during bottlenecks (Lynch 1996, Reed and Frankham 2001). And finally, following a bottleneck, quantitative trait variation should recover more quickly than either microsatellite or chloroplast polymorphisms (Willis and Orr 1993).

However, considering the high levels of among-population gene flow, owing to Pacific dogwood's great capacity for long distance seed dispersal, this result was especially interesting as gene flow is often thought to limit local adaptation (Lenormand 2002). It should also be noted that although linear regressions of climate data on quantitative traits were significant ($p<0.05$), the relationships were relatively weak, and Q_{ST} estimates for Pacific dogwood were relatively low when compared with other tree species (Howe *et al*. 2003). It is known that some tree species display steep latitudinal clines for quantitative traits (Savolainen *et al*. 2007), while others display much weaker differentiation (e.g. *Larix occidentalis* (Rehfeldt 1995)). Similarly, estimates of Q_{ST} have been found to be only slightly higher than F_{ST} estimates for a number of species and traits

(see Figure 4 of Savolainen *et al*. 2007). Although it is unclear why this is the case, it is suspected that critical, yet difficult to measure parameters such as dispersal distance, strength of selection, and level of additive genetic variation on which selection can act are important and influential factors.

3.5 Conclusions and implications for conservation

The cumulative findings of this study suggest that low genetic diversity is ubiquitous throughout the native range of *Cornus nuttallii*. Although genetically depauperate species are often of great conservation concern, results from this study suggest Pacific dogwood may have a relatively long history of low levels of diversity. Furthermore, despite this presumed handicap, this species appears capable of successfully sustaining itself and adapting to novel environmental conditions, challenging some important assumptions of conservation genetics. This observation may illustrate the ambiguous correlations of neutral diversity, quantitative trait variation and adaptability. However, although some species are capable of thriving following a bottleneck (e.g., red pine, Walter and Epperson 2001), it is dangerous to assume this would be the case for all genetically depauperate species. For this reason it would be advisable for baseline data to be collected regarding population sizes and densities, in the event that Pacific dogwood should show signs of decline in the face of climate change or new or intensified biotic challenges.

Furthermore, weak phylogeographic structure and low levels of among population differentiation suggest this species possesses great capacity for long distance dispersal. In the face of a rapidly changing climate, this feature of Pacific dogwood could be considered advantageous as it appears predisposed to move quickly into new, climatically

favorable habitats. Furthermore, in British Columbia the biogeoclimatic zone in which *C. nuttallii* is most commonly found, the Coastal Douglas-fir (CDF) zone (Klinka *et al*. 2000), is predicted to expand northwards, based on current predictions of climate change (Hamann and Wang 2006). Therefore, observations of trees establishing outside of the current range may be considered a harbinger of shifting climate envelopes, and have subsequent implications for forest management.

This genetic study is the first of its kind for Pacific dogwood and we acknowledge there is still much to be learned. However, we believe this species provides an excellent opportunity to clarify relationships between estimates of diversity from different marker types and phenotypic traits, and to explore assumptions regarding genetically depauperate species.

Table 3.1. Summary of all sampled populations.

Code	Sampled From	State/ Province	Latitude (N)	Longitude (W)	Elevation (m)	n
SB[1]	San Bernadino NF[2]	California	34.24	117.20	1684	30
SQ[1]	Sequoia NF	California	35.72	118.54	1722	28
ST	Stanislaus NF	California	37.81	119.92	1153	30
PL	Plumas NF	California	39.91	121.01	1245	30
MD[1]	Mendocino NF	California	39.43	122.99	725	30
SH	Shasta-Trinity NF	California	40.84	122.01	418	30
KL[1]	Klamath NF	California	41.88	123.41	543	30
SK	Siskiyou NF	Oregon	42.80	123.83	818	26
UM[1]	Umpqua NF	Oregon	43.38	122.75	425	30
WL	Willamette NF	Oregon	44.16	122.26	366	30
SI[1]	Siuslaw NF	Oregon	44.46	123.51	408	30
CL[1]	Clearwater NF	Idaho	46.23	115.46	523	31
GP	Gifford-Pinochet NF	Washington	46.48	121.87	640	30
OL	Olympic NF	Washington	47.36	123.16	60	30
CW[1]	Lake Cowichan, Vancouver Island	British Columbia	48.79	123.89	231	30
CM	Cameron Lake, Vancouver Island	British Columbia	49.29	124.58	215	30
SC[1]	Sechelt	British Columbia	49.53	123.75	89	30
YL	Yale	British Columbia	49.70	121.40	149	30
BT[1]	Buttle Lake, Vancouver Island	British Columbia	49.83	125.62	242	30
PM[1]	Pemberton	British Columbia	50.29	122.84	402	30

[1] Sampled for both foliage for genetic marker analysis and seeds for common garden.
[2] NF = United States Forest Service National Forest

Table 3.2. Summary of nested chloroplast primers.

Name for nested fragment	**Chloroplast region**	**Primer pair**	**Fragment length in *C. nuttallii* (bp)**
S16n	*rpS16*	**S16FCn** CAA AGA TAA AGG ATC CCC AGA A **rpS16R** AAC ATC WAT TGC AAS GAT TCG ATA	445
S12n	*5' rpS12-rpL20*	**S12RCn** CCC ATG AAT TAT CCA GTA ATA GGT C **5'rpS12** ATT AGA AAN RCA AGA CAG CCA AT	498
BHn	*psbB-psbH*	**BHFCn** TAG TCC CCA TGT TCC TCG AA **psbB** TCC AAA AAN KKG GAG ATC CAA C	249
L16-A	*rpL16*	**L16RCn** CCC ATC GCT TCT TGC TTA AT **rpL16F71** GCT ATG CTT AGT GTG TGA CTC GTT G	496
L16-B	*rpL16*	**L16FCn** CAA TTC AAT ACG ATA AGG GAC AAA **rpL16R1516** CCC TTC ATT CTT CCT CTA TGT TG	380
S4R2T-B	*rpS4R2-trnT^{UGU}*	**S4R2TFCn** TTC CTG ATA TAGTTG GGA GTT CCT **rpS4R2** CTG TNA GWC CRT AAT GAA AAC G	194

Table 3.3. Chloroplast haplotype frequencies for 20 populations of *C. nuttallii.*

Population Code	Haplotype 1 frequencies	Haplotype 2 frequencies	n
SB[3]	0.533	0.467	15
SQ[3]	1	0	15
ST[3]	1	0	15
PL[3]	0.929	0.071	15
MD[3]	0.933	0.067	15
SH[3]	1	0	15
KL[3]	0.733	0.267	15
SK[3]	0.867	0.133	15
UM[3]	0.800	0.200	15
WL	0.600	0.400	15
SI	0.533	0.467	15
CL[3]	1	0	15
GP	0.467	0.533	15
OL	0.733	0.267	15
CW	0.333	0.667	15
CM	0.467	0.533	15
SC	0.667	0.333	15
YL	0.400	0.600	15
BT	0.867	0.133	15
PM	0.867	0.133	15
Total	**0.735**	**0.265**	**298**

[3] Population has been included in southern clade.

Table 3.4. Microsatellite allele frequencies for two loci in 20 range-wide natural populations of Pacific dogwood.

Populations																				
Locus	**Allele**	**SB**[4]	**SQ**	**ST**	**PL**	**MD**	**SH**	**KL**	**SK**	**UM**	**WL**	**SI**	**CL**	**GP**	**OL**	**CW**	**CM**	**SC**	**YL**	**BT**	**PM**
Cn-G4	1	0.53	0.44	0.63	0.35	0.29	0.33	0.40	0.19	0.28	0.41	0.52	0.18	0.40	0.38	0.64	0.25	0.30	0.54	0.28	0.35
	2	0.43	0.56	0.35	0.65	0.71	0.67	0.60	0.81	0.72	0.59	0.48	0.82	0.60	0.62	0.36	0.75	0.70	0.46	0.72	0.65
	3	0.03	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cn-J7	1	0.15	0.20	0.22	0.13	0.03	0.17	0.23	0.17	0.02	0	0.14	0	0	0.03	0.05	0.05	0.03	0.13	0.02	0.10
	2	0.75	0.70	0.65	0.69	0.78	0.80	0.73	0.79	0.94	1	0.86	0.90	1	0.92	0.76	0.88	0.95	0.88	0.90	0.90
	3	0.07	0.09	0.07	0.15	0.17	0.03	0.02	0.04	0.04	0	0	0.10	0	0.05	0.19	0.07	0.02	0	0.08	0
	4	0.03	0	0.07	0.02	0.02	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0

[4] See Table 3.1 for full names and details of populations.

Table 3.5. Estimates of within-population genetic diversity parameters for 20 natural populations of Pacific dogwood.

Code[5]	A_R[6]	H_O[7]	H_E[8]	F_{IS}[9]
SB	3.5 ± 0.5	0.367 ± 0.033	0.476 ± 0.06	0.195
SQ	2.5 ± 0.5	0.407 ± 0.037	0.483 ± 0.02	0.143
ST	3.5 ± 0.5	0.367 ± 0.133	0.507 ± 0.023	0.251
PL	3.0 ± 1.0	0.288 ± 0.019	0.475 ± 0.013	0.379
MD	3.0 ± 1.0	0.207 ± 0.034	0.397 ± 0.024	0.463
SH	2.5 ± 0.5	0.133 ± 0.0	0.394 ± 0.058	0.649
KL	3.0 ± 1.0	0.167 ± 0.033	0.451 ± 0.037	0.628
SK	2.5 ± 0.5	0.173 ± 0.058	0.335 ± 0.018	0.462
UM	2.5 ± 0.5	0.22 ± 0.1	0.264 ± 0.147	0.079
WL	1.5 ± 0.5	0.241 ± 0.241	0.247 ± 0.247	0.005
SI	2.0 ± 0.0	0.180 ± 0.060	0.378 ± 0.132	0.510
CL	2.0 ± 0.0	0.150 ± 0.150	0.244 ± 0.061	0.499
GP	1.5 ± 0.5	0.167 ± 0.167	0.244 ± 0.244	0.306
OL	2.5 ± 0.5	0.3 ± 0.133	0.320 ± 0.161	0.008
CW	2.5 ± 0.5	0.190 ± 0.017	0.431 ± 0.039	0.545
CM	2.5 ± 0.5	0.267 ± 0.033	0.299 ± 0.082	0.052
SC	2.5 ± 0.5	0.283 ± 0.183	0.262 ± 0.165	-0.076
YL	2.0 ± 0.0	0.232 ± 0.125	0.365 ± 0.142	0.396
BT	2.5 ± 0.5	0.117 ± 0.05	0.299 ± 0.114	0.612
PM	2.0 ± 0.0	0.350 ± 0.150	0.323 ± 0.140	-0.105
Overall	**2.475**	**0.24**	**0.36**	**0.308**
SE	**± 0.124**	**± 0.022**	**± 0.024**	**± 0.047**

[5]See Table 3.1 for full names and details of populations.
[6]Allelic richness
[7]Observed heterozygosity
[8]Expected heterozygosity
[9]Inbreeding coefficient

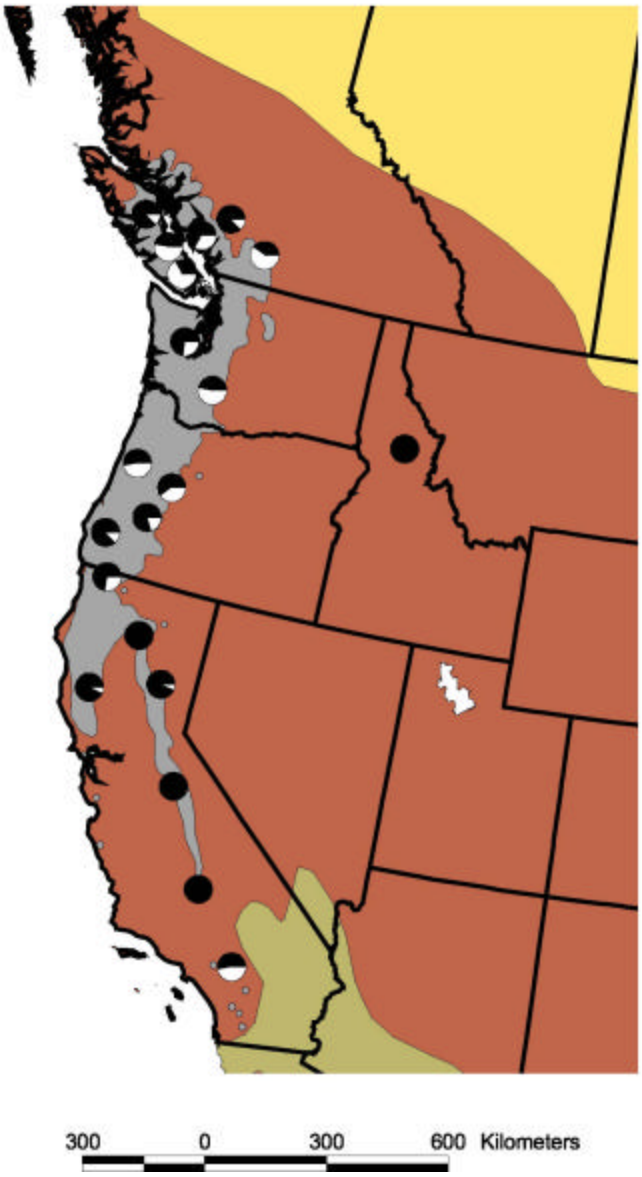

Figure 3.1. Geographic distribution and frequencies of haplotypes (black indicates haplotype 1, white indicates haplotype 2) in each sampled population throughout the native range of *Cornus nuttallii* (grey shading). The distribution of American robin (*Turdus migratorius)* is indicated by yellow for summer breeding grounds, brown for year round populations and green for wintering grounds) (Ridgely *et al*. 2003) *(Data provided by NatureServe in collaboration with Robert Ridgely, James Zook, The Nature Conservancy - Migratory Bird Program, Conservation International - CABS, World Wildlife Fund - US, and Environment Canada - WILDSPACE.)*

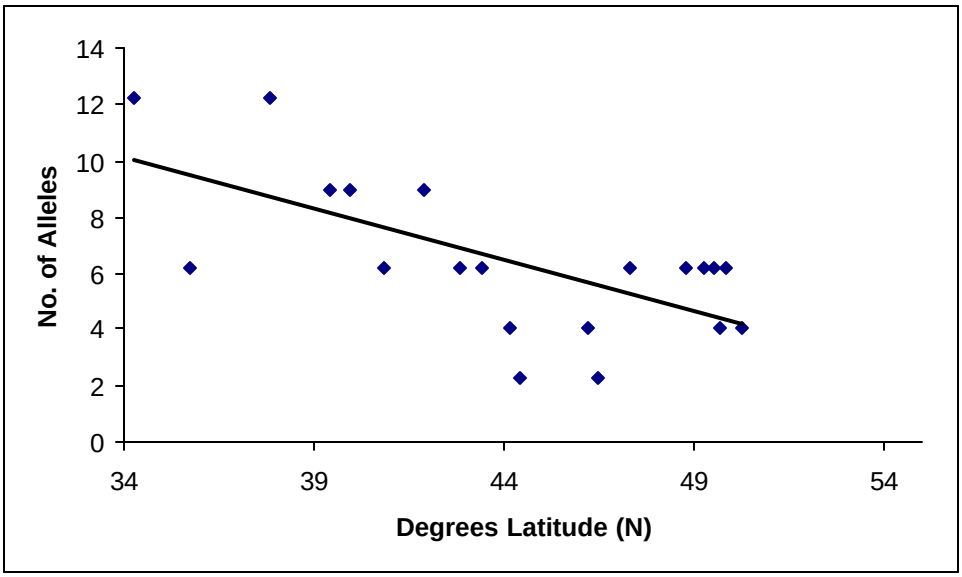

Figure 3.2. Regression of allelic richness on latitude based on average values from two microsatellite markers for 20 populations of *C. nuttallii* ($r^2 = 0.42$, $p<0.01$).

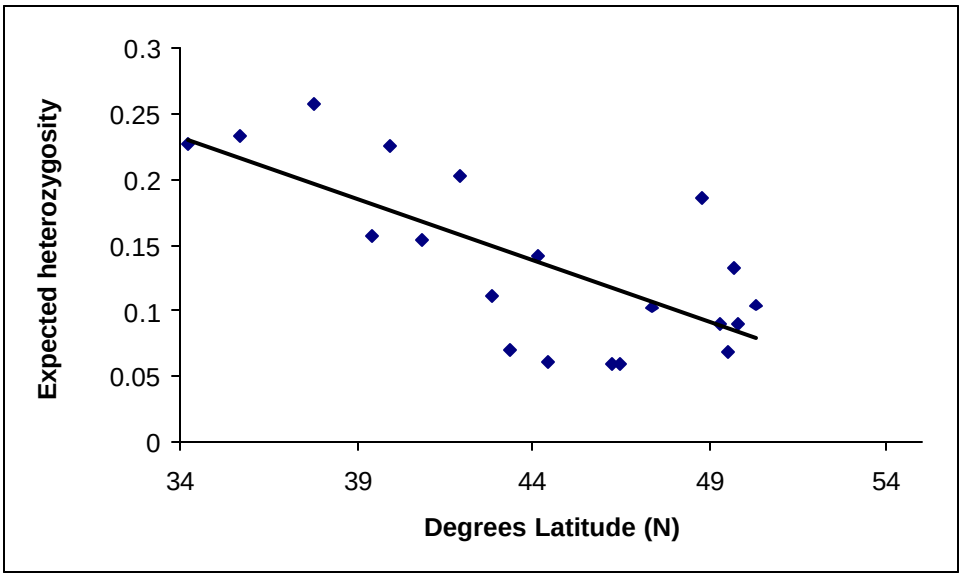

Figure 3.3. Regression of expected heterozygosity on latitude based on average values from two microsatellite markers for 20 populations of *C. nuttallii* (r^2 = 0.51, $p<0.01$).

.

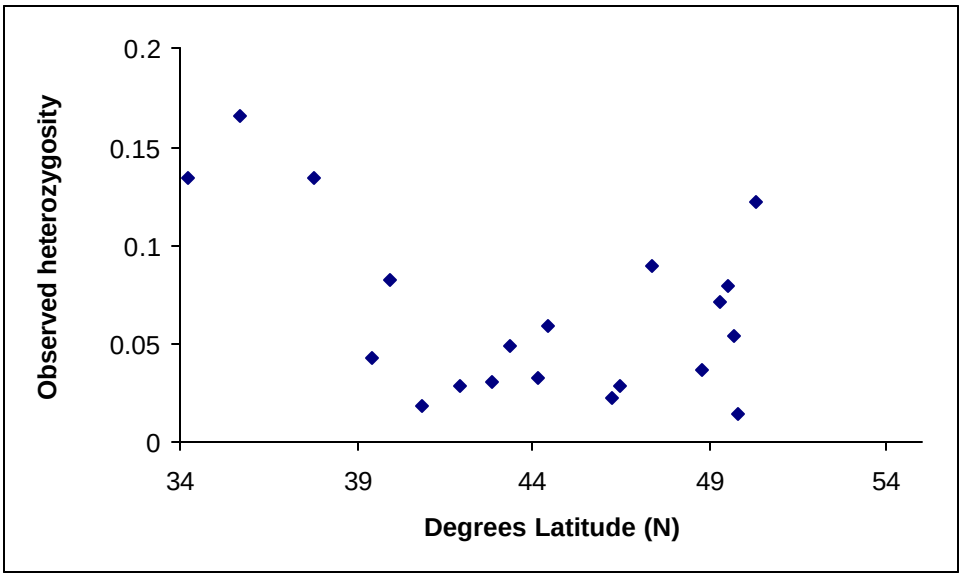

Figure 3.4. Regression of observed heterozygosity on latitude based on average values from two microsatellite markers for 20 populations of *C. nuttallii* (r^2 = 0.17, p = 0.06).

.

CHAPTER 4 – Conclusions and future directions

4.1 Conclusions and conservation

In this thesis I succeeded in the discovery of useful nuclear and chloroplast molecular markers for a species that had never been subject to genetic studies. Hopefully this feat will encourage others to do the same, so that we may increase our breadth of knowledge in the field of tree evolutionary dynamics and limit the tendency to study only those species deemed economically valuable. In the application of these markers, combined with preliminary results of quantitative trait variation, we began to piece together the story of *Cornus nuttallii*.

Uniform low diversity at both microsatellite and chloroplast loci suggests this species once existed as a single population which was subjected to a prolonged bottleneck. Bottleneck events were prevalent in glacial refugia and therefore it seems likely that the event which dramatically reduced levels of diversity in Pacific dogwood occurred during the last episode of glaciation in the Pacific Northwest. As the climate warmed, creating new climatically favourable habitats for numerous plant and animal species to colonize, *C. nuttallii* may have lingered in its refugium longer than other species, awaiting a suitable disperser for its seeds. This would prolong the bottleneck imposed on the species, which is evidenced by the low levels of diversity that continues to exist today in Pacific dogwood populations. Once a suitable and effective dispersal agent was available, Pacific dogwood appears to have rapidly filled its current range, assuming a post-glacial migration strategy similar to the 'leading edge hypothesis'. This is also supported by the significant decline in allelic richness and genetic diversity observed south to north in microsatellite loci. Low levels of population differentiation,

especially between populations that in theory should be quite genetically dissimilar (e.g. northern Idaho and coastal populations), support this hypothesis.

These observations of low genetic differentiation could reflect the high levels of contemporary gene flow between populations of Pacific dogwood. This would be further supported by the weak phylogeographic structure observed in chloroplast haplotypes but would explain the low levels of diversity. However, it could be a rapid and recent recolonization, combined with high levels of between population gene flow, which have together influenced current weak patterns of genetic variation.

Although species found to have low levels of genetic diversity are often considered to be of conservation concern, this research suggests Pacific dogwood may have a history of such levels. Furthermore, weak yet significant levels of differentiation for quantitative traits suggest this species has recovered or retained sufficient quantitative trait variation for selection to act. However, it is dangerous to assume that such a history guarantees success in the future. Current threats to *C. nuttallii* include habitat loss and fragmentation from human development, as well as an introduced fungal parasite, *Discula destructiva*, which has already severely impacted disjunct populations in northern Idaho. Furthermore, as is the case with most species, Pacific dogwood is now being forced to adapt to a rapidly changing climate. For this reason, it would be advisable for baseline data to be collected regarding population sizes and densities, in the event that Pacific dogwood should show signs of decline.

Should restorative efforts be required, it would be prudent to collect seeds from the nearest southern source as a precautionary measure. I specify southern source here, despite the general assumption that local genotypes are considered optimally adapted to

current climate conditions. There is mounting evidence that the earth is in a period of unprecedented warming. The rate at which this warming is occurring challenges this fundamental assumption as populations may quickly find themselves maladapted to current climate conditions. However, should seeds be transferred from southern sources, we may be able to establish seedlings that are more optimally suited for current and future climates, ensuring the future survival of species. In addition, as *C. nuttallii* is valued as an ornamental species, nurseries growing or selling these trees should also be careful to ensure seeds are from a suitable source, and seedlings are planted locally.

With climate change, a capacity for long distance dispersal may predispose Pacific dogwood to future success colonizing new, climatically favorable habitats. Furthermore, this trait may enable *C. nuttallii* to act as a harbinger of shifting climate envelopes at the northern part of this species' range and have subsequent implications for forest management. At the southern part of the range the situation is much different. Without anywhere to go, these high elevation populations, currently found over 1000m, will be forced to adapt to the changing climate or face elimination. With genetic variation being the precursor for adaptation, Pacific dogwood may not have the tools necessary. For this reason, it is important to monitor and consider moving these populations.

4.2 Future directions

As is the case with most studies aimed at unlocking the secrets that lie within the genetics of a previously unstudied species, there is still much to be learned about Pacific dogwood. Unfortunately, the lack of genetic variation made the application of the

usual arsenal of analytical methods for population genetic inferences difficult or impossible. For this reason, a continued search for polymorphic markers would be advisable. This would allow for more formal hypothesis testing, and greater confidence in results. Furthermore, information regarding the mating system of *C. nuttallii*, population densities, and average dispersal distance would be of great value. Finally, the results from this study indicate Pacific dogwood may provide an excellent opportunity to clarify relationships between estimates of diversity, and to explore assumptions regarding the ability of species that are genetically depauperate for genetic markers to adapt to new conditions.

Literature Cited

Aitken SN, Libby WT (1994) Evolution of the pygmy-forest edaphic subspecies of *Pinus contorta* across and ecological staircase. *Evolution*, **48**, 1009-1019.

Allen GA, Antos JA, Worley AC, Suttill TA, Hebda RJ (1996) Morphological and genetic variation in disjunct populations of the avalanche lily *Erthronium montanum*. *Canadian Journal of Botany*. **74**: 403-412.

Arno S (1977) *Northwest Trees*. pp. 222. Mountaineers, Seattle, WA.

Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts

Bridle JR, Vines TH (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, **22**, 140-147.

Brown JH (1984) On the relationship between abundance and distribution of species. *American Naturalist*, **124**, 255-279.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PE (2001) Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J) pp.319-339. Blackwell Science Ltd, Oxford.

Brunsfeld SJ, Sullivan J (2005) A multi-compartmented glacial refugium in the northern Rocky Mountains: Evidence from the phylogeography of *Cardamine constancei* (Brassicaceae). *Conservation Genetics*, **6**, 895-904.

Brunsfeld SJ, Miller TR, Carstens BC (2007) Insights into the Biogeography of the Pacific Northwest of North America: Evidence from the Phylogeography of *Salix malanopsis*. *Systematic Botany*, **32**(1), 129-139.

Brussard PF (1984) Geographic patterns and environmental gradients: the central-marginal model in *Drosophila* revisited. *Annual Review of Ecology and Systematics*, **15**, 25-64.

Bower AD, Aitken SN (2007) Geographic and seasonal variation in cold hardiness of whitebark pine. *Canadian Journal of Forest Research*, **36**(7), 1842-1850.

Cabe PR, Liles JS (2002) Dinucleotide microsatellite loci isolated from flowering dogwood (*Cornus florida* L.). *Molecular Ecology Notes*. **2**: 150-152.

Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2005) Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: Hypothesis testing within a comparative phylogeographic framework. *Evolution*, **59**, 1639-1652.

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results from over 200 angiosperms. *American Journal of Botany*, **75**(10), 1443-1458.

Crispo E, Hendry AP (2005) Does time since colonization influence isolation by distance? A meta analysis. *Conservation Genetics*, **6**, 665-682.

Cwynar LC, MacDonald GM (1987) Geographical variation of lodgepole pine in relation to populations history. *American Naturalist*. **129**: 463-469.

Daubenmire R (1952) Plant geography in Idaho. In: *Flora of Idaho* (ed Davis RJ) pp 1-17. Brigham Young University Press.

Davis MB, Shaw RG (2001) Range sifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-679.

Detling LE (1968) Historical background of the Flora of the Pacific Northwest. **13**, University of Oregon Press, Eugene.

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: *Molecular systematics of plants* (eds Soltis PE, Soltis DE, Doyle JJ) pp 14-35, Chapman Hall, New York.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11-15.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-1188.

Elliot-Fisk DL (1988) The boreal forest. In: *North American terrestrial vegetation* (eds Barbour MG, Billings WD) pp.33-62. Cambridge University Press, New York.

Fowler DP, Morris RW (1977) Genetic diversity in red pine: evidence of low genetic heterozygosity. *Canadian Journal of Forest Research*, **7**, 343-347.

Garciá-Ramos G, Kirkpatrick M (1997) Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, **51**, 21-28.

Grivet D, Petit RJ (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: Evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics*, **4**, 47-56.

Guries RP, Ledig FT (1982) Genetic diversity and population structure in pitch pine (*Pinus rigida* Mill.). *Evolution*, **36**, 387-402.

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**, 521-523.

Hamann A, El-Kassaby YA, Koshy MP, Namkoong G (1998) Multivariate analysis of allozymic and quantitative trait variation in *Alnus rubra* Bong.: geographic patterns and evolutionary implication. *Canadian Journal of Forest Research*, **10**, 1557-1565.

Hamann A, Wang T (2006) Potential effects of climate change on ecosystem and tree species distributions in British Columbia. *Ecology*, **87**(11), 2773-2786.

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Sciences*, **351**, 1291-1298.

Hengeveld R, Haeck J (1982) The distribution of abundance. 1. Measurements. *Journal of Biogeography*, **9**, 303-316.

Heusser CJ (1960) Late-Pleistocene environments of North Pacific North America. *American Geographic Society Miscellaneous Publications*, **35**.

Heusser CJ (1985) Quaternary pollen records from the Pacific northwest: Aleutians to the Oregon-California boundary. In: *Pollen Records of the Late-Quaternary North American Sediments* (eds Bryant Jr. VM, Holloway RG) pp. 141-165. American Association of Stratigraphic Palynologists Foundation, Dallas, Texas.

Heusser CJ (1989) North Pacific coastal refugia – the Queen Charlotte Islands in perspective. In: *Queen Charlotte Islands, B.C.* (eds Scudder GGE, Gessler N) pp. 91-106. Queen Charlotte Island Museum Press.

Hewitt GM (1993) Postglacial distribution and species substructures: lessons from pollen insects and hybrid zones. In: Evolutionary patterns and processes (eds Lees DR, Edwards D) pp 97-123. London. Academic Press.

Hewitt GM (1999) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH (2003) *Canadian Journal of Botany*, **81**, 1247-1266.

Iddrisu MN, Ritland K (2004) Genetic variation, population structure, and mating system in bigleaf maple (*Acer macrophyllum* Pursh). *Canadian Journal of Botany*, **82**, 1817-1825.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383-2394.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1-23.

Klinka K, Worrall J, Skoda L, Varga P (2000) The distribution and synopsis of ecological and silvical characteristics of tree species of British Columbia's forests. Canadian Cartographics Ltd., Vancouver. 180 p.

Ledig RT, Conkle MT (1983) Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex Carr.). *Evolution*, **37**, 79-85.

Lehman N (1998) Conservation biology: genes are not enough. *Current Biology*, **8**, R722-R724

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**(4), 183-189.

Lesica P, Allendorf FW (1995) When are peripheral-populations valuable for conservation? *Conservation Biology*. **9**: 753-760.

Li P, Adams WT (1989) Range-wide patterns of allozyme variation in Douglas fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research*, **19**,

Little EL Jr. (1976) Atlas of United States trees, volume 3, minor western hardwoods. U.S. Department of Agriculture Miscellaneous Publication 1314. 13 p., 290 maps.

Lynch M (1996) A quantitative-genetic perspective on conservation issues. In: *Conservation genetics: case histories from nature* (eds. Avise JC, Hamrick JL) pp. 471-501, Chapman and Hall, New York.

Magni CR, Ducousso A, Caron H, Petit RJ, Kremer A (2005) Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology*, **14**, 513-524.

McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution*, **10**(5), 198-202.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, **17**, 285-291

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, **14**, 892-903.

Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with post-glacial migration in *Picea sitchensis*. *Heredity*, **99**, 224-232.

Muona O, Paule L, Szmidt AE, Karkkainen K (1990) Mating system analysis in a central and northern European populations of *Picea abies*. *Scandinavian Journal of Forest Research*, **5**, 97-102.

Nybom H (2004) Comparison of different nuclear markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**(5), 1143-1155.

O'Connell L (2003) The evolution of inbreeding in western redcedar (*Thuja plicata: Cupressaceae)*. Ph.D. Thesis, University of British Columbia, p. 163.

Oxelman B, Liden M, Berglund D (1997) Chloroplast *rps16* intron phylogeny of the tribel Sileneae (Carophyllaceae). *Plant Systematics and Evolution*, **206**, 393-410.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

Peterson MA, Denno RF (1998) The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. *American Naturalist*, **152**, 429-446.

Petit RJ, Aguinagalde I, Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty P, Müller-Starck G, Demsure-Musch B, Palme A, Martin JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**(5625), 1563-1565.

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution and Systematics*, **37**, 187-214.

Pielou EC (1991) After the ice age: the return of life to glaciated North America. Chicago, University of Chicago Press.

Pojar J (1980) Brooks Peninsula: possible Pleistocene glacial refugium on northwestern Vancouver Island. *Botany Society American Miscellaneous Serial Publications*, **158**, 89.

Primmer CR, Moller AP, Ellegren H (1996) A wide-ranging survey of cross-species amplification in birds. *Molecular Ecology*, **5**, 365-378.

Rajora OP, DeVeerno L, Mosseler A, Innes DJ (1998) Genetic diversity and population structure of disjunct Newfoundland and central Ontario populations of eastern white pine (*Pinus strobes*). *Canadian Journal of Botany*, **76**, 500-508.

Reed DH, Frankham R (2001) How closely related are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095-1103.

Rehfeldt GE (1995) Genetic variation, climate models and the ecological genetics of *Larix occidentalis*. *Forest Ecology and Management*, **78**, 21-37.

Richardson BA, Brunsfeld SJ, Klopfenstein NB (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology*, **11**, 215-227.

Ridgely RS, Allnutt TF, Brooks T, McNicol DK, Mehlman DW, Young BE, Zook JR (2003) Digital Distribution Maps of the Birds of the Western Hemisphere, version 1.0. NatureServe, Arlington, Virginia, USA.

Ritland K (2000) Marker-inferred relatedness as a tool for detesting heritability in nature. *Molecular Ecology*, **9**(9), 1195-1204.

Rozen S, Skaletsky HJ (2000) Primer on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S) pp 365-386, Humana Press, Totawa, New Jersey.

Sagarin RD, Gaines SD (2002) The 'abundant centre' distribution: to what extents is it a biogeographical rule? *Ecology Letters*, **5**, 137-147.

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA plylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, **84**, 1120-1136.

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution and Systematics*, **38**, 595-619.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**(4), 465-474.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences fro phylogenetic analysis. *American Journal of Botany*, **92**(1), 142-166.

Shaw J, Small RL (2005) Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). *American Journal of Botany*, **92**, 2011-2030.

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**(3), 275-288.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Small A (1994) California birds: their status and distribution. Ibis, Visa, California.

Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF (1998) The Tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeneitic reconstruction in a recently diverged plant group. *American Journal of Botany*, **85**, 1301-1315.

Smith JP, Sawyer Jr. JO (1988) Endemic vascular plnats of northwestern California and southwestern Oregon. *Madroño,* **35**, 54-69.

Soltis DE, Soltis PS, Ness BD (1989) Chloroplast DNA variation and multiple origins of autopolyploidy in *Heuchera micrantha* (Saxifragaceae). *Evolution*, **43**, 650-656.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*. **206**: 353-373.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology,* **15**, 4261-4293.

Taberlet PL, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplication of three non-coding regions of chloroplast DNA. *Plant Molecular Biology,* **17**, 1105-1109.
Wakasugi TM, Sugita M, Tsudzuki T, Sugiura M (1998) Updated gne map of tobacco chloroplast DNA. *Plant Molecular Biology Reporter*, **16**, 231-241.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cossens JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology,* **7**, 453-464.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**(8), 1596-1599.

Tate JA, Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploidy species. *Systematic Botany*, **28**, 723-737.

Vendramin GG, Fady B, Gonzáles-Martinez SC, Hu FS, Scottie I, Sebastiani F, Soto A, Petit RJ (2008) Genetically depauperate but widespread: The case of an emblematic Mediterranean pine. *Evolution*, **62**(3), 680-688.

Waddell KL, Hiserote B (2005) The PNW – FIA Integrated Database User Guide and Documentation: Version 2.0. Internal Publication: Forest Inventory and Analysis program, Pacific Northwest Research Station. Portland, Oregon.

Walter R, Epperson BK (2001) Geographic pattern of genetic variation in *Pinus resionosa*: area of greatest diversity is not the origin of postglacial populations. *Molecular Ecology*, **10**, 103-111.

Webb SL (1986) Potential role of passenger pigeons and other vertebrates in the rapid Holocene migrations of nut trees. *Quaternary Research USA*, **26**(3), 367-375.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Whitlock MC (2008) Evolutionary Inference from Q_{ST}. *Molecular Ecology*, **17**(8), 1885-1896.

Whittaker RK (1961) Vegetation history of the Pacific coast states and the "central" significance of the Klamath region. *Madroño*, **16**, 5-17.

Willis JH, Orr HA (1993) Increased heritable variation following popuation bottlenecks: the role of dominance. *Evolution*, **47**, 949-957.

Willson, MF (1994) Fruit choice by captive American robins. *Condor* **96**, 494–502.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.

Xiang QY, Brunsfeld SJ, Soltis DE, Soltis PS (1996) Phylogenetic Relationships in Cornus Based on Chloroplast DNA Restriction Sites: Implications for Biogeography and Character Evolution. *Systematic Botany*. **21** (4): 515-525.

Xiang QY, Thomas DT, Zhang W, Manchester SR, Murrell Z (2006) Species level phylogeny of the genus Cornus (Cornaceae) based on molecular and morphological evidence-implications for taxonomy and Tertiary intercontinental migration. *Taxon*, **55**(1), 9-30.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1-16.

Idaho Rare Plant Conference (2004) The Idaho Native Plant Society rare plant list: State rare species list, [Online]. In: Results of the 20th annual Idaho rare plant conference. Idaho Native Plant Society (Producer). Available: http://www.idahonativeplants.org/rpc/RarePlantList.aspx

Appendix I – Chloroplast sequences from *C. nuttallii*

trnH^{GUG}-psbA

TAGAGAAGTAACGACGTTGTAAAACGACCGCGCATGGTGGATTCACAATCCA
CTGCCTTGATCCACWTGGCTACATCCGCCCCTATACTATATTACATTACAAAT
GATTCAATTTCGACCATTCATCATTATTTCTTTCTTATCTTATTTCTTTTCTGAG
ATACAAATCTGAAGCAATTTTATCTGTTATTTTAAATGTAAAATAACAACTTA
ACATTAGGGAGACGATATATAAATTAATAAAAAATGAATAAAGAAGTAAAG
CACAATACTCAATCATGAACCAATCTATAAGAATCCTTTTTCTTTTTATGTAA
AAAAAAGTATCTATAAGAAAAAGACTACTAAATAAAAATAAAGGAGCAATA
CCACCCTCTTGGTCTTGATAGAACAAGAAATTGGTTATTGCTCCTTTACTTTC
AAGAACTCATATACACTAAGAAAAGGTCTTATCCATTTTGTAGATGGAGCTTC
AATAGCAGCTAGGTCTAGAGGGAAGTTATGAGCATTACGTTCATGCATAACC
CTGTGTGAAA

rpS16

TGTGGTAGAWWGCAACGTGCGACTTGAAAGACACGATCCGTTGTGGATTCTT
ACATCCATCATTTTATATAGGAATGAAGGTGCTCCTGGCTCGACATCGTTTGT
TCTGTTCCACTAGAACCCCTCCTTTTTGTTGGGTTGTAATGTAAATAGTACAT
GGTGGAGCTCGAGCAGAAAGTATTGATTCATTTCTCGGGGGCAGGGATCTAG
GGTTAATGCTAATCAATAAATTGGAACAACTTCGTAAGTATATCTTCGATATA
GAAATAGAAATCGAAAGAATCCAATTCGAGCAAGTTTCCGCCCCAAAAGGA
AAAATTGTTGGAATTGATAAAACTCTTTCGATCCAAAGTGTATCGCGCGGGA
ATCCACCGTTCATATGATTCTTTGATAGAAAGAAATCACAAAAAGGGTATGT
TGCTGCCGTTTTGAAAGGATTAAGGATCGCCGAAGTAATGTCTAAACCTAAT
GATTCAAAACAAAGATAAAGGAWCCCCAGAACAAGGAAACGCCGTTTGAGT
TGTCYCAATAACTGGATCAGAATGAAGAATTAAAATTGTTTTTAAATGAGAC
AAACAAAAAAGGGGGTTAGAGACCACTCAATAAATGAAATAAATGCCTAAA
GATTTTTCTTTGAGCTATTGGGGAGTTATCCAACTTGAGTTATGAGTACAAAT
GATTTTTTTCTTTTTCAGTAAGGAAGAGAAGAAGAAAAAGGGCTTAAATCAT
AGCCTAATTGATTTGATGATTTTATGGATCCATTTGCCATTAGAATTCTATATT
ATAATTCGATACATCGAAATCACTTCGAATCATTTTTTCTTGAGCCGTACGAG
GAGAAAACTTCCTATACGTTTCTAAGGGGGGGGGTWHTGTTSMTCTMCATCT
ATCCCAATGAGCCGTCTATCGAAYCCTTGCAATTG

5' rpS12-rpL20

AGAAAGGCAAGACARCCAATCAGRAATGTCACAAAATCCCCTGCTCTTCGGG
GAGTGCCCTCAGCGTCGRGGAACATGTACTAGGGTGTAGTGTGCGACTCGTT
CAGATCATGGGTCTGGGACAAAGGAAAGAAAACCAATTTTCCAGTAACAACG
WTCAGTACCGATGWATAGGATAGACTGGAAGAACTCCATTCGSTATCTAAAA
ATAAAAAATCTATCGTATCCCTCTATTGTGTATGAAATTTATGGTTTCCGTTG
GTGCAAATCCAATTACCTCCATTTTAAATGAAAAGCAATTCCCCATTGGTAAC
AAWKRGTTATTCATTAAGCGGGGAAAATCCTATTTAAAATTTAAAAAAACAG
AAAGATTAGGTTCCCACTCCTTGCAAGAACGGACTCACAGGGTCAGCTACCC
CGCTAACTTTTATAATTAAATACCGTTACTGTATAGGTAGATCCTCATTGTGA
AAGACCTATTACTGGATAATTCATGGGTAGAGCCAAAGAGTGTGAACTGTAC
AAGTTACCAATAACATTGATTAAATCAAGTAAGGGGCTCCGGTGTATAGAGA
AGGCCTCACCGTTTAAGAAGTAACCATAGAAACGATGGAACCCACTATTTCT
TTATCCATTTATATTTACTACTTTATTTATTTACTATGTTTTTGATACCTAGGA
GAATACAATTGATTATTTCTAGGTGAAATGCCTAGAAAAAATCGTGGTCGGG
AAGGTTATAGTAGCCAAAGCCATTGGAATTTTTATTTTATACATTGGAAAAAT
CCGTTTTGTTATTAATAGGCTAGGAAGGGGGAAAAGAATAACTGAAAGAAAG
GAAATCAATTAGTTATTCGTCAAAGTTTCAATTATTCAATGACCAGAATTAAG
CGAGGATATATAGCTCGATA

psbB-psbH

AAAAAGGGGGAGATCCAACTACACGAAGACAAGTAGTCTGATACAACATTTC
TCTGGTAGTTTTCGCCTCTATTTTCTTTTTGTGATTTGGCATAGGGTCCCAGAG
AAAGCTTGATTTGAATCACTGCCTTTCTTTATCCGGTAGATGATCCCCAATAA
AATAAAAAGAAACAGGTATGGAAGCTATAATTGTAAACCACGATCAAATCTA
TGGAAGCATTGGTTTATACATTCCTCTTAGTCTCGACTCTAGGGATAATTTTTT
TCGCTATCTTTTTTCGAGAACCGCCTAAAGTTCCAACTAAAAAGATGAAATGA
TTTTTCATTATCTCAATTGAAGTAATGAGCCTTCCCAATATTGGAAGGCTCAT
TACTTCAACTAGTCCCCATGTTCCTCGAATGGATCTCTTAGTTGTTGAGAAGG
CTGCCCAAAAGCGGTATATAAGGCGTACCCAGTAAAACTTACAAGTAAACCA
GATATAAAGATGGCGACTAGGGTTGCTGTTTCCATTATTATATAATTTCAAGA
CCAAAATAGATCTATGATAAGATCGTTTATTTACAACARRAATMGTATACAA
AGTCAACAGATCTCAATGAATACAATAGGATTTATGGCTACACAAACA

rpL16

TGCTTAGTGTGTGTCTCGTTGATTTTTTTCGGGTTAGGATTAAAAAAAAATGA
CCAGCCCATAGTTTCTAATAATAAACAACTCATCACTTCGCATTATCTGGATC
TAAAGAACCAGTCAAGATATGATATATCAATCATATCATTGTAGCAACTGAA
TCTTTTTTGCATAAACAAAATAAAAATCAACCTGATTCTAAATTGTGAAGCGA
AATATAAAAGAAGGATGTGGATAAATGGAAGGACGAGAGAAAGAGAGAAA
AAGAATATCAATGATATATGAGTCCAATATGTAAGGTCTATGAATTATCTCAT
AAAAGGCAGTGTAATAAAGCATCAATACTGTGATTCATCCATAATTAAATGA
AATTGTTCATAGAACAAAAAAATCAAGAGCCTCGAGCCAATAAAGACTGAG
AAAATTGACTCAAGAAAAAATTTCATTAAGAGCTCCATTGTAGAATTCAGAC
CTAACCATTAAGCAAGAAGCGATGGGAACGACGGAACCCTATGAAATGCGN
NGATGGCGGAACGAACGGAGAACCAATTCATCTATTCTGAGAAGTCATGAGC
TAACCCTACAACTGAAATAGATATTGAAAGAGTAAATATTCGCCCGCGAAAA
CTTTATTTTCTTAGATTGATAAATTTTGAACAATTCAATACGATAAGGGACAA
AATAAAAAAAAGATTCGCTATAGCCCTATAAAAAACAATATTATCTATAAATA
GAAATATATATTTTATAGGTTTAGTTATATATCCAAATAAGATATACAAATTA
CTAATAAATTAAATGAAATCTCAAAGAATCCCATGATTCAATGTATTATTCAT
TAAATATTAAATACCTGTATATCTTCAATTCTATTTAAATATTTTT(T/G)AATTT
GAATCCTTTTATTCGCGAGGGTTCTGGATGAGAAGAAAC TCTCACGTCCGGTT
CTGTAGTAGGGATGGGTTTGAGAAAACAACCATCAACTATAACCCCAAAAGA
ACCAGATTCCGTAAACAACATAGAGGAAGAATGAAGGG

5'trnL^{UAA}- trnT^{UGU} (within 5'trnL^{UAA}- trnS^{UGA})

CGATGCTCTTTCCGTCGTGTGCTTTGTGGGGCGTCTCATTTTCAGTTTTGGGGA
CTCGTGGCGTGTGTTTTTTTTTCTGTTGGGGTTTTGGCGGTTCCCTGGTGTTTT
CGGCGTTAGCTATTGACCCTTTTTGTGATTTGTAGTGAACCGTAGTTTATAGTT
TTTTTTTTCAATATTGAATATTTATAGAGCATAACGATGAATATAGCGTTATA
GAATTTCGATTTATTTATCACAATTAGAATTCTAAATTTAAAAAAATCGCTAG
TCAAATTTGACTTTTCGTTTTTGAATTCAAATGTCATTTGAAATTCTTTTTATT
ACATATCATATAGATTTCTAATTCTAATTATTTTC TAATTATGAAATTAGTCTA
ATTAGACATATACAATTATACATTCTATATATATATATTATATATAGTCTATA
TTCTATATAGACATTAGTTACATTATTACTTTACTTATGTATTATATATACATA
TTGCTTATATATTATTAGATTATGAGTTTAATCTAATTAATTAATTAGAATTAT
GTAATGATTAATTAAAACTAATAAGACATTCCTCCGCTTTCATTCGTAAAGGC
GGAAGTTAAGACGAAAAAAAGAGAATCGACCGTTCAAGTATTCTAAATTGCA
TGGAAAAGTTAAAAGAAAGAGAGACATATATGGGGTATATGCCCATCTATAT
TGAATTGCGGATACAGAAATGATAGAATCATTTTTGATTGGACCAAATAGGT
GTCTCCTATAGAAGATGTAAAACGATCCCAAAAAAAATCTTTTTTTATGTTGT

GTGAACAGATCTGCCATATATCTTCTGTATGTAGTCACGAGATATTTCTGGCT
CAATACTCTCATATTGGGAACAATTGAGGAAAAGAGGGAACGGCTTTTTCAT
TTTGGACAAAATCTTAAAACATACAAGGGGATATGCGAAATCGG

rpS4R2-trnT^{UGU}(within 5'trnL^{UAA}- trnS^{UGA})

AGTGTAAAACACCGTTGAGWCCGTWATGWAAACGCAATTTTTGTTTTTCTTC
TAGACGAATACGATATTGAGATCTTTTMCCGGAACGCGATTGGTTTCTAAGA
TCGCTTCCGGCTCTAGGCCTTTTATTAGTTAGTCCCGGTAAAGCCCCCAGGCG
GCGTATTTTTTTGAAACGAGGTCCTCGGTAACGCGACATAAAGACTCCTTATT
CTTATTTATATTTAATTTTTTTTATTTATTGAAATTTCATTTTACAGAATAAAC
CTAAACTAAAACTGAACTAAATGATAAATGAATCGAAGTCTACTGAAGTATT
GTACTATAAAGAATAATGAGATGAATTGTATAAATATTCAGACCCCTTTGTAT
TATATATACAGAACAAGAAAAAGATCCTTTTCCTGATATAGTTGGGAGTTCCT
ATAACATAATAAATCGGCGATTTTTGGAAAAAAGAGAAGATTCTTTTTCAAT
ATTCTTTGAATTTCMAAGGGGCATTATCAATCATGTAAAAAGTCGAACAAAT
GGAGAAAAAAAGCCGGCTATCGGAATCGAACCGATGACCATCGCATTACAA
ATGCGATGCTCTAACCTC