PHYLOGEOGRAPHY OF DOUGLAS-FIR: TESTING HYPOTHESES FROM THE FOSSIL RECORD

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

PAUL FRANCIS GUGGER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ADVISED BY

JEANNINE CAVENDER-BARES
SHINYA SUGITA

JUNE 2010

Acknowledgments

I am grateful to my advisors, Jeannine Cavender-Bares and Shinya Sugita, for their continuous support and encouragement. I also thank my committee members, Ruth Shaw, Peter Tiffin, and Bob Zink, for helpful conversations and comments on manuscripts. Keith Barker provided invaluable help with analyses and software. Thank you to Nick Deacon, Kari Koehler, Jessica Savage, and Gina Quiram for feedback on manuscripts, lab assistance, and friendship. Special thanks to those who have helped with field work: Antonio González Rodríguez, Barry Jaquish, Michael Stoehr, Carlos Acosta Gómez, Ken Oyama, and Victor Rocha Ramírez. Laura Strickland assisted with the Packrat Midden Database (Chapter 1), Cathy Whitlock and Feng Sheng Hu commented on manuscripts (Chapters 1 and 2, respectively), Art Dyke and Louis Robertson provided ice sheet data used in Figures 1.2, 1.3 and 2.1 (Chapters 1 and 2), Molly Bowen and the Cornell University Computational Biology Service Unit provided computer resources for IMa analyses (Chapter 2 and 3), and Antonio González Rodríguez and Hernando Rodríguez Correa performed lab work on chloroplast microsatellites and ecological niche modeling (Chapter 3).

Finally, I express gratitude to those who were influential leading up to the dissertation: John Brick (Montville Township High School), Peter Mohler, Jason McLachlan, and Paul Manos.

My dissertation research was made possible by funding from the National Science Foundation Graduate Research Fellowship, Lewis and Clark Field Grant (American Philosophical Society), Sigma Xi Grants-In-Aid, and these sources from the University of Minnesota: Doctoral Dissertation Fellowship, Dayton and Wilkie Natural History Research Grants, Elmer Birney Fellowship, Carolyn Crosby Fellowship, Thesis Research Grant, Center for Community Genetics Research Grant, Florence Rothman Fellowship, and Block Grants.

Dedication

To Mom and Dad for always inspiring me to follow my passion.

Abstract

Paleobotanical records and molecular data from modern forests can provide a synergistic understanding of the ecological and evolutionary history of an organism. I used the fossil record to generate hypotheses that I tested with statistical phylogeographic methods for Douglas-fir (*Pseudotsuga menziesii*). In Chapter 1, I describe alternative scenarios of glacial refugia and postglacial migration based on compiled fossil pollen and macrofossil evidence from the late Quaternary. In Chapter 2, I test those hypotheses using coalescent analyses of mitochondrial and chloroplast DNA sequence data. I also test the paleobotanical hypothesis that Douglas-fir's two varieties diverged coincident with the Cascade orogeny in the late Pliocene. Finally in Chapter 3, I test whether Mexican Douglas-fir diverged from U.S. populations in the Miocene or Pleistocene, consistent with alternative interpretations of limited fossil evidence in the region. The present patterns of molecular variation in Douglas-fir are well-described by Pliocene (or early Pleistocene) divergence of its varieties, mid-Pleistocene colonization of Mexico, and restriction to multiple glacial refugia in the late Quaternary. Holocene expansion into Canada resulted in recontact among varieties and hybridization driven entirely by pollen dispersal but not seed dispersal. Douglas-fir populations have responded individualistically to past climatic and geologic change, such that some underwent expansions while others contracted to higher elevation and some diverged while others coalesced. These findings highlight the complementary insights that fossil and molecular data provide and can be used to inform the conservation and taxonomy of Douglas-fir.

Table of Contents

Acknowledgments …… i

Dedication ……..…… ii

Abstract ……... iii

Table of Contents …… iv

List of Tables ……..... v

List of Figures ……..……. vi

Introduction …….. 1

Chapter 1: Glacial populations and postglacial migration of Douglas-fir based on fossil pollen and macrofossil evidence ……..... **3**

Chapter 2: Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record …….. **37**

Chapter 3: Southward Pleistocene migration of Douglas-fir into Mexico: phylogeography, ecological niche modeling, and conservation of 'rear edge' populations ……... **76**

Conclusion ……...……. **111**

References ……. **113**

List of Tables

Chapter 1

Table 1.1. Conversion of ^{14}C yr BP to cal yr BP …… 27

Table 1.2. Migration rates ……... 27

Table 1.S1. Fossil pollen site information …….. 36

Table 1.S2. Macrofossil site information …… 36

Chapter 2

Table 2.1. Primers developed for this study ……... 63

Table 2.2. Sequence divergence, mutation rates and molecular clock tests ….. 63

Table 2.3. Diversity and neutrality measures for mtDNA and cpDNA …….. 64

Table 2.4. SAMOVA tables …… 66

Table 2.5. *P*-values for coalescent tests of population models ……... 66

Table 2.6. Divergence time estimates …… 67

Table 2.S1. Sample site information and haplotype frequencies …… 75

Table 2.S2. Mitotype and chlorotype definitions …… 75

Chapter 3

Table 3.1. Diversity measures for chloroplast data ……..... 99

Table 3.2. Diversity measures for mtDNA ……... 100

Table 3.3. Divergence time estimates ……... 101

Table 3.S1. Sampling site information and haplotype frequencies …….. 108

Table 3.S2. Mitotype and chlorotype definitions …….. 108

Table 3.S3. Definitions of cpSSR haplotypes ……... 109

Table 3.S4. SAMOVA and AMOVA tables ……. 110

List of Figures

Chapter 1

Figure 1.1. Map of study area and fossil sites ……. 28
Figure 1.2. Presence/absence maps of Douglas-fir at 3,000-year intervals …… 30
Figure 1.3. Hypothesized barriers among glacial populations …… 32
Figure 1.4. Change in elevation through time for six regions …… 33
Figure 1.5. Change in elevation through time for Eastern Grand Canyon ……. 35
Figure 1.S1. Distribution of fossil pollen sites ……... 36
Figure 1.S2. Distribution of macrofossil sites …… 36

Chapter 2

Figure 2.1. Hypotheses from the fossil record …… 68
Figure 2.2. Sampling sites and population groupings ……. 70
Figure 2.3. Population models for coalescent simulations ……. 71
Figure 2.4. Maps of haplotype distributions with haplotype networks ……... 72
Figure 2.5. Frequency distributions of *s* for tests of refugia hypotheses …… 73
Figure 2.6. Glacial refugia and postglacial migration routes …….. 74

Chapter 3

Figure 3.1. Map of sample sites and range of Douglas-fir in Mexico …….. 102
Figure 3.2. Maps of haplotype distributions with haplotype networks ……. 103
Figure 3.3. Correlation of diversity and latitude ……... 104
Figure 3.4. Bayesian skyline plot of N_e through time ……... 105
Figure 3.5. Population structure from SAMOVA and UPGMA …….. 106
Figure 3.6. Potential distributions based on ecological niche modeling ……... 107

Introduction

Understanding the ecological and evolutionary history of trees is critical to understanding how forests might respond to future climate change, how to best design conservation programs, and what are the major forces driving evolutionary change. For example, some models projecting potential distributions of plants in response to future warming are parameterized with past rates of population expansion (Morin *et al*. 2008). More generally, the past causes of natural population divergence are central to the study of evolution.

Fossil and molecular data offer complementary insights into the history of an organism. The fossil record can provide direct, dated evidence of a species presence, and sometimes abundance, at a particular site. However, fossil records are often incomplete through time and in space, most fossils are only identifiable to species or genus, and species-specific biases in pollen production, dispersal, and preservation mean that arbitrary thresholds must be used to determine presence. On the other hand, molecular data from modern forests can be used to infer historical changes in distribution or population size at the population level across an entire distribution but yield imprecise date estimates. The opposing strengths and weaknesses bring about a synergy of fossil and molecular data that can reveal population-level responses to particular past climatic and geologic events.

Two major insights have come from combined fossil and molecular analysis (Petit *et al*. 2008; Hu *et al*. 2009). First, populations within a species responded individualistically to changing environmental conditions; some remained small and isolated for long durations, whereas others expanded and contracted dramatically in response to cyclical glaciations (Magri *et al*. 2006). Second, molecular data suggest that some populations survived the Last Glacial Maximum (21 ka) far closer to the ice sheets than observed in the fossil record (McLachlan *et al*. 2005; Anderson *et al*. 2006). Northern glacial refugia mean low migration rates can explain poleward population expansion since the Last Glacial Maximum and suggest that trees might respond too slowly to track future climate change.

To build on this work, I have formally integrated fossil and molecular data into a statistical phylogeographic hypothesis-testing framework. The abundant plant fossil record is a rich source of *a priori* hypotheses that can be tested using coalescent simulations and other hypothesis-driven analyses of molecular variation. Similarly, others have used ecological niche models to estimate past species distributions, from which hypotheses of population divergence and expansion were derived (Carstens & Richards 2007). That approach works well for organisms with poor fossil records. However, direct evidence from the fossil record for well-represented taxa will be the best source of hypotheses as phylogeography turns to more statistical approaches.

In Chapter 1, I describe alternative scenarios of glacial refugia and postglacial migration based on compiled fossil pollen and macrofossil evidence from the late Quaternary. In Chapter 2, I test those hypotheses using coalescent analyses of mitochondrial and chloroplast DNA sequence data. I also test the paleobotanical hypothesis that Douglas-fir's two varieties diverged coincident with the Cascade orogeny in the late Pliocene. Finally in Chapter 3, I test whether Mexican Douglas-fir diverged from U.S. populations in the Miocene or Pleistocene, consistent with alternative interpretations of limited fossil evidence in the region.

Each chapter is divided into Abstract, Introduction, Materials and methods, Results, Discussion, Tables, Figures, and Supporting information. Some figures and tables listed in the Supporting information are provided as separate supplementary files because of their large sizes. Also, Figures 1.2 and 1.4 are duplicated in full size as supplementary files.

Chapter 1

Glacial populations and postglacial migration of Douglas-fir based on fossil pollen and macrofossil evidence

A version of this was published:
Gugger PF, Sugita S. *Quaternary Science Reviews*, doi:10.1016/j.quascirev.2010.04.022.

Abstract

To understand how temperate forests might respond to future episodes of global warming, it is important to study the effects of large-scale climate change brought about by rapid postglacial warming. Compilations of fossil evidence have provided the best evidence of past plant range shifts, especially in eastern North America and Europe, and provide a context for interpreting new molecular datasets from modern forests. In western North America, however, such reviews have lagged even for common, widespread taxa. Here, we synthesize fossil evidence for Douglas-fir (*Pseudotsuga menziesii*) from nearly 550 fossil pollen, sedimentary macrofossil, and packrat midden macrofossil sites to develop hypotheses about the species' late Quaternary history that can be tested with molecular phylogeographic studies. For both the coastal and interior varieties, we identified alternative hypotheses on the number of glacial populations and postglacial migration patterns that can be characterized as single-population versus multiple-population hypotheses. Coastal Douglas-fir may have been subdivided into two populations at the Last Glacial Maximum (LGM) and colonized British Columbia from populations in Washington and Oregon. Interior Douglas-fir could have been subdivided along major topographic barriers into at least three LGM populations and colonized British Columbia and Alberta from populations in northwest Wyoming and/or northeast Utah. For both varieties, we calculated migration rates lower than previous studies, which could have been as high as 100 – 220 m/yr if Douglas-fir reached its modern distribution 9000 cal yr BP, or as low as 50 m/yr if it reached its modern range at present. The elevational range of populations in California and the southern Rockies shifted upslope by 700 – 1000 m. If there were multiple LGM populations, these elevational shifts suggest that those populations did not contribute to the colonization of Canada. Our findings emphasize the possibility of low-density northern LGM populations and that populations within species react individualistically in response to large-scale climate change.

Introduction

Dated fossil pollen and macrofossil records from late Quaternary lake sediments have been compiled for many European and eastern North American species indicating that species ranges shifted dramatically as they tracked favorable climates. For example, Davis (1981; 1983) marked the first arrival of tree taxa in deglaciated regions in eastern North America and found that thermophilous trees spread at rates of 100 – 1000 m/yr into these landscapes from small populations restricted south of the ice sheets at the Last Glacial Maximum [LGM; 21 000 – 18 000 calendar years before present (cal yr BP)]. Similarly, Huntley and Birks (1983) reviewed the late Quaternary history of European trees, finding that many taxa were restricted to the Iberian, Italian and Balkan peninsulas during the LGM and expanded to their present distributions at rates up to 2000 m/yr. Because these rates are higher than those considered possible based on empirical estimations in modern forests (Clark *et al.* 2001), rare long-distance seed dispersal by weather or animal vectors has been proposed as a mechanism (Davis 1987; Johnson & Webb 1989; Clark *et al.* 1998).

For many species, this relatively simple scenario in which one or a few restricted populations concurrently and rapidly expanded northward in response to warming has been challenged as fossil data are re-interpreted and other independent types of data become available for comparison (Bennett & Provan 2008). In particular, molecular genetic data from modern populations have been used to infer past changes in distribution because the effects of past isolation and expansion on the distribution of genotypes often endure to the present (Hewitt 2000; Petit *et al.* 2003).

One key insight has been the detection of additional source populations farther north than previously thought (Stewart & Lister 2001). For example, American beech (*Fagus grandifolia*) and red maple (*Acer rubrum*) might have colonized their northern ranges from scattered, low-density populations near the ice margin (McLachlan *et al.* 2005) and not from a single, small refugium in the South (Davis 1981). Low-density populations could have escaped detection in the fossil record altogether or could have been overlooked where pollen percentages were below thresholds used to define

presence. Similarly, molecular data suggest low-density populations of white spruce (*Picea glauca*) persisted the LGM in Alaska in addition to the known LGM populations south of the ice (Anderson *et al.* 2006). In Alaska, *Picea* pollen accounts for less than 2% of pollen assemblages dating to the LGM (Brubaker *et al.* 2005). If populations were farther north during the LGM than previously thought, postglacial migration rates were likely much lower than previously thought (< 100 m/yr), more consistent with empirical observations of seed dispersal (Clark *et al.* 2001).

Another important insight is that often only a subset of populations expands, while others remain stable or even shrink. For example, the northern latitudes of the range of red maple (McLachlan *et al.* 2005; Gugger *et al.* 2008) were only colonized by populations near the ice margin and not by populations to the south. A comparison of fossil and molecular data for European beech (*Fagus sylvatica*; Magri *et al.* 2006) revealed that populations from Mediterranean countries did not contribute to the colonization of central and northern Europe, as was inferred in early fossil studies in the region (Huntley & Birks 1983). And in some cases, those populations may have been isolated for more than one glacial cycle. Instead, populations from central Europe expanded to colonize middle and high latitudes.

Fossil surveys and their comparison to molecular datasets in western North America have lagged those in Europe and eastern North America because the climate history and topography are more complex and generally fewer fossil sites have been analyzed. Phylogeographic studies in the region suggest migration into western Canada from multiple northern source populations in Washington and the northern U.S. Rocky Mountains (Richardson *et al.* 2002; Carstens *et al.* 2005; Mimura & Aitken 2007; Godbout *et al.* 2008; O'Connell *et al.* 2008), and some fossil studies concur (Tsukada 1982; Rosenberg *et al.* 2003). In the Southwest, more complex histories of long-term isolation, elevational shifts, and migration are suggested in fossil (Spaulding *et al.* 1983) and molecular (Mitton *et al.* 2000; Galbreath *et al.* 2009) data. For temperate trees set in a topographically complex region like western North America, LGM populations could have been scattered and discontinuous and only some may have expanded in response to

postglacial climate warming, while others retreated to higher elevation or migrated regionally.

Combined fossil and molecular studies have highlighted some limitations of the fossil record, which can include difficulty detecting low-density populations and the inability to resolve the independent histories of different populations within a species. However, fossils still provide the most direct, dated evidence of a species presence or absence through time compared to the imprecise date estimates inferred from molecular data. Without the backdrop of fossil evidence, patterns of molecular variation cannot confidently be attributed to particular geologic or climatic events. Moreover, the fossil record is an important source of hypotheses as phylogeography turns to more sophisticated statistical methods requiring *a priori* hypotheses (Knowles & Maddison 2002; Carstens & Richards 2007).

We synthesize late Quaternary fossil records of Douglas-fir (*Pseudotsuga menziesii*) for the specific purpose of estimating the location of populations during the LGM and patterns of postglacial range shifts that can be tested in future surveys of molecular variation. To do so, we have compiled fossil pollen, sedimentary macrofossil, and packrat (*Neotoma* spp.) midden macrofossil data from hundreds of previously published records spanning almost the entire modern range of Douglas-fir and extending back to just before the LGM. Our approach emphasizes the possibility of low-density glacial populations and the possibility that populations had independent histories. The various interpretations are discussed as alternative hypotheses that can be tested with future molecular phylogeographic studies and additional fossil evidence.

Study organism

Douglas-fir is a wide-ranging economically and ecologically important conifer in western North America. Two varieties are commonly recognized (Hermann & Lavender 1990): coast Douglas-fir (*P. menziesii* var. *menziesii*) and interior or Rocky Mountain Douglas-fir (*P. menziesii* var. *glauca*). Coast Douglas-fir is found from central California to central and coastal British Columbia at elevations from sea level near the coast to 2300 m

in the Sierra Nevada. Interior Douglas-fir ranges from British Columbia and Alberta to central Mexico, though Mexican populations are occasionally regarded as a different variety, or even species (Martínez 1949). Compared to coast Douglas-fir, interior Douglas-fir inhabits drier, higher elevation sites primarily in the montane zone but can be found from 500 m to 3200 m depending upon latitude and local conditions. The range becomes increasingly fragmented in the Southwest and Mexico. The two varieties meet in a transition zone in southern, central British Columbia (Li & Adams 1989; Ponoy *et al.* 1994).

Douglas-fir's late Quaternary history has been partially reviewed in two previous publications (Tsukada 1982; Hermann 1985). Tsukada (1982) focused on the Pacific Northwest and using fossil pollen records found that Douglas-fir migrated north from the Willamette Valley of northern Oregon into southern British Columbia at a rate of about 450 m/yr. Subsequent studies cast doubt on the precise location of populations at the LGM (Barnosky 1985a; Worona & Whitlock 1995). Hermann's (1985) survey was broader in temporal and spatial scope, encompassing most of the range and extending back to the earliest fossil records in the Miocene and Oligocene. Hermann's (1985) interpretation generally agreed with that of Tsukada (1982) in the Northwest. In the interior, he concluded that Douglas-fir rapidly migrated into southwestern Alberta from populations in the southern Rockies. Since then, a number of new fossil records have been published, which may shed new light on the late Quaternary history of Douglas-fir.

Regional setting

The topography of western North America is characterized by two sets of major north-south mountain chains: the combination of the Rocky Mountains and Sierra Madre and the combination of the Coast Range, Sierra Nevada, and Cascade Mountains (Fig. 1.1). The arid Great Basin and Columbia Plateau separate these two sets of mountains. Today, the climate varies substantially from cool, moist Pacific coastal habitats to the drier, summer-monsoonal Rocky Mountains to the intervening arid deserts.

Late Quaternary climate change in western North America is as complex as the geography, owing to changes in circulation in a topographically diverse region. Generally, the LGM was attained about 18 000 cal yr BP during the maximum extent of the Fraser (Northwest), Tioga (California), and Pinedale (Rockies) advances of the Cordilleran ice sheet and mountain glaciers, some 3000 years after the Laurentide glacial maximum in eastern North America (Porter *et al.* 1983). Dry valleys and glaciated mountains predominated. Climate was generally cooler and drier than present in the Pacific Northwest with subalpine parkland growing close to the ice margin (Whitlock 1992). The northern Rockies (north of Great Divide Basin; Fig. 1.1) were also cooler and drier and valleys were largely treeless with permafrost. In both regions, temperate trees could have grown along mountain flanks. The southern Rockies and nearby arid lands were generally cooler and wetter-than-present (Barry 1983; Thompson *et al.* 1993; Bartlein *et al.* 1998). By the early Holocene (~9000 cal yr BP), warmer and wetter-than-present conditions prevailed in the Southwest and drier-than-present conditions prevailed in the Pacific Northwest (Whitlock 1992; Thompson *et al*. 1993).

Materials and methods

Fossil pollen and macrofossil data were compiled for the parts of western North America that currently or historically may have supported Douglas-fir populations (Tables 1.S1 and 1.S2; Figs. 1.S1 and 1.S2). These include British Columbia and Alberta, all of central and northern Mexico, and the western states of Washington, Oregon, California, Arizona, Utah, Idaho, Montana, Wyoming, Colorado, New Mexico and western Texas. To simplify the data set for the purposes of presentation and analysis, we recorded Douglas-fir fossils as "present," "rare," or "absent" at 3000 calendar year intervals from before the Last Glacial Maximum to the present (40 000 – 21 000; 21 000 – 18 000; 18 000 – 15 000; 15 000 – 12 000; 12 000 – 9000; 9000 – 6000; 6000 – 3000; 3000 - 0 cal yr BP), where for example, 3000 – 0 cal yr BP includes all samples with ages, ignoring error, that fall within that interval. We chose 3000-year intervals because many fossil

sites only contained partial records through time, many radiocarbon dates had confidence intervals of 500 -1000 years, and 3000 years is considerably higher resolution than that attainable in molecular phylogeographic studies.

Most dates of fossils were originally reported in radiocarbon years (^{14}C yr BP), so conversions to calendar years were made according to Table 1.1. The radiocarbon age cutoffs were first estimated by eye using the IntCal04 curve (Reimer *et al.* 2004) and then verified by entering that year with a 100-year standard error into each of two calendric age conversion programs: Calib version 5.0.1 (http://calib.qub.ac.uk/calib) and CalPal (http://www.calpal.de).

Fossil pollen

For the above-mentioned regions, all 96 fossil pollen records from lake or bog sediments were extracted from the North American Pollen Database (NAPD; http://www.ncdc.noaa.gov/paleo/napd.html). Douglas-fir pollen percentages were calculated for each sample by dividing the Douglas-fir pollen (plus *Larix*; more on this below) count by the pollen sum, not including aquatic/bog plants or indeterminable pollen. These were then converted to the three categories, where "present" is greater than 0.5%, "rare" is greater than 0% but less than 0.5%, and "absent" is 0%.

To this list, we added published fossil pollen records not found in the database, bringing the total number of records to 316 (Table 1.S1, Fig. 1.S1). All records had date controls, either radiocarbon dates or tephra with a widely accepted date (*e.g.* Mazama ash, 7600 cal yr BP; (Zdanowicz *et al.* 1999). When tephra were the only source of date estimates (especially, Hansen 1947; Heusser 1960), we did not extrapolate beyond that date to be conservative, but when tephra were accompanied by other chronological evidence presented by the author, we deferred to the author's extrapolations and interpolations (especially, Heusser 1960, 1964). No attempt was made to acquire the original pollen counts; rather, we simply converted pollen percentage diagrams to present, rare, or absent for each time period by visual inspection. Every effort was made to use the same pollen percentage thresholds in recording presence/absence as used for

the NAPD. However, because different authors used different presentations, we were occasionally forced to define rare as between 0% and 1%.

Regardless of how rare is defined, very low Douglas-fir pollen representation might indicate Douglas-fir's presence. Douglas-fir pollen is usually underrepresented in the fossil record due to relatively poor production, dispersal, and preservation. Douglas-fir produces lower quantities of pollen compared to other conifers (*e.g. Pinus*) and many other wind-dispersed trees (Mack *et al.* 1978; Baker 1983). Most pollen falls close to the tree: the 90% characteristic radius (Prentice 1988; Sugita 2007) of Douglas-fir pollen is 2320 m based on Prentice's model (Prentice 1985) and 1440 m based on Sugita's model (Sugita 1993), assuming that basin size is 10 ha, fall speed of its pollen 0.127 m/s (Eisenhut 1961), and wind speed 3.0 m/s. Some evidence suggests that, once deposited, it preserves poorly because it is highly susceptible to fungal decomposition (Goldstein 1960), though some do not consider this an issue (Whitlock, personal communication). Thus only one or a few pollen grains might be considered evidence of Douglas-fir's presence in the area (though not necessarily locally), especially when considering the range-wide scale of interest here (Tsukada 1982).

In other poorly represented species, such as American beech, a 0.5% threshold is often used to define presence, and analyses of surface sediments have shown that false negatives are more common than false positives with that threshold (McLachlan & Clark 2004). Of course, false-presences do occur at low rates due to the occasional long-distance transport of pollen. In modern surface samples, Douglas-fir pollen is often less than 1 – 2 % when present in surrounding forests (Whitlock 1993; Williams *et al.* 2006). In most cases where Douglas-fir pollen was rare (< 0.5%) in our study, it appears in multiple samples within a given time interval, which in rare pollen types may be considered evidence of a species presence (Smith & Pilcher 1973; McLachlan & Clark 2004). For these reasons, most of our analyses consider low pollen percentages evidence of Douglas-fir's regional presence, but maps distinguish very low percentages (< 0.5%) to capture uncertainty throughout.

In most fossil pollen studies, Douglas-fir and larch (*Larix*) pollen are not distinguished. Their pollen grains are spherical and inaperturate, and similar in size and

appearance under light microscopy. This introduces uncertainty in the Canadian and northernmost U.S Rocky Mountains, where ranges overlap. Western larch (*L. occidentalis*) and subalpine larch (*L. lyallii*) are currently and probably have been restricted to the interior northwest. Western larch co-occurs with Douglas-fir in the montane zones of Montana, Idaho, eastern Oregon and Washington, and the southern Canadian Rockies. This presents the most likely source of ambiguity, and macrofossil evidence suggests both can be represented in a pollen assemblage from a particular site (Whitlock 1995). Subalpine larch is rarer and restricted to very high elevation in subalpine zones. An eastern, boreal larch (*L. laricina*), enters the study area in Alberta, east and north of the Rocky Mountains, but does not overlap the range of Douglas-fir. Records that are thought to reflect *Larix* and not *Pseudotsuga* pollen are noted in Table 1.S1 and were decided based on the claims of the original publication and/or on an unusual geographic position (*e.g.* high elevation, high latitude) relative to Douglas-fir's modern range. *Larix* pollen is also underrepresented in the fossil record and thus would be unlikely to drown out the signal of Douglas-fir (Brubaker *et al*. 2005).

Macrofossils

Two-hundred sixteen packrat midden macrofossil records from over 100 localities across western North America were retrieved from the North American Packrat Midden Database (http://esp.cr.usgs.gov/data/midden/; Table 1.S2, Fig. 1.S2). These include all records mentioning Douglas-fir (present or absent) plus most of the sites that do not mention Douglas-fir but clearly would have reported it if it were present. Those not explicitly mentioning Douglas-fir fall into three categories: 1) complete taxa lists reported, but no Douglas-fir, 2) partial taxon lists with no Douglas-fir, but reported taxa clearly limit midden site to habitat not favorable to Douglas-fir, such as low elevation deserts, or in one case, 3) partial taxa lists where it is unclear if Douglas-fir was or was not present (see Table 1.S2). Some midden sites that did not report Douglas-fir and were redundant in well-sampled areas, such as southern Nevada in particular, are not included here.

In all cases, Douglas-fir is identified to the species-level (*P. menziesii*). Its presence in a given midden site indicates that Douglas-fir was likely a component of the plant community within about 30 - 50 m (rarely over 100 m) of the site (Finley 1990). A survey of modern middens by Lyford *et al*. (2004) revealed that in nearly 90% of cases where Douglas-fir was living within 50 m of a modern midden site, it was present in the midden, and when representing over 25% of the nearby forest cover it was always present. False-absences occurred in several cases when Douglas-fir was a minor component (<20%) of the surrounding forests. False-presences are unlikely, except if contamination across time horizons is considered.

The Packrat Midden Database already assigns samples as absent (0), rare (1), or present (2), where rare is designated by the original author and/or database administrators (Strickland *et al.* 2001) to control for the possibility of contamination from another time horizon. We consider these roughly comparable to our three pollen categories, where rare indicates a high degree of uncertainty. All samples included here were radiocarbon dated, though the material dated varies widely including various plant fragments, packrat fecal pellets, debris, and other midden components. In cases where multiple dates are given for a single taxa list, we chose the author-preferred date or a date directly from a Douglas-fir macrofossil over other dates. Where multiple dates seemed plausible, we averaged the dates, though these were usually fairly close to each other (noted in Table 1.S2).

To the midden macrofossil list, we added 14 fossil sites for which Douglas-fir macrofossils were found in dated sediment. Seven of these sites are from the North American Plant Macrofossil Database (NAPMD; http://www.ncdc.noaa.gov/paleo/plantmacros.html) and seven are from various published sources (Table 1.S2, Fig. 1.S2). All of these cases strongly suggest Douglas-fir's presence, thus rare is not distinguished. We have not reported Douglas-fir absences from other sites reported in the database or elsewhere because macrofossils in lake sediments are often rare or absent even when the plant of interest is abundant as inferred from pollen abundance or as observed in comparisons of local plant composition with surface sediments (Jackson *et al.* 1997).

Migration patterns

Patterns of migration were observed by mapping Douglas-fir fossils as present, rare, or absent at 3000-year intervals. Clusters of sites with Douglas-fir fossils and solitary sites with macrofossils or high pollen percentages were presumed good evidence for glacial populations.

We assumed separate glacial populations for each of the commonly recognized varieties because their divergence must predate the last glacial cycle (Li & Adams 1989; Hermann & Lavender 1990). Within varieties, we attempted to approximate a number of independent LGM populations by delineating LGM populations along large geographic gaps and/or major topographic barriers, which assumes these geographic features isolated those populations. This does not exclude the possibility that additional population substructure existed.

Limited spatial and temporal resolution precluded the ability to construct meaningful latitudinal transects beyond those already published (Tsukada 1982). Nonetheless approximate migration rates from one time interval to the next were calculated by measuring the geographic distance (WGS84 reference system) from leading edge populations in the first interval to leading edge populations in the next and dividing by the duration of a time interval (3000 years). An overall minimum mean rate was estimated assuming no barriers and was calculated from the distance from the ice margin at 18 000 cal yr BP to the present-day northern limit divided by 18 000 years. A "maximum" mean rate was calculated assuming Douglas-fir started from the observed northernmost LGM population (where fossils present, not rare) at 18 000 cal yr BP and tracked the receding ice until reaching its present range limit at 9000 cal yr BP. This ignores possible barriers and indirect routes, so is not necessarily a true maximum.

To examine trends in elevational migration, we displayed elevation at which Douglas-fir was present versus time in boxplots for each of six regions: California, western Oregon/Washington (coast to Cascades), western British Columbia (west of Canadian Rockies), southern Rockies (south of Great Divide Basin; Fig. 1.1), northern

Rockies (north of Great Divide Basin to Canadian border), and Canadian Rockies. First, all data where Douglas-fir fossils were present and rare were included, and second only data points where Douglas-fir fossils were present were included. Rare was excluded in the latter approach because low amounts of wind or water transported pollen can settle at elevations much higher or lower than the source. To complement our regional analyses, we created an elevational transect marking Douglas-fir presence and absence through time according to macrofossil data in the eastern Grand Canyon, the only locality with enough data to meaningfully do so. Several scenarios might be meaningful: 1) Elevational range shift should manifest as increasing minima, means (or medians), and maxima through time; 2) Elevational range expansion would show rising elevation maxima, constant minima, and either constant or slightly rising means (or medians); and 3) No change (elevational stasis) would show no relationship of minima, means, or maxima through time. Scenarios 1 and/or 2 might be expected in regions continuously inhabited since the LGM, whereas scenarios 2 and/or 3 might be expected in regions colonized after the LGM.

In all of the above analyses, we excluded sites suspected of having recorded *Larix* pollen (Table 1.S1), but this did not alter the results in any meaningful way (data not shown).

Results and Discussion

Glacial populations: 40 000 – 18 000 cal yr BP

Rocky Mountain variety. Glacial populations were located throughout much of the southern and central Rockies (Fig. 1.2), including strong macrofossil evidence for populations near the Grand Canyon, along the Mogollon Rim in central Arizona, in southern New Mexico and in eastern Utah. It is also possible that Douglas-fir was as far north as the Yellowstone region as suggested by very low pollen percentages in two glacial sediment samples at Hedrick Pond (Whitlock 1993) and in samples from 16 700

cal yr BP at nearby Cub Lake in Idaho (Baker 1983). Long-distance pollen transport cannot be ruled out, but when considering the underrepresentation of Douglas-fir pollen, it seems unlikely that pollen found in multiple basins and multiple sediment layers came from glacial populations in northeast Utah over 350 km away (Sugita 1993; Sugita 1994), the nearest Douglas-fir population supported by macrofossil evidence (Dutch John Mountain, Jackson *et al.* 2005). Moreover, low Douglas-fir pollen percentages (<1%) have been accompanied by macrofossils at other time horizons at other nearby sites, suggesting low pollen percentages sometimes indicate local presence (Fallback Lake; Whitlock 1993). Therefore, Douglas-fir populations may have persisted the LGM in the Yellowstone region or northern Rockies more generally. Montane species could have survived in the northern Rockies along unglaciated mountain flanks (Thompson *et al*. 1993; Whitlock 1995).

Some mountains may have been barriers to dispersal that separated LGM populations, and certain populations that are presently isolated might have been continuous during the LGM (Fig. 1.3) if Douglas-fir was 700 m below its present elevational range (Fig. 1.4). Some high elevation barriers could include the glaciated mountains of northwest Wyoming, Utah's Uinta Mountains, and the Colorado Rockies, and some prominent low elevation barriers could include the Great Divide Basin in southwest Wyoming and the Colorado and Green River system. These putative barriers would have partitioned the Hedrick Pond site (northwest WY) from the Dutch John Mountain site (northeast UT), the Dutch John Mountain site from all sites to the south, and the sites on either side of the Grand Canyon (Fig. 1.3). Unfortunately, the absence of Douglas-fir between these putatively independent populations is uncertain because intervening LGM fossil sites have not been investigated. Phylogeographic studies in other species have found genetic breaks along the Great Divide Basin (Demboski & Cook 2001; Barrett & Freudenstein 2009), Colorado Rockies (Galbreath *et al.* 2009; Jaramillo-Correa *et al.* 2009), and Colorado River system (Mitton *et al.* 2000; Conroy & Cook 2000). The implication of such barriers is that only a subset of populations would have expanded into Canada, while others remained farther south.

Several external considerations support a hypothesis of subdivided LGM populations for Rocky Mountain Douglas-fir. First, in a survey of allozyme variation, Li and Adams (1989) found that northern and southern Rocky Mountain populations differed, suggesting they had been isolated through at least the last glacial cycle. Similarly, Critchfield (1984) argues that discontinuities in leaf terpene chemistry among geographic regions suggest multiple Pleistocene populations (Zavarin & Snajberk 1973). Finally, response surface models projecting the past distribution of Douglas-fir based on simulated climate from a general circulation model suggest that the Rocky Mountain variety was fragmented (Bartlein *et al.* 1998), though it is not clear where the exact partitions among LGM populations may have been. However, this model should be interpreted cautiously because populations were also predicted far outside the current range (*e.g.* Alaska). Nonetheless, it is interesting that the model seems to suggest low-density populations close to the ice margin which may have been separated from the various pockets of populations in the southern Rockies - a scenario that may be confirmed as more fossil sites are analyzed in the northernmost U.S. Rockies.

Coastal variety. The fossil evidence presented here suggests two distinct populations during and before the LGM (40 000 - 18 000 cal yr BP): one in unglaciated western Oregon and Washington and the other in the San Francisco Bay area of California (Fig. 1.2). Those regions are separated by the Klamath Mountains of northern California and southern Oregon. The absence of Douglas-fir fossils in-between at Twin Lakes (CA; Wanket 2002; West *et al*. 2007), Grass Lake (CA; Hakala & Adam 2004), and Caledonia Marsh (OR; Hakala & Adam 2004) could be considered corroborating evidence. Consistent with this split, leaf terpene composition shows a break along the Klamath Mountains, with high diversity to the south and uniformity to the north (von Rudloff 1972; Zavarin & Snajberk 1973; von Rudloff & Rehfeldt 1980; Critchfield 1984). Moreover, a genetic break is found along the Klamath Mountains in many plant and animal taxa (Soltis *et al.* 1997; Swenson & Howard 2005; Jaramillo-Correa *et al.* 2009).

However, allozyme (Li & Adams 1989) and nuclear DNA (Krutovsky *et al.* 2009) variation in coastal populations is not strongly geographically partitioned. In addition,

the models of Bartlein *et al*. (1998) suggest Douglas-fir was both abundant and continuously represented in the U.S. Pacific coastal states from southern California to central Washington. Therefore, it is also possible that the two putative LGM populations are an artifact of the limited number of fossil sites (especially low elevation) in northern California and southern Oregon from before 15 000 cal yr BP, and rather, there was a continuous population from coastal southern California to Washington. The more consistent coastal climate since the LGM (Bartlein *et al*. 1998) and much reduced topographic complexity compared to the Rockies argue that Douglas-fir's range was not nearly as fractured as in the Rockies.

Regardless, the presence of glacial populations relatively close to the ice margin is evident. Tsukada (1982) observed Douglas-fir pollen in the Willamette and unglaciated Puget lowland in full-glacial sediments and based on his more limited dataset argued that Douglas-fir did not likely exist north of about 46°N. This conclusion was based on low pollen percentages in LGM sediment at Fargher Lake, Onion Flats, and Silverton Bone Site Bog. Recurring low pollen percentages (< 1.5 %) during the LGM have also been found at Davis Lake (Barnosky 1981) and Battle Ground Lake (Barnosky 1985a) in the Puget Trough. None of these sites has high pollen percentages, which suggests either small, local populations or larger populations elsewhere in the region.

In the region, the LGM location of large Douglas-fir populations remains elusive. Very low pollen percentages were observed at Carp Lake in the southwestern Columbia Basin (Barnosky 1985b) and possibly at Bogachiel River Bog along the coast (Heusser 1978). Coastal Washington has been identified as a glacial refugium for tree taxa that currently co-occur with Douglas-fir (Heusser 1972). In the central Coast Range pollen and macrofossils were absent during the LGM, but present just before and after (Little Lake, Worona & Whitlock 1995). Sites in the central Cascades seem to suggest a similar pattern, though no pre-LGM data are available (Indian Prairie Fen, Gordon Lake; Sea & Whitlock 1995; Grigg & Whitlock 1998). These patterns agree with those at Battle Ground Lake, where a macrofossil marks the presence of Douglas-fir before the Vashon Stade and high pollen percentages mark its high abundance afterwards (Barnosky 1985a). Low pollen percentages throughout the region at glacial times followed by higher pollen

percentages might suggest expansion from low-density populations. Some of these small populations could have been farther north than suggested by Tsukada (1982), for example at Davis Lake or Bogachiel Bog. A large body of new research suggests that low-density, northern LGM populations were more common than previously realized (McLachlan *et al.* 2005; Anderson *et al.* 2006; Petit *et al.* 2008).

Post-LGM migration: 18 000 – 9000 cal yr BP

Rocky Mountain variety. The presence of glacial populations in the Yellowstone area or northern Utah modifies our previous understanding of Interior Douglas-fir's postglacial migration. Tsukada (1982) and Hermann (1985) have argued that Douglas-fir from the southern Rockies colonized Canada, reaching northeastern Washington and northern Montana and Idaho by 13 000 – 11 000 cal yr BP and southwestern Alberta by 10 000 cal yr BP, implying a migration rate of nearly 200 m/yr. However, both the Dutch John Mountain (Uinta Mtns.; Jackson *et al.* 2005) and Hedrick Pond/Cub Lake (Yellowstone area; Baker 1983; Whitlock 1993) evidence were unknown at the time, and now offer strong support for more northern glacial populations that colonized the Canadian Rockies, which would reduce the rate to about 120 m/yr over that interval. Moreover, the migration rate could be further lowered if glacial populations are found closer to the ice margin as more sites are analyzed, though they also could be raised if the northern limit was reached sooner than assumed here (9000 cal yr BP). Considering all the uncertainties, we estimate a minimum northward migration rate of 50 m/yr and a maximum of 220 m/yr in the Rockies (Table 1.2).

Douglas-fir LGM populations also appear to have expanded eastward into the Bighorn Mountains (Wyoming) and Highwood Mountains (Montana) in the northern Rockies and into the northern Colorado Rockies in the southern Rockies. Mean rates ranged from 10 – 38 m/yr, but that migration was mostly observed 12 000 – 9000 cal yr BP at rates of 61 - 114 m/yr (Table 1.2).

If multiple LGM populations existed, then the data presented here suggest that southern Rockies populations primarily retreated to higher elevations rather than

expanding northward (Figs. 1.4 and 1.5). An elevational shift occurred from 18 000 – 6 000 cal yr BP as summarized in our regional data (Fig. 1.4) showing an elevation shift of 700 – 900 m in the full data set, 300 m in the fossil pollen data, and 600 m the macrofossil data. The pattern in the macrofossil data is much clearer, shows less elevational variation at each time period, and reflects the lower end of the elevational range of Douglas-fir in the southern Rockies. This is consistent with the generally low elevation habitat of packrats relative to Douglas-fir (Betancourt *et al.* 1990) and the higher abundance of lakes for sediment cores at higher elevations in the Southwest. Macrofossil evidence from the eastern Grand Canyon locality also shows that Douglas-fir's elevational range shifted upward by about 500 m (Fig. 1.5), consistent with our regional observations (Fig. 1.4). This shift is consistent with other taxa in the region (Maher 1961; Maher 1963; Spaulding *et al.* 1983; Hall 1985).

In contrast, northern Rocky Mountain and, once colonized, Canadian Rocky Mountain populations may have persisted at a similar elevational range during northward migration (Fig. 1.4). Climate along elevational gradients in the northern Rockies is determined by the balance of summer monsoonal and subtropical high-pressure climate regimes (Whitlock & Bartlein 1993). Although the strength of each has varied, the boundary between has remained relatively stable throughout the Holocene, which might explain some of the apparent stability in the elevational distribution of Douglas-fir in the northern Rockies. However, confusion with larch pollen and sparse data in the region limit the strength of this conclusion.

Larix pollen also confounds our ability to detect the time of first colonization in the Canadian Rockies. On the basis of unusually high elevation (> 2000 m) and unusual coordinates (*e.g.* in the Plains) relative to the modern range of Douglas-fir, we assume that *Pseudotsuga*/*Larix* pollen found in Alberta and the Canadian Rockies from 21 000 – 12 000 cal yr BP was *Larix* pollen except for Twin Lakes, BC (1100 m; Hazell 1979) at 12 000 cal yr BP. Even if we are wrong about the *Larix* sites in the Canadian Rockies, it would not alter the overall mean migration rate or migration pathway we conclude, because more northern sites are eventually colonized by Douglas-fir (*e.g.* Jasper; Heusser 1956; also see present range). Douglas-fir is not thought to have ever occupied the Plains

(Hermann 1985), which is supported by the lack of Douglas-fir macrofossils in Alberta's plains in general and the presence of *Larix* macrofossils in at least one site, Muskiki Lake (Kubiw *et al.* 1989).

Coastal variety. The migration of coastal populations is better documented and not confounded by *Larix* pollen misidentification, yet the overall pattern is consistent with that of the Rockies. California populations underwent a pronounced elevational expansion/shift of over 1000 m (Fig. 1.4) as they colonized the Sierra Nevada (12 000 – 9000 cal yr BP) and various mountain chains of northern California, presumably responding to the warming and drying of the region (Barry 1983; Bartlein *et al*. 1998). Migrating populations from Washington/Oregon LGM populations closely tracked the receding Cordilleran ice northward into coastal and central British Columbia (Fig. 1.2), accompanied by some expansion into higher elevations (Fig. 1.4). Accounting for uncertainties, we estimate minimum and maximum northward mean migration rates of 50 – 100 m/yr, though higher rates could have been achieved at particular time intervals (Table 1.2). If glacial populations in California and western Washington/Oregon were isolated from one another, it remains unclear from which source the intervening area in northern California and southern Oregon was colonized. Douglas-fir macrofossils at the Willow Creek site on Santa Cruz Island 16 600 cal yr BP suggests limited southward migration with temporary inhabitation when the island was connected to mainland California during the last glacial period (Anderson *et al.* 2008). Finally, some eastward migration into the Sierra Nevada and east Cascades also occurred from 12 000 – 9 000 cal yr BP (Fig. 1.2; Table 1.2).

Tsukada (1982) estimated the rate of Douglas-fir migration in the Pacific Northwest to be 450 m/yr, more than four times that estimated here. Several considerations could explain this discrepancy. First, Tsukada (1982) only considers northern Oregon to the southernmost edge of British Columbia spanning a roughly 2000-year period. The resolution of his analyses was finer than ours, and thus for certain periods of time higher rates could have been possible. Our 3000-year intervals mask the fact that forests probably shifted back and forth as ice expanded and contracted on

decadal and centurial scales. However, Tsukada (1982) assumed LGM populations south of 46°N, south of our observed northernmost LGM populations, which would produce a greater migration distance and rate. In part, this is because the threshold value of pollen percentage when marking time of first arrival was set higher than ours, which means possible more northern low-density populations during the full glacial were undetected (*e.g.* Hall Lake, Bellingham Bog). Additionally, we infer that populations closely tracked ice sheet dynamics and could have persisted at the glacial maximum in nearby Washington because populations of Douglas-fir are found on Vancouver Island before the LGM during the Olympic Interglacial in both pollen and macrofossil data (*e.g.* Cowichan Head, Dashwood; 40 000 to < 25 000 cal yr BP; Alley 1979). Finally, if our migration rates for the coast are off, then it seems that they are likely too high. For example, only the Rainbow Lake site (Heusser 1960) marks the arrival of Douglas-fir from 12 000 – 9000 cal yr BP at what is near its current northernmost coastal site. Chronological determinations for this site were based on correlations with other radiocarbon dated sites in the region, which may not hold after direct dating, and could reduce the migration rate. As a result, we think that Tsukada's (1982) estimates are too high, and considering our upper bound migration rate estimates, certainly could not have been maintained over millennia as Douglas-fir expanded into Canada.

Our migration rate estimates for both varieties are roughly consistent with rates predicted from theoretical models (80 m/yr) based on average modern seed dispersal patterns and survival rates (Thompson & Schorn 1998). In addition, they are consistent with rates estimated in western North American species with extensive macrofossil records, such as Utah juniper (*Juniperus osteosperma*; 110 m/yr; Lyford *et al.* 2003) and single-leaf (*Pinus monophylla*; 40 – 130 m/yr) and two-leaf (*Pinus edulis*; 60 – 80 m/yr) pinyon pines (Thomas 1983; Spaulding 1984). Low migration rate estimates are becoming more common as cryptic northern glacial populations are discovered in combined analyses of fossil and molecular data (McLachlan *et al.* 2005; Anderson *et al.* 2006; Godbout *et al.* 2008; Petit *et al.* 2008).

Modern range: 9000 – 0 cal yr BP

By 9000 cal yr BP the modern range took shape. The northern limit appears roughly defined, the major coastal and interior mountains were well populated, and even isolated mountain chains, such as those in central Wyoming (*e.g.* Sherd Lake, Bighorn Mtns.; Burkart 1976) and central Montana (Lost Lake, near Highwood Mtns.; Barnosky 1989) were inhabited by Douglas-fir (Fig. 1.2). Moreover, Tsukada (1982) and Critchfield (1984) suggest that the two varieties may have come into contact around this time (7700 cal yr BP) in British Columbia, though we think this remains unclear due to insufficient sampling and lack of variety-level resolution. Also by this time (9000 – 6000 cal yr BP), elevational migration appears to have plateaued (Fig. 1.4). Finally, in some sites (*e.g.* Dog Lake, BC, Hallett & Hills 2006; Poulsbo Bog, WA, Tsukada 1982; Swan Lake, ID, Bright 1966), pollen abundances reach modern levels suggesting that modern Douglas-fir population densities were achieved.

However, this period was not entirely static, and in fact, there has been a slight cooling trend since 6000 - 4000 cal yr BP (Wanner *et al.* 2008). From 9000 – 5000 cal yr BP, a number of northern records show higher Douglas-fir pollen abundance than the last 6000 – 4000 years. Examples include Misty Lake on northern Vancouver Island (BC; Lacourse 2005), which peaks 11 000 – 7500 cal yr BP; Burnt Knob Lake (ID; Brunelle & Whitlock 2003) from 9700 to 6000 cal yr BP; Lost Trail Pass Bog (MT; Mehringer *et al.* 1977) from 8000 – 5000 cal yr BP; and Lily Lake (WY; Whitlock 1993) before 7000 cal yr BP. Together, these imply some late Holocene range contraction or reduction in abundance at higher elevations and northern latitudes as climate cooled. In contrast, many records in the Pacific Northwest show a marked increase of Douglas-fir pollen after 5000 – 4000 cal yr BP, which corresponds with increasing moisture in the region. These include Gold Lake Bog (Sea & Whitlock 1995), Bolan Lake (Briles *et al.* 2005), and Gordon Lake (Grigg & Whitlock 1998) in Oregon; Battle Ground Lake (Barnosky 1985a), Carp Lake (Barnosky 1985b), and Bonaparte Meadows (Tsukada 1982) in Washington; and Little Qualicum Falls (Heusser 1960) in British Columbia. One exception is the decline in Douglas-fir's pollen abundance of coastal Washington which

is thought to reflect the reduced role of fire and consequent rise of western hemlock (Rosenberg *et al.* 2003) and western redcedar forest (*e.g.* Hoh Valley, Tsukada 1982; Heusser 1974). In the Southwest, records are mixed with some showing increased abundance (*e.g.* Posy Lake, UT, Shafer 1989; Bear Lake, AZ, Weng & Jackson 1999) and some showing decreased abundance (*e.g.* Fracas Lake, AZ; Weng & Jackson 1999) depending on local changes in climate. Pollen percentages are often difficult to use for quantifying actual changes in vegetation abundance (Sugita 1994). Taken together, however, these changes in Douglas-fir pollen percentages suggest that the major changes in Douglas-fir populations over the last 9000 years were primarily changes in abundance, rather than changes in distribution.

Support for our approach

The data from 3000 – 0 cal yr BP seem to capture the modern distribution of Douglas-fir fairly well (Fig. 1.2). Poorly sampled regions excluded, all major mountains and even many isolated minor mountains where Douglas-fir currently resides and where fossil sites are available show the presence of Douglas-fir. Exceptions include some isolated localities in southern New Mexico and western Texas. Moreover, the elevational range of Douglas-fir recorded 3000 – 0 cal yr BP accurately reflects the known modern elevational ranges for western British Columbia and western Washington/Oregon (Hermann & Lavender 1990). In the remaining four regions Douglas-fir was observed as many as a few hundred meters above modern elevational limits. In all cases, these were pollen records, suggesting that pollen may have blown upslope from nearby stands, rather than the range actually being higher than present. In the northern U.S. and Canadian Rockies, larch (especially subalpine larch) pollen may also explain this discrepancy. Nonetheless, the effect is not large, and the overall congruence of modern distribution with recent fossil evidence lends confidence to our approach for detecting Douglas-fir's presence at the broad scale of interest here.

Significance

Models projecting the range of Douglas-fir at the end of the 21st century show that Douglas-fir will need to shift or expand into the Great Basin and northern British Columbia/southeast Alaska to survive (Shafer *et al.* 2001). Douglas-fir is projected to be absent from much of its current range, including Mexico, parts of the southern Rockies and coast, and much of central British Columbia, indicating range shift is critical to avoid regional extinction. However, that would require a migration distance of 100 – 1000 km or greater in less than 100 years, which equals a migration rate of 1000 – 10 000 m/yr, far higher than anything considered possible naturally. In other regions, this may be less of an issue, as range shifts might be primarily elevational (Bartlein *et al.* 1997; Shafer *et al.* 2005). Future research should emphasize understanding the different responses to past and future climate at the population level.

Conclusion: testable hypotheses and future research

The combined fossil pollen and macrofossil dataset leaves us with a set of testable hypotheses for the late Quaternary history of Douglas-fir. Alternative hypotheses can be characterized as single-population or multiple-population during the LGM. In the former, each variety formed a contiguous interbreeding population that expanded northward into Canada. In the latter, two distinct glacial populations were present on the coast and three or more were present in the Rockies, each separated by prominent geographic barriers. Multiple glacial populations would suggest that only northernmost populations colonized Canada, while southern populations contracted to higher elevations. Furthermore, we hypothesize that Douglas-fir may have persisted the LGM in the Yellowstone region or northern Rockies based on very limited fossil evidence.

Some of the uncertainties presented here could be resolved with additional sampling of sites dating back to the LGM. These include some large regions that remain sparsely studied, such as eastern Oregon, the northern U.S. Rockies, central Utah, and southwestern Wyoming; but also, well-sampled, topographically complex regions that

require intensive sampling to distinguish among some of the hypotheses posed here; and finally, central British Columbia to refine northward migration rate estimates and better estimate the time of first contact among varieties. A number of California fossil data remain unpublished (Adam 1985) and would be of great value distinguishing hypotheses on the number and distribution of glacial populations and patterns of migration for the coast. Finally, future fossil investigations in Mexico might reveal the Quaternary dynamics of those poorly understood populations.

These hypotheses can also be tested using molecular phylogeographic approaches, which can detect population structure, migration routes, and provide further insight into the individualistic responses of populations. Moreover, molecular methods could reveal the relative contribution of each variety to Canadian populations of Douglas-fir. Our findings from the fossil record provide an important dated backdrop for interpreting genetic variation in modern Douglas-fir forests, and for understanding the ecological and evolutionary consequences of past and present large-scale climate change on western North American forests.

Tables

Table 1.1. Equivalence of calendar and radiocarbon year cutoffs for each 3000 calendar year time interval.

Cal yr BP	^{14}C yr BP
0	0
3000	2880
6000	5260
9000	8050
12 000	10 260
15 000	12 640
18 000	14 670
21 000	17 600

Table 1.2. Migration rates (m/yr) during each time interval along major migration routes. Italicized rates were calculated across multiple time intervals and evenly distributed accordingly, a situation forced by limited intervening data.

Migration direction (source) Time interval (cal yr BP)	north (coast)	north (N. Rockies)	east (N. Rockies)	east (S. Rockies)	east (CA)
21 000 – 18 000	-	-	-	0	0
18 000 – 15 000	42	0 (*139*)	0	-	0
15 000 – 12 000	144	278 (*139*)	0 - 18	0	0
12 000 - 9000	129	87	61 - 86	114	89
9000 - 6000	0	*64*	0	0	0
6000 – 3000	0	*64*	0	0	0
3000 – 0	0	*64*	0	0	0
Overall mean	52	93	10 - 17	19	15
Mean 18 000 - 9000	105	122	20 - 35	38	30
Minimum	50	58	-	-	-
"Maximum"	~100	~220	-	-	-

Figures

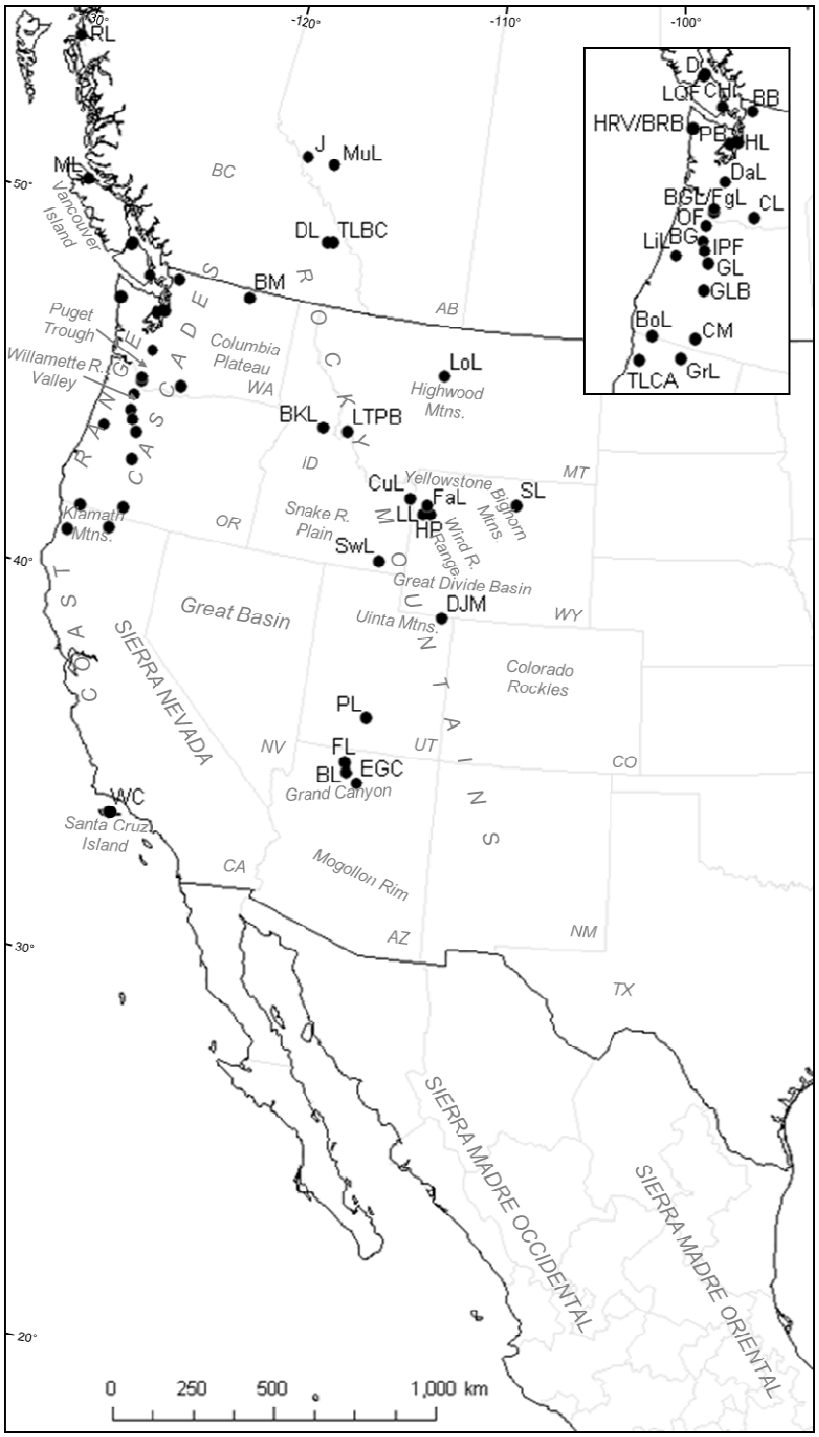

Figure 1.1. Geographic features and fossil sites mentioned in the text (BB = Bellingham Bog; BGL = Battle Ground Lake; BKL = Burnt Knob Lake; BL = Bear Lake; BM =

Bonaparte Meadows; BoL = Bolan Lake; BRB = Bogachiel River Bog; CH = Cowichan Head; CL = Carp Lake; CM = Caledonia Marsh; CuL = Cub Lake; D = Dashwood; DaL = Davis Lake; DJM = Dutch John Mountain; DL = Dog Lake; EGC = Eastern Grand Canyon; FaL = Fallback Lake; FgL = Fargher Lake; FL = Fracas Lake; GL = Gordon Lake; GLB = Gold Lake Bog; GrL = Grass Lake; HL = Hall Lake; HP = Hedrick Pond; HRV = Hoh River Valley; IPF = Indian Prairie Fen; J = Jasper; LL = Lily Lake; LiL = Little Lake; LoL = Lost Lake; LDF = Little Qualicum Falls; LTPB = Lost Trail Pass Bog; ML = Misty Lake; MuL = Muskiki Lake; OF = Onion Flats; PB = Poulsbo Bog; PL = Posy Lake; RL = Rainbow Lake; SB = Silverton Bone Site Bog; SL = Sherd Lake; SwL = Swan Lake; TLBC = Twin Lakes, BC; TLCA = Twin Lakes, CA; WC = Willow Creek). Inset shows labeled sites in the Pacific Northwest.

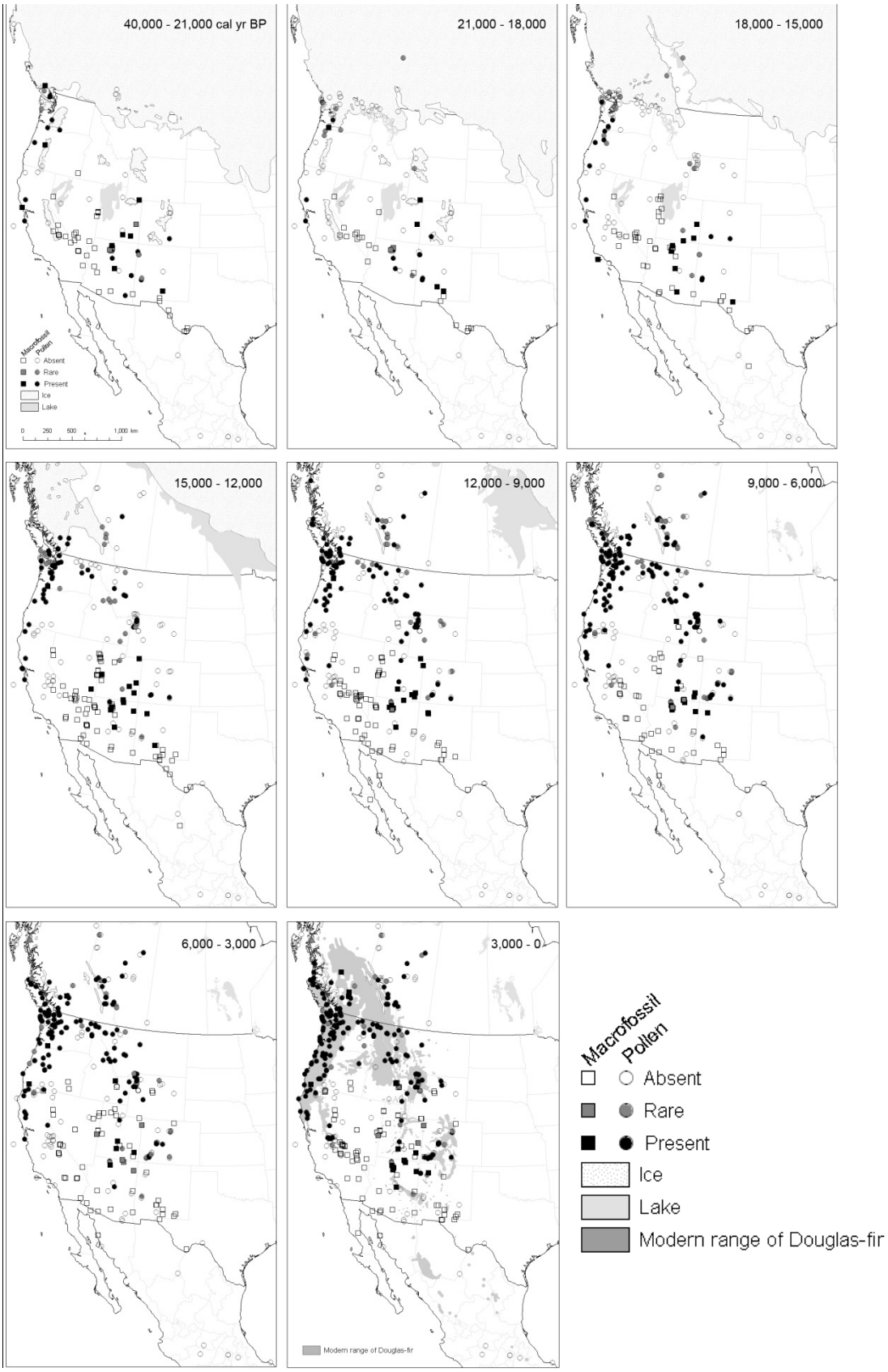
40,000 - 21,000 cal yr BP
21,000 - 18,000
18,000 - 15,000
15,000 - 12,000
12,000 - 9,000
9,000 - 6,000
6,000 - 3,000
3,000 - 0
Macrofossil Pollen
Absent
Rare
Present
Ice
Lake
0 250 500 1,000 km
Modern range of Douglas-fir

Figure 1.2. Maps marking Douglas-fir fossils present, rare, or absent at each fossil site for each time interval. The extent of the ice sheets at each time period within a few hundred years of the end of that interval (*e.g.* ice limits shown in 21 000 – 18 000 cal yr BP map correspond to 18 300 cal yr BP). Because ice limits are approximate and fluctuate on finer scales than those depicted here, occasional data points appear on top of the ice, but this is not the case. For example, the point showing low pollen percentages (for *Larix* in this case) that appears on top of the ice 21 000 – 15 000 cal yr BP is known to have been ice-free since at least 19 000 cal yr BP (Holloway *et al.* 1981). The congruence of fossils dated 3000 – 0 cal yr BP with Douglas-fir's present range (Little 1971) is shown in the last panel. (Full-size version included as supplemental file Figure 1.2.tif)

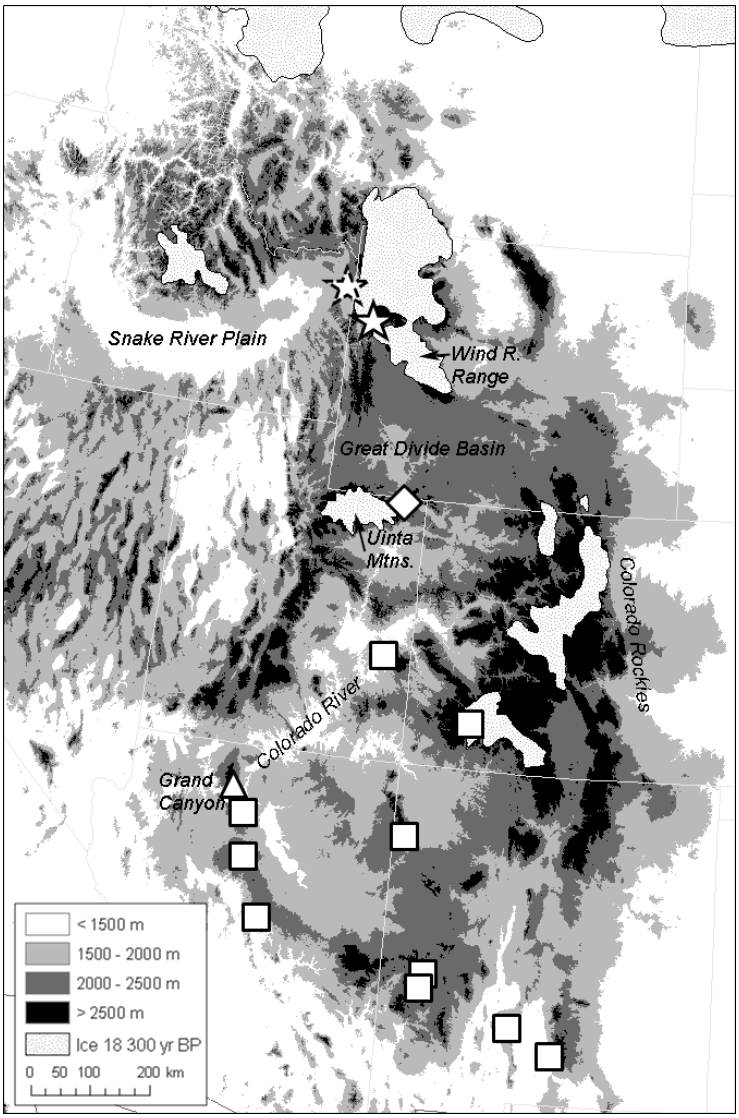

Figure 1.3. Topography of western North America is shown to highlight possible late Pleistocene barriers (labeled features and most areas < 1500 m or > 2500 m) between different glacial populations of Douglas-fir (different shapes). For example, the population on the north rim of the Grand Canyon (triangle) could have been isolated from populations south of the Colorado River (squares), and populations south of the Colorado River could have been more interconnected during the LGM if Douglas-fir was 700 m below its present elevation. Marked sites (stars, triangles, squares and circles) are based on fossils present 21 000 – 18 000 cal yr BP. Star with dotted border represents fossil from 16 700 cal yr BP (Cub Lake, ID), and is shown to better illustrate a possible glacial population in the region.

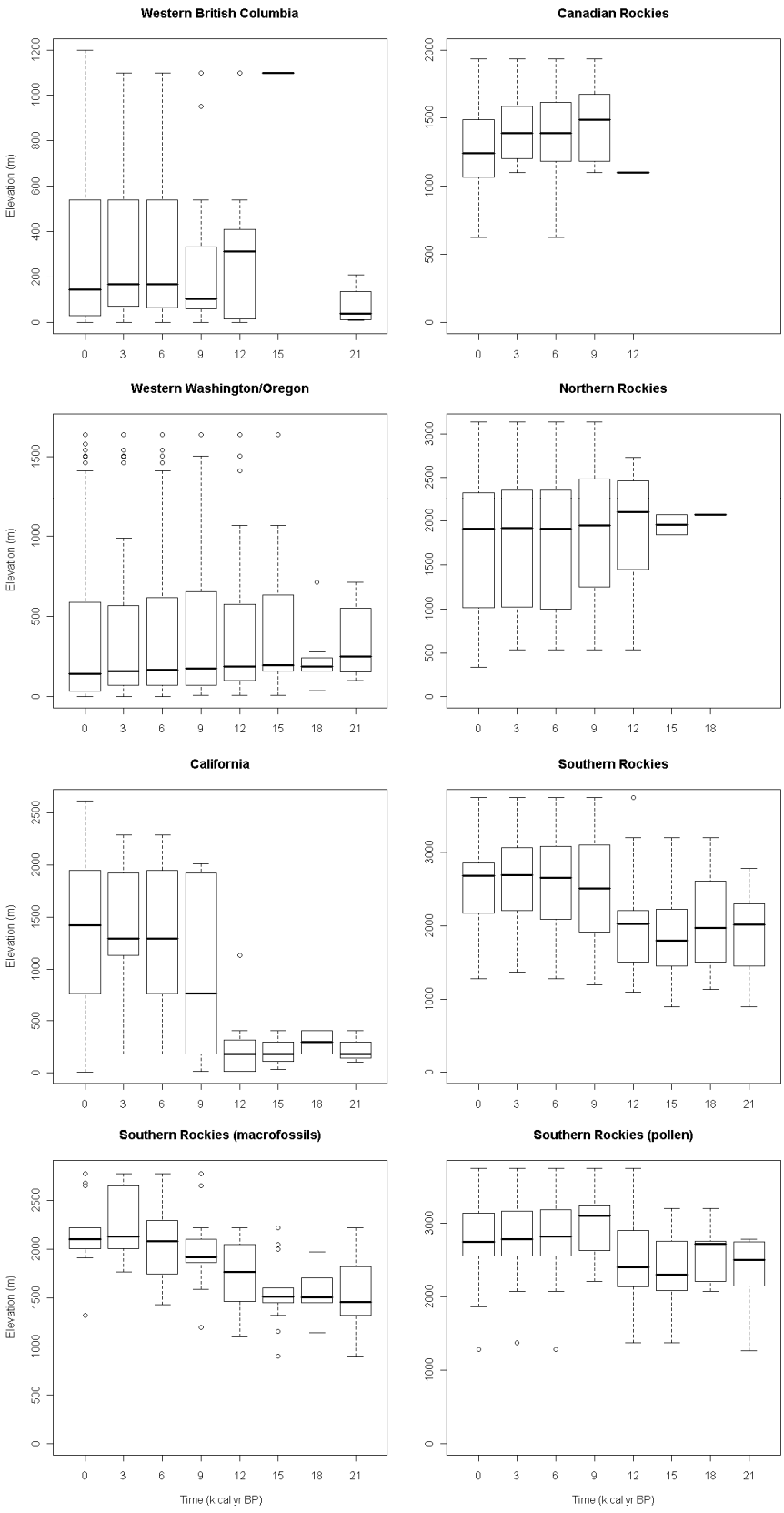
Western British Columbia
Canadian Rockies
Western Washington/Oregon
Northern Rockies
California
Southern Rockies
Southern Rockies (macrofossils)
Southern Rockies (pollen)
Elevation (m)
Time (k cal yr BP)

Figure 1.4. Change in elevational range of Douglas-fir through time (in thousands of cal yr BP) for six regions in the study area shown in boxplots. These patterns did not depend on whether or not "rare" sites were included, so we only report tests where both rare and present are included. Lines mark medians, boxes capture the 25th to 75th quantiles, and the whiskers capture the most extreme points not more than 1.5 times the range of the box. The *y*-axis scale differs among graphs. (Full-size version included as supplemental file Figure 1.4.tif)

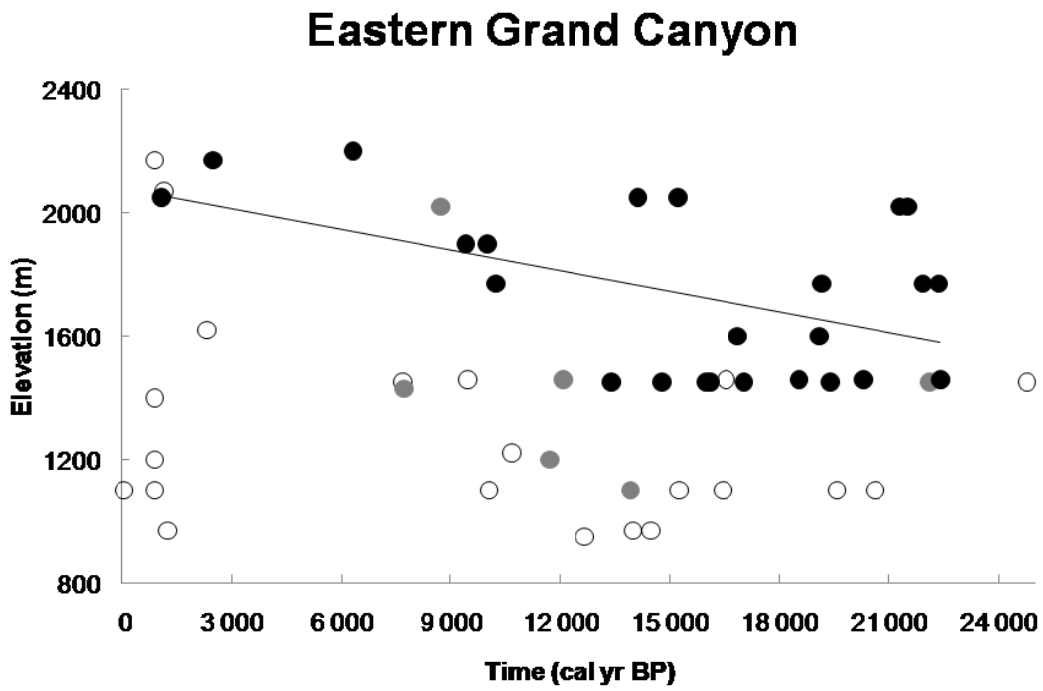

Figure 1.5. The elevational range of Douglas-fir in the eastern Grand Canyon increased through time. Black circles = present; gray circles = rare; open circles = absent. Regression line is based on sites at which Douglas-fir fossils were marked "present."

Supporting information

Figure 1.S1. Distribution of fossil pollen sites referenced in Table 1.S1.

(Included as supplementary file Figure 1.S1.tif)

Figure 1.S2. Distribution of macrofossil sites referenced in Table 1.S2.

(Included as supplementary file Figure 1.S2.tif)

Table 1.S1. Fossil pollen site coordinates, elevation, references, age of oldest sample, and comments with numerical references to Figure 1.S1.

(Included as supplementary file Table 1.S1.xlsx; references within table that are not cited in the text of the dissertation are included in Table 1.S1 and 1.S2 references.docx)

[a] *Larix* pollen according to original author
[b] Under our assumption, these sites would be considered to contain *Larix* pollen if any were found.
[c] *Larix* pollen according to this study

Table 1.S2. Macrofossil site/locality coordinates, elevation, references, age of oldest sample, and comments with numerical references to Figure 1.S2.

(Included as supplementary file Table 1.S2.xlsx; references within table that are not cited in the text of the dissertation are included in Table 1.S1 and 1.S2 references.docx)

[a] A = Reported taxa clearly limit midden site to habitat not favorable to Douglas-fir; B = Unclear if Douglas-fir was or was not present

Chapter 2

Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record

A version of this was published:
Gugger PF, Sugita S, Cavender-Bares J (2010) *Molecular Ecology,* **19**, 1877-1897.

Abstract

The integration of fossil and molecular data can provide a synthetic understanding of the ecological and evolutionary history of an organism. We analyzed range-wide maternally inherited mitochondrial DNA (mtDNA) and paternally inherited chloroplast DNA (cpDNA) sequence data with coalescent simulations and traditional population genetic methods to test hypotheses of population divergence generated from the fossil record of Douglas-fir (*Pseudotsuga menziesii*), an ecologically and economically important western North American conifer. Specifically, we tested 1) the hypothesis that the Pliocene orogeny of the Cascades and Sierra Nevada coincided with the divergence of coastal and Rocky Mountain Douglas-fir varieties and 2) the hypothesis that multiple refugia existed on the coast and in the Rocky Mountains during the Last Glacial Maximum (21 ka). We found that Douglas-fir varieties diverged about 2.11 Ma (4.37 Ma – 755 ka), which is consistent with a Pliocene divergence. Rocky Mountain Douglas-fir likely resided in three or more glacial refugia. More variable molecular markers would be required to detect the two coastal refugia suggested in the fossil record. Comparison of mtDNA and cpDNA variation revealed that gene flow via pollen linked populations isolated from seed exchange. Postglacial colonization of Canada from coastal and Rocky Mountain refugia near the ice margin at the Last Glacial Maximum produced a wide zone of hybridization among varieties that formed almost exclusively by pollen exchange and cpDNA introgression, not seed exchange. Postglacial migration rates were 50 – 165 m/yr, insufficient to track projected 21st century warming in some regions. Although fossil and genetic data largely agree, each provides unique insights.

Introduction

The integration of fossil and molecular genetic data is critical to understanding the ecological and evolutionary history of an organism (Petit *et al.* 2008; Hu *et al.* 2009). Fossil data offer direct, dated evidence of species presence (and sometimes abundance), but suffer from lack of population (or even species) resolution and limited sampling. In contrast, inferences of population history from molecular data from modern forest trees can provide population and species level resolution from across an entire distribution, but yield estimates of population divergence that are imprecise. Hypotheses from the fossil record can be tested with molecular data, which in turn produce new insights that can be tested with fossil data as more sites are analyzed. In this way, fossil and molecular data offer complementary information that jointly indicate major geological, climatological, or ecological influences on population divergence.

Range-wide syntheses of the fossil record in Europe and North America have revealed that temperate and boreal forests responded to the climate warming after the Last Glacial Maximum [21 - 18 thousand years ago (ka)] by northward range expansion (postglacial migration) from glacial refugia south of the ice sheets (Davis 1976, 1981, 1983; Huntley & Birks 1983). This process had a profound effect on the population genetic structure of forests that can still be detected today (Petit *et al.* 1997, 2003; Hewitt 2000). Isolated from one another since at least the start of the Wisconsinan glaciation (115 ka), glacial refugia are characterized by within-refugium genetic diversity and among-refugium divergence due to the effects of isolation, mutation, drift, and selection. Within-refugium diversity is lost along postglacial migration routes due to the effects of successive dispersal bottlenecks and founding events (Petit *et al.* 1997; Bialozyt *et al.* 2006; Gugger *et al.* 2008). Although fossil and molecular studies largely corroborate one another (Petit *et al.* 2002a, 2003), molecular data have revealed cryptic northern refugia and identified migration pathways that suggest lower postglacial migration rates than the fossil record (McLachlan *et al.* 2005; Anderson *et al.* 2006; Magri *et al.* 2006; Godbout *et al.* 2008; Petit *et al.* 2008). Moreover, in some cases, deep genetic divergence among some populations observed in molecular phylogeographic studies has invited the

reanalysis of the fossil record, suggesting pre-Wisconsinan population divergence (Magri *et al.* 2007).

The fossil record is a rich source of *a priori* hypotheses that can be tested using coalescent and other statistical phylogeographic methods that account for stochastic processes during population divergence (Knowles & Maddison 2002; Spellman & Klicka 2006). Despite the especially abundant fossil record for tree species, these methods have not commonly been used in plants. One persistent limitation has been the low sequence variation of plant mitochondrial (mtDNA) and chloroplast DNA (cpDNA) compared to animal mtDNA. Here, we formally integrate fossil and phylogeographic data, and attempt to overcome limitations that have prevented the use of coalescent methods in phylogeographic studies of plants, to study the evolutionary history of Douglas-fir (*Pseudotsuga menziesii*), an ecologically and economically important western North American tree.

Douglas-fir is well represented in the fossil record from the early Miocene [~23 million years ago (Ma)] to the late Holocene (Fig. 2.1; Hermann 1985). During the Miocene and Pliocene, *Pseudotsuga* fossils are found primarily from central British Columbia to southern California, from the coast to the interior as far as Idaho (Fig 2.1a). By the Pleistocene, *Pseudotsuga* fossils remained near the coast, had disappeared from the Columbia Plateau and western Great Basin, and had appeared in the central and southern Rocky Mountains. Coincident with this transition was the rise of the Cascade Range and Sierra Nevada primarily during the Pliocene, which is thought to have imposed a rain shadow that dried the Columbia Plateau and Great Basin (Brunsfeld *et al.* 2001). Orogeny and xerification could have caused the vicariant separation of Douglas-fir populations into what we now recognize as its two varieties: coastal (*P. menziesii* var. *menziesii*) and Rocky Mountain (*P. menziesii* var. g*lauca*). These varieties are morphologically (Hermann & Lavender 1990), chemically (von Rudloff 1972; Zavarin & Snajberk 1973), and genetically (Li & Adams 1989; Aagaard *et al.* 1995) distinct with a transition zone in British Columbia thought to be restricted to the east slope of the Coast Range (Critchfield 1984; Li & Adams 1989; Hermann & Lavender 1990; Ponoy *et al.* 1994).

Subsequent population divergence within each variety likely occurred during Pleistocene glacial cycles (Fig. 2.1b). In a recent review of late Quaternary fossil pollen and packrat midden macrofossil data from over 500 sites in the U.S. and Canada, Gugger & Sugita (in press; Chapter 1) found that Douglas-fir resided in as many as two coastal refugia and three to four Rocky Mountain refugia. Refugia were defined broadly to mean regions or localities that were the sources of modern populations (Bennett & Provan 2008). Some of these refugia are uncertain because they were based on a limited number of fossil sites (southern Utah) or very low fossil Douglas-fir pollen abundances (Yellowstone National Park area). Moreover, because fossils cannot definitively delimit the ranges of distinct populations, the hypothesis of a single refugium for each variety is also plausible. Canada was colonized by an unknown mixture of each variety from a coastal refugium in Washington near the ice margin at the Last Glacial Maximum and a Rocky Mountain refugium in the northern Rockies near Yellowstone. The southernmost populations in California and the Rockies retreated upslope 700 - 1000 m and probably did not contribute to the colonization of the northern part of the range.

Other hypotheses for the number of glacial refugia and postglacial migration routes have been proposed. Li & Adams (1989) observed differences in allozyme allele composition and frequency between the northern and southern Rockies and thus proposed two glacial refugia. St. Clair *et al*. (2005) proposed the postglacial migration of the Rocky Mountain variety southward along the east side of the Cascades via Canada to explain genecological dissimilarity of east and west Cascades populations and the apparent convergence of east Cascades populations with Rocky Mountain populations.

We used mtDNA and cpDNA sequence data to test 1) the hypothesis that the Pliocene orogeny of the Cascade Range and Sierra Nevada coincided with the divergence of Douglas-fir varieties and 2) hypotheses describing the number of glacial refugia and postglacial migration routes during the late Pleistocene and Holocene. Finally, we take advantage of the fact that in Douglas-fir mtDNA is maternally inherited and thus seed dispersed (Marshall & Neale 1992), whereas cpDNA is paternally inherited (Neale *et al.* 1986) and thus dispersed first in pollen and then as fertilized seed, to investigate differences in the history of seed and pollen dispersal.

Materials and methods

Sampling

Leaf or bud tissue was collected from 87 sites throughout the range of Douglas-fir in the U.S. and Canada (Fig. 2.2; Table 2.S1). Seven sites were sampled from a common garden near Sooke, BC (48.41667°N, -123.867°W, 140 m) and 17 from a common garden near Enderby, BC (50.5°N, -119°W, 600 m; Zhang *et al.* 1993), both under the direction of the British Columbia Ministry of Forests. Each provenance in the common gardens was seeded with multiple individuals from each of several maternal lines, thus we only used one individual to avoid the possibility of sampling the same tree (mother) twice. The remaining 63 sites were natural populations, where 1 - 11 individuals (mean = 3.1) separated from one another by at least 50 m were collected. We also collected four samples of *P. macrocarpa* and one of *Larix occidentalis* as outgroups.

DNA preparation

Total genomic DNA was extracted from leaf or bud tissue using Qiagen DNeasy Plant Mini Kit (Valencia, CA, USA) according to the manufacturer's instructions.

After surveying variation in several mtDNA and cpDNA sequences that were shown to vary neutrally in other species, we selected the two most variable segments from each genome for analyses. For cpDNA, we chose *rps7-trnL*, which contains intergenic spacers, the *ndhB* pseudogene and its intron; and *rps15-psaC*, which contains intergenic spacers and the *ndhH, ndhI, and ndhE* pseudogenes. *Ndh* genes have been lost or are non-coding in Pinaceae (Wakasugi *et al*. 1994; Braukmann *et al*. 2009). For each, we designed novel primers (Table 2.1) based on the *Pinus thunbergii* complete chloroplast genome sequence (Wakasugi *et al.* 1994) using Primer3 (Rozen & Skaletsky 2000).

The mtDNA segments we chose were variable region 7 (V7) of the small-subunit ribosomal RNA gene (*19S rDNA V7*; Duff & Nickrent 1997, 1999) and the first intron of

the *nad7* gene (*nad7i1*; Jaramillo-Correa *et al*. 2004). Both were shown variable in *Picea* spp. (Jaramillo-Correa *et al.* 2003, 2004) and the latter in *Pinus* spp. (Godbout *et al.* 2005, 2008). Douglas-fir-specific primers (Table 2.1) were designed for each based on published sequences (Chaw *et al.* 2000; Jaramillo-Correa *et al*. 2004). For *V7*, care was taken to produce primers that would not amplify similar small-subunit rDNA sequences in the chloroplast (16S; Wakasugi *et al*. 1994) and nuclear (18S; Chaw *et al.* 1997, 2000) genomes.

Polymerase chain reactions (PCR; 25 μL) using Qiagen *Taq* PCR Core Kit contained 2.5 μL 10× PCR buffer, 5 μL Q solution, and final concentrations of 200 μM dNTPs, 2.5 mM $MgCl_2$, 4 μM of each primer, and 0.5 U *Taq* DNA polymerase. Except for annealing temperature (Table 2.1), amplification programs were the same: 4 min at 94°C; 35 cycles of 1 min at 94°C, 45 s at 55 – 58°C, and 1.5 min at 72°C; and 10 min 72°C. PCR products were then cleaned with QIAquick PCR Purification Kit according to the manufacturer's instructions.

Purified products were sequenced on an ABI Prism 3730xl DNA Analyzer (PE Applied Biosystems). Each locus was sequenced using the forward primer yielding 300 – 700 base pairs (Table 2.1). Unique haplotypes were verified by resequencing with forward and reverse primers. Sequences were aligned and edited using Sequencher 4.8 (Gene Codes). Because organellar genomes in plants generally do not recombine and are uniparentally inherited (Reboud & Zeyl 1994), they each act as a single locus. Accordingly, we combined the two cpDNA sequences to produce a single cpDNA sequence and the two mtDNA sequences to produce a single mtDNA sequence. Here, their variants are referred to as chlorotypes and mitotypes, respectively.

Parsimony network and mutation rate

Parsimony networks were produced separately for mitotypes and chlorotypes using TCS 1.21 (Clement *et al.* 2000) with insertion-deletions (indels) coded as a fifth state.

The rate of mutation to neutral alleles was estimated from the rate of nucleotide substitution. Using DnaSP 5.0 (Librado & Rozas 2009), per base pair per year (μ_y) and

per base pair per generation (μ_g) substitution rates for each locus were calculated as the mean number of nucleotide substitutions per site (d_{XY}; Nei 1987) with Jukes and Cantor (1969) correction among all *Pseudotsuga* and *Larix* haplotypes divided by two times the divergence time in years or generations. We assumed the divergence of *Pseudotsuga* and *Larix* occurred about 50 Ma. The oldest *Larix* fossil dates to the Eocene (~45 Ma; Axelrod 1990; LePage & Basinger 1991; Schorn 1994) and oldest unambiguous *Pseudotsuga* fossil dates to the early Oligocene (~32 Ma; Lakhanpal 1958; Schorn 1994), though older fossils have been proposed (Penhallow 1902, 1907; Axelrod 1966; Hermann 1985). Consistent with our assumption, a molecular clock enforced on a *matK* gene phylogeny of Pinaceae suggests that *Larix* and *Pseudotsuga* split in the Paleocene (~60 Ma; Wang *et al.* 2000). An approximate average generation time (T) can be calculated according to $T = \alpha + [s/(1-s)]$ (Lande *et al*. 2003; Spellman & Klicka 2006), where α is the time to maturity (~15 yr; Herman & Lavender 1990) and s is the adult annual survival rate (~0.99 – 0.995; Franklin & DeBell 1988; Hann *et al*. 2003). Based on this, estimates for T range from 114 – 214 yr. If we consider periodic major disturbances (*e.g.* fire or disease outbreaks) that were not observed in the studies estimating the above adult annual survival rates, a lower survival rate (Mathiasen *et al*. 1990) and therefore a lower average generation time would be more likely. We primarily assumed an average generation time of 100 yr, but also explored the effects of generation times as high as 200 yr.

Mutation rates calculated among genera (*Larix* and *Pseudotsuga*) should be valid within *P. menziesii* if the sequences have evolved under a molecular clock model. To test this null hypothesis, we compared the likelihood scores under a model with and without a molecular clock (Muse & Weir 1992) using PAUP* 4.0b10 (Swofford 2003). We conducted two tests per locus, one including one representative from each species and variety and one including only haplotypes observed in *P. menziesii* (all individuals, both varieties). If the molecular clock could not be rejected at either evolutionary scale, we inferred a constant rate of substitutions. These tests require that a particular nucleotide substitution model be chosen; we chose one using hierarchical likelihood ratio tests implemented in PAUP* and Modeltest 3.7 (Posada & Crandall 1998).

Genetic diversity and neutrality

Because many of our sample sites only contained one individual, we quantified general patterns of genetic diversity by grouping data *a priori* from each sample site into 16 populations and four regions based on geography (Fig. 2.2). Using Arlequin 3.11 (Excoffier *et al.* 2005), we estimated the haplotype richness (h), number of segregating sites (S), haplotype diversity (H; Nei 1987), nucleotide diversity (π; Tajima 1983; Nei 1987), and D (Tajima 1989) for each population, for each region, and for the whole dataset. We also estimated the haplotype richness after correcting for unequal sample sizes with rarefaction (h_r; Hurlbert 1971) using Contrib 1.01 (Petit *et al.* 1998), and R_2 (Ramos-Onsins & Rozas 2002) using DnaSP. Latitudinal trends in the diversity (H, π, h_r) of populations were investigated with linear regression. D and R_2 were used to test for departures from the neutral expectations of constant population size and selective neutrality, and their significance was assessed with coalescent simulations (Hudson 1990). More negative values of D suggest population expansion or purifying selection. The R_2 test statistic tests the null hypothesis of constant population size (Ramos-Onsins & Rozas 2002). We computed genetic differentiation among populations standardized by diversity as G'_{ST} (Hedrick 2005).

Population structure

To identify clusters of genetically similar populations, we analyzed our data with a spatial analysis of molecular variance (SAMOVA). Geographically close populations are grouped into a user-defined number of groups (K) using a simulated annealing approach to maximize the variance (F_{CT}) among those groups (Dupanloup *et al.* 2002). We performed this analysis in SAMOVA 1.0 (Dupanloup *et al.* 2002) using $K = 2 – 12$ for each locus, and chose the number of groups that gave the highest F_{CT} or the number of groups for which F_{CT} began to plateau.

Testing refugia hypotheses with coalescent methods

Hypotheses for the number and extent of Rocky Mountain glacial refugia (Fig. 2.1b, 2.3) were tested using coalescent simulations in Mesquite 2.6 (www.mesquiteproject.org). We did not use coalescent simulations to test hypotheses for the coastal variety due to lack of molecular variation (see Results). For the Rocky Mountain variety, we tested the following hypotheses: one-refugium, two-refugia, three-refugia, and four-refugia, each with multiple Pleistocene divergence times among refugia: beginning of the Wisconsinan glaciation (115 ka) and beginning of the late Illinoisan glaciation (190 ka; Richmond & Fullerton 1986a; Lisiecki & Raymo 2005).

First, 1000 coalescent genealogies were simulated under a hypothesized model of population divergence (Fig. 2.3) with effective population size, N_e (different numerical values were specified for different runs; more below). The fit of the simulated genealogies to the hypothesized model of population divergence was then assessed using the test statistic *s* (Slatkin & Maddison 1989), generating a null distribution for that hypothesis. The *s* value is the number of parsimony steps on a gene tree on which source population has been mapped as a multistate character. Typically, the *s* value for a single maximum likelihood tree from empirical data is compared to this null distribution to assess statistical significance of deviation from the hypothesized model (Knowles 2001; Spellman & Klicka 2006; Carstens & Richards 2007). However, as a way of dealing with the limited sequence variation and consequent poorly resolved tree, we computed *s* for 1000 trees drawn from the posterior distribution of a Bayesian Markov chain Monte Carlo (MCMC) analysis implemented in BEAST 1.4.8 (Drummond & Rambaut 2007). This generated an empirical distribution for *s* that accounted for the range of possible true trees. If the empirical *s* values were significantly higher (*i.e.* worse fit) than those from the simulated data, we rejected the hypothesis. Significance ($\alpha = 0.05$) was assessed conservatively by calculating the probability of observing an *s* value from the null distribution at least as high as the lowest empirical *s*. We could not assess the single-refugium hypothesis in this way because $s = 0$ for a single population calculated on any gene tree. Instead, we simulated 1000 genealogies under the single-refugium (null)

hypothesis, and then measured their fit to each alternative hypothesis of two, three, and four refugia, respectively (Knowles 2001). We rejected the null hypothesis in favor of the respective alternative if the highest empirical value of s was lower (*i.e.* better fit) than 95% of the simulated s values.

To estimate simultaneously the empirical trees and N_e, we ran BEAST for 20 – 30 million steps assuming a constant population size, a strict molecular clock, and a Hasegawa-Kishino-Yano substitution model (HKY; Hasegawa *et al.* 1985) with empirical base frequencies and no site heterogeneity. We repeated this twice for each locus to verify convergence upon the same values and combined replicate outputs. N_e was calculated for each locus from the median estimate of θ and the upper and lower bounds of its 95% highest posterior density (HPD) interval according to $\theta = 2N_e\mu_g$. We report these as the effective population sizes of females (mtDNA; N_{ef}) and males (cpDNA; N_{em}).

Coalescent simulations are sensitive to different N_e, and θ estimates from single loci have large errors (Edwards & Beerli 2000), so we tested each hypothesis in Mesquite using N_e calculated from the median point estimate of θ and the high and low values of its 95% HPD interval. We repeated this for N_e values calculated assuming a generation time of 200 yr instead of 100 yr, which had the effect of doubling μ_g and therefore halving N_e.

Divergence time among varieties

Divergence time among coastal (populations 1 - 3; Fig. 2.2) and Rocky Mountain (populations 8 - 16) Douglas-fir varieties, excluding Canadian populations, was investigated using an isolation with migration (Nielson & Wakeley 2001) model as implemented in IMa (Hey & Nielson 2007). Under the full model, IMa simultaneously estimates six parameters (scaled by substitution rate): divergence time (t), migration from population one to two (m_1), migration from population two to one (m_2), effective population size of each population (N_1 and N_2), and effective population size of the ancestor (N_A). IMa assumes constant population size, neutral molecular markers, no recombination within loci, free recombination among loci, and a particular mutation

model. We assessed selective neutrality and constant population size with D and R_2 as described earlier. Recombination is not common in organellar genomes (Chiu & Sears 1985; Reboud & Zeyl 1994; Birky 2001), but has been observed in *Picea* mtDNA (Jaramillo-Correa & Bousquet 2005). A four-gamete test (Hudson & Kaplan 1985) showed no evidence of mtDNA recombination in our data (not shown). We chose an infinite sites mutation model (Kimura 1969) for mtDNA and an HKY model (Hasegawa *et al.* 1985; Palsbøll *et al.* 2004) for cpDNA to accommodate homoplasy (see Results).

In IMa, 25 geometrically heated ($g_1 = 0.95$, $g_2 = 0.85$) Metropolis-coupled MCMC chains with 10 swaps per step and a maximum prior for $t = 5$ were run for 10 - 50 million steps beyond a 2 million step burn-in. The prior for t was chosen because values over five are biologically unreasonable, suggesting a divergence time of over 9 Ma. Maximum priors for m_1 and m_2 and scalars for the effective population size parameters (q_1, q_2, q_A) varied depending upon the comparison and were chosen based on preliminary runs of the program with larger parameter intervals (Won & Hey 2005). This analysis was performed for each locus and with both loci combined. To verify convergence upon the same parameter values, we ran this analysis three times for each comparison with different random seeds. Only estimates whose posterior distribution dropped to zero within the prior intervals investigated were trusted. To scale the outputs to demographic units, IMa uses the generation time and per locus per year substitution rate, which we calculated from μ_y by multiplying by the number of base pairs in the locus under consideration.

We also ran IMa in L-mode with 100,000 - 250,000 genealogies to test the null hypothesis that the likelihood of our data under the full model described above equals the likelihood under a simpler, nested model without migration (*i.e.* t, N_1, N_2 , N_A, $m_1 = m_2 = 0$). This was conducted as a two part test, where first the full model was evaluated against a nested model in which $m_1 = m_2$. The test statistic is negative two times the natural log of the ratio of the estimated likelihoods of the nested model to the full model (-2Λ), which is X^2-distributed with a one degree of freedom (difference in number of parameters between models). If the nested model could not be rejected, the second test compared a model in which $m_1 = m_2$ against a yet simpler model, in which $m_1 = m_2 = 0$.

Here, -2Λ has 0.5 probability of taking on a value of zero and 0.5 probability of taking on a value from a X^2 distribution with one degree of freedom (Chernoff 1954; Hey & Nielson 2007). If neither test was significant, the model with $m_1 = m_2 = 0$ was not rejected, so we ran the IMa analysis again with migration set to zero.

Results

We obtained sequences for both mtDNA segments (aligned, edited length of 943 bp) in 190 individuals from 82 sites and from both cpDNA segments (1537 bp) in 219 individuals from 87 sites (Table 2.S1). We observed two *nad7i1* haplotypes and six *V7* haplotypes, which combined, yielded 7 mitotypes containing 10 base substitutions and five indels (Fig. 2.4; GenBank accessions in Table 2.S2). Of those, four were common (M1, M3, M4, M6) and three were rare (M2, M5, M7), occurring in four or fewer individuals. We also observed nine *rps7-trnL* haplotypes and 11 *rps15-psaC* haplotypes, which combined to 20 chlorotypes (Table 2.S2) containing 18 base substitutions and three indels (Fig. 2.4). Two of these were common (C1, C5) and the rest were rare, occurring in one to nine individuals.

Visual inspection of the data suggested strong geographic structure in both mtDNA and cpDNA (Fig. 2.4). The two varieties were well delineated for both loci throughout much of the range, except in Canada, where cpDNA haplotypes were shared among varieties. The coastal variety lacked genetic structure for the sequences investigated here. Within the Rockies, additional population structure was clear in mtDNA data and suggestive in the cpDNA, where some mitotypes and chlorotypes had southern (*e.g.* M6, C20), central (*e.g.* M4, C16), or northern (*e.g.* M3, C7, C10) Rocky Mountain distributions.

Mutation rate

The per base pair per year substitution rate (μ_y) was estimated as 5.26×10^{-10} for mtDNA and 4.41×10^{-10} for cpDNA (Table 2.2). Sequence divergence and other mutation rates are reported in Table 2.2. For cpDNA, we rejected a molecular clock among species (P = 0.034; TVM) and within *P. menziesii* chlorotypes (P = 0.012; F81; Felsenstein 1981). For mtDNA, a molecular clock was not rejected among species (P = 0.25; JC; Jukes & Cantor 1969) nor within *P. menziesii* (P = 1; F81; Table 2.2).

Genetic diversity and neutrality

Overall, haplotype diversity was similarly high for both mtDNA and cpDNA data (H = 0.74 and 0.64, respectively), and nucleotide diversity was low for both loci (π = 0.00280 and 0.00073; Table 2.3). Haplotype richness, haplotype diversity and nucleotide diversity varied across populations and regions, with the coast having the lowest diversity by all measures (Table 2.3). However, no significant latitudinal trends were observed (mtDNA: $r_{hr|Lat}$ = -0.18, P = 0.49; $r_{H|Lat}$ = -0.18, P = 0.50; $r_{\pi|Lat}$ = 0.26, P = 0.34; cpDNA: $r_{hr|Lat}$ = -0.10, P = 0.70; $r_{H|Lat}$ = -0.01, P = 0.98; $r_{\pi|Lat}$ = 0.14, P = 0.61). Diversity was more strongly partitioned among populations in mtDNA (G'_{ST} = 0.94) than in cpDNA (G'_{ST} = 0.64).

The mtDNA data showed little evidence of violating the assumptions of neutrality. Only for the combined coastal data was a significantly negative D observed (Table 2.3), suggesting possible recent population expansion (though R_2 was not significant). The cpDNA data had a significantly negative D value overall and for the coast a significantly positive R_2 values for the northern and southern U.S. Rockies. Within the Rockies, only two D and one R_2 were significant, and two of these were in the northern Rockies. Population expansion in the northern Rockies and coast is consistent with the fossil record for Douglas-fir (Gugger & Sugita, in press; Chapter 1); however, this does not exclude the possibility of both expansion and selection.

Population structure

SAMOVA of mtDNA data showed that F_{CT} began to plateau at 0.92 for 5 – 6 groups of populations (Table 2.4). The five groups corresponded to the coast (populations 1 - 5), northern Rockies (7 – 10), central Rockies (11 - 13), southern Rockies (14 – 16), and northern British Columbia (6). The six groups were the same except northern Utah (11) was split from the central Rockies. The separation of northern British Columbia as a distinct group in both cases probably occurred because that population is a mixture of mitotypes from neighboring populations that are fixed for divergent coastal and Rocky Mountain mitotypes. Less than 1% of the variation was among populations within groups, and 7.8% was within populations.

In the SAMOVA of cpDNA, the highest F_{CT} (0.68) was for two groups, corresponding to the coast (1 – 4) and Rockies (5 – 16; Table 2.4). In this case, central British Columbia (5) was grouped with the Rockies instead of coast. For cpDNA, 4.1% of the variation was among populations within groups and 27.5% was within populations. SAMOVA of cpDNA compared to mtDNA showed that in both cases most of the variation was explained among groups, but that proportion was lower in the cpDNA than in the mtDNA data. Also, cpDNA had a much higher percentage of variation within populations than mtDNA.

Testing hypotheses with coalescent methods

The median estimates of θ were 0.000491 (95% HPD: 0.000104, 0.001060) for mtDNA and 0.001252 (95% HPD: 0.000585, 0.002086) for cpDNA. These corresponded to N_{ef} of 4667 (989, 10,076) and N_{em} of 14,185 (6628, 23,635), which we used in coalescent simulations. Because these estimates seemed low for a common and widespread tree (Millar & Libby 1991), we also ran simulations with the median N_e estimate times ten.

Overall, simulation results did not clearly point to a single best hypothesis (Fig. 2.5; Table 2.5). For the mtDNA, simulations at the lowest N_{ef} (95% HPD low) for all divergence times rejected all multiple-refugia hypotheses but not the single-refugium

hypothesis. For the median N_{ef} estimate, the single-refugium hypothesis was rejected in favor of the two-refugia hypothesis, and the two-refugia hypothesis could not be rejected for divergence dating to the beginning of the Wisconsinan glaciation (115 ka; Table 2.5; Fig. 2.5). For the 95% HPD high N_{ef}, the two-refugia hypothesis was not rejected at any divergence time, and the one-refugium hypothesis was rejected in favor of two- and three-refugia hypotheses. At N_{ef} ×10, no multiple-refugia hypothesis was rejected, and all one-refugium hypotheses were rejected in favor of their alternative multiple-refugia hypotheses. For the cpDNA, no single-refugium hypothesis could be rejected for any N_{em}. In addition, all multiple-refugia hypotheses were rejected at N_{em}, 95% HPD low N_{em}, and 95% HPD high N_{em}. For N_{em}×10, two- and three-refugia hypotheses were not rejected at any divergence, and four-refugia hypotheses were not rejected only at 115 ka. These general patterns of significance were similar when considering a generation time of 200 yr (not shown).

Divergence time among varieties

We ignored the possibility that some populations may have undergone recent expansions (Table 2.3) to avoid more complicated models (*e.g.* IM, Hey & Nielson 2004). Also, we excluded Canadian populations, which were largely a mixture of the two varieties (Fig. 2.4), and several individuals near Canada that were presumed recent migrants from the range of one variety to the other during postglacial contact (three C1s and two C3s from the northern U.S. Rockies and one C5 from northern Washington).

Divergence time among coastal and Rocky Mountain varieties could not be estimated under the full IMa model for both loci combined nor for each locus independently (posterior distribution never dropped to zero). However, per generation migration rates were very low for both loci combined, mtDNA, and cpDNA, and we could not reject the null hypothesis that both migration parameters equaled zero (Table 2.6). Moreover, each variety in our analysis was fixed for divergent haplotypes, suggesting migration among varieties in the U.S. has not been pronounced for a long time. Therefore, we set $m_1 = m_2 = 0$ and repeated the analyses. After this adjustment, we

were able to estimate divergence time of about 2.11 Ma with the 90% HPD interval ranging 4.37 Ma - 755 ka, which spans the early Pliocene through middle Pleistocene (Table 2.6). Divergence time estimates from mtDNA and cpDNA differed substantially (4.41 Ma and 491 ka, respectively).

Discussion

The distribution and divergence of Douglas-fir populations (Fig. 2.4) is well explained by both Pliocene geology and late Quaternary climate cycles. The divergence of coastal and Rocky Mountain varieties coincided with the uprising of the Cascades and Sierra Nevada (Fig. 2.1a; Table 2.6). Much of the variation within the Rocky Mountains is likely the product of isolation and recontact induced by glacial-interglacial cycles (Fig. 2.1, 2.5, 2.6; Table 2.4, 2.5). Phylogeographic patterns of mtDNA and cpDNA variation were broadly consistent (Fig. 2.4; Table 2.3, 2.4), but differed in ways suggesting that gene flow via pollen dispersal might have played an important role in connecting populations that were otherwise isolated from seed exchange. The most striking example of this is the ongoing cpDNA introgression observed primarily in Canadian populations east of the Coast Range that likely began in the early Holocene (Fig. 2.4).

Divergence of coastal and Rocky Mountain varieties

The divergence of the coastal and Rocky Mountain Douglas-fir varieties as estimated by IMa (2.11 Ma; Table 2.6) is consistent with the Pliocene divergence hypothesis and the fossil record (Fig. 2.1b), but it does not exclude the possibility of divergence times well into the Pleistocene (90% HPD interval, 4.37 Ma – 755 ka). Part of the reason for the large interval is that the mtDNA and cpDNA data when run separately gave quite different estimates (4.41 Ma and 491 ka, respectively; Table 2.6), and the mtDNA estimate itself has a very broad HPD interval. This discrepancy could be understood in two ways: 1) the estimates reflect the variability expected among different loci evolving

and sorting under stochastic processes (Edwards & Beerli 2000; Arbogast *et al.* 2002; Carstens & Knowles 2007), or 2) the estimates for each locus reflect different processes. Without data from more loci it is conservative to assume the former; however, the latter has interesting implications, if true. For example, the mtDNA divergence time estimate of 4.41 Ma could reflect the time at which gene flow via seed dispersal was cut off, and the later cpDNA estimate of 491 ka could indicate that substantial gene flow via pollen continued until the middle Pleistocene. Under this scenario, pollen exchange could have been continuous over that interval or could have been in one large homogenizing burst just before final divergence.

However, additional uncertainties should be considered for all estimates. First, mtDNA mutations among varieties seemed to come in clusters that could have occurred as one event, and therefore, might not have fit the substitution model employed by IMa. The outgroup, *P. macrocarpa*, is positioned two mutational steps from the coastal variety (M1), but the coastal and Rocky Mountain varieties are seven mutational steps apart (Fig. 2.4). Under a molecular clock assumption, this implies that the varietal divergence preceded the species divergence, which is not supported by cpDNA (Fig. 2.4), morphology (Hermann & Lavender 1990), or the fossil record (Fig. 2.1; Hermann 1985). Overall, this could result in an overestimate of the divergence time among varieties based on mtDNA and both loci. Second and more generally, our mutation rate estimates may not be accurate. In calculating the per locus per year substitution rate used in IMa, we assumed a divergence time for *Pseudotsuga* and *Larix* of 50 Ma consistent with the fossil record (LePage & Basinger 1991; Schorn 1994; Gernandt & Liston 1999) and other molecular evidence (Wang *et al.* 2000) and that the mutation rate estimated among genera is applicable within species. The latter was assumed despite the fact that a molecular clock model was rejected within and among species for the cpDNA (Table 2.2), and its effect on our conclusions is unknown. If the true divergence time among genera was actually 100 Ma, for example, then our estimated mutation rates would have been halved and all our varietal divergence time estimates would have been doubled.

Pushing back the divergence date from 2.11 Ma to 4.22 Ma for the analysis with both loci combined would not affect our main conclusion of divergence associated with

the Cascade and Sierra Nevada orogeny and subsequent xerification of the Columbia Plateau and Great Basin between the present ranges of each variety. Moreover, our mutation rate estimates are consistent with the most recent fossil-calibrated neutral substitution rates for cpDNA observed in *Pinus* ($\mu_y = 2.2 - 4.2 \times 10^{-10}$; Willyard *et al.* 2007). In the Pacific Northwest, Douglas-fir shares this history with a number of similarly distributed taxa (Carstens *et al.* 2005), including tailed frogs (*Ascaphus* spp.; Nielson *et al.* 2001), Van Dyke's salamanders (*Plethodon* spp.; Carstens *et al.* 2004), the giant Pacific salamander (*Dicamptodon* spp.; Steele *et al.* 2005), and Constance's bittercress (*Cardamine constancei*; Brunsfeld & Sullivan 2005).

Glacial refugia

Glacial refugia for the coastal and Rocky Mountain varieties were distinct (Fig. 2.1b, 2.4, 2.6; Table 2.6). One chlorotype and one mitotype dominated the coastal region, and the few rare haplotypes were wide-ranging across the region. This low diversity and lack of substructure within the coastal variety is consistent with nuclear microsatellite and allozyme variation (Li & Adams 1989; Krutovsky *et al.* 2009) and suggests we cannot reject the hypothesis of one coastal refugium spanning central California to unglaciated western Washington. One coastal refugium is also consistent with models that used modern climatic tolerances to predict Douglas-fir's distribution during the Last Glacial Maximum (Bartlein *et al.* 1998). Even so, the lack of structure in the coastal variety is surprising considering the strong suggestion of two refugia in the fossil record (Fig. 2.1b; Gugger & Sugita, in press; Chapter 1) and that leaf terpene chemistry differs in California compared to Oregon/Washington (von Rudloff 1972; Zavarin & Snajberk 1973; von Rudloff & Rehfeldt 1980; Critchfield 1984). Moreover, a diverse array of organisms form contact zones near the California/Oregon border (Soltis *et al.* 1997; Swenson & Howard 2005). However, given the very limited number of haplotypes observed in the region, we cannot exclude the possibility of multiple coastal refugia that more variable molecular markers with the signature of more recent events might reveal.

For the Rocky Mountains, in contrast, multiple glacial refugia were likely (Fig. 2.6). Coalescent tests for mtDNA and high N_{ef} estimates (95% HPD high and $N_{ef}\times10$) tended to reject the single-refugium hypotheses in favor of multiple-refugia alternatives (Fig. 2.5; Table 2.5). Some two-refugia hypotheses were not rejected, and at $N_{ef}\times10$ no multiple-refugia hypothesis was rejected. The tendency to not reject two-refugia hypotheses lends some support to the two refugia proposed by Li & Adams (1989) based on allozyme data. Nonetheless, there was limited power to distinguish among multiple-refugia hypotheses with the mtDNA coalescent tests, especially considering that most multiple-refugia hypotheses were not rejected at lower N_{ef}.

Aside from some coalescent tests, other evidence supports at least three glacial refugia. For example, mtDNA SAMOVA defines three to four Rocky Mountain population clusters: northern, southern, and one to two central (Table 2.4). The southern and northern mtDNA SAMOVA groups conform well to the refugia identified in the fossil record, and fall along topographic barriers such as the Grand Canyon and the high peaks of the Colorado Rockies and Yellowstone region that harbored extensive mountain glaciers at the Last Glacial Maximum and would have formed formidable barriers to seed dispersal (Figs. 2.1b, 2.6; Gugger & Sugita, in press; Chapter 1).

The broadly defined northern refugium indentified by mtDNA SAMOVA is further supported by, and perhaps further subdivided by, uniquely northern chlorotypes (C6 - C11). In particular, C10 and C6 (private to the Yellowstone region) support the contention that low fossil *Pseudotsuga* pollen abundances during the full glacial (Baker 1983; Whitlock 1993) could represent a refugium in the region (Gugger & Sugita, in press; Chapter 1). Furthermore, the distribution of C7, which spans northern Montana to east-central British Columbia and C11, which is private to northern Montana, could reflect a small refugium even farther north, with which fossil data are not available to compare (Fig. 2.6). Finally, two private chlorotypes (C8, C9) in eastern Oregon suggest that this geographically distinct region could have also been a distinct glacial refugium.

Based on other phylogeographic data, the northern Rockies are thought to have been a refugium for diverse mesic plant (Richardson *et al.* 2002; Brunsfeld & Sullivan 2005, Brunsfeld *et al.* 2007; Godbout *et al.* 2008) and animal taxa (Nielson *et al.* 2001;

Carstens *et al.* 2004; Steele *et al.* 2005). Moreover, this is consistent with fossil evidence for a Douglas-fir refugium close to the ice margin in western Washington (Barnosky 1981, 1985a; Gugger & Sugita, in press; Chapter 1). Molecular phylogeographic studies and reinterpretations of the fossil record for a variety of North American and European taxa have revealed cryptic northern refugia (McLachlan *et al.* 2005; Anderson *et al.* 2006; Petit *et al.* 2008; Hu *et al.* 2009; Gugger & Sugita in press; Chapter 1). Douglas-fir could have survived on the unglaciated portions at the base of mountain ranges and ice-free ridges in the northern Rockies (Thompson *et al*. 1993; Whitlock 1995).

Similarly, the southern region defined by mtDNA SAMOVA and the distribution of some chlorotypes (C13 – C15, C20) fit with expectations of a southern refugium from the fossil record (Fig. 2.1b). Again, chlorotypes suggest that the southern Rocky Mountain region may have been subdivided into multiple small refugia, consistent with isolation on distinct mountain chains.

Defining the number and extent of refugia in the central Rockies is less clear. Although at least one to two central Rocky Mountain refugia are observed in both fossil data (Fig. 2.1b) and the mtDNA SAMOVA groupings, their boundaries do not align as expected (Fig. 2.6). For example, the northern Utah group appears to have a distribution associated with the north and east shores of glacial Lake Lahontan (present-day Great Salt Lake), and the other central group populations are less well defined, spanning Nevada and the North Rim of the Grand Canyon to northern Colorado and Wyoming. In contrast, the limited fossil record in the region suggested that southern Utah was distinct from the remaining central Rocky Mountain populations (Gugger & Sugita, in press; Chapter 1). The primarily central Rocky Mountain cpDNA clade (C16 – C19) might also support a single central refugium. However, private chlorotypes in isolated mountain chains in Nevada (C18), northern Utah (C12, C17), and western Colorado (C19), and a private, genetically divergent mitotype (M7) in western Colorado suggest a more complicated scenario with multiple small refugia scattered throughout the region.

Overall, the absence of southern and central mitotypes and chlorotypes in the north supports the hypothesis that migration was primarily elevational (Gugger & Sugita, in press; Chapter 1), not latitudinal, in the Southwest, as has been inferred for white fir

(*Abies concolor*), bristlecone pine (*Pinus longaeva*), limber pine (*Pinus flexilis*, Mitton *et al.* 2000), and many other taxa in the region (Maher 1961, 1963; Spaulding *et al.* 1983; Hall 1985).

Differences in mitotype and chlorotype distributions that may reflect differences in seed and pollen dispersal are apparent. Coalescent tests with mtDNA rejected one-refugium hypotheses in many cases, whereas tests with cpDNA never rejected a one-refugium hypothesis (Table 2.5). Similarly, mtDNA SAMOVA split the Rockies into northern, central and southern groups, whereas cpDNA SAMOVA did not split the Rockies at all, even though some chlorotypes have northern, central, or southern Rocky Mountain distributions. The sharing of chlorotypes across boundaries defined by mitotypes could reflect exchange of pollen, but not seed, among regions. This pattern is widely observed in studies of conifers that compare maternally and paternally inherited organellar DNA data (Dong & Wagner 1994; Tsumura & Suyama 1998; Latta & Mitton 1999; Liepelt *et al.* 2002; Burban & Petit 2003; Petit *et al.* 2005; Jaramillo-Correa *et al.* 2006; Du *et al.* 2009).

Among other reasons (Knowles & Maddison 2002), this cautions against overinterpreting the location of refugia from haplotype distributions alone. However, our analysis poses hypotheses that could be tested with new fossil data from those regions (Fig. 2.6).

Finally, we caution that much of the observed sequence variation could predate the last glacial cycle. The mutation rates for cpDNA and mtDNA suggest one base substitution per locus per 800,000 to 1 million years (Table 2.2). If so, it is striking that any regional structure does exist, given the repeated range expansions and contractions expected across multiple glacial cycles. That it does suggests that particular regions may have repeatedly served as refugia for the same populations across multiple climate cycles.

Postglacial migration, Holocene secondary contact and introgression

The postglacial colonization of Canada is characterized by the northward expansion and secondary contact of Douglas-fir's two long-isolated varieties (Fig. 2.4, 2.6; Table 2.3).

Maternally inherited (*i.e.* seed-dispersed) mtDNA indicates that the Rocky Mountain variety migrated primarily along the west slope of the Canadian Rockies, while the coastal variety spread north along the coast and northeast into the interior of British Columbia, east of the Coast Range. Both varieties seem to have stopped at the same barrier, which from north to south, begins as the Rocky Mountain Trench, then follows the Columbia River downstream until Revelstoke, BC, where it veers west (possibly along the route of the Trans-Canada Highway), until regaining a southward trajectory along the Okanagan Valley into central Washington (Fig. 2.4, 2.6). Much of this narrow barrier is set apart from the surrounding areas by its low elevation, long north-south lakes, and relatively dry conditions favoring grass communities. There is some evidence of intermixing in the northernmost part of the range in British Columbia, where the Trench is less sharply cut into the landscape, the terrain settles to the more uniform Interior Plateau, and Douglas-fir inhabits lower elevations.

Pollen appears to have readily dispersed across the seed dispersal barrier and spread throughout British Columbia east of the Coast Range, generating a broad hybrid zone 450 km wide. Chloroplast DNA introgression was bi-directional among varieties, despite the prevailing westerly wind. Evidence for this effect is clear in Fig. 2.4, where coastal and Rocky Mountain chlorotypes are common on both sides of the barrier to seed dispersal as defined by mtDNA. Additional support comes from the higher genetic diversity in the contact zone (central, northern, and eastern BC) compared to surrounding "pure" regions (western BC, northern U.S. Rockies; Table 2.3), and the fact that SAMOVA of mtDNA grouped the central British Columbian population with the coast, whereas the cpDNA SAMOVA grouped it with the northern Rockies. Overall, this process could not have begun prior to the start of the Holocene (11.7 ka) because ice still covered much of British Columbia, but the precise contact time remains unclear for lack of fossil sites in the area (Gugger & Sugita, in press; Chapter 1).

Some have argued that the hybrid zone extends down the east slope of the Cascades, possibly due to the southward migration of the Rocky Mountain variety (Sorensen 1979; St. Clair *et al.* 2005). The existence of coastal, but no Rocky Mountain, mitotypes in the east Cascades of Washington and Oregon argues against southward

migration and in favor of migration from the coast (Fig. 2.4). However, the east Cascades site in Washington contains a mixture of chlorotypes from both varieties, suggesting hybridization by pollen dispersal from the Rockies. This could explain the similarity of phenology and growth traits of those populations to the Rocky Mountain variety (St. Clair *et al*. 2005). Interestingly, allozymes (Li & Adams 1989) and putatively neutral nuclear SNPs display weak differentiation across the crest of Cascades in Washington, but some SNPs associated with cold-hardiness traits are strongly differentiated (Eckert *et al*. 2009). This might suggest only limited gene flow from the Rocky Mountain variety to the east Cascades and subsequent strong selection for cold-adapted Rocky Mountain alleles. Douglas-fir did not become abundant in the east Cascades until relatively recently (early to mid-Holocene; Whitlock & Bartlein 1997), consistent with the idea that opportunities for gene flow among varieties may have been limited. In Oregon, adaptive traits among varieties gradually transition from the Oregon Cascades to the Blue Mountains in eastern Oregon (Sorensen 1979), consistent with allozyme variation (Li & Adams 1989). However, studies of mitochondrial randomly amplified polymorphic DNA (RAPD) showed a sharp division among varieties in the region (Aagaard *et al.* 1995). Our data reconcile this discrepancy by showing that western and eastern Oregon have distinct maternal origins (mtDNA), but that the coastal variety in the Cascades made paternal contributions (cpDNA) to the Rocky Mountain variety in the Blue Mountains (Fig. 2.4).

Loss of diversity along migration routes is thought to depend on the width of the migration front, migration distance, the frequency of long-distance dispersal, and the presence of topographic obstacles (Davies *et al*. 2004; Bialozyt *et al*. 2006). European trees show severe losses in diversity at high latitudes (Petit *et al*. 2003) due to the action of founder effects from rare long-distance dispersal (Petit *et al*. 1997) and topographic barriers (Mátyás & Sperisen 2001) along narrow migration corridors. Those patterns have not been observed in eastern North America, where broad migration fronts lacking major topographic obstacles may have preserved diversity at northern range limits (McLachlan *et al*. 2005; Magni *et al*. 2005; Gugger *et al*. 2008). Given the relatively narrow, topographically complex corridors along the west coast and northern Rocky

Mountains, it is somewhat surprising that Douglas-fir shows no significant loss of diversity with latitude (Fig. 2.4; Table 2.3). This pattern cannot be explained alone by the admixture of the two varieties (Walter & Epperson 2001; Petit et al. 2002b) because almost all haplotypes from northern refugia of each variety are found in Canada. Instead, either dispersal and topographic bottlenecks were not important, or low diversity in the region precluded our ability to observe this effect.

With the postglacial migration pathways defined by mtDNA and cpDNA, we are able to revise postglacial migration rate estimates previously based solely on fossil evidence (Tsukada 1982; Gugger & Sugita, in press; Chapter 1). Radiocarbon-dated fossil records suggest that Douglas-fir migration tracked the ice sheets as they began to recede starting about 18 ka (Gugger & Sugita, in press; Chapter 1). The precise timing of northernmost colonization is not known for lack of fossil sites, but a site whose dating was approximated based on correlations with other nearby radiocarbon dated sites suggests the coastal variety could have reached the present-day northern range limit as early as 9 ka (Heusser 1960). Dividing migration distance from refugium to northernmost site by time, we estimate a mean migration rate from coastal sources near the ice margin to central British Columbia of about 55 – 110 m/yr, depending on the time of northernmost colonization (present-day or 9 ka). From northern Rocky Mountain sources to central British Columbia, the rate was about 60 – 165 m/yr, where the lower bound is estimated from a present-day colonization of the northern limit from a refugium near the ice margin (limited cpDNA evidence, no fossil evidence) and the upper bound is estimated from colonization of the northern limit 9 ka from a source near Yellowstone (mtDNA, cpDNA, and limited fossil evidence). These agree with recent estimates from the fossil record (50 – 220 m/yr; Gugger & Sugita, in press; Chapter 1) and are consistent with migration rates estimated from modern seed dispersal models (80 m/yr; Thompson & Schorn 1998). However, they are substantially less than earlier estimates (450 m/yr; Tsukada 1982), and more importantly, they are far lower than migration rates required to keep pace with projected 21st century climate change in some regions. For example, in the absence of adaptive evolution, Douglas-fir is expected to disappear from much of its present range in Mexico, parts of the southern Rockies and coast, and central British

Columbia, but could survive in parts of the Great Basin and northern British Columbia/southeast Alaska (Shafer *et al*. 2001). To colonize those regions in the next 100 years could require a migration rate greater than 1000 m/yr, far higher than any estimate considered likely naturally. This suggests human-assisted migration could be considered in some regions to help Douglas-fir reach some potential future habitats (McLachlan *et al.* 2007).

Synergy of fossil and molecular data

Combining fossil and molecular data provides insight into the geological and climatological causes of and patterns of population divergence that neither can provide independently (McLachlan *et al.* 2005; Magri *et al.* 2006; Hu *et al.* 2009). Here, we demonstrated that molecular data can be used to test alternative hypotheses suggested by the fossil record, which poorly resolves populations. When molecular and fossil datasets agree, we can attribute patterns of modern population structure to the timing of geologic and climatologic processes implied by fossils. When they disagree, we are forced to reassess our understanding of the causes or consequences of population divergence. Discordance among datasets was limited. However, lack of molecular variation on the coast prevented formal tests of expected population structure there. We identified concordance of molecular and fossil datasets for the Pliocene divergence of varieties, some of the multiple Pleistocene refugia in the Rockies, and the general postglacial migration patterns (Fig. 2.6). Molecular data added value by constraining the range of postglacial migration rate estimates and generating new hypotheses for possible additional glacial refugia in eastern Oregon, the central Rockies, and near the ice margin at the Last Glacial Maximum in the northern U.S. Rockies. Future fossil pollen studies in each region could test those hypotheses. Finally, molecular data exposed an interesting hybrid zone in Canada driven by paternal gene flow, a pattern that could not have been shown from fossil data alone.

Tables

Table 2.1. Primers developed for this study.

Locus	Segment	Amplified length (bp)	Primer name	Primer sequence (5' - 3')	Annealing temperature (°C)
mtDNA	*19S rDNA V7*	690	V7_3f	GAGCCAAGGAGGCAGATTG	58
			V7_3r	ATCCTTGGTCTGATGCTTCG	
	nad7i1	300	nad7i1_2f	ACCTAACAGAACGCACAAGG	55
			nad7i1_2r	TTCCAACCAAGAATTGATCC	
cpDNA	*rps7 - trnL*	1810	rps7f	GGTTATTAGGGGCATCTCG	58
	(*ΨndhB*)		trnLr	CGTGTCTACCATTTCACCATC	
	rps15 - psaC	1630	rps15f	GGTATCCGTGGGCTAAAAAC	58
	(*ΨndhH/I/E*)		psaCr	CAATACATCTGTGGGACAAGC	

Table 2.2. Jukes-Cantor-corrected sequence divergence (d_{XY}) with standard deviation (SD), mutation rates among *Larix* and *Pseudotsuga*, and results of molecular clock tests among species and within *P. menziesii*.

		mtDNA	cpDNA
d_{XY} (SD)		0.05260 (0.04332)	0.04413 (0.03636)
μ_y*		5.26×10^{-10}	4.41×10^{-10}
μ_g†		5.26×10^{-8}	4.41×10^{-8}
μ_{IMa}‡		4.81×10^{-7}	6.66×10^{-7}
Indel rate*		2.55×10^{-10}	1.49×10^{-10}
Model among species		JC	TVM
Molecular clock hypothesis	df	2	2
	-2Λ	2.77	6.75
	P	0.25	0.034
Model within P. menziesii		F81	F81
Molecular clock hypothesis	df	188	217
	-2Λ	64.60	266.63
	P	1	0.012

* Rate per site per year

† Rate per site per generation ($T = 100$)

‡ Rate per locus per year

Table 2.3. Diversity and neutrality measures for 16 populations, 4 regions, and overall dataset for mtDNA and cpDNA (next page)

		mtDNA							
	Population	***n***	***h***	***h_r****	***S***	***H* (SD)**	**$\pi \times 10^3$ (SD)**	***D*†**	***R_2*†**
1	Northern CA	12	2	1.4	1	0.17 (0.13)	0.18 (0.293)	-1.14	0.28
2	Western OR	14	1	1	0	0	0	0	0
3	Western WA	16	1	1	0	0	0	0	0
	U.S. Coast	**42**	**2**	**2**	**1**	**0.05 (0.05)**	**0.052 (0.142)**	***-1.12***	**0.15**
4	Western BC	9	2	1.6	1	0.22 (0.17)	0.241 (0.354)	-1.09	0.31
5	Central BC	14	1	1	0	0	0	0	0
6	Northern BC	6	2	2	4	0.53 (0.17)	2.272 (1.688)	-1.18	0.27
7	Eastern BC	19	2	1.3	4	0.11 (0.09)	0.448 (0.476)	-1.86	0.22
	Canada	**48**	**3**	**2.9**	**5**	**0.53 (0.03)**	**2.192 (1.388)**	**2.05**	**0.2**
8	Northern U.S. Rockies	15	2	1.3	1	0.13 (0.11)	0.143 (0.252)	-1.16	0.25
9	Eastern OR	10	1	1	0	0	0	0	0
10	Yellowstone area	17	2	1.9	1	0.44 (0.10)	0.472 (0.496)	0.95	0.22
	Northern U.S. Rockies	**42**	**2**	**2**	**1**	**0.25 (0.08)**	**0.269 (0.343)**	**0.11**	**0.13**
11	Northern UT	5	2	2	1	0.60 (0.18)	0.642 (0.704)	1.22	0.3
12	NV/southern UT	8	1	1	0	0	0	0	0
13	Northern CO	6	1	1	0	0	0	0	0
14	Southern CO	14	3	2.4	3	0.56 (0.13)	0.953 (0.793)	-0.17	0.15
15	AZ	13	2	2	1	0.54 (0.06)	0.577 (0.572)	1.48	0.27
16	NM	12	2	1.4	1	0.17 (0.13)	0.178 (0.289)	-1.14	0.28
	Southern U.S. Rockies	**58**	**4**	**3.9**	**4**	**0.59 (.033)**	**0.796 (0.663)**	**-0.31**	**0.09**
	Overall	**190**	**7**	**-**	**10**	**0.74 (0.01)**	**2.804 (1.663)**	**1.27**	**0.13**

Table 2.3. (cont.)

		cpDNA							
	Population	***n***	***h***	***h_r*‡**	***S***	***H* (SD)**	***$\pi\times10^3$* (SD)**	***D*†**	***R_2*†**
1	Northern CA	14	2	1.6	1	0.26 (0.14)	0.175 (0.229)	-0.34	*0.13*
2	Western OR	15	2	1.3	1	0.13 (0.11)	0.088 (0.156)	-1.16	0.25
3	Western WA	20	3	1.5	3	0.20 (0.12)	0.198 (0.241)	*-1.72*	0.16
	U.S. Coast	**49**	**4**	**4**	**4**	**0.19 (0.07)**	**0.158 (0.205)**	***-1.67***	**0.08**
4	Western BC	9	1	1	0	0	0	0	0
5	Central BC	17	4	2.5	4	0.60 (0.10)	0.797 (0.598)	0.06	0.15
6	Northern BC	6	2	2	2	0.60 (0.13)	0.793 (0.674)	1.75	0.3
7	Eastern BC	24	4	2.4	3	0.59 (0.08)	0.697 (0.534)	0.79	0.18
	Canada	**56**	**6**	**5.6**	**5**	**0.59 (0.04)**	**0.761 (0.554)**	**0.14**	**0.11**
8	Northern U.S. Rockies	21	5	2	6	0.35 (0.13)	0.378 (0.356)	*-2.06*	0.12
9	Eastern OR	12	4	2.5	3	0.56 (0.15)	0.511 (0.450)	-0.73	0.14
10	Yellowstone area	18	6	2.6	4	0.56 (0.13)	0.424 (0.386)	-1.35	*0.09*
	Northern U.S. Rockies	**51**	**#**	**9.8**	**8**	**0.41 (0.09)**	**0.508 (0.420)**	**-1.54**	***0.05***
11	Northern UT	8	5	3.8	5	0.86 (0.11)	1.015 (0.771)	-0.92	0.15
12	NV/southern UT	8	3	2.6	2	0.68 (0.12)	0.543 (0.491)	0.24	0.22
13	Northern CO	7	3	2.7	2	0.67 (0.16)	0.567 (0.516)	0.21	0.23
14	Southern CO	14	3	1.7	2	0.28 (0.15)	0.189 (0.239)	-1.48	0.17
15	AZ	14	2	1.8	2	0.36 (0.13)	0.479 (0.426)	0.42	0.18
16	NM	12	4	2.3	5	0.46 (0.17)	0.551 (0.474)	*-1.83*	0.15
	Southern U.S. Rockies	**63**	**#**	**8.2**	**12**	**0.52 (0.07)**	**0.851 (0.600)**	**-1.41**	***0.04***
	Overall	**219**	**#**	**-**	**18**	**0.64 (0.02)**	**0.733 (0.532)**	***-1.65***	**0.03**

* Rarefaction to 5 for populations and 42 for regions

† Italicized values were significant under coalescent simulations

‡ Rarefaction to 5 for populations and 49 for regions

Table 2.4. SAMOVA results that gave highest F_{CT} for each locus. All variance components were significant ($P < 0.001$).

mtDNA (K = 5)	df	Sum of Squares	Variance components	Percentage of variation
Among groups	4	343.18	2.474	91.5
Among pops. within groups	11	4.74	0.018	0.7
Within pops.	174	36.89	0.212	7.8

mtDNA (K = 6)	df	Sum of Squares	Variance components	Percentage of variation
Among groups	5	344.51	2.471	91.8
Among pops. within groups	10	3.41	0.01	0.4
Within pops.	174	36.89	0.212	7.9

cpDNA (K = 2)	df	Sum of Squares	Variance components	Percentage of variation
Among groups	1	92.47	1.067	68.3
Among pops. within groups	14	18.17	0.065	4.1
Within pops.	203	87.36	0.43	27.5

Table 2.5. *P*-values less than 0.05 indicate empirical and simulated distributions of *s* differed significantly, rejecting the null hypothesis. The rejection of a null one-refugium hypothesis suggests the multiple-refugia alternative in parentheses, whereas the rejection of a null multiple-refugia hypothesis with divergence time in parentheses suggests no particular alternative. Tests suggesting multiple refugia are in bold.

	mtDNA				cpDNA			
Null hypothesis	N_{ef} low	N_{ef}	N_{ef} high	N_{ef}×10	N_{em} low	N_{em}	N_{em} high	N_{em}×10
One (H_A: two)	1	**0.034**	**<0.001**	**<0.001**	1	1	1	0.985
One (H_A: three)	1	0.900	**0.023**	**<0.001**	1	1	1	0.942
One (H_A: four)	1	0.999	0.337	**<0.001**	1	1	1	0.995
Two (115 ka)	<0.001	**0.157**	**0.897**	**1**	<0.001	<0.001	<0.001	**0.947**
Two (190 ka)	<0.001	0.006	**0.386**	**1**	<0.001	<0.001	<0.001	**0.582**
Three (115 ka)	<0.001	<0.001	0.020	**1**	<0.001	<0.001	<0.001	**0.734**
Three (190 ka)	<0.001	<0.001	<0.001	**0.990**	<0.001	<0.001	<0.001	**0.131**
Four (115 ka)	<0.001	<0.001	<0.001	**0.996**	<0.001	<0.001	<0.001	**0.420**
Four (190 ka)	<0.001	<0.001	<0.001	**0.838**	<0.001	<0.001	<0.001	0.016

Table 2.6. IMa results showing t, m_1, and m_2 for a run with the full model, followed by results of the likelihood ratio tests, and then t with its 90% HPD interval for a run with $m_1 = m_2 = 0$. In each case, the values for the run with the highest effective sample size (ESS; Hey & Nielson 2007) are shown.

	Full model Priors (m_1, m_2, q_1, q_2, q_A)	t (ka)	m_1	m_2	$m_1 = m_2$ -2Λ	P	$m_1 = m_2 = 0$ -2Λ	P	**Model without migration ($m_1 = m_2 = 0$)** t (ka)	90% HPD low	90% HPD high
Both loci	**2.5. 0.5, 10, 10, 10**	**6580***	**0.0000055**	**0**	**-0.1101 ≈ 0**	**1**	**0.001**	**0.49**	**2113**	**755**	**4372**
mtDNA	2.5, 1.5, 10, 10, 10	4360*	0.0000001	0	0.1846	0.67	-0.0014 ≈ 0	≈ 0.5	4410	1813	8826
cpDNA	6, 2, 8, 30, 8	413*	0.0000002	0.0000001	0.4594	0.5	0.002	0.48	491	176	1832

* Posterior distribution for this parameter never dropped to zero, suggesting unreliable estimates.

Figures

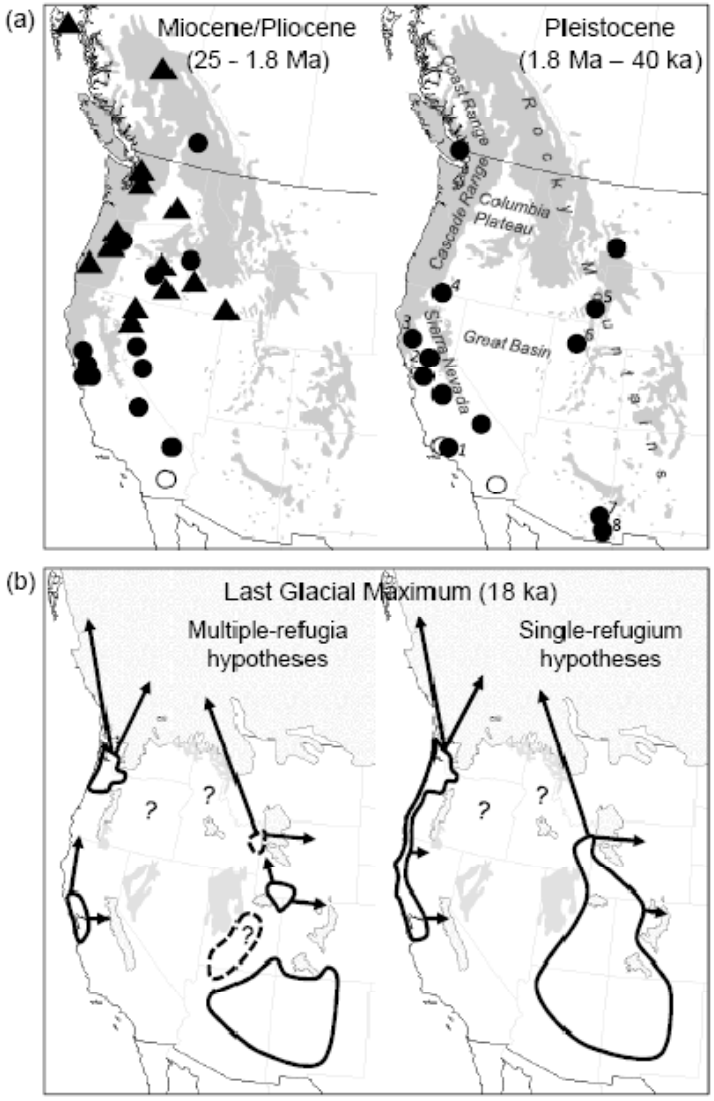

Figure 2.1. Hypotheses from the fossil record. (a) *Pseudotsuga* fossil data are consistent with the Pliocene vicariance hypothesis. Fossils indicate that it was present near the coast and in the Great Basin and Columbia Plateau during the Miocene and Pliocene, but fossils indicate that it was absent from the Great Basin and Columbia Plateau and present near the coast and in the Rockies during the Pleistocene. Miocene and Pliocene are grouped on one panel because when many of the fossil records were first published the geological time scale delineated those time periods differently with a boundary at 15 Ma. This means that many fossils originally attributed to the Pliocene were in fact from the

Miocene. We marked fossils from 25 – 15 Ma as triangles and those from 15 – 1.8 Ma as circles. Both show that *Pseudotsuga* was present in the Great Basin. Filled shapes represent fossils from the genus *Pseudotsuga*, and empty circles indicate *P. macrocarpa* in particular. Modern range of Douglas-fir is in grey (Little 1971). This figure was modified from Hermann (1985) with our additions numbered: 1 = Seacliff (Axelrod 1983), 2 = Teichert Site (Rymer 1981), 3 = Cache Formation (Ritter & Hatoff 1977), 4 = Tule Lake (Adam & Vagenas 1990), 5 = Grays Lake (Beiswenger 1991), 6 = Great Salt Lake (Davis & Moutoux 1998), 7 = Lake Cochise (Martin 1963), 8 = Safford Valley (Gray 1961). (b) Based on a synthesis of late Quaternary fossil records, Gugger & Sugita (in press; Chapter 1) proposed multiple-refugia hypotheses for the coastal and Rocky Mountain varieties (left), which are shown with ice sheets (stippled) and glacial lakes (grey) at the Last Glacial Maximum. Their analysis suggested two coastal refugia and 3 – 4 Rocky Mountain refugia (encircled). Less certain refugia are marked with a dotted line and regions where late Quaternary fossil evidence is lacking are marked with question marks. The two-refugia hypothesis for the Rocky Mountains proposed by Li & Adams (1989) would approximately cluster the three southernmost refugia, but retain the northernmost refugium shown here. Because the fossil records do not offer population-level resolution, single-refugium hypotheses were also proposed for each variety (right). For either set of hypotheses, postglacial migration (arrows) was primarily northward into Canada with some eastward migration.

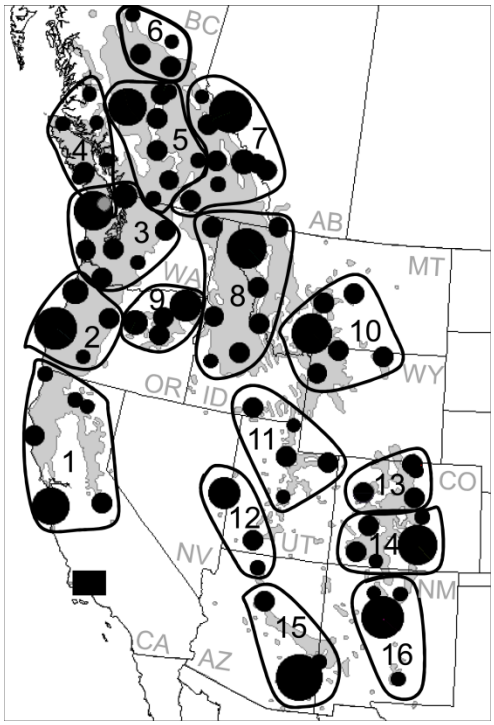

Figure 2.2. Sampling strategy. We grouped 219 individuals from 87 sites (black/grey circles) into 16 populations (lines). These populations were defined primarily along major geographic breaks, such as the Klamath Mountains (populations 1/2 boundary), Columbia River Valley (2/3), Coast Range (4/5), Okanagan Valley (5/7), Deschutes River Valley (2/9), Hells Canyon (8/9), Snake River Plain (8/11), Great Divide Basin (10/11), Grand Canyon (12/15), other wide, arid valleys in the Southwest (*e.g.* 11/13, 15/16), and previously glaciated high peaks in Colorado (13/14; Fig. 2.1). Several boundaries were already recognized as probable varietal boundaries (2/9, 4/5, 5/7) and others were hypothesized subdivisions from the fossil record (*e.g.* 1/2, 10/11, ~11/12). Other groupings (not shown) based on geography were also evaluated, but did not lead to meaningful differences in our results or conclusions. Circles are scaled to the sampling intensity (1 - 11 individuals per site). One circle is grey to set it apart from overlapping points. Rectangle marks sampling location of four *P. macrocarpa*. The modern range of Douglas-fir is shown in grey (Little 1971). States and provinces are labeled with abbreviations: AB = Alberta; AZ = Arizona; BC = British Columbia; CA = California; CO = Colorado; ID = Idaho; MT = Montana; NM = New Mexico; NV = Nevada; OR = Oregon; UT = Utah; WA = Washington; WY = Wyoming.

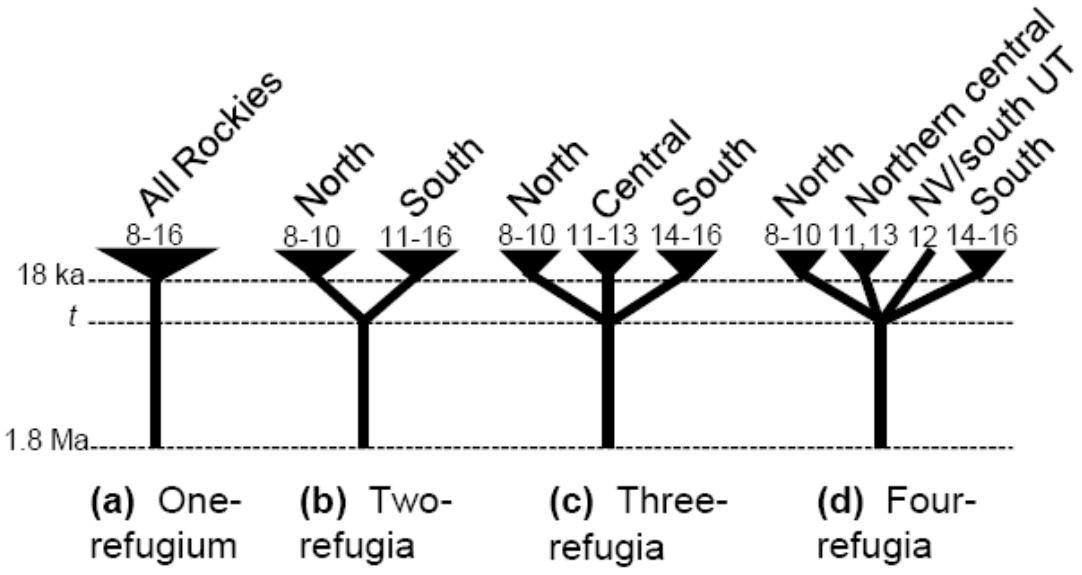

Figure 2.3. (a) One-refugium (null), (b) two-refugia, (c) three-refugia, and (d) four-refugia population models used to test hypotheses with coalescent simulations in Mesquite. Simulations were run so that a single ancestral population of constant size N_e split at time t into equal fractions of N_e. Dotted lines correspond to the Last Glacial Maximum (18 ka), the onset of recent glaciations (t; Wisconsinan = 115 ka, late Illinoisan = 190 ka), and the beginning of the Pleistocene (1.8 Ma). Population numbers matching those on Fig. 2.2 are shown above population groups.

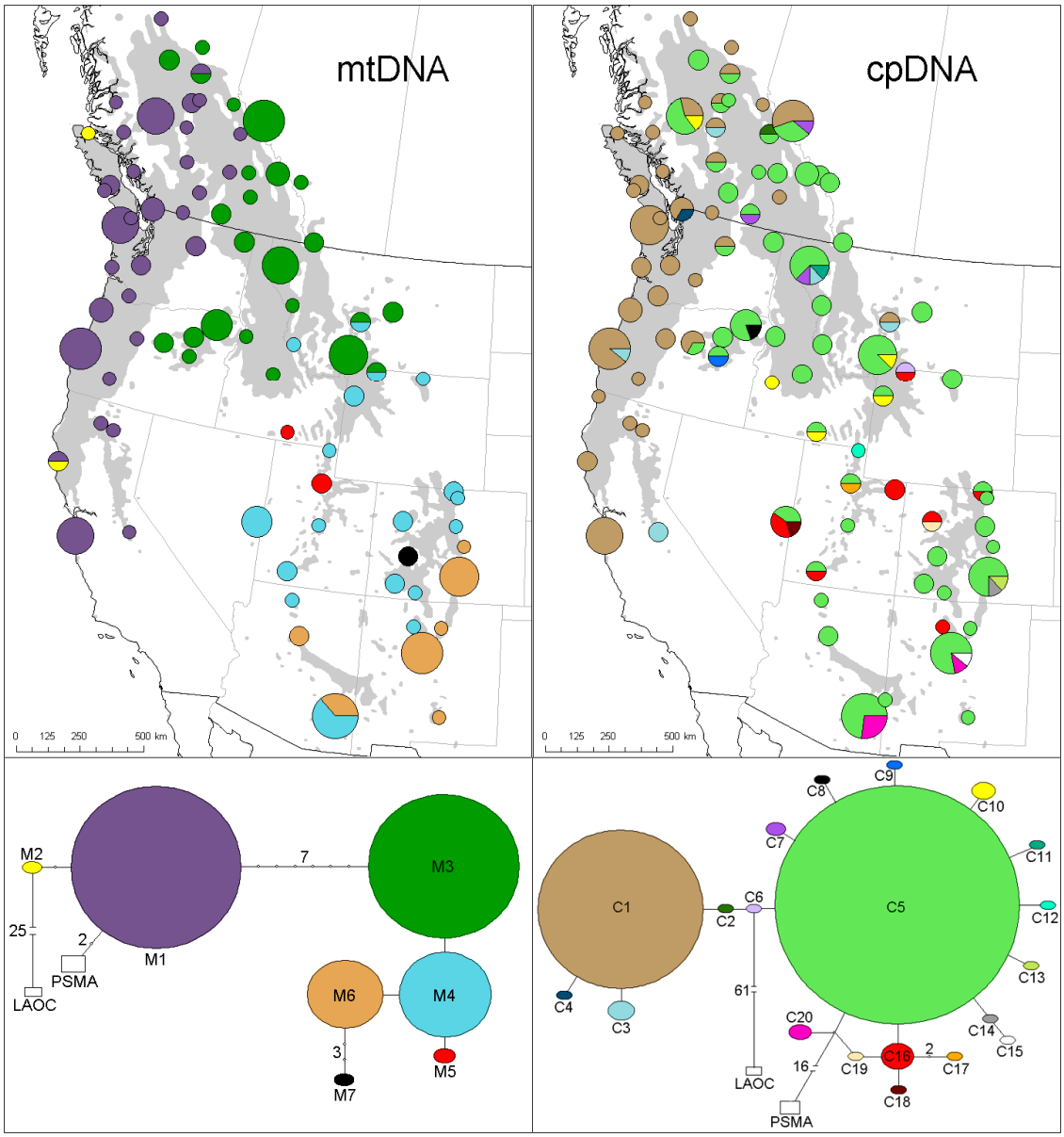

Figure 2.4. Maps with sample sites colored according to mitotype (left) and chlorotype (right) composition and proportioned according to sampling intensity. Below, parsimony networks show the relation of those mitotypes (left) and chlorotypes (right), with the size of the circles proportional to the frequency in the overall dataset. Modern range of Douglas-fir is shown in grey (Little 1971).

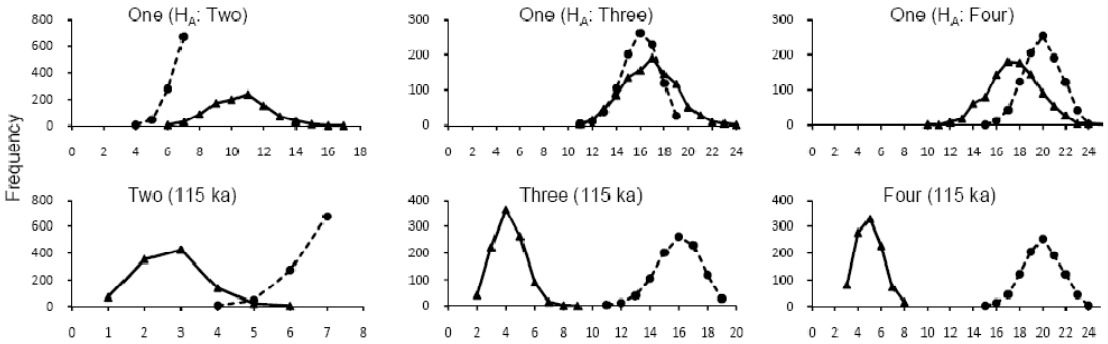

Figure 2.5. Examples of distributions of *s* for simulated (solid line) and empirical (dashed line) trees for each refugial hypothesis for mtDNA data with N_{ef} = 4667 and a splitting time of 115 ka. The one-refugium hypothesis was rejected in favor of the two-refugia hypothesis but was not rejected in favor of three-refugia and four-refugia alternatives (top row of graphs). Three-refugia and four-refugia hypotheses were rejected, but the two-refugia hypothesis was not rejected (bottom row of graphs). *P*-values are reported in Table 2.5.

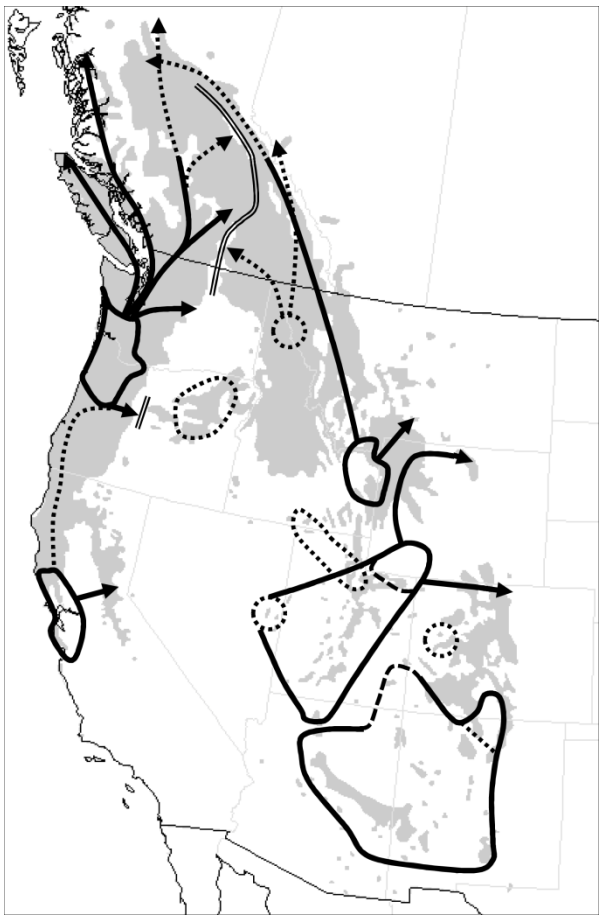

Figure 2.6. Summary of glacial refugia and postglacial migration routes based on fossil and genetic data. Solid lines indicate refugia and postglacial migration routes supported by fossil and molecular data, dashed lines indicate those supported by fossil data only, and dotted lines indicate those supported by molecular data only. Double line marks some barriers to seed dispersal as suggested by mtDNA.

Supporting information

Table 2.S1. Population and sample site information with mitotype and chlorotype frequencies.

(Included as supplementary file: Table 2.S1.xlsx)

Table 2.S2. Mitotype and chlorotype definitions in terms of *V7*, *nad7i1*, *rps7-trnL*, and *rps15-psaC* haplotypes reported to GenBank (accessions in parentheses).

Mitotype	*V7* haplotype	*nad7i1* haplotype
M1	H1 (GQ999615)	H1 (GQ999625)
M2	H8 (GQ999622)	H1 (GQ999625)
M3	H1 (GQ999615)	H2 (GQ999626)
M4	H3 (GQ999617)	H2 (GQ999626)
M5	H5 (GQ999619)	H2 (GQ999626)
M6	H2 (GQ999616)	H2 (GQ999626)
M7	H4 (GQ999618)	H2 (GQ999626)
PSMA	(GQ999623)	(GQ999628)
LAOC	(GQ999624)	(GQ999629)

Chlorotype	*rps7-trnL* haplotype	*rps15-psaC* haplotype
C1	H1 (GQ999630)	H1 (GQ999644)
C2	H2 (GQ999631)	H1 (GQ999644)
C3	H3 (GQ000632)	H1 (GQ999644)
C4	H12 (GQ000641)	H1 (GQ999644)
C5	H2 (GQ999631)	H2 (GQ999645)
C6	H2 (GQ999631)	H8 (GQ999651)
C7	H2 (GQ999631)	H3 (GQ999646)
C8	H2 (GQ999631)	H6 (GQ999649)
C9	H2 (GQ999631)	H17 (GQ999660)
C10	H6 (GQ000635)	H2 (GQ999645)
C11	H11 (GQ000640)	H2 (GQ999645)
C12	H2 (GQ999631)	H10 (GQ999653)
C13	H4 (GQ000633)	H2 (GQ999645)
C14	H2 (GQ999631)	H9 (GQ999652)
C15	H2 (GQ999631)	H16 (GQ999659)
C16	H5 (GQ000634)	H2 (GQ999645)
C17	H5 (GQ000634)	H7 (GQ999650)
C18	H7 (GQ000636)	H2 (GQ999645)
C19	H5 (GQ000634)	H5 (GQ999648)
C20	H2 (GQ999631)	H4 (GQ999647)
PSMA	(GQ000642)	(GQ999664)
LAOC	(GQ000643)	(GQ999665)

Chapter 3

Southward Pleistocene migration of Douglas-fir into Mexico: phylogeography, ecological niche modeling, and conservation of 'rear edge' populations

Abstract

The southern range limit of temperate trees in the Northern Hemisphere has received less attention than the northern range limit, yet southern populations may be an important reserve of genetic diversity in changing climates. Douglas-fir (*Pseudotsuga menziesii*) is a widespread, ecologically and economically important tree in western North America, but the taxonomy, biogeography, and Pleistocene history of the small, isolated populations at its southern range limit in Mexico are largely unknown. We used mtDNA sequences, cpDNA sequences, and chloroplast microsatellites (cpSSR) to (i) test Miocene versus Pleistocene colonization/divergence times proposed based on fossils, (ii) assess the taxonomic status of Mexican populations, and (iii) explore geographic patterns of molecular variation in relation to Pleistocene climate history. We found that Mexican Douglas-fir likely diverged from Rocky Mountain Douglas-fir (*P. menziesii* var. *glauca*) in the mid-Pleistocene [958 ka (1.6 Ma – 491 ka)], consistent with paleobotanical records in the southern Rockies. Mexican Douglas-fir was largely genetically distinct from U.S. and Canadian populations, but more closely related to the Rocky Mountain variety than the coastal variety (*P. menziesii* var. *menziesii*). We concur with others who argue for Mexican Douglas-fir to be considered a third variety within *P. menziesii*. Within Mexico, genetic diversity was high (H = 0.59 - 0.91), strongly partitioned among populations (G_{ST} = 0.55 – 0.77 for sequences, 0.16 for microsatellites), and positively correlated with latitude (though not significantly). Spatial patterns of molecular variation did not correspond to major geographic regions, suggesting a complex Pleistocene history. There was evidence for modest late Pleistocene population expansion overall and in the Sierra Madre Occidental, but populations in central Mexico and the Sierra Madre Oriental did not show evidence of expansion or bottleneck, and some populations showed evidence of long-term isolation. Ecological niche models revealed mountain corridors during the Last Glacial Maximum (21 ka) that could have led to the complex patterns of haplotype distributions in Mexico. These results provide critical insights into conservation proposals for Mexican Douglas-fir.

Introduction

Phylogeographic studies of trees have largely focused on temperate species (Petit *et al.* 2003; McLachlan *et al.* 2005), especially on 'leading edge' populations that have responded most dramatically to glacial cycles (Petit *et al.* 1997; Gugger *et al.* 2008). Less attention has been paid to 'rear edge' populations or to species that span temperate and subtropical zones. Rear edge populations have often been long-isolated, display strong differentiation, and may exhibit local adaptation in response to strong selection and lack of gene flow (Hampe & Bairlein 2000; Chang *et al.* 2004; Martin & McKay 2004; Hampe & Petit 2005; Parisod & Joost 2010). Thus the conservation of rear edge populations may be especially important under changing climates.

Douglas-fir (*Pseudotsuga menziesii*) is a widespread, ecologically and economically important conifer in western North America, but our understanding of the small, isolated 'sky island' populations in Mexico is limited. In Mexico, Douglas-fir inhabits the Madrean pine-oak (*Pinus*-*Quercus*) biodiversity hotspot, which covers the middle to high elevations of the Sierra Madre Occidental, Sierra Madre Oriental, Trans-Mexican Volcanic Belt, and Sierra Madre del Sur (Fig. 3.1). This region contains nearly 4,000 endemic plant species, of which at least 20 tree species or subspecies in Pinaceae, including Douglas-fir, are considered threatened (Conservation International; Norma Oficial Mexicana 1994, 2001; Farjon & Page 1999). Intensive logging and projected climate warming pose a continued threat to many of these taxa (Gómez-Mendoza & Arriaga 2007).

Tertiary *Pseudotsuga* fossils are relatively common in the northern Rocky Mountains and Pacific coastal states (Hermann 1985) and support a northern North American (or Asian) origin for the genus (Schorn 1994; Gernandt & Liston 1999). The colonization of Mexico may have occurred in the early Miocene (~20 Ma) when many other temperate taxa are thought to have arrived (Graham 1999) or in the Pleistocene in response to glaciation at high latitudes (Deevey 1949; Dressler 1953; Perry *et al.* 1998). A few putative *Pseudotsuga* pollen grains found in Miocene sediment from Chiapas could suggest Miocene colonization (Palacios-Chavez & Rzedowski 1993), however,

Pseudotsuga and *Larix* pollen cannot be distinguished (Barnosky 1985). Alternatively, Pleistocene colonization is supported by the much later first appearance of fossil *Pseudotsuga* pollen in the southern Rocky Mountains in the early Pleistocene (Gray 1961) or late Pleistocene (Martin 1963) in southern Arizona. In Mexico, the limited Pleistocene fossil record contains no evidence of Douglas-fir (Brown 1985; Gugger & Sugita, in press; Chapter 1).

Studies have suggested to varying degrees that some Mexican populations are ecologically (Vargas-Hernández *et al.* 2004; Acevedo-Rodríguez *et al.* 2006), morphologically (Reyes-Hernández *et al.* 2005, 2006), and genetically (Li & Adams 1989) distinct from those in the United States and Canada. Consequently, their taxonomic status remains uncertain. Flous (1934a, b) and Martínez (1949) proposed multiple species of Douglas-fir in Mexico on the basis of leaf and cone morphology. Others argued that Mexican populations of Douglas-fir belong to the Rocky Mountain variety of Douglas-fir (*P. menziesii* var. *glauca*), whose northern limit extends into central British Columbia (Little 1952; Hermann & Lavender 1990). Still others have proposed an intermediate classification, arguing that Mexican Douglas-fir is best treated as a distinct variety within *P. menziesii* (Reyes-Hernández *et al.* 2006; Earle 2009).

In a range-wide survey of allozyme variation in Douglas-fir that included two populations in northeastern Mexico, Li & Adams (1989) found that one population was very distinct from U.S. populations, whereas the other population was very similar to Rocky Mountain Douglas-fir. This suggested the possibility of the Rocky Mountain variety coexisting with a possible Mexican species. However, there are no other studies of molecular genetic variation in Mexican Douglas-fir populations to date.

Morphological studies also suggest differences among U.S. and some Mexican populations. Principal components analysis of leaf and cone traits showed that central Mexican populations were most different from U.S. populations and that intervening populations in northern Mexico were intermediate or similar to Rocky Mountain Douglas-fir (Reyes-Hernández *et al*. 2006). These data also showed that northern populations exhibited more heterogeneity of trait values than those in the central region (Reyes-Hernández *et al*. 2005, 2006). When grown in a common garden, a similar

pattern of more heterogeneous northern populations than central populations and clear differences among Mexican and U.S. populations was observed for phenology of budset and budburst (Acevedo-Rodríguez *et al.* 2006). Trait variation within Mexico was best described by geographically defined populations or regions (Reyes-Hernández *et al*. 2005, 2006) and not among putative taxa proposed by Flous (1934a,b) and Martínez (1949). Together, these studies suggest that Mexican populations, especially central Mexican populations, could be considered a separate variety from Rocky Mountain Douglas-fir, with the possibility of a gradual transition zone among them.

Compared to most U.S. and Canadian populations, all the Mexican populations are small and fragmented, ranging from a few dozen to a few thousand individuals (Mápula-Larreta *et al.* 2007; Velasco-García *et al.* 2007). Not surprisingly, a number of studies suggest low fertility and seedling recruitment rates due to inbreeding depression. For example, Mexican populations had a lower proportion of fully developed seeds, lower weight seeds, and therefore lower seed efficiency compared to U.S. populations (Gashwiler 1969; Owens *et al.* 1991; Webber & Painter 1996; Vargas-Hernández *et al.* 2004; Mápula-Larreta *et al.* 2007; Velasco-García *et al.* 2007). Within Mexico, the lowest seed efficiencies and seedling recruitment rates were found in central populations, which are among the smallest and most isolated in Mexico (Juárez-Agis *et al.* 2006; Mápula-Larreta *et al.* 2007; Velasco-García *et al.* 2007). A study using isozymes to detect mating patterns revealed that all Mexican populations showed high rates of inbreeding due to local pollination, and those in the central region exhibited the highest rates of selfing (Cruz-Nicolás *et al.* 2008).

Accordingly, Mexican Douglas-fir populations are listed as "rare" and "subject to special protection" by the Mexican Government (Norma Oficial Mexicana 1994, 2001), and conservation strategies have been proposed. For example, Cruz-Nicolás *et al*. (2008) proposed crossing central Mexican populations to alleviate the more severe effects of inbreeding there. Vargas-Hernández *et al*. (2004) considered the possibility of using genetic reserves in Mexican populations to populate parts of the United States in the face of future climate change.

These recommendations have been proposed in the absence of an understanding of the history and genetic structure of these populations. Here, we investigate mitochondrial (mtDNA) and chloroplast DNA (cpDNA) sequence and cpDNA microsatellite (cpSSR) variation in 11 populations throughout Mexico to (i) test Miocene versus Pleistocene colonization/divergence times proposed based on the fossils, (ii) assess the taxonomic status of Mexican populations, and (iii) explore geographic patterns of molecular variation in relation to Pleistocene climate history. We also expect diversity to be positively correlated with latitude because latitude is correlated with morphological and phenological diversity (Reyes-Hernández *et al*. 2005, 2006), fertility rates (Acevedo-Rodríguez *et al.* 2006), and population size (Mápula-Larreta *et al.* 2007; Velasco-García *et al.* 2007).

Materials and methods

Sampling

Leaf tissue was collected from 9 – 16 individuals (mean = 11.6) from 11 natural populations throughout most of the range of Douglas-fir in Mexico (Fig. 3.1; Table 3.S1). Samples were kept on ice and then stored at -80°C. In addition, pressed herbarium vouchers were made for each sample. Representative vouchers from each population were deposited in Herbario Nacional de México (MEXU; Universidad Nacional Autónoma de México, México, DF) and Herbario del Centro Regional del Bajío (IEB; Pátzcuaro, Michoacán), and the remaining vouchers are stored in the A. González-Rodríguez lab.

DNA preparation

Total genomic DNA was extracted from leaf tissue using QIAGEN DNeasy Plant Mini Kit according to the manufacturer's instructions.

Two mtDNA and two cpDNA segments shown to neutrally vary in U.S. and Canadian populations of Douglas-fir were amplified, sequenced, and aligned according to previously established procedures (Gugger *et al.* 2010; Chapter 2). The cpDNA segments were *rps7-trnL* and *rps15-psaC*, containing *ndh* pseudogenes, intergenic spacers, and an intron (Wakasugi *et al.* 1994; Braukmann *et al.* 2009). The mtDNA segments were variable region 7 of the small-subunit ribosomal RNA gene (*V7*; Duff & Nickrent 1999; Duff & Nickrent 1997) and the first intron of the *nad7* gene (*nad7i1*; Jaramillo-Correa *et al.* 2004). Douglas-fir mtDNA is maternally inherited (Marshall & Neale 1992) and cpDNA is paternally inherited (Neale *et al.* 1986), thus neither generally undergoes heterologous recombination (Birky 2001), and each acts as a single locus. Therefore, we concatenated the two cpDNA sequences to produce a single cpDNA sequence and the two mtDNA sequences to produce a single mtDNA sequence.

Three chloroplast DNA simple sequence repeats (cpSSR) shown variable in Canadian populations of Douglas-fir were assayed (*Pt*26081, *Pt*63718, *Pt*71936; (Vendramin *et al.* 1996; Viard *et al.* 2001). The three loci were simultaneously amplified using the QIAGEN Multiplex PCR kit in 5 µL reactions as follows: 1× multiplex PCR master mix, 2 µM each primer, deionized water, and 20 ng DNA (Cortés-Palomec *et al.* 2008). The thermal cycling program consisted of one cycle at 94 ºC for 2 min, and then 35 cycles, each at 94 ºC for 1 min, 55 ºC for 1 min, and 72 ºC for 2 min. A final extension step at 72 ºC for 15 min was included. PCR products were diluted 1:1 in deionized water, combined with the GenScan-500 LIZ size standard (Applied Biosystems) and analyzed in an ABI-PRISM 3100-avant sequencer. Peak Scanner 1.0 (Applied Biosystems) was used for fragment analysis and final sizing.

Parsimony networks for all Mexican sequences plus representative sequences of haplotypes found only north of Mexico and outgroups (Gugger *et al*. 2010; Chapter 2) were produced separately for mtDNA and cpDNA using TCS 1.21 (Clement *et al.* 2000) with insertion-deletions coded as a fifth state. A network for cpSSR haplotypes was constructed manually in TCS assuming a stepwise mutation model (SMM; Ohta & Kimura 1973). Because cpSSRs likely undergo extensive homoplasious mutation, this network may not be phylogenetically informative.

Genetic diversity

Patterns of genetic diversity were quantified for each population and for three geographic regions commonly defined for Mexican Douglas-fir: Sierra Madre Occidental (I), Sierra Madre Oriental (II), and central Mexico (III). For each population and region, we estimated haplotype richness (h), haplotype richness after correcting for unequal sample sizes with rarefaction (h_r; Hurlbert 1971), and haplotype diversity (H; Nei 1987) using Contrib 1.01 (Petit *et al.* 1998). For sequence data, we calculated nucleotide diversity (π; Tajima 1983; Nei 1987) using Arlequin 3.11 (Excoffier *et al.* 2005) and for cpSSR data, we calculated $\overline{D}^2_{SH}$, which is the mean pairwise genetic distance among individuals within a population under a SMM (Goldstein *et al.* 1995a; Vendramin *et al.* 1998). Latitudinal trends in the diversity (h_r and H for combined cpDNA and cpSSR; π for cpDNA sequences; $\overline{D}^2_{SH}$ for cpSSR) of populations were investigated with linear regression.

Changes in population size

To test for recent population bottlenecks, observed haplotype diversity (H from above) was compared to expected haplotype diversity (H_{eq}) under an infinite alleles model (IAM; (Kimura & Crow 1964) or SMM at mutation-drift equilibrium using coalescent simulations in BOTTLENECK 1.2 (Piry *et al.* 1999). H significantly greater than H_{eq} indicates a recent genetic bottleneck (Cornuet & Luikart 1996).

To test for population expansion with DNA sequence data, we calculated R_2 (Ramos-Onsins & Rozas 2002) and F_S (Fu 1997) and assessed their significance with coalescent simulations (Hudson 1990) in DnaSP 5.0 (Librado & Rozas 2009). For cpSSR, we calculated F_S in Arlequin by binary coding haplotypes following Navascués *et al*. (2006). Significant negative values of F_S or significant positive values of R_2 suggest population expansion or directional selection.

Finally, a Bayesian skyline plot (Drummond *et al.* 2005) of changes in effective population size (N_e) through time based on all Mexican cpDNA sequence data was estimated using BEAST 1.5.3 (Drummond & Rambaut 2007). We chose the Hasegawa-

Kishino-Yano substitution model (Hasegawa *et al.* 1985) with empirical base frequencies, a strict molecular clock, and a piecewise-constant coalescent Bayesian skyline tree prior with 10 starting groups (results same for 5 groups; not shown). Two runs of 20 million steps and ESS > 200 were compared to ensure convergence. Outputs were combined in LogCombiner 1.5.3 and visualized in Tracer 1.5 (Drummond & Rambaut 2009). To convert the *x*-axis (subs/site) to demographic units (years), we used a previously published mutation rate for this locus (4.41×10^{-10} subs/site/year; Gugger *et al*. 2010; Chapter 2).

Population structure

We computed genetic differentiation among populations for all loci as G_{ST} (Nei 1973; Pons & Petit 1995).

For mtDNA and cpDNA sequence data, we performed a spatial analysis of molecular variance (SAMOVA) in SAMOVA 1.0 (Dupanloup *et al.* 2002) to identify groups of genetically similar populations. SAMOVA uses a simulated annealing approach to group geographically close populations to maximize the variance (F_{CT}) among a user-defined number of groups (K). We performed this analysis for $K = 2 – 6$, and chose the number of groups that gave the highest F_{CT}.

We also performed hierarchical analysis of molecular variance (AMOVA) with populations within regions to test for significant structure among regions and to determine whether regional variation comprised a similar proportion of total variation as variation among groups defined by SAMOVA.

To identify clusters of populations sharing similar cpSSR compositions, we computed pairwise genetic distance among populations as $(\delta\mu)^2$ (Goldstein *et al.* 1995b), and used this distance matrix to create a UPGMA dendrogram (Sneath & Sokal 1973) in PAUP* 4.0b10 (Swofford 2003). Given that all cpSSRs are linked, we modified $(\delta\mu)^2$ from Goldstein *et al*. (1995b) for the case of multiple microsatellite markers (m) for a single nonrecombining locus: $(\delta\mu)^2 = \left(\sum_{k=1}^{m} |\mu_{Ak} - \mu_{Bk}|\right)^2$, where μ_{Ak} and μ_{Bk} are the mean allele size in populations A and B at the kth microsatellite marker. Bootstrap support is

always 100% based on one microsatellite locus, so we approximated support values for major branches in the dendrogram assuming that each of the three cpSSR markers was a separate locus in POPTREE2 (Takezaki *et al.* 2009).

Divergence from Rocky Mountain variety

Divergence time between Mexican and Rocky Mountain populations (Gugger *et al*. 2010; Chapter 2) was estimated using an isolation-with-migration model (Nielson & Wakeley 2001) in IMa (Hey & Nielson 2007). The full IMa model simultaneously estimates six parameters scaled by substitution rate: divergence time (t), migration from population one to two (m_1), migration from population two to one (m_2), effective population size of each population (N_1 and N_2), and effective population size of the ancestor (N_A). IMa assumes constant population size, neutral molecular markers, no recombination within loci, free recombination among loci, and a particular mutation model. We chose mutation models and priors as described in Gugger *et al*. (2010; Chapter 2), and at least three runs with ESS > 50 were compared to ensure convergence. We only trusted estimates whose posterior distribution dropped to zero within the prior intervals investigated. To scale the outputs to demographic units, we used a generation time of 100 years and per locus per year substitution rates of 4.81×10^{-7} for mtDNA and 6.66×10^{-7} for cpDNA (Gugger *et al*. 2010; Chapter 2). We also tested whether or not migration among Rocky Mountain and Mexican populations was important during divergence by conducting a two step likelihood ratio test described in Gugger *et al*. (2010; Chapter 2). If neither test was significant, the model with $m_1 = m_2 = 0$ was not rejected, so we ran the IMa analysis again with migration set to zero.

Ecological niche modeling

The potential distribution of *P. menziesii* was modeled using a maximum entropy model in Maxent 3.3.1 (Phillips *et al.* 2006) for present climate conditions, future climate conditions in the years 2050 and 2080 under the Hadley Centre Coupled Model 3

(HadCM3; Gordon *et al*. 2000; Pope *et al*. 2000), and two Last Glacial Maximum (LGM; ~21 ka) scenarios: Community Climate System Model (CCSM; Collins *et al.* 2006) and Model for Interdisciplinary Research on Climate (MIROC; Hasumi & Emori 2004). Maxent employs a maximum entropy-based machine-learning method used for making predictions when incomplete data are available. Its basic principle is to predict species distribution from occurrence localities over a finite geographical space. The goal is to find a model in which the values of the empirical environmental variables are the same as the expectations of the theoretical distribution (Ciliberti *et al.* 2009). This approach assumes that modern distributions are well-described by climate variables included in the model, biotic factors such as competition do not affect distributions, modern climatic tolerances have not changed or will not change (*i.e*. no evolution and no unobserved plasticity), individuals throughout Mexico exhibit the same climatic tolerances, there is no dispersal limitation, and models of past and future climates are accurate (Wiens *et al*. 2009).

We used a set of 74 presence points for *P. menziesii* in Mexico obtained from the Global Biodiversity Information Facility (GBIF; http://data.gbif.org/species/browse/taxon/), the Herbario Nacional de México collections, and personal observations. Climatic data were obtained from the WorldClim dataset (Hijmans & Graham 2006) for LGM and current climate scenarios with 30 arcsec and 2.5 arcmin resolutions, respectively (available at http://www.worldclim.org/download). Climatic data used for the future scenarios were provided by the CIAT downscaled GCM Data Portal (http://gisweb.ciat.cgiar.org/GCMPage/) with a 30 arcsec resolution. As a threshold independent method for model validation, we used the area under the receiver operating characteristic curve (AUC). LGM reconstructions were developed independently with respect to the present-day distribution, but models for the future were projected using the present climate layers with the same resolution (30 arcsec) and using the "Projection" option in Maxent. Finally, a jackknife test was performed to measure the relative importance of climatic variables on the occurrence prediction for every distribution model.

Results

Genetic diversity

Mexican populations were genetically distinct from U.S. and Canadian populations, but more closely related to the Rocky Mountain variety than the coastal variety (Fig. 3.2; Gugger *et al*. 2010; Chapter 2).

We observed four *rps7-trnL* and nine *rps15-psaC* haplotypes that combined for 12 cpDNA haplotypes based only on cpDNA sequences (GenBank accessions in Table 3.S2). These formed two major clades, one found primarily in northern Mexico (C21-C24) and the other found throughout the rest of Mexico (C25-C32; Fig. 3.2). None of these haplotypes were observed in the U.S. or Canada but one (C20) from southern Arizona and New Mexico fell into the clade from northwestern Mexico. In central Mexico, two geographically structured subclades were also observed (C26/C28 and C30/C31).

The mtDNA data also support the distinction of northwestern populations (Fig. 3.2). We observed three *V7* and two *nad7i1* haplotypes that combined for four mtDNA haplotypes (Table 3.S2). The most common (M4) was also common in the southwestern U.S. and predominates in Durango and northeast Mexico. The three other combined mtDNA haplotypes were private to populations in Chihuahua (M8, M9) and the northernmost population in Durango (M10). No individuals from central Mexico could be PCR-amplified for either mtDNA marker nor could many individuals from the rest of Mexico. Chloroplast DNA was easily amplified from those same individuals, suggesting the problem was not due to poor quality DNA extract. We varied PCR conditions and used alternative primer pairs without improved success. Therefore, we believe that the mitochondrial genome in those individuals may have undergone major rearrangements or insertion-deletion events, rather than point mutations at the primer site. Such rearrangements are thought to be common in the plant mitochondrial genome, perhaps even more common than point mutations (Palmer & Herbon 1988; Birky 2001).

Fragment lengths for each cpSSR marker were 99-100 and 102-105 bp for *Pt*26081, 90-93 bp for *Pt*63718, and 148-152 bp for *Pt*71936. These combined for 21 cpSSR haplotypes (Table 3.S3). Some cpSSR haplotypes were common to all three geographic regions (S5, S11, S12, S15), some distinguished regions (S10, S16), and many were private to populations (especially in Durango; Fig. 3.2). cpSSR variation in Mexico was highly divergent from that in the coastal variety of Douglas-fir (*P. menziesii* var. *menziesii*) in British Columbia, with each marker containing approximately ten fewer repeats (20 bp) in Mexico than British Columbia (Table 3.S3; Viard *et al*. 2001).

Overall, genetic diversity was high for cpDNA and cpSSR (Table 3.1) and moderately high for mtDNA (Table 3.2). Genetic diversity was positively correlated with latitude for all cpDNA and cpSSR diversity measures investigated, but those correlations were not statistically significant (Fig. 3.3). Trends in diversity were not measured for mtDNA due to small sample size.

Changes in population size

For sequence data, there was no evidence of population expansion based on F_S and R_2 and no evidence of a population bottleneck based on H versus H_{eq} for any population or region (Tables 3.1 and 3.2). For cpSSR data, no populations showed evidence of a bottleneck under SMM, and no populations showed significant evidence of expansion based on F_S. However, both populations in Chihuahua had marginally significant F_S, and Region I had a significant F_S, suggesting expansion (F_S = -7.57, P = 0.002). Moreover, H was significantly *less* than H_{eq} for Region I (P = 0.005), counter to expectations under a bottleneck and consistent with population expansion (Cornuet & Luikart 1996). Overall F_S values for cpSSR also suggested population expansion across Mexico. Finally, a Bayesian skyline plot of N_e through time constructed using all cpDNA sequence data for Mexico shows an increase in population size starting in the late Pleistocene (~100 ka; Fig. 3.4).

Population structure

G_{ST} values for mtDNA, cpDNA, and cpSSR were 0.77 ± 0.12, 0.56 ± 0.10, and 0.16 ± 0.03, respectively.

SAMOVA of mtDNA defined four groups (F_{CT} = 0.84): (i) El Largo, (ii) Chureachi, (iii) Guanaceví, and (iv) the rest of Mexico (Fig. 3.5). The remaining variation resided within populations (17.6%) rather than among populations within groups (-2.0%; Table 3.S4). SAMOVA of cpDNA also defined four groups (F_{CT} = 0.56): (i) El Largo, Chureachi, Cerro Potosí; (ii) Estanzuela; (iii) Cerro Pingüica; (iv) the rest. Most of the remaining variation was within populations (37.4%) with some among populations within groups (6.5%). All variance components were significant ($P < 0.025$).

AMOVA of populations within regions (I, II, III) for cpDNA and cpSSR revealed significant within-population and among-population within-region variance; however, variation among regions was not significant (F_{CT} = 0.09, P = 0.16 and F_{CT} = 0.01, P = 0.29, respectively; Table 3.S4).

UPGMA of cpSSR data based on $(\delta\mu)^2$ defined two major groups: northwestern Mexico plus Villa Real, Tlaxcala and eastern Mexico plus Las Flores, Durango (Fig. 3.5).

Divergence from Rocky Mountain variety

The posterior distribution for divergence time (t) among Mexican and Rocky Mountain populations did not converge under the full model, though a clear peak corresponding to 1.01 Ma was observed (Table 3.3). The zero migration model could not be rejected, so the analysis was repeated without migration. With $m_1 = m_2 = 0$, divergence time converged at 958 ka with the 90% highest posterior density interval ranging from 1.63 Ma to 464 ka, which falls entirely within the Pleistocene. When run separately, mtDNA and cpDNA gave similar divergences of 878 ka and 866 ka, respectively.

Ecological niche modeling

The AUC values for the training and test data in the LGM based on MIROC, LGM based on CCSM, present, and future models were 0.973/0.986, 0.972/0.986, 0.981/0.985, and 0.957/0.979, respectively. These values indicate a good performance for the four models.

The distribution models suggest a reduction in suitable distribution area in the present-day distribution model of Mexican Douglas-fir in comparison with the models under both LGM climate scenarios (Fig. 3.6). Some areas in the western portion of the Trans-Mexican Volcanic Belt, the Central Plateau, and the Sierra Madre Oriental appear as suitable for the species during the LGM, particularly under the CCSM model. Also it is probable that larger areas in the Sierra Madre Occidental were occupied by Douglas-fir during the LGM. However, the MIROC model suggests that the northernmost portion of the Sierra Madre Occidental was less suitable for the species during the LGM conditions that under present-day conditions. The future projections under the HadCM3 climate change model suggest a severe reduction in the area suitable for Mexican Douglas-fir by 2050 and the almost complete loss of the species by 2080.

The jackknife analysis indicated that the variables with the highest relative contributions to the present-day and to LGM models were mean annual temperature, minimum temperature of the coldest month, mean temperatures of the warmest, coldest and driest quarters, annual precipitation, and precipitation of the driest quarter. In contrast, for the future projections the annual precipitation had the highest effect in explaining the distribution. This suggests that the predicted reduction in suitable area for the species will be due mainly to a reduction in precipitation rather than by an increase in temperature.

Discussion

Colonization of Mexico and divergence from Rocky Mountain variety

Mexican Douglas-fir populations probably diverged from Rocky Mountain populations in the mid-Pleistocene (958 ka; Table 3.3), though uncertainties in these estimates (90% HPD: 1.6 Ma – 464 ka), in mutation rate estimates (Gugger *et al*. 2010; Chapter 2), and due to missing mtDNA data suggest that other Pleistocene divergence times are possible. Mexican populations did not share any cpDNA haplotypes with U.S. and Canadian populations, and three mtDNA haplotypes were confined to Mexico (Fig. 3.2). For both mtDNA and cpDNA, Mexican haplotypes were more similar to Rocky Mountain haplotypes than coastal ones. In addition, the most common mtDNA haplotype in Mexico (M4) is also common in the central Rockies. The complete divergence of cpDNA haplotypes and divergence of some mtDNA haplotypes suggests that the shared mtDNA haplotype, M4, is ancestral and remains in common due to low mutation rates and incomplete lineage sorting. The different modes of inheritance and dispersal of mtDNA (maternal, seed-dispersed) and cpDNA (paternal, pollen- then seed-dispersed) could not explain this pattern because a fertile seed bearing a particular mtDNA haplotype could not disperse without also bringing its pollen donor's cpDNA haplotype. Moreover, divergence models without gene flow for each locus or both loci combined could not be rejected (Table 3.3). Mexican populations also differed from the Rocky Mountain populations in that mtDNA could not be amplified in nearly 60% of the samples (especially from central Mexico), presumably due to major insertion-deletions or rearrangements at that locus.

The Pleistocene divergence time, near-complete differentiation between Mexico and the Rockies, and Pleistocene fossil evidence from southern Arizona (Gray 1961; Martin 1963) support the hypothesis of Pleistocene colonization of Mexico and subsequent isolation from the Rocky Mountain variety. Continuous connectivity with populations in the Rockies from the Miocene until Pleistocene divergence is highly unlikely, given dynamic climate and general lack of fossils in the region. This conclusion

is consistent with early interpretations of the fossil record (Deevey 1949; Dressler 1954); however, it contrasts with more recent interpretations (Palacios-Chavez & Rzedowski 1993) that *Pseudotsuga* arrived in southern Mexico in the early to middle Miocene (~20 Ma) when many temperate taxa are thought to have first colonized Mexico (Graham 1999). The Miocene colonization hypothesis was based on the limited evidence of several putative *Pseudotsuga* pollen grains found in sediment in Chiapas (Palacios-Chavez & Rzedowski 1993). However, *Pseudotsuga* and *Larix* pollen cannot be distinguished under optical microscopes even in modern pollen (Tsukada 1982; Barnosky 1985a). Either *Larix*, *Pseudotsuga*, or both could have been plausible associates of the other temperate taxa found in the fossil bed, which included *Abies*, *Picea*, *Pinus*, *Alnus*, *Populus*, *Quercus*, and *Salix* (Palacios-Chavez & Rzedowski 1993). Given that *Larix* is not found in Mexico currently and our data suggest a Pleistocene arrival for *P. menziesii* in Mexico, that Miocene *Larix* or *Pseudotsuga* species is most likely extinct. However, we were not able to sample a very small, isolated, morphologically divergent population (Debreczy & Rácz 1995) in the Sierra Madre del Sur of Oaxaca, which could be a relict population.

The colonization of Mexico could have occurred in multiple waves or there could have been one colonization with varying durations of isolation for central compared to northern populations. Within Mexico, populations are drawn from two cpDNA clades: one is exclusively Mexican and mostly to the south and east (C25-C32), and the other is primarily northern (C21-C24) with one related chlorotype found only in Arizona/New Mexico (C20; Gugger *et al.* 2010; Chapter 2). The southern/eastern clade, has several derived subclades that are restricted to populations in central Mexico (C26 and C30-C31). To some extent, the mtDNA parallel this pattern: Chihuahuan populations are different, M4 reaches farther south and east, and the southernmost populations failed to amplify. Modest divergence (Fig. 3.2) and lower diversity (Fig. 3.3) in central populations compared to northern populations is consistent with the long-term isolation of small populations in central Mexico. Moreover, central populations are more distinct morphologically from the Rocky Mountain variety than northern populations (Reyes-Hernández *et al.* 2006). Collectively, the evidence suggests longer isolation of central

compared to northern populations, either due to an older colonization or an earlier loss of connectivity to the Rocky Mountain populations.

If Mexico was colonized in the mid-Pleistocene, it was likely associated with southward retreat from glaciation(s) at high latitudes and the expansion of more favorable cool, moist conditions in montane Mexico. Our divergence estimate coincides with the 'mid-Pleistocene revolution' when glacial cycles transitioned from 41,000 yr to deep 100,000 yr periods (~900 ka; Head & Gibbard 2005). The divergence of an isolated population of MacGillivray's warbler (*Oporornis tolmiei*) in north-eastern Mexico from its primarily north-west North American range also occurred about this time (Milá *et al.* 2000), supporting the contention that this transition could have promoted the colonization of cool temperate species into Mexico. Alternatively, colonization could have been associated with some of the later deep pre-Illinoisan glaciations that fall within the HPD intervals of our estimates (*e.g.* ~650 ka; Richmond & Fullerton 1986b; Lisiecki & Raymo 2005).

Late Pleistocene history

The post-colonization Pleistocene history of Douglas-fir in Mexico was likely complicated, consisting of long-term population isolation and repeated phases of expansion and contraction. Population structure identified by SAMOVA and UPGMA differed for each molecular marker (Fig. 3.5) and was not consistent with geographic regions defined in other studies (Table 3.S4). In each case, some groupings spanned multiple regions that are now separated by up to 500 km of arid lands uninhabited by Douglas-fir, whereas others were restricted to small areas within a region.

One consistent pattern was the subdivision of populations in the Sierra Madre Occidental. Populations in Chihuahua and the bordering area of Durango were fixed or nearly-fixed for private mtDNA haplotypes and thus mtDNA SAMOVA identified them as separate groups from one another and the remaining Durangan populations. The cpDNA SAMOVA identified Chihuahuan and Durangan populations as two separate groups in the region, but the two common cpDNA haplotypes in the region formed

opposing gradients with C21 most common in Chihuahua and C25 most common in Durango. Similarly, cpSSR UPGMA distinguished the southernmost population in Durango from the remaining Sierra Madre Occidental populations. Common cpSSR haplotypes (S11, S12, S15, S16) are shared in Durango and Chihuahua, but Durangan populations have many private haplotypes. Finally, the Bayesian skyline analysis of cpDNA supports modest population expansion during the Wisconsinan glaciation (Fig. 3.4), and F_S for cpSSR suggests expansion only in the Sierra Madre Occidental (Table 3.1). The strong groupings found in maternally inherited mtDNA, shared haplotype compositions observed in paternally inherited cpDNA and cpSSR, and signature of demographic expansion in cpSSR suggest that Douglas-fir pollen spread genes across populations isolated from seed exchange. Some of the deep east-west canyons in the region (*e.g.* Río Fuerte and Río Culiacán basins) may function as barriers to seed dispersal.

A similar, but more striking, pattern of distinct mtDNA haplotypes and shared cpDNA haplotypes in the Sierra Madre Occidental has also been described in Chihuahua spruce (*Picea chihuahuana*; Jaramillo-Correa *et al.* 2006). This pattern was attributed to Wisconsinan population expansion permitting north-south pollen exchange and Holocene contraction subsequently isolating populations. Late Pleistocene expansion along opposing north-south routes was also observed in Chihuahua pine (*Pinus leiophylla*; Rodríguez-Banderas *et al.* 2009) and southwestern white pine (*Pinus strobiformis*; Moreno-Letelier & Piñero 2009) based on cpSSRs, and population expansion was observed in the region in Mexican pine beetles (*Dendroctonus mexicanus*) based on mtDNA (Anducho-Reyes *et al.* 2008). Although the precise delineations among populations differ, these data suggest that conifer forests in the Sierra Madre Occidental expanded during the Wisconsinan glaciation permitting north-south gene flow, especially by pollen. Population expansion is counter to expectations based on ecological niche models (Fig. 3.6) and the expected sensitivity of Douglas-fir to historical trends of warming and drying throughout the region (González-Elizonado *et al.* 2005).

Other populations throughout Mexico are better described by recent histories of stability and isolation. In the Sierra Madre Oriental and central Mexico, no significant

signature of population expansion or contraction was observed (Table 3.1). Despite sharing some haplotypes with other regions (Fig. 3.2), each retains a unique set of private haplotypes (*e.g.* C27, C29, C32, S10 in Sierra Madre Oriental). In particular, central populations are delineated in cpDNA SAMOVA. The isolated Cerro Pingüica population in Querétaro is fixed for a derived cpDNA subclade (C30, C31) and has thus been evolving independently for a long time. Farther south, the Estanzuela, Hidalgo population is fixed for C26, which it shares with the Villa Real, Tlaxcala population. These populations may have last been connected during a severe glaciation recorded in the region about 195 ka (Vázquez-Selem & Heine 2004). Mountain glaciers reached over 1000 m lower in than present and climate was cooler and moister.

Ecological niche models suggest that regions were not connected during the LGM (Fig. 3.6). Possible connections among regions are apparent under one of the two models at the LGM, but these models clearly overestimate distributions in the present. Nonetheless, these models suggest that populations within regions may have been more continuous and that corridors among regions along the Sierra Madre Occidental, Sierra Madre Oriental and Trans-Mexican Volcanic belt could have opened during some Pleistocene glaciations and closed during interglacials. This might explain the complex pattern of haplotype sharing throughout Mexico.

Taxonomic implications

Mexican Douglas-fir represents at least one evolutionarily significant unit (Ryder 1986; Crandall *et al.* 2000) and might be treated best as a third variety (Earle 2009). There is strong evidence for molecular genetic differentiation (Fig. 3.2; Gugger *et al*. 2010; Chapter 2) and long-term isolation without gene flow from the Rocky Mountain variety (Table 3.3). This divergence (958 ka) is of similar magnitude to the divergence time estimated between coastal and Rocky Mountain varieties: roughly half as long when considering both loci (2.11 Ma) or twice as long when considering only cpDNA (491 ka; Gugger *et al*. 2010; Chapter 2). Leaf and cone morphology of central Mexican populations also differs from the Rocky Mountain variety (Reyes-Hernández *et al.* 2005,

2006). Finally, Mexican Douglas-fir exhibits phenological and ecological differentiation with earlier budburst and later budset associated with the generally warmer climate (Acevedo-Rodríguez *et al.* 2006).

Some evidence suggests a clinal gradient from the Rocky Mountain populations to central Mexican populations. Northern Mexican populations shared similar leaf and cone morphology with Rocky Mountain populations, whereas central Mexican populations were quite distinct (Reyes-Hernández *et al.* 2005, 2006). Mitochondrial DNA variation (M4) is shared among populations in the Rockies and those in northern Mexico, whereas central populations (and some northern individuals) may be very different based on failure to PCR-amplify. Finally, a cpDNA haplotype (C20) in an otherwise Mexican clade is found in Arizona and New Mexico, and central Mexican populations have some cpDNA haplotypes from more derived, purely Mexican clades. Such a gradient could be due to different degrees and durations of isolation and recontact for populations at different latitudes and in different regions. Therefore, a varietal delineation might equally separate only central Mexican populations, for example. Nonetheless, we demonstrate that Mexican populations have been long-isolated from Rocky Mountain populations (Table 3.3) and that molecular variation within Mexico is largely shared among regions (Figs. 3.2 and 3.5; Table 3.S4). These results provide support for the delineation of a separate variety that includes all Mexican populations.

The lack of deep genetic divergences within Mexico and shared variation across regions argue against the hypothesis of multiple species within Mexico (Flous 1934a,b; Martínez 1949). Morphological and phenological studies also fail to support the multiple species hypothesis (Vargas-Hernández *et al.* 2004; Reyes-Hernández *et al.* 2005, 2006; Acevedo-Rodríguez *et al.* 2006).

Conservation implications

Conservation efforts for Mexican Douglas-fir should consider patterns of genetic diversity and population history. Mexican Douglas-fir has high haplotype diversity for all markers investigated here (Tables 3.1 and 3.2) and populations are strongly

differentiated at mtDNA (G_{ST} = 0.77) and cpDNA (G_{ST} = 0.55) sequence markers. About 40 % of all mtDNA and cpDNA haplotypes observed across the range of Douglas-fir are found in Mexico and all but one are endemic (Gugger *et al*. 2010; Chapter 2). Haplotype diversity and richness varied widely from one population to the next (Tables 3.1 and 3.2), and was correlated (though not significantly) with latitudinal gradients in morphological heterogeneity, population size, and inbreeding rates (Fig. 3.3).

Previous studies have demonstrated reduced fecundity associated with inbreeding depression in many Mexican populations, particularly in the smallest, most isolated ones in central Mexico (Mápula-Larreta *et al.* 2007; Velasco-García *et al.* 2007; Cruz-Nicolás *et al.* 2008). One proposal to restore genetic variation by crossing central Mexican populations (Cruz-Nicolás *et al.* 2008) could be revised to consider the patterns of variation shown here. For example, the Cerro Pingüica, Querétaro population is fixed for unique cpDNA haplotypes that suggest it has evolved independently from other central Mexican populations; therefore, it should not be crossed with Hidalgo and Tlaxcala populations that have different compositions. However, Hidalgo and Tlaxcala populations likely share a more recent history and might be reasonably crossed. Additional investigation into the genetic structure of more populations in the region would reveal the precise locations of genetic breaks among populations and could be used to choose the best candidates for breeding programs to bolster fertility rates. In other regions, breeding programs should also consider the genetic structure identified here (Fig. 3.2). Considering the long separation of Mexican and Rocky Mountain populations, we further recommend that Mexican populations not be used to populate the U.S. (*cf*. Vargas-Hernández *et al.* 2004), except under dire circumstances.

Models project range contractions for pine and oak species across Mexico by up to 60 % (Gómez-Mendoza & Arriaga 2007) under projected warmer and drier conditions (Liverman & O'Brien 1991). For Douglas-fir, ecological niche models also suggest population decline under projected future climates (Fig. 3.6), consistent with the observed sensitivity of Mexican Douglas-fir growth rates to precipitation and maximum temperature (González-Elizonado *et al.* 2005). With temperatures already increasing

(Pavia *et al.* 2009), careful conservation measures will be necessary to protect the unique biodiversity of Douglas-fir and other species in Mexico.

Although Mexican Douglas-fir once must have formed the leading edge as it colonized Mexico, those populations have since been long-isolated and sit at the subtropical rear edge of a wide-ranging temperate species. Hampe & Petit (2005) suggest that such rear edge populations far from the dynamic northern range limit are an important genetic resource because they contain much of neutral and adaptive diversity within species, form reserves under changing climates, and are a principal site of speciation. Because rear edge populations are often strongly differentiated, they propose conserving the maximum number of local populations, rather than focusing on high diversity regions. Conservation efforts for Douglas-fir might target such goals, while still protecting high diversity regions (Frankham *et al.* 2002) such as in the Sierra Madre Occidental.

Tables

Table 3.1. Diversity measures for cpDNA, cpSSR, and combined cpDNA and cpSSR for each population, region, and the overall dataset

			cpDNA						cpSSR					Combined		
	Population	***n****	***h***	***H* (SD)**	H_{eq} **(IAM)†**	**$\pi \times 10^3$ (SD)‡**	**R_2†**	**F_S†**	***h***	***H* (SD)**	H_{eq} **(SMM)†**	**$\bar{D}^2_{SH}$‡**	**F_S†**	***h***	**h_r‡§**	***H* (SD)‡**
1	El Largo, Chi.	13	2	0.28 (0.14)	0.32	0.186 (0.239)	0.14	0.24	8	0.86 (0.09)	0.91	4.27	-2.68	8	4.6	0.86 (0.09)
2	Chureachi, Chi.	10	3	0.60 (0.13)	0.56	0.749 (0.599)	0.18	0.72	5	0.80 (0.10)	0.82	0.59	-1.99	6	5.2	0.84 (0.10)
3	Guaneceví, Dur.	12	2	0.17 (0.13)	0.33	0.220 (0.265)	0.27	0.43	4	0.64 (0.13)	0.71	2.10	0.24	4	3.3	0.64 (0.13)
4	Guaneceví 2, Dur.	16	3	0.49 (0.12)	0.48	0.688 (0.541)	0.17	1.09	7	0.85 (0.06)	0.86	1.40	-2.26	8	5.2	0.86 (0.06)
5	Altares, Dur.	12	2	0.30 (0.15)	0.32	0.401 (0.384)	0.15	1.38	5	0.73 (0.11)	0.79	1.35	-0.72	6	4.6	0.80 (0.10)
6	Las Flores, Dur.	11	1	0	-	0	0	0	3	0.69 (0.09)	0.59	0.38	0.24	3	2.9	0.69 (0.09)
	I: Sierra Madre Occidental	**74**	**5**	**0.51 (0.04)**	**0.53**	**0.701 (0.521)**	**0.13**	**0.41**	**17**	**0.86 (0.03)**	***0.92***	**2.33**	***-7.57***	**23**	**11.3**	**0.91 (0.02)**
7	Jamé, Coa.	10	4	0.73 (0.12)	0.70	0.734 (0.590)	0.17	0.68	6	0.84 (0.10)	0.87	3.08	-1.66	8	6.8	0.96 (0.06)
8	Cerro Potosí, N.L.	11	4	0.49 (0.18)	0.68	0.817 (0.631)	0.18	1.02	6	0.89 (0.08)	0.88	3.23	-1.23	7	6.0	0.91 (0.08)
	II: Sierra Madre Oriental	**21**	**4**	**0.70 (0.07)**	**0.57**	**0.932 (0.663)**	**0.17**	**2.17**	**9**	**0.88 (0.04)**	**0.89**	**3.25**	**-2.54**	**14**	**14**	**0.96 (0.03)**
9	Cerro Pingüica, Que.	10	2	0.20 (0.15)	0.35	0.132 (0.201)	0.30	-0.34	3	0.75 (0.10)	0.64	0.46	0.14	3	3.0	0.75 (0.10)
10	Estanzuela, Hid.	10	1	0	-	0	0	0	3	0.64 (0.10)	0.61	3.03	1.67	3	2.9	0.64 (0.13)
11	Villa Real, Tla.	14	3	0.56 (0.13)	0.51	0.327 (0.333)	0.24	1.14	3	0.67 (0.08)	0.56	0.45	0.64	6	4.5	0.79 (0.09)
	III: central Mexico	**34**	**5**	**0.74 (0.04)**	**0.61**	**0.661 (0.508)**	**0.17**	**0.5**	**7**	**0.84 (0.03)**	**0.81**	**1.75**	**-0.77**	**11**	**9.6**	**0.91 (0.02)**
	Overall	**129**	**12**	**0.79 (0.07)**	**0.76**	**0.897 (0.618)**	**0.07**	**-1.28**	**21**	**0.91 (0.02)**	**0.93**	**2.63**	***-9.16***	**44**	**-**	**0.97 (0.02)**

* Total sample size. No data could be obtained for one cpSSR sample in Population 8, two cpSSR samples in Population 9, and one cpDNA sample in Population 10

† Italicized values were significant under coalescent simulations

‡ Used in regressions of latitude on diversity (Fig. 3.3)

§ Rarefaction to 8 for populations and 20 for regions

Table 3.2. Mitochondrial DNA diversity measures for populations, regions, and overall dataset

	Population	*n*	*h*	*H* (SD)	H_{eq} (IAM)*	$\pi \times 10^3$ (SD)	R_2*	F_S*
1	El Largo, Chihuahua	8	2	0.43 (0.17)	0.38	0.459 (0.523)	0.21	0.54
2	Chureachi, Chihuahua	4	1	0	-	0	0	0
3	Guaneceví, Durango	8	2	0.54 (0.12)	0.38	0.574 (0.601)	0.27	0.87
4	Guaneceví 2, Durango	11	1	0	-	0	0	0
5	Altares, Durango	9	1	0	-	0	0	0
6	Las Flores, Durango	9	1	0	-	0	0	0
	I: Sierra Madre Occidental	**49**	**4**	**0.50 (0.08)**	**0.48**	**0.599 (0.553)**	**0.09**	**-0.58**
7	Jamé, Coahuila	6	1	0	-	0	0	0
	Overall	**55**	**4**	**0.59 (0.18)**	-	**0.539 (0.516)**	**0.08**	**-0.75**

* No values were statistically significant under coalescent simulations

Table 3.3. IMa results showing t, m_1, and m_2 for a run with the full model, followed by results of the likelihood ratio tests, and then t with its 90% highest posterior density interval for a run with $m_1 = m_2 = 0$. In each case, the values for the run with the highest effective sample size (ESS; Hey & Nielson 2007) are shown

	Full model Priors (t, m_1, m_2, q_1, q_2, q_A)	t (ka)	m_1	m_2	**$m_1 = m_2$** -2Λ	P	**$m_1 = m_2 = 0$** -2Λ	P	**Model without migration ($m_1 = m_2 = 0$)** t (ka)	90% HPD low	90% HPD high
Both loci	**5, 2, 3, 10, 10, 10**	**1011***	**0.0000001**	**0.0000001**	**0.3506**	**0.55**	**-0.188 ≈ 0**	**≈ 0.5**	**958**	**464**	**1628**
mtDNA	5, 10, 6, 10, 10, 10	878*	0.0000002	0.0000001	0.702	0.40	-0.0032 ≈ 0	≈ 0.5	878	265	2789
cpDNA	5, 1, 3, 20, 7, 7	806*	0	0.0000001	0.3729	0.54	0.405	0.26	866	378	1765

* Posterior distribution for this parameter never dropped to zero, suggesting unreliable estimates.

Figures

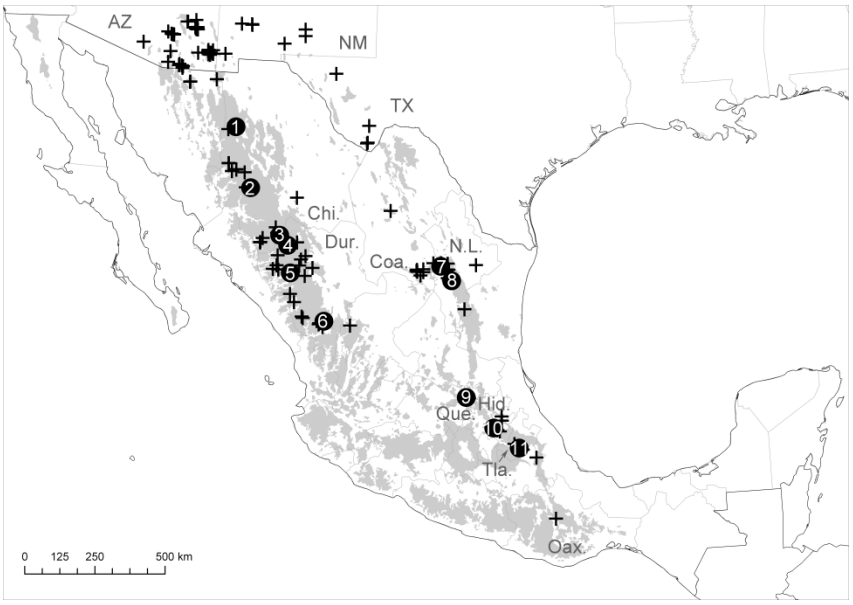

Figure 3.1. Map of Mexico with Madrean pine-oak region shown in light gray (Conservation International Foundation), *Pseudotsuga* range approximated by location herbarium samples (from GBIF) and shown as crosses, and sample sites shown as black points numbered according to Table 3.1. States mentioned in the text are labeled with abbreviations: AZ = Arizona, Chi. = Chihuahua, Coa. = Coahuila, Dur. = Durango, Hid. = Hidalgo, N.L. = Nuevo León, NM = New Mexico, Oax. = Oaxaca, Que. = Querétaro, Tla. = Tlaxcala, TX = Texas.

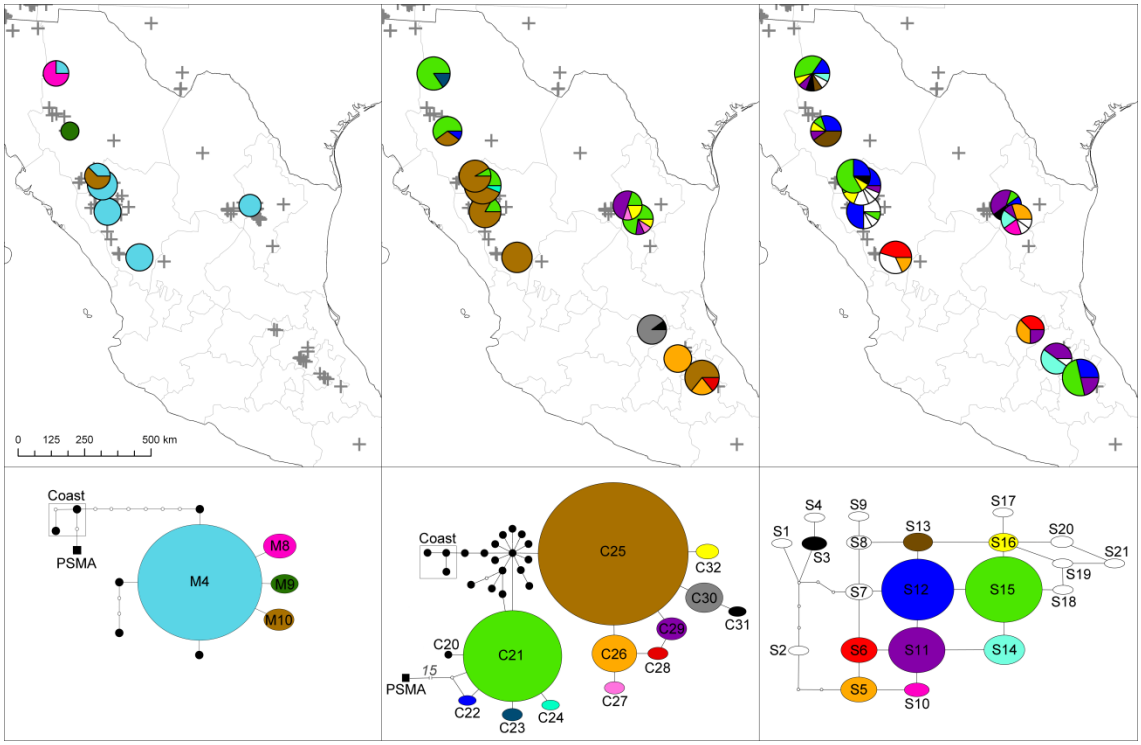

Figure 3.2. Maps with sample sites colored according to mtDNA (left), cpDNA (center), and cpSSR (right) haplotype composition and proportioned according to sampling intensity. Below each, parsimony networks show the relationship among haplotypes observed in Mexico (colored elliptical nodes, proportional to the frequency in the dataset) and haplotypes observed in the U.S. and Canada (small circular black nodes; Gugger *et al*. 2010; Chapter 2). Haplotypes from the coastal variety (*P. menziesii* var. *menziesii*) are marked; all other U.S. or Canadian haplotypes are the Rocky Mountain variety (*P. menziesii* var. *glauca*). Black square nodes indicate the outgroup, *P. macrocarpa* (PSMA), and small, open points indicate inferred, unobserved haplotypes. Internodes are one mutational step unless otherwise noted. In the cpSSR panel, white indicates haplotypes private to one population. Colors do not relate from one map/network to the next or to those in Gugger *et al*. (2010; Chapter 2), except M4. Haplotype numbering is consistent with Gugger *et al*. (2010; Chapter 2).

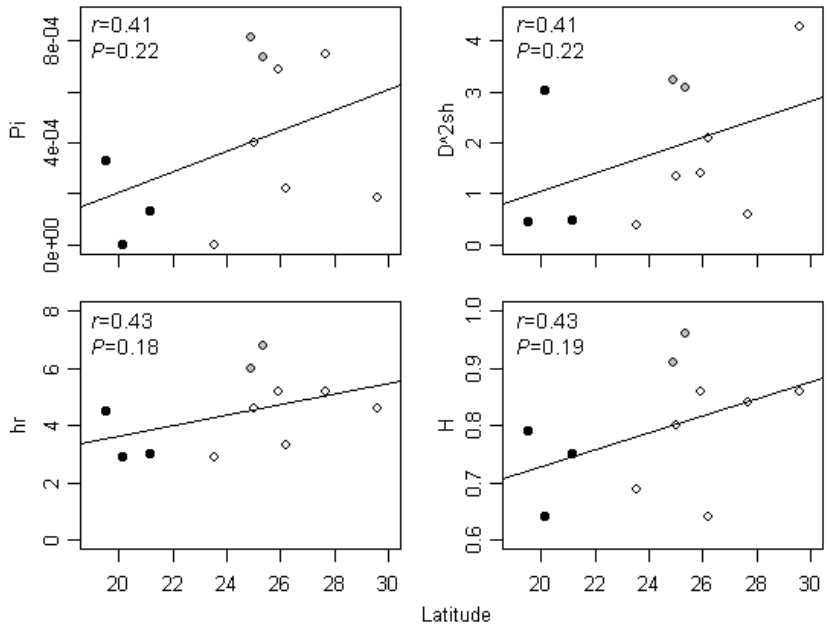

Figure 3.3. Diversity was positively correlated with latitude for cpDNA sequence (π), cpSSR ($\overline{D}^2_{SH}$), and combined diversity (h_r, H), but these relationships were not statistically significant. Open circles = Region I, gray = Region II, black = Region III.

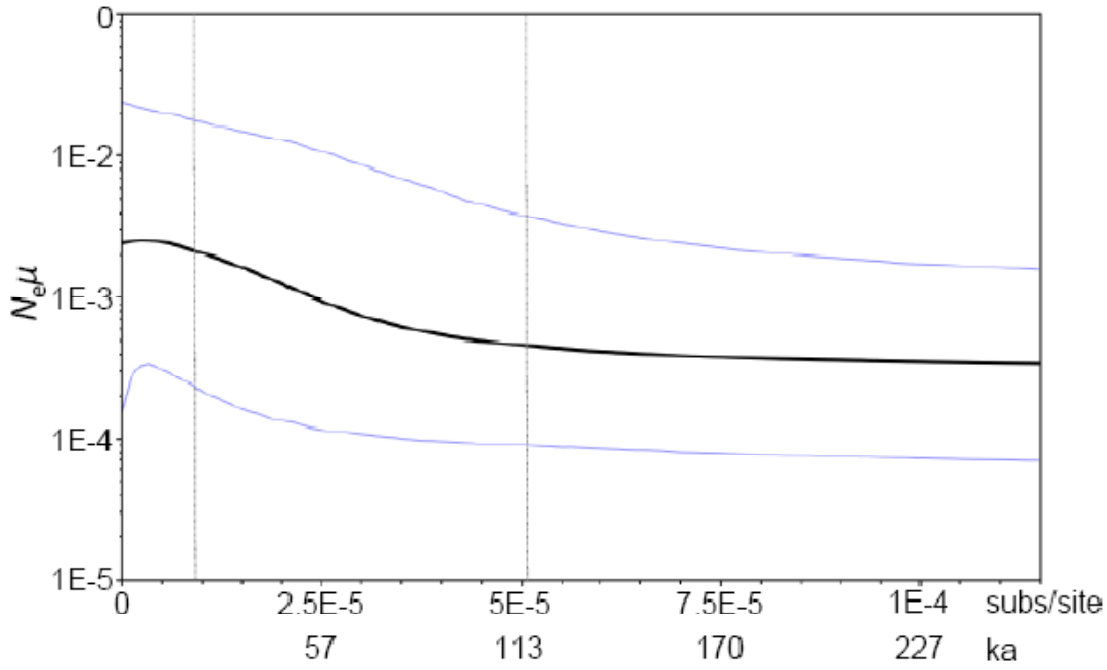

Figure 3.4. Bayesian skyline semi-log plot of median $N_e\mu$ through time based on cpDNA sequence data, where μ is the mutation rate per generation. Confidence intervals are shown as gray lines. The *x*-axis is shown in two scales: substitutions/site and calendar time in thousands of years before present (ka). A mutation rate of 4.41×10^{-10} subs/site/yr was used to convert subs/site to years (Gugger *et al*. 2010; Chapter 2). Vertical lines mark the start of the Wisconsinan glaciation (~115 ka) and Last Glacial Maximum (~21 ka).

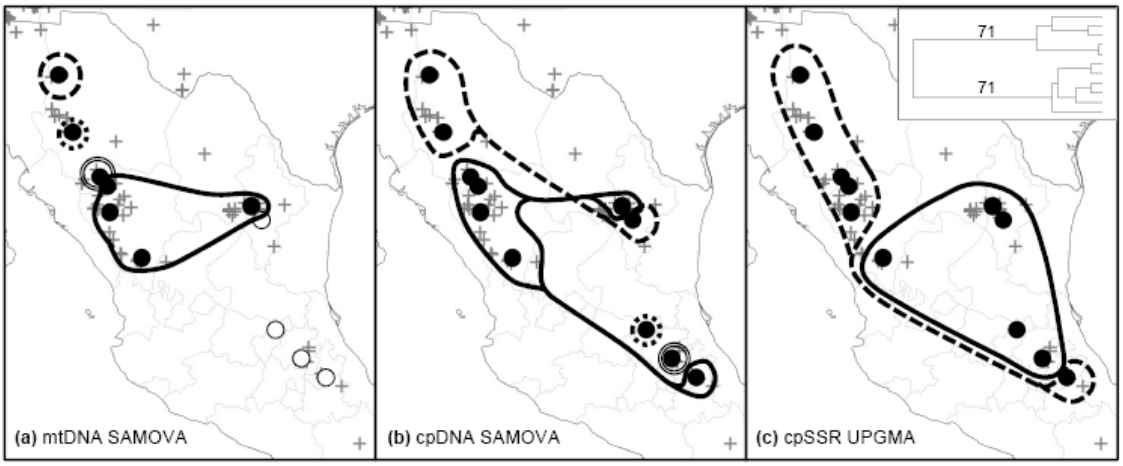

Figure 3.5. Population structure identified in (a) mtDNA SAMOVA (b) cpDNA SAMOVA, and (c) UPGMA clustering analysis of cpSSR $(\delta\mu)^2$ values (dendrogram with approximate bootstrap support for two major groups shown as inset). Each group is delineated with different line styles. For (a), open circles indicate populations in which no samples could be PCR-amplified.

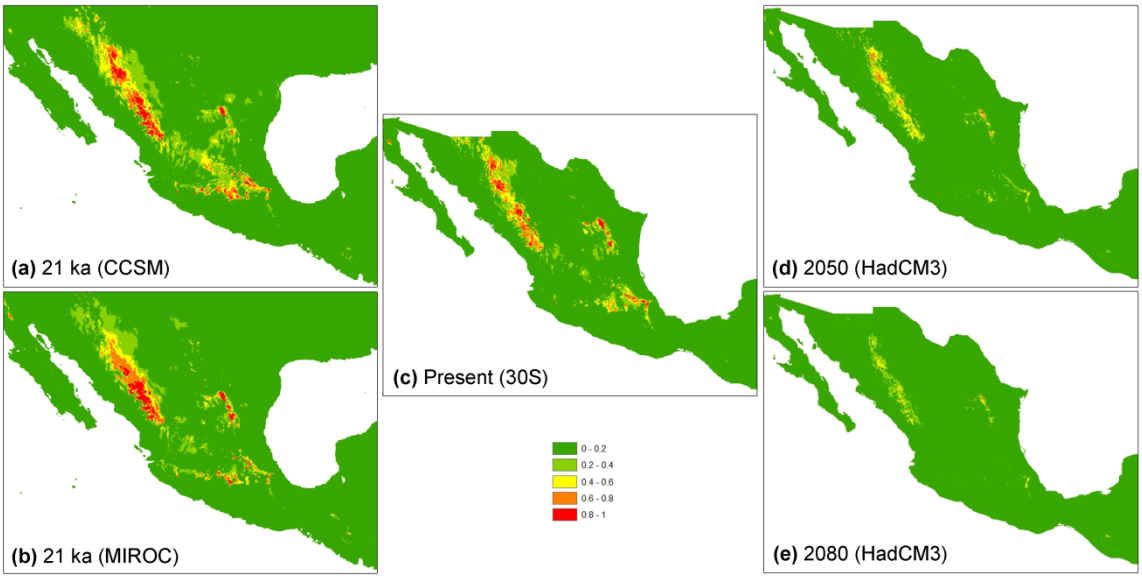

Figure 3.6. Maps showing potential distribution as probability of occurrence (green = 0 - 0.2, red = 0.8 - 1.0) for *Pseudotsuga* in Mexico (a) at the Last Glacial Maximum (21 ka) based on the CCSM model, (b) at the LGM based on the MIROC 3.2 model, (c) the present, and (d) the year 2050 and (e) 2080 based on the HadCM3 model. At the LGM, populations within regions may have been more continuous, but regions likely remained isolated. In the future, further range contraction is predicted.

Supporting information

Table 3.S1. Sampling site information and haplotype frequencies for each marker and each population

(Included as supplementary file Table 3.S1.xlsx)

Table 3.S2. Mitotype and chlorotype definitions in terms of *V7*, *nad7i1*, *rps7-trnL*, and *rps15-psaC* haplotypes reported to GenBank (accessions in parentheses)

Mitotype	*V7* haplotype	*nad7i1* haplotype
M4	H3 (GQ999617)	H2 (GQ999626)
M8	H3 (GQ999617)	H3 (GQ999627)*
M9	H6 (GQ999620)*	H2 (GQ999626)
M10	H7 (GQ999621)*	H2 (GQ999626)

Chlorotype	*rps7-trnL* haplotype	*rps15-psaC* haplotype
C21	H2 (GQ999631)	H5 (GQ999648)
C22	H8 (GQ999637)*	H5 (GQ999648)
C23	H2 (GQ999631)	H20 (GQ999663)*
C24	H2 (GQ999631)	H19 (GQ999662)*
C25	H2 (GQ999631)	H11 (GQ999654)*
C26	H2 (GQ999631)	H14 (GQ999657)*
C27	H7 (GQ999636)	H14 (GQ999657)*
C28	H2 (GQ999631)	H18 (GQ999661)*
C29	H2 (GQ999631)	H15 (GQ999658)*
C30	H2 (GQ999631)	H13 (GQ999656)*
C31	H2 (GQ999631)	H12 (GQ999655)*
C32	H10 (GQ999639)*	H11 (GQ999654)*

* New to this study

Table 3.S3. Definitions of cpSSR haplotypes based on fragment lengths of each cpSSR marker and binary coding used to calculate F_S (Table 3.1) and do AMOVA (Table 3.S4)

cpSSR haplotype	*Pt*26081	*Pt*63718	*Pt*71936	Binary coding
S1	99	91	150	0000000010011
S2	100	91	148	0000010010000
S3	100	90	150	0000010000011
S4	100	90	151	0000010000111
S5	102	92	148	0001110110000
S6	102	92	149	0001110110001
S7	102	92	150	0001110110011
S8	102	92	151	0001110110111
S9	102	92	152	0001110111111
S10	103	92	148	0011110110000
S11	103	92	149	0011110110001
S12	103	92	150	0011110110011
S13	103	92	151	0011110110111
S14	104	92	149	0111110110001
S15	104	92	150	0111110110011
S16	104	92	151	0111110110111
S17	104	92	152	0111110111111
S18	104	93	150	0111111110011
S19	104	93	151	0111111110111
S20	105	92	151	1111110110111
S21	105	93	151	1111111110111

Table 3.S4. SAMOVA tables for groupings that gave the highest F_{CT} for mtDNA and cpDNA sequence data, and AMOVA of populations within regions for cpDNA and cpSSR data

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation index
mtDNA SAMOVA					
Among groups	3	10.225	0.33819*	84.47	F_{CT}=0.84473*
Among populations within groups	3	0	-0.00815**	-2.04	F_{SC}=-0.13109**
Within populations	48	3.375	0.07031**	17.56	F_{ST}=0.82438**
Total	54	13.6	0.40036		
cpDNA SAMOVA					
Among groups	3	43.362	0.54718**	56.08	F_{CT}=0.56082**
Among populations within groups	7	7.938	0.06354**	6.51	F_{SC}=0.14829**
Within populations	117	42.7	0.36496**	37.41	F_{ST}=0.62595**
Total	127	94	0.97568		
cpDNA AMOVA					
Among regions (I, II, III)	2	6.433	0.03323^{NS}	9.19	F_{CT}=0.09190^{NS}
Among populations within regions	8	16.485	0.16221**	44.86	F_{SC}=0.49399**
Within populations	117	19.411	0.16616**	45.95	F_{ST}=0.54049**
Total	127	42.359	0.36161		
cpSSR AMOVA					
Among regions (I, II, III)	2	2.619	0.00550^{NS}	1.21	F_{CT}=0.01215^{NS}
Among populations within regions	8	9.179	0.06659**	14.71	F_{SC}=0.14891**
Within populations	115	43.766	0.38057**	84.07	F_{ST}=0.15925**
Total	125	55.563	0.45266		

* $P < 0.05$

** $P < 0.001$

^{NS} not significant

Conclusion

Phylogeographic tests were able to reject some interpretations of the fossil record. Divergence among varieties was consistent with the Pliocene Cascade orogeny hypothesis, but Pleistocene divergence could not be ruled out. Miocene colonization of Mexico was rejected in favor of Pleistocene colonization from the Rocky Mountains. Finally, some molecular analyses reject the single-refugium hypothesis for Rocky Mountain Douglas-fir, although others do not. Where fossil and molecular data were concordant we were able to associate particular climatic and geologic events with observed patterns of molecular variation.

Fossil and molecular data each offered unique and complementary advantages. For example, dated fossil evidence showed clear evidence for retreat to higher elevations in the southern Rocky Mountains in response to postglacial warming. Molecular data showed that those populations are genetically distinct and did not contribute to the northward range expansion observed in northern parts of the distribution. In addition, molecular data suggested possible glacial refugia in the northern Rockies, where fossil data are lacking. Holocene expansion from those refugia into British Columbia was observed in the fossil record, and a unique pattern of hybridization among varieties by pollen dispersal was revealed through phylogeographic analysis. Finally, molecular data provided some insight into the complex Pleistocene history of Douglas-fir in Mexico, where no paleobotanical evidence was available. Those hypotheses, in turn, can be tested with future fossil investigations. Together, these findings revealed the individualistic response of populations to past environmental change.

These findings further suggest that Douglas-fir populations will continue to respond individualistically to changing climates. Considering patterns of genetic variation should improve conservation programs to preserve the unique biodiversity in Mexico. Furthermore, the rates of northward expansion estimated here should help determine the capacity of Douglas-fir to respond to future warming by range shift.

In sum, the statistical phylogeographic framework in which I combined fossil and molecular data should prove useful in organisms with rich fossil records. However, a limitation in distinguishing among some hypotheses was low molecular variation. For

example, tests of multiple refugia with coalescent simulations were sensitive to effective population size priors, which only could be roughly estimated. Moreover, hypotheses on the number of glacial refugia could not be tested on the coast for lack of genetic variation. Finally, broad confidence intervals on divergence time estimates among varieties indicated only tentative support for the divergence associated with Pliocene orogeny. To improve this framework, future research should focus on multi-locus or genomic approaches to reveal the molecular variation required to test more nuanced hypotheses. Moreover, these approaches may offer insight into the adaptive evolutionary response of trees to climate change, a pressing concern if trees are to survive projected climate change.

References

Aagaard JE, Volmer SS, Sorensen FC, Strauss SH (1995) Mitochondrial DNA products among RAPD profiles are frequent and strongly differentiated between races of Douglas-fir. *Molecular Ecology,* **4**, 441-447.

Acevedo-Rodríguez R, Vargas-Hernández JJ, López-Upton J, Mendoza JV (2006) Effect of geographic origin and nutrition on shoot phenology of Mexican Douglas-Fir (*Pseudotsuga* spp.) seedlings. *Agrociencia,* **40**, 125-137.

Adam DP (1985) Quaternary pollen records from California. In: *Pollen Records of Late-Quaternary North American Sediments* (eds Bryant VM Jr, Holloway RG), pp. 125-140. American Association of Stratigraphic Palynologists Foundation, Dallas, TX.

Adam DP, Vagenas GL (1990) Pollen data for a 3-m.y. core record from Tulelake, Siskiyou County, California. *United States Geological Survey Open-File Report,* **90-65**, 1-307.

Alley NF (1979) Middle Wisconsin stratigraphy and climate reconstruction, southern Vancouver Island, British Columbia. *Quaternary Research,* **11**, 213-237.

Anderson LL, Hu F-S, Nelson DM, Petit RJ, Paige KN (2006) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences,* **103**, 12447-12450.

Anderson RL, Byrne R, Dawson T (2008) Stable isotope evidence for a foggy climate on Santa Cruz Island, California at similar to 16,600 cal. yr. BP. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **262**, 176-181.

Anducho-Reyes MA, Cognato AI, Hayes JL, Zúñiga G (2008) Phylogeography of the bark beetle *Dendroctonus mexicanus* Hopkins (Coleoptera:Curculionidae: Scolytinae). *Molecular Phylogenetics and Evolution,* **49**, 930-940.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics,* **33**, 707-740.

Axelrod DI (1966) The Eocene Copper Basin flora of northeastern Nevada. *University of California Publications in Geological Science,* **59**, 1-83.

Axelrod DI (1983) New Pleistocene conifer records, coastal California. *University of California Publications in Geological Science,* **107**, 1-108.

Axelrod DI (1990) Environment of the Middle Eocene (45 Ma) Thunder Mountain flora, central Idaho. *National Geographic Research,* **6**, 355-361.

Baker, RG (1983) Holocene vegetational history of the western United States. In: *Late-Quaternary Environments of the United States: The Holocene* (ed Wright HE Jr), pp. 109-127. University of Minnesota Press, Minneapolis, MN.

Barnosky CW (1981) A record of late Quaternary vegetation from Davis Lake, southern Puget lowland, Washington. *Quaternary Research,* **16**, 221-239.

Barnosky CW (1985a) Late Quaternary vegetation near Battle Ground Lake, southern Puget Trough, Washington. *Geological Society of America Bulletin,* **96**, 263-271.

Barnosky CW (1985b) Late Quaternary vegetation in the southwestern Columbia Basin, Washington. *Quaternary Research,* **23**, 109-122.

Barnosky CW (1989) Postglacial vegetation and climate in the northwestern Great Plains of Montana. *Quaternary Research,* **31**, 57-73.

Barrett CF, Freudenstein JV (2009) Patterns of morphological and plastid DNA variation in the *Corallorhiza striata* species complex (Orchidaceae). *Systematic Botany,* **34**, 496-504.

Barry RG (1983) Late-Pleistocene climatology. In: *Late-Quaternary Environments of the United States: The Late Pleistocene* (eds Wright HE Jr, Porter SC), pp. 390-407. University of Minnesota Press, Minneapolis, MN.

Bartlein PJ, Whitlock C, Shafer SL (1997) Future climate in the Yellowstone National Park region and its potential impact on vegetation. *Conservation Biology,* **11**, 782-792.

Bartlein PJ, Anderson KH, Anderson PM, *et al*. (1998) Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Reviews,* **17**, 549-585.

Braukmann TWA, Kuzmina M, Stefanović S (2009) Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Current Genetics*, **55**, 323-337.

Beiswenger JM (1991) Late Quaternary vegetational history of Grays Lake, Idaho. *Ecological Monographs,* **61**, 165-182.

Bennett KD, Provan J (2008) What do we mean by 'refugia'? *Quaternary Science Reviews*, **27**, 2449-2455.

Betancourt JL, Van Devender TR, Martin PS, Van Devender TR, Martin PS (1990) *Packrat Middens: The Last 40,000 Years of Biotic Change.* University of Arizona Press, Tucson, AZ.

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology,* **19**, 12-20.

Birky WC Jr (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics,* **35**, 125-148.

Bright RC (1966) Pollen and seed stratigraphy of Swan Lake, southeastern Idaho: Its relation to regional vegetational history and to Lake Bonneville history. *Tebiwa,* **9**, 1-47.

Briles CE, Whitlock C, Bartlein PJ (2005) Postglacial vegetation, fire, and climate history of the Siskiyou Mountains, Oregon, USA. *Quaternary Research,* **64**, 44-56.

Brown RB (1985) A summary of late-Quaternary pollen records from Mexico west of the Isthmus of Tehuantepec. In: *Pollen Records of Late-Quaternary North American Sediments* (eds Bryant VM Jr, Holloway RG), pp. 71-93. American Association of Stratigraphic Palynologists Foundation, Dallas, TX.

Brubaker LB, Anderson PM, Edwards ME, and Lozhkin V (2005) Beringia as a glacial refugium for boreal trees and shrub: new perspectives from mapped pollen data. *Journal of Biogeography,* **32**, 833-848.

Brunelle A, Whitlock C (2003) Postglacial fire, vegetation, and climate history in the Clearwater Range, Northern Idaho, USA. *Quaternary Research,* **60**, 307-318.

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, Sullivan J, Soltis DE, Soltis PS (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 Publishing, Williston, VT.

Brunsfeld SJ, Miller TR, Carstens BC (2007) Insights into the biogeography of the Pacific Northwest of North America: evidence from the phylogeography of *Salix melanopsis*. *Systematic Botany,* **32**, 129-139.

Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology,* **12**, 1487-1495.

Burkart MR (1976) *Pollen biostratigraphy and late Quaternary vegetation history of the Bighorn Mountains, Wyoming.* Ph.D. Dissertation, University of Iowa.

Carstens BC, Knowles LL (2007) Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology,* **16**, 619-627.

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution,* **61**, 1439-1454.

Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004) Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Systematic Biology,* **53**, 781-792.

Carstens BC, Brunsfeld SJ, Dembroski JR, Good JM, 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.

Chang CS, Kim H, Park TY, Maunder M (2004) Low levels of genetic variation among southern peripheral populations of the threatened herb, *Leontice microrhyncha* (Berberidaceae) in Korea. *Biological Conservation,* **119**, 387-396.

Chaw S-M, Zharkikh A, Sung H-M, Lau T-C, Li W-H (1997) Molecular phylogeny of extant Gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. *Molecular Biology and Evolution,* **14**, 56-68.

Chaw S-M, Parkinson CL, Cheng Y, Vincent TM, Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences,* **97**, 4086-4091.

Chernoff H (1954) On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics,* **25**, 573-578.

Chiu W-L, Sears BB (1985) Recombination between chloroplast DNAs does not occur in sexual crosses of *Oenothera*. *Molecular and General Genetics,* **198**, 525-528.

Ciliberti P, de Jong H, Schouten MA, Verweij PA, Heitmans WRB (2009) Ecological niche models for four native cockroach species (Blattaria: Blattellidae: Ectobiinae) in the Netherlands. *Contributions to Zoology,* **78**, 149-156.

Clark JS, Fastie C, Hurtt G, *et al*. (1998) Reid's paradox of rapid plant migration. *Bioscience,* **48**, 13-24.

Clark JS, Lewis M, Horvath L (2001) Invasion by extremes: variation in dispersal and reproduction retards population spread. *American Naturalist,* **157**, 537-554.

Clement M, Posada D, Crandall KA (2000) TCS: A computer program to estimate gene genealogies. *Molecular Ecology,* **9**, 1657-1660.

Collins WD, Bitz CM, Blackmon ML, *et al*. (2006) The Community Climate System Model version 3 (CCSM3). *Journal of Climate,* **19**, 2122-2143.
Conroy CJ, Cook JA (2000) Phylogeography of a post-glacial colonizer: *Mocrotus longicaudus* (Rodentia: Muridae). *Molecular Ecology,* **9**, 165-175.
Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics,* **144**, 2001-2014.
Cortés-Palomec AC, McCauley RA, Oyama K (2008) Isolation, characterization and cross-amplification of polymorphic microsatellite loci in *Laelia speciosa* (Orchidaceae). *Molecular Ecology Resources,* **8**, 135-138.
Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution,* **15**, 290-295.
Critchfield WB (1984) Impact of the Pleistocene on the genetic structure of North American conifers. In: *Proceedings of the Eighth North American Forest Biology Workshop* (ed Lanner RM), pp. 70-118. Logan, UT.
Cruz-Nicolás J, Vargas-Hernández JJ, Ramírez-Vallejo P, López-Upton J (2008) Mating pattern in natural populations of *Pseudotsuga menziesii* (Mirb.) Franco in Mexico. *Agrociencia,* **42**, 367-378.
Davies S, White A, Lowe A (2004) An investigation into effects of long-distance seed dispersal on organelle population genetic structure and colonization rate: a model synthesis. *Heredity,* **93**, 566-576.
Davis MB (1976) Pleistocene biogeography of temperate deciduous forests. *Geoscience and Man,* **13**, 13-26.
Davis MB (1981) Quaternary history and the stability of forest communities. In: *Forest Succession: Concepts and Application* (eds West DC, Shugart HH, Botkin DB), pp. 132-153. Springer-Verlag, New York, NY.
Davis MB (1983) Holocene vegetational history of the eastern United States. In: *Late-Quaternary Environments of the United States: The Holocene* (ed Wright HE Jr), pp. 166-181. University of Minnesota Press, Minneapolis, MN.
Davis MB (1987) Invasions of forest communities during the Holocene: beech and hemlock in the Great Lakes region. In: *Colonization, Succession and Stability: The 26th Symposium of the British Ecological Society* (eds Gray AJ, Crawley MJ, Edwards PJ), pp. 373-393. Blackwell Scientific Publications, Oxford.
Davis OK, Moutoux TE (1998) Tertiary and quaternary vegetation history of the Great Salt Lake, Utah, USA. *Journal of Paleolimnology,* **19**, 417-427.
Debreczy Z, Rácz I (1995) New species and varieties of conifers from Mexico. *Phytologia,* **78**, 217-243.
Deevey ES Jr (1949) Biogeography of the Pleistocene. *Geological Society of America Bulletin,* **60**, 1315-1416.
Demboski JR, Cook JA (2001) Phylogeography of the dusky shrew, *Sorex monticolus* (Insectivora, Soricidae): insight into deep and shallow history in northwestern North America. *Molecular Ecology,* **10**, 1227-1240.
Dong J, Wagner DB (1994) Paternally inherited chloroplast DNA polymorphism in

Pinus: estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics,* **136**, 1187-1194.
Dressler RL (1953) Some floristic relationships between Mexico and the United States. *Rhodora,* **56**, 81-96.
Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology,* **7**, 214-221.
Drummond AJ, Rambaut A (2009) Tracer 1.5. http://tree.bio.ed.ac.uk/software/tracer/.
Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution,* **22**, 1185-1192.
Du FK, Petit RJ, Liu J-Q (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and a comparison with other Conifers. *Molecular Ecology,* **18**, 1396-1407.
Duff RJ, Nickrent DL (1997) Characterization of mitochondrial small-subunit ribosomal RNAs from holoparasitic plants. *Journal of Molecular Evolution,* **45**, 631-639.
Duff RJ, Nickrent DL (1999) Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. *American Journal of Botany,* **86**, 372-386.
Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology,* **11**, 2571-2581.
Earle CJ (2009) *Pseudotsuga lindleyana*. http://www.conifers.org/pi/ps/lindleyana.htm.
Eckert AJ, Bower AD, Wegrzyn JL, *et al*. (2009) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics,* **182**, 1289-1302.
Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution,* **54**, 1839-1854.
Eisenhut G (1961) Untersuchungen über die Morphologie und Ökologie der Pollenkörner heimischer und fremdländischer Waldbäume. *Forstwissenschaftliche Forschungen,* **15**, 1-68.
Excoffier L, Laval G, Schneider S (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.
Farjon A, Page CN (1999) *Conifers: Status Survey and Conservation Action Plan.* IUCN/SSC Conifer Specialist Group, Gland, Switzerland.
Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution,* **17**, 368-376.
Finley RB Jr (1990) Woodrat ecology and behavior and the interpretation of paleomiddens. In: *Packrat Middens: The Last 40,000 Years of Biotic Change* (eds Betancourt JL, Van Devender TR, Martin PS, Van Devender TR, Martin PS), pp. 28-42. University of Arizona Press, Tucson, AZ.
Flous F (1934a) Deux espèces nouvelles de *Pseudotsuga* Américains. *Bulletin de la Société d'histoire naturelle de Toulouse,* **66**, 211-224.
Flous F (1934b) Diagnoses d'espèces et variétés nouvelles de *Pseudotsuga* Américains.

Bulletin de la Société d'histoire naturelle de Toulouse, **66**, 329-346.
Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.
Franklin JF, DeBell DS (1988) Thirty-six years of tree population change in an old-growth *Pseudotsuga-Tsuga* forest. *Canadian Journal of Forest Research,* **18**, 633-639.
Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics,* **147**, 915-925.
Galbreath KE, Hafner DJ, Zamudio KR (2009) When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution,* **63**, 2848-2863.
Gashwiler JS (1969) Seed fall of three conifers in west-central Oregon. *Forest Science,* **15**, 290-295.
Gernandt DS, Liston A (1999) Internal transcribed spacer region evolution in *Larix* and *Pseudotsuga* (Pinaceae). *American Journal of Botany,* **86**, 711-723.
Godbout J, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology,* **14**, 3497-3512.
Godbout J, Fazekas A, Newton CH, Yeh FC (2008) Glacial vicariance in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Molecular Ecology,* **17**, 2463-2475.
Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW (1995a) An evaluation of genetic distances for use with microsatellite loci. *Genetics,* **139**, 463-471.
Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences,* **92**, 6723-6727.
Goldstein S (1960) Degradation of pollen by Phycomycetes. *Ecology,* **41**, 543-545.
Gómez-Mendoza L, Arriaga L (2007) Modeling the effect of climate change on the distribution of oak and pine species of Mexico. *Conservation Biology,* **21**, 1545-1555.
González-Elizonado M, Jurado E, Návar J, *et al*. (2005) Tree-rings and climate relationships for Douglas-fir chronologies from the Sierra Madre Occidental, Mexico: a 1681-2001 rain reconstruction. *Forest Ecology and Management,* **213**, 39-53.
Gordon C, Cooper C, Senior CA, Banks HT, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147-168.
Graham A (1999) The Tertiary history of the northern temperate element in the northern Latin American biota. *American Journal of Botany,* **86**, 32-38.
Gray J (1961) Early Pleistocene paleoclimatic record from Sonoran Desert, Arizona. *Science,* **133**, 38-39.
Grigg LD, Whitlock C (1998) Late-glacial vegetation end climate change in western Oregon. *Quaternary Research,* **49**, 287-298.

Gugger PF, Sugita S (in press) Glacial populations and postglacial migration of Douglas-fir based on fossil pollen and macrofossil evidence. *Quaternary Science Reviews*, **29**, xxxx-xxxx.

Gugger PF, McLachlan JS, Manos PS, Clark JS (2008) Inferring long-distance dispersal and topographic barriers during post-glacial colonization from the genetic structure of red maple (*Acer rubrum* L.) in New England. *Journal of Biogeography,* **35**, 1665-1673.

Gugger PF, Sugita S, Cavender-Bares J (2010) Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Molecular Ecology,* **19**, 1877-1897.

Hakala KJ, Adam DP (2004) Late Pleistocene vegetation and climate in the southern Cascade Range and the Modoc Plateau region. *Journal of Paleolimnology,* **31**, 189-215.

Hall SA (1985) Quaternary pollen analysis and vegetational history of the Southwest. In: *Pollen Records of Late-Quaternary North American Sediments* (eds Bryant VM Jr, Holloway RG), pp. 95-124. American Association of Stratigraphic Palynologists Foundation, Dallas, TX.

Hallett DJ, Hills L (2006) Holocene vegetation dynamics, fire history, lake level and climate change in the Kootenay Valley, southeastern British Columbia, Canada. *Journal of Paleolimnology,* **35**, 351-371.

Hampe A, Bairlein F (2000) Modified dispersal-related traits in disjunct populations of bird-dispersed *Frangula alnus* (Rhamnaceae): a result of its Quaternary distribution shifts? *Ecography,* **23**, 603-613.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters,* **8**, 461-467.

Hann DW, Marshall DD, Hanus ML (2003) Equations for predicting height-to-crown-base, 5-year diameter-growth rate, 5-year height-growth rate, 5-year mortality rate, and maximum size-density trajectory for Douglas-fir and western hemlock in the coastal region of the Pacific Northwest. Research Contribution 40, Forest Research Laboratory, Oregon State University, Corvallis.

Hansen HP (1947) Postglacial forest succession, climate, and chronology in the Pacific Northwest. *Transactions of the American Philosophical Society,* **37**, 1-130.

Hasegawa MH, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* **22**, 160-174.

Hasumi H, Emori S (2004) *K-1 coupled GCM (MIROC) description.* Center for Climate System Research, University of Tokyo, Tokyo.

Hazell S (1979) *Late Quaternary vegetation and climate of the Dunbar Valley, British Columbia.* M.S. Thesis, University of Toronto.

Head MJ, Gibbard PL (2005) Early-Middle Pleistocene transitions: an overview and recommendation for the defining boundary. *Geological Society of London Special Publications,* **247**, 1-18.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633-1638.

Hermann RK (1985) *The Genus Pseudotsuga: Ancestral History and Past Distribution*

Forest Research Laboratory, Oregon State University, Corvallis, OR.
Hermann RK, Lavender DP (1990) Douglas-fir. In: *Silvics of North America: Conifers* (eds Burns RM, Honkala BH), pp. 1080-1108. USDA Forest Service, Washington, DC.
Heusser CJ (1956) Postglacial environments in the Canadian Rocky Mountains. *Ecological Monographs,* **26**, 263-302.
Heusser CJ (1960) Late-Pleistocene environments of north Pacific North America: an elaboration of late-glacial and postglacial climatic, physiographic, and biotic changes. *American Geographical Society Special Publication,* **35**, 1-308.
Heusser CJ (1964) Palynology of four bog sections from the western Olympic Peninsula, Washington. *Ecology,* **45**, 23-40.
Heusser CJ (1972) Palynology and phytogeographical significance of a late-Pleistocene refugium near Kalaloch, Washington. *Quaternary Research,* **2**, 189-201.
Heusser CJ (1974) Quaternary vegetation, climate, and glaciation of the Hoh River Valley, Washington. *Geological Society of America Bulletin,* **85**, 1547-1560.
Heusser CJ (1978) Palynology of Quaternary deposits of the lower Bogachiel River area, Olympic Peninsula, Washington. *Canadian Journal of Earth Sciences,* **15**, 1568-1578.
Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature,* **405**, 907-913.
Hey J, Nielson R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics,* **167**, 747-760.
Hey J, Nielson R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences,* **104**, 2785-2790.
Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology,* **12**, 2272-2281.
Holloway RG, Bryant VM Jr, Valastro S (1981) A 16,000 year pollen record from Lake Wabamun, Alberta, Canada. *Palynology*, **5**, 195-207.
Hu F-S, Hampe A, Petit RJ (2009) Paleoecology meets genetics: deciphering past vegetational dynamics. *Frontiers in Ecology and the Environment,* **7**, 371-379.
Hudson RR (1990) Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology,* **7**, 1-44.
Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics,* **111**, 147-164.
Huntley B, Birks HJB (1983) *An Atlas of Past and Present Pollen Maps for Europe: 0-13000 Years Ago.* Cambridge University Press, Cambridge.
Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology,* **52**, 577-586.
Jackson ST, Betancourt JL, Lyford ME, Gray ST, Rylander KA (2005) A 40,000-year woodrat-midden record of vegetational and biogeographical dynamics in north-eastern Utah, USA. *Journal of Biogeography,* **32**, 1085-1106.
Jackson ST, Overpeck JT, Webb T III, Keattch SE, Anderson KH (1997) Mapped plant-

macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews,* **16**, 1-70.
Jaramillo-Correa JP, Bousquet J (2005) Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics,* **171**, 1951-1962.
Jaramillo-Correa JP, Bousquet J, Beaulieu J, Isabel N, Perron M, Bouillé M (2003) Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of polymorphism and variation within and among species in *Picea*. *Theoretical and Applied Genetics,* **106**, 1353-1367.
Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology,* **13**, 2735-2747.
Jaramillo-Correa JP, Beaulieu J, Ledig FT, Bousqet J (2006) Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican-endemic conifer. *Molecular Ecology,* **15**, 2787-2800.
Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousqet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: seeing the forest for the genes. *Canadian Journal of Forest Research,* **39**, 286-307.
Johnson WC, Webb T III (1989) The role of blue jays (*Cyanocitta cristata* L.) in the postglacial dispersal of Fagaceous trees in eastern North America. *Journal of Biogeography,* **16**, 561-571.
Juárez-Agis A, López-Upton J, Vargas-Hernández JJ, Saénz-Romero C (2006) Geographic variation in germination and initial seedling growth of *Pseudotsuga menziesii* of Mexico. *Agrociencia,* **40**, 783-792.
Jukes TH, Cantor C (1969) Evolution of protein molecules. In: *Mammalian Protein Metabolism* (ed Munro MN), pp. 21-132. Academic Press, New York, NY.
Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics,* **61**, 893-903.
Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics,* **49**, 725-738.
Knowles LL (2001) Did Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology,* **10**, 691-701.
Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology,* **11**, 2623-2635.
Krutovsky KV, St. Clair JB, Saich R, Hipkins VD, Neale DB (2009) Estimation of population structure in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii*] using allozyme and microsatellite markers. *Tree Genetics and Genomes,* **5**, 641-658.
Kubiw H, Hickman M, Vitt DH (1989) The development of peatlands at Muskiki and Marguerite lakes, Alberta. *Canadian Journal of Botany,* **67**, 3534-3544.
Lacourse T (2005) Late Quaternary dynamics of forest vegetation on northern Vancouver Island, British Columbia, Canada. *Quaternary Science Reviews,* **24**, 105-121.
Lakhanpal RE (1958) The Rujuda flora of west central Oregon. *University of California Publications in Geological Science,* **35**, 1-66.
Lande R, Engen S, Saether BE (2003) *Stochastic population dynamics in ecology and*

conservation. Oxford University Press, Oxford, UK.
Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution,* **53**, 769-776.
LePage BA, Basinger JF (1991) A new species of *Larix* (Pinaceae) from the early Tertiary of Axel Heiberg Island, Arctic Canada. *Review of Palaeobotany and Palynology,* **70**, 89-111.
Li P, Adams WT (1989) Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research,* **19**, 149-161.
Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* **25**, 1451-1452.
Liepelt S, Bialozyt R, Ziegenhagen B (2002) Wind-dispersed pollen mediates postglacial gene flow among refugia. *Proceedings of the National Academy of Sciences,* **99**, 14590-14594.
Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleooceanography,* **20**, PA1003.
Little EL Jr (1952) The genus *Pseudotsuga* (Douglas-fir) in North America. *Leaflets Western Botany,* **6**, 181-198.
Little EL Jr (1971) *Atlas of United States Trees, Volume 1: Conifers and Important Hardwoods.* U.S. Department of Agriculture, Washington, DC.
Liverman DM, O'Brien KL (1991) Global warming and climate change in Mexico. *Global Environmental Change,* **1**, 351-364.
Lyford ME, Jackson ST, Betancourt JL, Gray ST (2003) Influence of landscape structure and climate variability on a late Holocene plant migration. *Ecological Monographs,* **73**, 567-583.
Lyford ME, Jackson ST, Gray ST, Eddy RG (2004) Validating the use of woodrat (*Neotoma*) middens for documenting natural invasions. *Journal of Biogeography,* **31**, 333-342.
Mack RN, Rutter NW, Valastro S, Bryant VM Jr (1978) Late Quaternary vegetation history at Waits Lake, Colville River Valley, Washington. *Botanical Gazette,* **139**, 499-506.
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.
Magri D, Vendramin GG, Comps B, *et al.* (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist,* **171**, 199-221.
Magri D, Fineschi S, Bellarosa R, *et al.* (2007) The distribution of *Quercus suber* chloroplast haplotypes matches palaeogeographical history of the western Mediterranean. *Molecular Ecology,* **16**, 5259-5266.
Maher LJ Jr (1961) *Pollen analysis and postglacial vegetation history in the Animas Valley Region, Southern San Juan Mountains, Colorado.* Ph.D. Dissertation, University of Minnesota.
Maher LJ Jr (1963) Pollen analysis of surface materials from the southern San Juan Mountains, Colorado. *Geological Society of America Bulletin,* **74**, 1485-1504.
Mápula-Larreta M, López-Upton J, Vargas-Hernández JJ, Hernández-Livera A (2007)

Reproductive indicators in natural populations of Douglas-fir in Mexico. *Biodiversity and Conservation,* **16**, 727-742.
Marshall KA, Neale DB (1992) The inheritance of mitochondrial DNA in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research,* **22**, 73-75.
Martin PR, McKay JK (2004) Latitudinal variation in genetic divergence of populations and the potential for speciation. *Evolution,* **58**, 938-945.
Martin PS (1963) Geochronology of pluvial Lake Cochise, southern Arizona. II. Pollen analysis of a 42-meter core. *Ecology,* **44**, 436-444.
Martínez M (1949) Las Pseudotsugas de México. *Anales del Instituto de Biología,* **20**, 129-184.
Mathiasen RL, Hawksworth FG, Edminster CB (1990) Effects of dwarf mistletoe on growth and mortality of Douglas-fir in the Southwest. *Great Basin Naturalist,* **50**, 173-179.
Mátyás G, Sperisen C (2001) Chloroplast DNA polymorphisms provide evidence for postglacial re-colonisation of oaks (*Quercus* spp.) across the Swiss Alps. *Theoretical and Applied Genetics,* **102**, 12-20.
McLachlan JS, Clark JS (2004) Reconstructing historical ranges with fossil data at continental scales. *Forest Ecology and Management,* **197**, 139-147.
McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology,* **86**, 2088-2098.
McLachlan JS, Hellman JJ, Schwartz MW (2007) A framework for debate of assisted migration in an era of climate change. *Conservation Biology,* **21**, 297-302.
Mehringer PJ Jr, Arno SF, Petersen KL (1977) Postglacial history of Lost Trail Pass Bog, Bitterroot Mountains, Montana. *Arctic and Alpine Research,* **9**, 345-368.
Milá B, Girman DJ, Kimura M, Smith TB (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society B,* **267**, 1033-1040.
Millar CI, Libby WJ (1991) Strategies for conserving clinal, ecotypic, and disjunct population diversity in widespread species. In: *Genetics and Conservation of Rare Plants* (eds Falk DAI, Holsinger KE), pp. 149-170. Oxford University Press, Oxford, UK.
Mitton JB, Kreiser BR, Latta RG (2000) Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Molecular Ecology,* **9**, 91-97.
Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity,* **99**, 224-232.
Moreno-Letelier A, Piñero D (2009) Phylogeographic structure of *Pinus strobiformis* Engelm. across the Chihuahuan Desert filter-barrier. *Journal of Biogeography,* **36**, 121-131.
Morin X, Viner D, Chuine I (2008) Tree species range shifts at a continental scale: new predictive insights from a process-based model. *Journal of Ecology*, **96**, 784-794.
Muse S, Weir B (1992) Testing for equality of evolutionary rates. *Genetics,* **132**, 269-276.
Navascués M, Vaxevanidou Z, González-Martínez SC, Climent J, Gil L, Emerson BC

(2006) Chloroplast microsatellites reveal colonization and metapopulation dynamics in the Canary Island pine. *Molecular Ecology,* **15**, 2691-2698.
Neale DB, Wheeler NC, Allard RW (1986) Paternal inheritance of chloroplast DNA in Douglas-fir. *Canadian Journal of Forest Research,* **16**, 1152-1154.
Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences,* **70**, 3321-3323.
Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York, NY.
Nielson R, Wakeley JW (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics,* **158**, 885-896.
Nielson MK, Lohman K, Sullivan J (2001) Phylogeography of the tailed frog (*Ascaphus truei*): implications for the biogeography of the Pacific Northwest. *Evolution,* **55**, 147-160.
Norma Oficial Mexicana (1994) *NOM-059-ECOL-1994: Que Determina las Espeies y Subespecies de Flora y Fauna Silvestres Terrestres y Acuáticas en Penligro de Extinción, Amenazadas, Raras y las Sujetas a Protección Especial, y que Establece Especificaciones para su Protección.* Secretaría de Desarrollo Social, México, DF.
Norma Oficial Mexicana (2001) *NOM-059-SEMARNAT-2001: Protección Ambiental – Especies Nativas de México de Flora y Fauna Silvestres - Categorías de Riesgo y Especificaciones para su Inclusión, Exclusión o Cambio - Lista de Especies en Riesgo.* Secretaría de Medio Ambiente y Recursos Naturales, México, DF.
O'Connell LM, Ritland K, Thompson SL (2008) Patterns of post-glacial colonization by western redcedar (*Thuja plicata*, Cupressaceae) as revealed by microsatellite markers. *Botany,* **86**, 194-203.
Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research,* **22**, 201-204.
Owens JN, Colangeli AM, Morris SJ (1991) Factors affecting seed set in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Botany,* **69**, 229-238.
Palacios-Chavez R, Rzedowski J (1993) Estudio palinológico de las floras fósiles del Mioceno inferior y principios del Mioceno medio de la región de Pichucalco, Chiapas, México. *Acta Botánica Mexicana,* **24**, 1-96.
Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *Journal of Molecular Evolution,* **28**, 87-97.
Palsbøll PJ, Bérubé M, Aguilar A, Notarbartolo-Di-Sciara G, Nielsen R (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. *Evolution,* **58**, 670-675.
Parisod C, Joost S (2010) Divergent selection in trailing- versus leading-edge populations of *Biscutella laevigata*. *Annals of Botany,* **105**, 655-660.
Pavia EG, Graef F, Reyes J (2009) Annual and seasonal surface air temperature trends in Mexico. *International Journal of Climatology,* **29**, 1324-1329.
Penhallow DP (1902) Notes on Cretaceous and Tertiary plants of Canada. *Royal Society of Canada Transactions, Series 2,* **8**, 31-91.

Penhallow DP (1907) *Manual of North American Gymnosperms*. The Athenaeum Press, Boston, MA.
Perry JP Jr, Graham A, Richardson DM (1998) The history of pines in Mexico and Central America. In: *Ecology and Biogeography of Pinus* (ed Richardson DM), pp. 137-149. Cambridge University Press, Cambridge, UK.
Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducousso A, Kremer A (1997) Chloroplast DNA footprints of postglacial recolonization by oaks. *Proceedings of the National Academy of Sciences,* **94**, 9996-10001.
Petit RJ, Mousadik AE, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology,* **12**, 844-855.
Petit RJ, Brewer S, Bordács S, *et al.* (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management,* **156**, 49-74.
Petit RJ, Csaikl UM, Bordács S, *et al*. (2002b) Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management,* **156**, 5-26.
Petit RJ, Aguinagalde I, de Beaulieu JL, *et al*. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science,* **300**, 1563-1565.
Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology,* **14**, 689-701.
Petit RJ, Hu F-S, Dick CW (2008) Forests of the past: a window to future changes. *Science,* **320**, 1450-1452.
Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190**, 231-259.
Piry S, Luikart G, Cornuet J- (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity,* **90**, 502-503.
Ponoy B, Hong Y-P, Woods J, Jaquish B, Carlson JE (1994) Chloroplast DNA diversity of Douglas-fir in British Columbia. *Canadian Journal of Forest Research,* **24**, 1824-1834.
Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity. I. Haploid locus. *Theoretical and Applied Genetics,* **90**, 462-470.
Pope V, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parameterisations in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, **16**, 123-146.
Porter SC, Pierce KL, Hamilton TD (1983) Late Wisconsin mountain glaciation in the western United States. In: *Late-Quaternary Environments of the United States: The Late Pleistocene* (eds Wright HE Jr, Porter SC), pp. 71-111. University of Minnesota Press, Minneapolis, MN.
Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**, 817-818.
Prentice IC (1985) Pollen representation, source area, and basin size: toward a unified theory of pollen analysis. *Quaternary Research,* **23**, 76-86.
Prentice IC (1988) Records of vegetation in time and space: the principles of pollen

analysis. In: *Vegetation History: Handbook of Vegetation Science* (eds. Huntley B, Webb T III), pp. 17-42. Kluwer, Dordrecht.
Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution,* **19**, 2092-2100.
Reboud X, Zeyl C (1994) Organelle inheritance in plants. *Heredity,* **72**, 132-140.
Reimer PJ, Baillie MGL, Bard E, *et al.* (2004) IntCal04 terrestrial radiocarbon age calibration, 0-26 cal kyr BP. *Radiocarbon,* **46**, 1029-1058.
Reyes-Hernández VJ, Vargas-Hernández JJ, López-Upton J, Vaquera-Huerta H (2005) Variación morfológica y anatómica en poblaciones mexicanas de *Pseudotsuga* (Pinaceae). *Acta Botánica Mexicana,* **70**, 47-67.
Reyes-Hernández VJ, Vargas-Hernández JJ, López-Upton J, Vaquera-Huerta H (2006) Phenotypic similarity among Mexican populations of *Pseudotsuga* Carr. *Agrociencia,* **40**, 545-556.
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.
Richmond GM, Fullerton DS (1986a) Introduction to Quaternary glaciations in the United States of America. *Quaternary Science Reviews,* **5**, 3-10.
Richmond GM, Fullerton DS (1986b) Summation of Quaternary glaciations in the United States of America. *Quaternary Science Reviews,* **5**, 183-196.
Ritter EW, Hatoff BW (1977) Late Pleistocene pollen and sediments: and analysis of a central California locality. *Texas Journal of Science,* **29**, 195-207.
Rodríguez-Banderas A, Vargas-Mendoza CF, Buonamici A, Vendramin GG (2009) Genetic diversity and phylogeographic analysis of *Pinus leiophylla*: a post-glacial range expanansion. *Journal of Biogeography,* **36**, 1807-1820.
Rosenberg SM, Walker IR, Mathewes RW (2003) Postglacial spread of hemlock (*Tsuga*) and vegetation history in Mount Revelstoke National Park, British Columbia, Canada. *Canadian Journal of Botany,* **81**, 139-151.
Rozen S, Skaletsky HJ (2000) Primer3 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, Totowa, NJ.
Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution,* **1**, 9-10.
Rymer MJ (1981) Stratigraphic revision of the Cache Formation (Pliocene and Pleistocene), Lake County, California. *United States Geological Survey Bulletin,* **1502-C**, 1-35.
Schorn HE (1994) A preliminary discussion of fossil larches (*Larix*, Pinaceae) from the Arctic. *Quaternary International,* **22-23**, 173-183.
Sea DS, Whitlock C (1995) Postglacial vegetation and climate of the Cascade Range, central Oregon. *Quaternary Research,* **43**, 370-381.
Shafer DS (1989) *The timing of late Quaternary monsoon precipitation maxima in the southwest United States.* Ph.D. Dissertation, University of Arizona.
Shafer SL, Bartlein PJ, Thompson RS (2001) Potential changes in the distributions of

western North America tree and shrub taxa under future climate scenarios. *Ecosystems,* **4**, 200-215.
Shafer SL, Bartlein PJ, Whitlock C (2005) Understanding the spatial heterogeneity of global environmental change in mountain regions. In: *Global Change and Mountain Regions: An Overview of Current Knowledge* (eds Huber UM, Bugmann HKM, Reasoner MA), pp. 21-30. Springer, Dordrecht.
Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics,* **123**, 603-613.
Smith AG, Pilcher JR (1973) Radiocarbon dates and vegetational history of the British Isles. *New Phytologist,* **72**, 903-914.
Sneath PHA, Sokal RR (1973) *Numerical taxonomy: the principles and practice of numerical classification.* Freeman, San Franciso, CA.
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.
Sorensen FC (1979) Provenance variation in *Pseudotsuga menziesii* seedlings from the var. *menziesii*-var. *glauca* transition zone in Oregon. *Silvae Genetica,* **28**, 37-119.
Spaulding WG (1984) The last glacial-interglacial climatic cycle: its effects on woodlands and forests in the American West. In: *Proceedings of the Eighth North American Forest Biology Workshop* (ed Lanner RM), pp. 42-68. Logan, UT.
Spaulding WG, Leopold EB, Van Devender TR (1983) Late Wisconsin paleoecology of the American Southwest. In: *Late-Quaternary Environments of the United States: The Late Pleistocene* (eds Wright HE Jr, Porter SC), pp. 259-293. University of Minnesota Press, Minneapolis, MN.
Spellman GM, Klicka J (2006) Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). *Proceedings of the Royal Society B,* **273**, 3057-3063.
St. Clair JB, Mandel NL, Vance-Boland KW (2005) Genecology of Douglas fir in western Oregon and Washington. *Annals of Botany,* **96**, 1199-1214.
Steele CA, Carstens BC, Storfer A, Sullivan J (2005) Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterrimus* (Caudata: Dicamptodontidae). *Molecular Phylogenetics and Evolution,* **36**, 90-100.
Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Journal of Biogeography,* **16**, 608-613.
Strickland LE, Thompson RS, Anderson KH (2001) USGS/NOAA North American Packrat Midden Database data dictionary. *United States Geological Survey Open-File Report,* **01-022**, 1-29.
Sugita S (1993) A model of pollen source area for an entire lake surface. *Quaternary Research,* **39**, 239-244.
Sugita S (1994) Pollen representation of vegetation in Quaternary sediments: theory and methods in patchy vegetation. *Journal of Ecology,* **82**, 881-897.
Sugita S (2007) POLLSCAPE Model. In: *Encyclopedia of Quaternary Science* (ed Elias SA), pp. 2561-2570. Elsevier, Amsterdam.
Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and

phylogeographic breaks in North America. *The American Naturalist,* **166**, 581-591.
Swofford DL (2003) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4.0b10. Sinauer Associates, Sunderland, MA.
Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics,* **105**, 437-460.
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* **123**, 585-595.
Takezaki N, Nei M, Tamura K (2009) POPTREE2: Software for constructing population trees from allele frequency data and computing other statistics with Windows interface. *Molecular Biology and Evolution,* **27**, 747-752.
Thomas DH (1983) The archaeology of Monitor Valley 2. Gatecliff Shelter. *Anthropological Papers of the American Museum of Natural History,* **59**, 1-552.
Thompson A, Schorn HE (1998) Mechanisms of migration in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco]: a model. *Integrative Paleontology and the Future,* University of California, Museum of Paleontology, Berkeley, CA.
Thompson RS, Whitlock C, Bartlein PJ, Harrison SP, Spaulding WG (1993) Climatic changes in the western United States since 18,000 yr B.P. In: *Global Climates since the Last Glacial Maximum* (eds Wright HE Jr, Kutzbach JE, Webb T III, *et al*.), pp. 468-513. University of Minnesota Press, Minneapolis, MN.
Tsukada M (1982) *Pseudotsuga menziesii* (Mirb.) Franco: its pollen dispersal and late Quaternary history in the Pacific Northwest. *Japanese Journal of Ecology,* **32**, 159-187.
Tsumura Y, Suyama Y (1998) Differentiation of mitochondrial DNA polymorphisms in populations of five Japanese *Abies* species. *Evolution,* **52**, 1031-1042.
Vargas-Hernández JJ, López-Upton J, Reyes-Hernández VJ, Domínguez-Alvarez A, Mápula-Larreta M (2004) Natural populations of Douglas-fir in Mexico: current status and needs for conservation. In: *Silviculture and the Conservation of Genetic Resources for Sustainable Forest Management: Proceedings of the Symposium of the North American Forest Commission, Forest Genetic Resources and Silviculture Working Groups, and the International Union of Forest Research Organizations (IUFRO)* (ed Beaulieu J), pp. 26-36. Quebec City, Canada.
Vázquez-Selem L, Heine K (2004) Late Quaternary glaciation of Mexico. In: *Quaternary Glaciations Extent and Chronology: Part III: South America, Asia, Africa, Australasia, Antarctica* (eds Ehlers J, Gibbard PL), pp. 233-242. Elsevier, Amsterdam.
Velasco-García MV, López-Upton J, Angeles-Pérez G, Vargas-Hernández JJ, Guerra-de la Cruz V (2007) *Pseudotsuga menziesii* seed dispersion in populations of central Mexico. *Agrociencia,* **41**, 121-131.
Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology,* **5**, 595-598.
Vendramin GG, Anzidei M, Madaghiele A, Bucci G (1998) Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics,* **97**, 456-463.

Viard F, El-Kassaby YA, Ritland K (2001) Diversity and genetic structure in populations of *Pseudotsuga menziesii* (Pinaceae) at chloroplast microsatellite loci. *Genome,* **44**, 336-344.
von Rudloff E (1972) Chemosystematic studies in the genus *Pseudotsuga*. I. Leaf oil analysis of the coastal and Rocky Mountain varieties of the Douglas-fir. *Canadian Journal of Botany,* **50**, 1025-1040.
von Rudloff E, Rehfeldt GE (1980) Chemosystematic studies in the genus *Pseudotsuga*. IV. Inheritance and geographical variation in the leaf oil terpenes of Douglas-fir from the Pacific Northwest. *Canadian Journal of Botany,* **58**, 546-556.
Wakasugi T, Tsudzuki J, Ito S, Nakashima T, Tsudzuki T, Sugiura M (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proceedings of the National Academy of Sciences,* **91**, 9794-9798.
Walter R, Epperson BK (2001) Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. *Molecular Ecology,* **10**, 103-111.
Wang X-Q, Tank DC, Sang T (2000) Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Molecular Biology and Evolution,* **17**, 773-781.
Wanket, J (2002) *Late Quaternary vegetation and climate of the Klamath Mountains*. Ph.D. Dissertation, University of California, Berkeley.
Wanner H, Beer J, Bütikofer J, *et al.* (2008) Mid- to Late Holocene climate change: an overview. *Quaternary Science Reviews,* **27**, 1791-1828.
Webber JE, Painter RA (1996) *Douglas-fir Pollen Management Manual*, 2nd edn. British Columbia Ministry of Forests, Victoria, BC.
Weng CY, Jackson ST (1999) Late Glacial and Holocene vegetation history and paleoclimate of the Kaibab Plateau, Arizona. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **153**, 179-201.
West GJ, Woolfenden W, Wanket JA, Anderson, RS (2007) Late Pleistocene and Holocene environments. In: *California Prehistory: Colonization, Culture, and Complexity* (eds Jones TL, Klar KA), pp. 11-34. Altamira, Lanham, MD.
Whitlock C (1992) Vegetational and climatic history of the Pacific Northwest during the last 20,000 years: implications for understanding present-day biodiversity. *The Northwest Environmental Journal,* **8**, 5-28.
Whitlock C (1993) Postglacial vegetation and climate of Grand Teton and southern Yellowstone National Parks. *Ecological Monographs*, **63**, 173-198.
Whitlock C (1995) The history of *Larix occidentalis* during the last 20,000 years of environmental change. In: *Ecology and Management of* Larix *forests: A Look Ahead, General Technical Report GTR-INT-319* (eds Schmidt WC, McDonald KJ) pp. 83-90. USDA Forest Service, Intermountain Research Station, Ogden, UT.
Whitlock C, Bartlein PJ (1993) Spatial variations of Holocene climatic change in the Yellowstone region. *Quaternary Research,* **39**, 231-238.
Whitlock C, Bartlein PJ (1997) Vegetation and climate change in northwest America during the past 125 kyr. *Nature,* **388**, 57-61.
Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA (2009) Niches, models,

and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, **106**, 19729-19736.
Williams JW, Shuman B, Bartlein PJ*, et al.* (2006) *Atlas of Pollen-Vegetation-Climate Relationships for the United States and Canada.* American Association of Stratigraphic Palynologists, Dallas, TX.
Willyard A, Syring J, Gernandt DS, Liston A, Cronn R (2007) Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. *Molecular Biology and Evolution,* **24**, 90-101.
Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution,* **22**, 297-307.
Worona MA, Whitlock C (1995) Late Quaternary vegetation and climate history near Little Lake, central Coast Range, Oregon. *Geological Society of America Bulletin,* **107**, 867-876.
Zavarin E, Snajberk K (1973) Geographic variability of monoterpenes from cortex of *Pseudotsuga menziesii*. *Pure and Applied Chemistry,* **34**, 411-434.
Zdanowicz CM, Zielinski GA, Germani MS (1999) Mount Mazama eruption: calendrical age verified and atmospheric impact assessed. *Geology,* **27**, 621-624.
Zhang J, Marshall JD, Jaquish BC (1993) Genetic differentiation in carbon isotope discrimination and gas exchange in *Pseudotsuga menziesii*: a common garden experiment. *Oecologia,* **93**, 80-87.