Tree Genetics & Genomes (2011) 7:11–21
DOI 10.1007/s11295-010-0311-0

ORIGINAL PAPER

Genetic diversity and structure of western white pine (*Pinus monticola*) in North America: a baseline study for conservation, restoration, and addressing impacts of climate change

Mee-Sook Kim · Bryce A. Richardson · Geral I. McDonald · Ned B. Klopfenstein

Received: 21 October 2009 / Revised: 20 January 2010 / Accepted: 21 May 2010 / Published online: 22 June 2010

Abstract Western white pine (*Pinus monticola*) is an economically and ecologically important species in western North America that has declined in prominence over the past several decades, mainly due to the introduction of *Cronartium ribicola* (cause of white pine blister rust) and reduced opportunities for regeneration. Amplified fragment length polymorphism (AFLP) markers were used to assess the genetic diversity and structure among populations at 15 sites (e.g., provenances) across the native range of western white pine. The level of genetic diversity was different among 15 populations tested using 66 polymorphic AFLP loci. Nei's gene diversity (H_E) at the population level ranged from 0.187 to 0.316. Genetic differentiation (G_{ST}) indicated that 20.1% of detected genetic variation was explained by differences among populations. In general, populations below 45°N latitude exhibited a higher level of genetic diversity than higher latitude populations. Genetic distance analysis revealed two major clades between northern and southern populations, but other well-supported relationships are also apparent within each of the two clades. The complex relationships among populations are likely derived from multiple factors including migration, adaptation, and multiple glacial refugia, especially in higher latitudes. Genetic diversity and structure revealed by this study will aid recognition and selection of western white pine populations for species management and conservation programs, especially in consideration of current and future climate changes.

Keywords Genetic variation · Tree population structure · *Pinus monticola*

Communicated by S. Aitken

M.-S. Kim
Department of Forest Resources, Kookmin University,
Seoul 136-702, South Korea
e-mail: mkim@kookmin.ac.kr

B. A. Richardson
USDA Forest Service, RMRS,
735 N. 500 E.,
Provo, UT 84606, USA
e-mail: brichardson02@fs.fed.us

G. I. McDonald · N. B. Klopfenstein (✉)
USDA Forest Service, RMRS,
1221 S. Main St.,
Moscow, ID 83843, USA
e-mail: nklopfenstein@fs.fed.us

G. I. McDonald
e-mail: gimcdonald@clearwire.net

Introduction

Western white pine (*Pinus monticola*) is distributed across a large region of western North America. Thus, this species persists across widely ranging environmental conditions and diverse habitats from 55°N to 36°N latitude and 113°W to 126°W longitude. Western white pine (WWP) is the dominant seral species of mesic, montane forest ecosystems of interior northwestern North America that includes northern Idaho and adjacent regions in Montana, Washington, and British Columbia (Wellner 1962). West of the Cascade Range, WWP is a minor seral component of the western hemlock (*Tsuga heterophylla*) zone; however, it is also found as a minor seral species in drier forests on the eastern slopes of the Cascade Range in southern Oregon

(Franklin and Dyrness 1973). In the Sierra Mountains, WWP has a restricted distribution as a minor subalpine species typically at elevations above 2,500 m (Griffin and Critchfield 1972).

Rapid growth, regeneration ability, and occupancy of sites make WWP highly desirable throughout its range. Under optimal conditions, WWP can reach a height of 72 m and a diameter of 2.1 m. Before its decline starting in the 1930s, this species was the most economically valuable timber species in interior northwestern North America. The dramatic decline of WWP is attributed to several factors—white pine blister rust caused by the introduced pathogen *Cronartium ribicola*, mountain pine beetle, fire suppression, and logging (Mehes et al. 2009). In this region, the coverage of WWP has been reduced by 90% compared with its distribution 70 years ago (Neuenschwander et al. 1999). WWP also faces additional threats due to climate change (Rehfeldt et al. 2006).

Genetic studies of WWP became prominent during the white pine blister rust epidemic and the creation of a resistance-breeding program (Bingham 1983; McDonald et al. 2001). Western white pine is predominantly an outcrossing species, due to self-sterility (Bingham et al. 1974). A previous study showed WWP populations had a wide range of genetic variation based on isozyme analysis (Steinhoff et al. 1983). That previous study attributed most of the genetic variation to a latitudinal genetic discontinuity at the Oregon–California border (Fig. 1). Based on the same seed collections, a subsequent, common-garden study of quantitative traits also found that most genetic variation was attributable to growth and development along a latitudinal gradient in the same region (Rehfeldt et al. 1984). However, little adaptive trait variation was observed within these broad geographical regions. On this basis, Rehfeldt et al. (1984) proposed that WWP behaved as a generalist within the northern and southern groups. In addition to latitude, a separate common-garden study in the Cascade Range found that growth and development traits also varied with a longitudinal relationship caused by the crest of the southern Cascade Range (Campbell and Sugano 1989).

Understanding population structure is critical because of multiple interactions among hosts, pathogens, and the environment (Thompson 2005). Locations of genetic discontinuities, suture zones, and zones of secondary contact on landscapes provide insights about important adaptive variation and discordant evolutionary histories (Soltis et al. 1997). Such discontinuities have been demonstrated in ponderosa pine (*Pinus ponderosa*; Johansen and Latta 2003), red pine (*Pinus resinosa*; Walter and Epperson 2001), and whitebark pine (*Pinus albicaulis*; Richardson et al. 2002); however, this information is largely lacking for WWP.

While some adaptive variation has been identified in WWP, limited information is available about the historical demographics and impact from climatic fluctuations on neutral genetic variation. Genetic data from other temperate and boreal tree species suggest analogous patterns of constriction and isolation of northern populations during the last glacial period, followed by expansion during Holocene warming (Soltis et al. 1997; Newton et al. 1999). Such data allow inferences into the potential responses of these plant species under changing climates, the influence of climate on plant distributions and thus genetic structure and diversity. In this study, amplified fragment length polymorphism (AFLP) markers were utilized to (1) assess genetic diversity and structure across the major regions of WWP distribution and (2) develop inferences about the influence of past climatic change on genetic diversity and structure.

Materials and methods

Sampling and DNA extraction

We collected from 15 natural WWP populations representing range-wide distribution across six regions suggested by Steinhoff et al. (1983): 1 interior northern Rocky Mountains, 2 coastal, 3 north-central Cascade Range, 4 south-central Cascade Range, 5 Siskiyou Mountains, and 6 Sierra Mountains (Table 1; Fig. 1). Two or three populations were selected in each of the six regions (Table 1; Fig. 1). Populations were placed to maximize the probability of discovering heterogeneity by locating populations at either longitudinal/latitudinal extremes or across potential geographic boundaries (Fig. 1). One population was included to investigate the genetic relationships of an isolated central Oregon population, Indian Creek (*InCr*, Oregon, USA), to other WWP populations (Fig. 1). A total of 357 trees were sampled from the 15 populations (Table 1) during the summer season (July–September). All trees sampled were from natural stands that showed low mortality from white pine blister rust within sampled cohorts. In addition, most sampled trees were at least 75 years of age.

Approximately 100 mg (fresh weight) of WWP needles/buds were used for DNA extraction. DNA was extracted and purified from all samples using a Qiagen DNA extraction kit (Qiagen, Inc.) following the protocol of the manufacturer. DNA yield was quantified by fluorometry. High-quality DNA (average DNA concentration=560 ng/ul) was obtained from all tree samples (data not shown). A subset of DNA samples from diverse populations was selected and duplicated for checking repeatability of bands (peaks), based on the AFLP protocol and analyses of Kim et al. (2003).

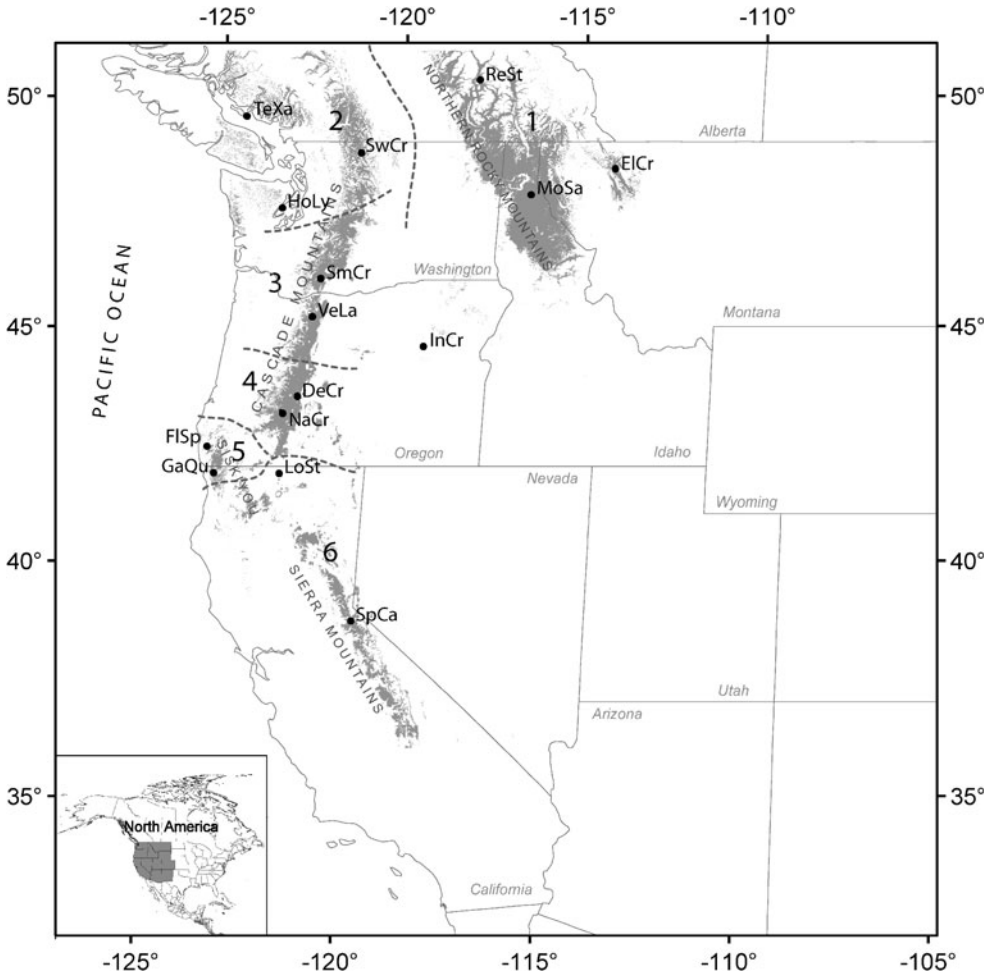

Fig. 1 Western white pine (*Pinus monticola*) population locations are overlaid with the predicted distribution of the species shown in *gray* based on bioclimate modeling (Rehfeldt et al. 2006). The *numbers* refer to geographic regions of western white pine distribution, as suggested by Steinhoff et al. (1983). The *inset* depicts the region of North America included in the map (see Table 1 for region and population information)

AFLP analyses

The AFLP analyses were performed following the protocol of Kim et al. (2003). For restriction digests, 500 ng of genomic DNA were digested with *Eco*RI and *Mse*I to serve as the template. Resulting DNA fragments were ligated to adapters and diluted 1:10 with sterile, distilled water prior to pre-amplification. Pre-amplification, polymerase chain reaction (PCR) mixtures (total 30 µl) contained 6 µl of diluted restriction/ligation mixture as template, 10× PCR buffer (Applied Biosystems, Inc., http://www.appliedbiosystems.com/), 3 mM $MgCl_2$, 200 µM dNTPs, 300 nM of +2 primers (E-AC and M-CC), and 1.5 U Ampli*Taq*® DNA polymerase (Applied Biosystems). Fluorescent dye-labeled *Eco*RI (+3 primer; E-ACG) and unlabeled *Mse*I (+4 primers; M-CCAG, M-CCAT, and M-CCTA) primers (IDT DNA, http://www.idtdna.com/) were used for the selective amplification. For selective amplification, reaction mixtures (total 25 µl) contained 5 µl diluted, pre-amplification products (1:40 with low TE buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 8.0)) as a template, 10× PCR buffer, 2.4 mM $MgCl_2$, 300 µM dNTPs, 100 nM of +3 E primer, 300 nM of +4 M primers, and 1 U of Ampli*Taq*® Gold polymerase. Amplifications were performed following the method of Remington et al. (1999). A negative control was included for both pre-selective and selective amplification. Selective amplification products (diluted 1:4 with sterile, distilled water) were separated in an ABI 3700 DNA automated sequencer (Applied Biosystems) along with known positive controls at the University of Wisconsin Biotechnology Center (http://www.biotech.wisc.edu/). Genotyper 3.7 NT (Applied Biosystems) was used to identify peaks with a fluorescent intensity greater than the threshold value (ca. 150 units) in at least one sample. Categories were made from these identified peaks for scoring samples. AFLP bands were scored as present (1) or absent (0) using Genotyper 3.7 NT, and a binary matrix was developed with molecular sizes ranging from 70 to 286 bp. Each band (peak) was checked visually using GeneScan 3.7 NT (Applied BioSystems) by three separate scientists (MS Kim, BA Richardson, and JE Stewart) to minimize genotyping errors. These data sets were cross-read by three scientists, and suspicious/ambiguous peaks (e.g., shoulder peaks) were revisited and eliminated from further analyses (Bonin et al. 2004).

Table 1 Geographic location and genetic diversity of western white pine (*Pinus monticola*) populations in this study

Region[a]	Population	Latitude	Longitude	*n*	Elev. (m)	*P*	PPL (%)	n_a	n_e	H_E	H'_E	*I*
1	Revelstoke, British Columbia, Canada (*ReSt*)	50º59′49	118º11′38	27	500	58	87.9	1.606	1.326	0.194	0.240	0.296
2	Texada Island, BC, Canada (*TeXa*)	48º15′48	113º52′37	25	1100	60	90.9	1.621	1.328	0.198	0.253	0.301
2	Swamp Creek, Washington, USA (*SwCr*)	47º35′19	116º01′41	29	1345	58	87.9	1.667	1.339	0.208	0.255	0.318
1	Elya Creek, Montana, USA (*ElCr*)	49º41′57	124º21′57	27	215	56	84.8	1.636	1.304	0.190	0.239	0.294
1	Moon Saddle, Idaho, USA (*MoSa*)	48º34′11	120º46′54	22	1190	58	87.9	1.576	1.307	0.187	0.231	0.285
2	Holly, Washington, USA (*HoLy*)	47º34′22	122º55′15	17	150	59	89.4	1.636	1.353	0.215	0.281	0.325
3	Smokey Creek, Washington, USA (*SmCr*)	46º1′28	121º40′59	21	1114	60	90.9	1.621	1.346	0.204	0.266	0.308
3	Veda Lake, Oregon, USA (*VeLa*)	45º15′34	121º45′45	26	1310	60	90.9	1.727	1.387	0.233	0.280	0.354
	Little Indian Creek, Oregon, USA (*InCr*)[b]	44º21′27	118º47′40	25	1585	61	92.4	1.758	1.486	0.274	0.325	0.404
4	Deer Creek, Oregon, USA (*DeCr*)	43º15′01	121º51′48	28	1740	64	97.0	1.894	1.545	0.316	0.353	0.471
4	National Creek, Oregon, USA (*NaCr*)	42º59′57	122º22′52	23	1150	63	95.5	1.697	1.379	0.225	0.284	0.341
5	Fly Catcher Spring, Oregon (*FlSp*)	42º21′17	124º17′42	27	800	59	89.4	1.773	1.465	0.275	0.322	0.412
5	Gasquet, California, USA (*GaQu*)	41º51′22	123º54′42	21	435	62	93.9	1.789	1.445	0.264	0.315	0.389
6	Lodge Pole Station, California, USA (*LoSt*)	41º49′54	122º12′53	22	1880	61	92.4	1.773	1.511	0.289	0.334	0.426
6	Spur Canyon, California, USA (*SpCa*)	38º42′14	120º06′10	17	2500	54	81.8	1.721	1.500	0.254	0.312	0.376
Mean				23.8		59.5	90.2	1.700	1.401	0.235	0.286	0.353
Total				357								

Elev elevation, *n* number of individuals sampled, *P* number of polymorphic loci, *PPL* percentage of polymorphic loci, n_a observed number of alleles per locus, n_e effective number of alleles per locus, H_E Nei's (1973) gene diversity, H'_E Nei's gene diversity estimated with the computer program AFLP-SURV 1.0 (Vekemans 2002), *I* Shannon's index of phenotypic diversity

[a] 1 1 interior northern Rocky Mountains, 2 coastal, 3 north-central Cascade Range, 4 south-central Cascade Range, 5 Siskiyou Mountains, and 6 Sierra Mountains (see Fig. 1)

[b] Not previously designated in the geographic regions of western white pine, as suggested by Steinhoff et al. (1983)

Data analyses

We analyzed AFLP data based on both allele and phenotypic frequencies. Polymorphic bands were selected at the 95% level (two-tailed test) for use in further analysis. Data matrices were analyzed using POPGENE version 1.32 (Yeh et al. 1999) with the assumption that the populations were in Hardy–Weinberg equilibrium. The following genetic parameters were determined: number of polymorphic loci (*P*), percentage of polymorphic loci (PPL), number of alleles per locus (n_a), effective number of alleles per locus (n_e), genetic diversity (H_E = expected heterozygosity), Shannon's index (*I*) of phenotypic diversity, genetic differentiation among populations (G_{ST}). In addition, Nei's gene diversity (H'_E) was also estimated using AFLP-SURV version 1.0 (Vekemans 2002). This value is calculated from allele frequencies derived from Lynch and Milligan's (1994) method with an option of a Bayesian approach with non-uniform prior distribution of allele frequencies (Zhivotovsky 1999). Genetic relationships among populations based on genetic distance (Nei 1978) were displayed with a neighbor-joining tree. This analysis was performed using 1,000 bootstrap replicates from AFLP-SURV and constructed with NEIGHBOR and CONSENSE of the software package PHYLIP version 3.6 (Felsenstein 2004). We also used *t* tests with equal variance to determine whether a significant difference existed between the genetic diversity parameters of northern and southern populations (based on results from the Nei's genetic distance) using SAS® (SAS Institute Inc 2000).

We examined hierarchical genetic variation across the geographic range of WWP using the analysis of molecular variance (AMOVA) determined with ARLEQUIN 3.01 (Excoffier et al. 2005). Pairwise F_{ST} values (θ statistic) obtained from AMOVA were used to measure the genetic differentiation between populations. Relationships among the members of the F_{ST} genetic distance matrix were represented by non-metric, multidimensional scaling (MDS) as suggested by Pinedo-Cancino et al. (2006). The MDS and a minimum spanning tree were conducted using SYSTAT ver. 9 (SPSS Inc. 1998). We compared four alternative population groupings with AMOVA to test which grouping explained the greatest proportion of variance. The tested groups were assembled on the basis of different combinations of geographic regions (Table 1; Fig. 1). To test the correlations among genetic distances and geographic distances of populations, Mantel's (1967) tests were conducted using tools for population genetic analyses (Miller 1997).

Bayesian model-based clustering was applied using STRUCTURE version 2.2 (Pritchard et al. 2000) to (a) infer the number of genetic clusters (K), (b) determine genetic cluster membership of WWP individuals without assuming predefined genetic clusters, and (c) detect genetic discontinuities. For analysis using STRUCTURE, no prior assumptions of population structure were imposed on the analysis. To estimate the number of genetic clusters, eight simulations were run with a model of 2 to 9 genetic clusters using the recessive alleles approach in STRUCTURE (Falush et al. 2007). Each model was run with admixture, and the recommended methods for recessive alleles. The models were run with a burn-in length of 10,000 followed by 10^5 iterations. The number of genetic clusters (K) was estimated by calculating the second-order rate of change in the log probability scores (Evanno et al. 2005).

Results

Genetic diversity

Analysis of three selective primer combinations yielded a total of 102 presumptive loci in 357 individuals from 15 populations, of which 66 (64.7%) loci were polymorphic. Each of the 357 sampled individuals had a unique AFLP profile. The level of genetic diversity was different among 15 populations (Table 1). In general, southern populations (south of 45°N latitude) exhibited a higher level of genetic variation than northern populations. The relationship between latitude and heterozygosity (H_E) is plotted (Fig. 2). Most of the genetic diversity parameters (e.g., n_a, n_e, H_E, H'_E, and I) were significantly different between northern and southern populations ($P<0.001$). The *DeCr* population had the highest values for all genetic diversity parameters (e.g., H_E, H'_E, I), whereas *MoSa* had the lowest genetic diversity values except for percent polymorphic loci and effective number of alleles per locus (Table 1).

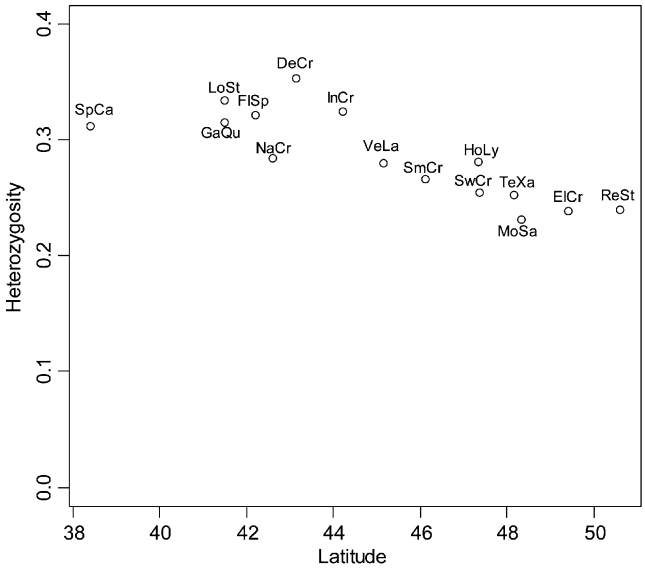

Fig. 2 The plotted relationship of within population heterozygosity (H_E) and latitude (see Table 1 for population information)

Genetic structure

Large-scale genetic relationships among populations are illustrated with a neighbor-joining tree based on Nei's (1978) genetic distances (Fig. 3). The major division in genetic distance among populations occurred across a relatively small geographic region. Oregon populations located east of the Cascade Range crest (i.e., *DeCr* and

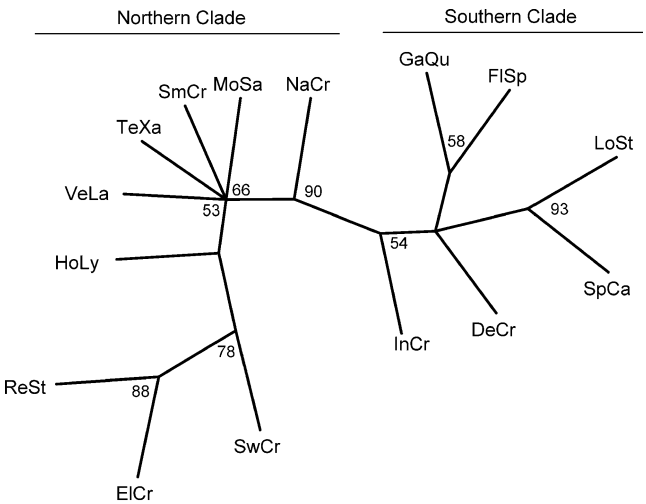

Fig. 3 The genetic relationships of populations based on Nei's pairwise genetic distance using neighbor-joining consensus tree. Bootstrap values above 50% are shown (see Table 1 for population information)

InCr) were included with a well-supported clade with other southern populations (Fig. 3). However, *NaCr* approximately 40 km to the southwest of *DeCr* and west of the Cascade crest was included in the clade with northern populations. Other well-supported relationships are also apparent within each of the two major clades. Within the northern clade, the two populations from the northeastern edge of the WWP range formed a distinct subclade. Within the southern clade, populations in Sierra Mountains of California (*LoSt* and *SpCa*) and the Siskiyou Mountains of California and Oregon (*GaQu* and *FlSp*) each formed distinct subclades (Fig. 3).

Genetic differentiation (G_{ST}) indicated that 20.1% of detected genetic variation was explained by differences among populations. Most pairwise F_{ST} values were significant and ranged from 0.387 between *SpCa* and *TeXa* to 0.023 between *ElCr* and *SwCr*. Statistically insignificant ($P>0.05$) F_{ST} values were obtained for pairwise comparisons among *HoLy*, *SmCr*, *SwCr*, and *VeLa* populations in the central and northern Cascade Range of Washington and Oregon (Table 2). The MDS and minimum spanning tree provide indications of potential population relationships (Stress=0.037). Non-significant differences among *HoLy*, *VeLa*, and *SmCr* populations of Washington and Oregon and between *HoLy* and *SwCr* populations of Washington are a striking feature of the genetic distance matrix (Fig. 4). Other notable features are a large genetic differentiation among the populations located in the Siskiyou (*FlSp* and *GaQu*), Sierra (*SpCa*), and southeastern Cascade Range (*DeCr*), and a large genetic distance between *MoSa* and *ReSt* populations in the northern Rocky Mountains (Fig. 4). However, small genetic differentiation is also detected between *ElCr* (geographically close to the *MoSa* site) and *ReSt* in the northern Rocky Mountains (Figs. 1 and 4). Overall, a Mantel test showed a significant correlation between genetic and geographic distances for the entire data set ($r=0.624$, $P<0.001$). However, several inconsistencies in this pattern are apparent. For example, several populations within the northern clade showed geographic distances are not correlated with genetic similarities (e.g., between *TeXa* and *SwCr*, between *MoSa* and *ElCr*; Figs. 1, 3, and 4). *TeXa* and *SwCr* are geographically closer to each other than any other populations in northern clade, but apparently they are genetically more distant compared to the other northern populations (Figs. 1, 3, and 4). Although *DeCr* grouped with populations in the Siskiyou and Sierra Mountains and *NaCr* grouped with northern populations, these are spatially closer than any other pair of populations. Similar results are found with AMOVA and Bayesian structure methods discussed below (Figs. 5 and 6).

Hierarchical genetic structure across the geographic range of WWP populations was analyzed with AMOVA. For AMOVA, a four-group model explained the highest genetic variation among groups (13.7%) while minimizing genetic variation within groups (8.1%). The remaining genetic variation (78.2%) was attributable to differences among individuals within populations. These groupings are illustrated in Fig. 5. Without prior assumptions about genetic structure, the Bayesian model assigned WWP individuals to populations in a pattern similar to that from the AMOVA model. The Bayesian model with four genetic clusters ($K=4$) had the highest probability (99.9%). The Bayesian-based analysis assigned nearly all individual samples from nine populations in the northern clade to one genetic cluster with high

Table 2 The pairwise F_{ST} and significant P values for all populations of western white pine (*Pinus monticola*)

Site	ReSt	TeXa	SwCr	ElCr	MoSa	HoLy	SmCr	VeLa	InCr	DeCr	NaCr	FlSp	GaQu	LoSt	SpCa
ReSt	0.000	+	+	+	+	+	+	+	+	+	+	+	+	+	+
TeXa	0.187	0.000	+	+	+	+	+	+	+	+	+	+	+	+	+
SwCr	0.077	0.111	0.000	+	+	−	+	+	+	+	+	+	+	+	+
ElCr	0.057	0.149	0.023	0.000	+	+	+	+	+	+	+	+	+	+	+
MoSa	0.171	0.146	0.127	0.146	0.000	+	+	+	+	+	+	+	+	+	+
HoLy	0.105	0.094	0.019	0.033	0.067	0.000	−	−	+	+	+	+	+	+	+
SmCr	0.093	0.106	0.034	0.038	0.082	0.000	0.000	−	+	+	+	+	+	+	+
VeLa	0.109	0.097	0.040	0.054	0.079	0.017	0.002	0.000	+	+	+	+	+	+	+
InCr	0.162	0.198	0.128	0.118	0.156	0.079	0.076	0.091	0.000	+	+	+	+	+	+
DeCr	0.235	0.191	0.171	0.179	0.177	0.123	0.116	0.119	0.088	0.000	+	+	+	+	+
NaCr	0.151	0.090	0.084	0.071	0.105	0.049	0.042	0.029	0.101	0.099	0.000	+	+	+	+
FlSp	0.262	0.273	0.234	0.211	0.264	0.196	0.170	0.178	0.156	0.169	0.150	0.000	+	+	+
GaQu	0.357	0.318	0.280	0.290	0.329	0.237	0.240	0.248	0.173	0.123	0.227	0.133	0.000	+	+
LoSt	0.303	0.283	0.264	0.261	0.226	0.192	0.174	0.205	0.101	0.107	0.163	0.181	0.159	0.000	+
SpCa	0.385	0.387	0.363	0.349	0.367	0.300	0.314	0.319	0.192	0.204	0.281	0.263	0.258	0.137	0.000

See Table 1 for population information, + $P< 0.05$

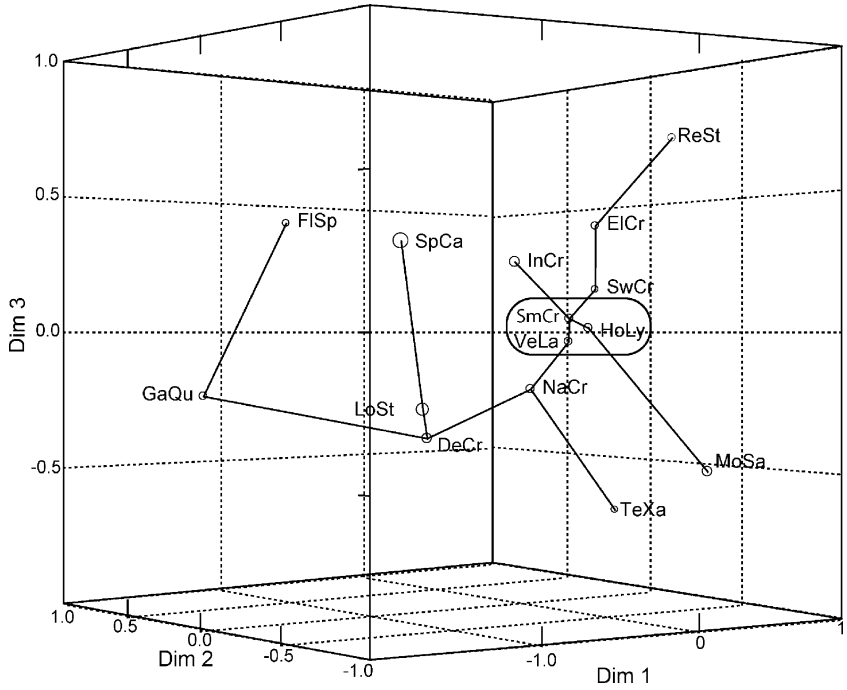

Fig. 4 Non-metric multidimensional scaling representation of F_{ST} genetic distance matrix (Table 2) showing minimum spanning tree relationships of non-significant pairwise comparisons (*circled*) to significant comparisons (see Table 1 for population information)

probability (Fig. 6). These populations were all located above 45°N with the exception of *NaCr*, which is located west of the Cascade crest in southern Oregon, USA. Remaining individuals from populations located below 45°N were generally assigned to one of three genetic clusters. Individuals from *SpCa*, *FlSp*, and *DeCr* populations generally had high assignment probabilities for three different genetic clusters. Individuals from the remaining populations, *LoSt*, *GaQu*, and *InCr*, were generally assigned to mixtures of two or more genetic clusters (Fig. 6). Interestingly, individuals from the most isolated site (*InCr*) located in between the northern and southern clades showed the most abundant proportions of mixed ancestry (Fig. 6).

Discussion

The present AFLP study found a large genetic discontinuity in WWP that occurs across southern Oregon. The previous allozyme study (Steinhoff et al. 1983) showed a similar pattern; however, the genetic differentiation is more prevalent in this region than previously reported and, as discussed below, is not entirely consistent with geographic breaks in the distribution of WWP. The present study detected considerably more genetic differentiation between populations located in the southern Cascade Range and Siskiyou Mountains than was found in the previous isozyme study (Steinhoff et al. 1983). In addition, significant genetic differentiation among several of the northern populations was detected by the present study. For example, marked genetic differentiation was observed among the northern Rocky Mountain populations (*MoSa, ElCr*, and *ReSt*), between the Siskiyou Mountain populations (*FlSp* and *GaQu*) and in the Pacific Northwest populations (*TeXa* and its closest geographic neighbors, *HoLy*, *SmCr*, *SwCr*, and *VeLa*).

Similarities in the partitioning of genetic variation can be expanded to quantitative traits. Quantitative genetic studies have shown that much of the genetic variation is associated with growth potential. Discontinuity in the quantitative variation is also confined to the region of southern Oregon and northern California (Rehfeldt et al. 1984). Subsequent quantitative studies in this region with more robust sampling have shown that variation of growth potential also forms a cline across the Cascade crest (Campbell and Sugano 1989). The Cascade crest also represents a boundary between recognized ecoregions (Olson et al. 2001). This pattern is also reflected in the genetic structure between populations on opposite sides of the Cascade crest in Oregon (i.e., *DeCr* versus *NaCr*). Apparent similarities in adaptive traits and AFLP data in WWP were the subject of a parallel study (Richardson et al. 2009).

For conifers, a large majority of genetic diversity is typically partitioned within populations with low genetic differentiation among populations (Hamrick and Godt 1996). In contrast, our analyses demonstrate that approximately 20% of genetic variation within WWP is distributed among populations. It is noteworthy that the majority of genetic variation among populations is attributable to populations in southern Oregon and northern California that represent a relatively small proportion of the total

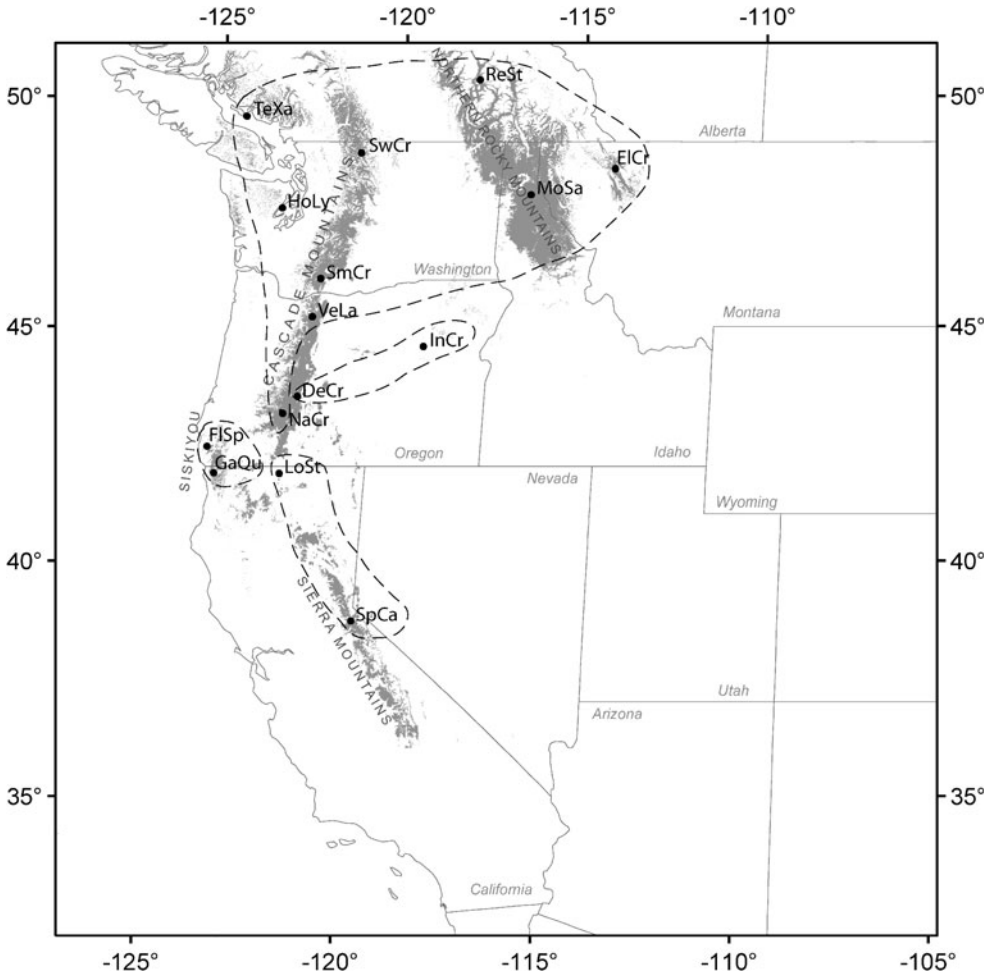

Fig. 5 Analysis of molecular variance (AMOVA)-based four grouping from AFLP phenotypes. Groups are depicted with the *dashed lines*. The genetic variation explained is 13.7% for among groups, 8.1% within groups, and 78.2% within populations (see Table 1 for population information). Fixation indices (Φ_{SC}=0.09337, P>0.001; Φ_{ST}=0.21781, P>0.001; Φ_{CT}=0.13726, P>0.001)

WWP distribution. The level of genetic differentiation in WWP across this region is high for a conifer species. These high levels of genetic diversity may be associated with refugial populations or populations that are closer to the evolutionary origin of the species.

The population structure found in WWP, with distinct northern and southern clades, is concordant to several other plant taxa distributed in this region (Fig. 3). For example, Soltis et al. (1997) studied the chloroplast (cp)DNA haplotypes of several temperate plant taxa distributed in the Pacific Northwest of the USA and Canada. Similar to WWP, the discontinuities in cpDNA haplotypes typically occurred in southern or central Oregon and were attributed to influences resulting from Pleistocence glaciation (Soltis et al. 1997). Moreover, a recent study of sugar pine (*Pinus lambertiana*) found an abrupt discontinuity in cpDNA haplotypes between the Cascade Range and Sierra Mountains in northern California (Liston et al. 2007).

Patterns of genetic structure have also been observed in other continents. Historic climatic processes in western Europe have had analogous impacts on genetic structure of plant species. When genetic studies from a number of temperate plant taxa were synthesized, some general trends were noted for genetic processes caused by past climate change. First, the highest genetic diversity was most often found in the lower to mid-latitudes of species, and reduced genetic diversity was typically found near the leading edges (e.g., high latitudes) of the species ranges. Second, genetic

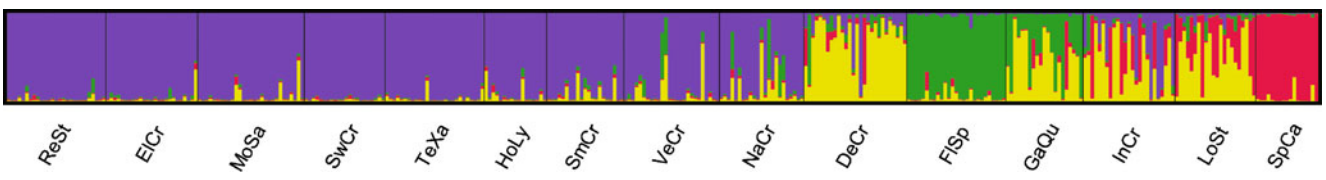

Fig. 6 Assignment probability averaged for each individual assignment using STRUCTURE version 2.2 (Pritchard et al. 2000). This result is based on a model of four genetic clusters (K=4), shown in different colors, that had the highest probability (0.99; see Table 1 for population information)

structure was highest between populations near the rear edges of the species ranges (Petit et al. 2003; Hampe and Petit 2005). Both principles fit the present data for WWP (Table 1, Fig. 2). Several hypotheses based on historic climatic fluctuations and glaciations have been proposed to explain geographic distributions of species and genetic structure (Soltis et al. 1997; Brunsfeld et al. 2001; Mehes et al. 2007, 2009; Richardson et al. 2002). The "clinal environment hypothesis" appears to a good fit for the life history and distribution WWP. This hypothesis assumes that widely adapted species like WWP were not limited to glacial refugia especially in the mid-latitudes of its range (e.g., the southern Cascades). Under such a scenario, frequencies in neutral markers should change gradually across geographic distance (Brunsfeld et al. 2001). Based on the results of the Mantel test, this hypothesis is concordant with WWP populations north of 45°N latitude. However, the genetic discontinuity in the southern Cascade Range does not fit this hypothesis. For example, the geographic distance between *NaCr* and *DeCr* is approximately 40 km with an F_{ST} of 0.099. This F_{ST} value is greater than many pairwise comparisons between *MoSa* and northern Cascade populations (e.g., *SwCr* and *SmCr*) that are geographically separated by more than 370 km. This genetic discontinuity across a small geographic region would invoke a hypothesis that previously isolated populations have only recently come into contact. However, such hypotheses are not supported by the paleoecological record. Based on lake sediment cores, forest types that support the presence of WWP were found west of the Cascade crest in central Oregon (Grigg and Whitlock 1998; MacDonald et al. 1998) and in northern California during the late-glacial period, ca. 17,000 to 11,000 years B.P. (Mohr et al. 2000). Thus, the paleoecological evidence suggests that WWP from both the northern and southern clade has been present in this region at least since the late Pleistocene. Other patterns of genetic differentiation are concordant with the paleoecological record. For example, populations at the northeastern edge of WWP distribution, *ReSt* and *ElCr*, are more divergent among other northern populations. The genetic distance dendrogram and multidimensional scaling suggest that these two populations are more closely related and separated from the remaining northern populations (Figs. 3 and 4). Fossil records near Glacier National Park in northern Montana suggest WWP was represented in a glacial refugium (MacDonald et al. 1998). Such a refugium was likely isolated at times from other populations to the west. Two other northern populations, *MoSa* and *TeXa*, also genetically differentiated from each other and the remaining northern populations (Fig. 4). These two WWP populations are perhaps derived from northern glacial refugia – "Clearwater" (northern Idaho, USA) (Brunsfeld and Sullivan 2005) for *MoSa* and west coast islands such as "Haida Gwaii (formerly known as the Queen Charlotte Islands)" and/or "the Brooks Peninsula" (British Columbia, Canada) (Hebda and Haggarty 1997; Swenson and Howard 2005) for *TeXa*. It is important to identify and prioritize populations for species conservation, and these populations may be derived from refugial areas.

Disease pressure from white pine blister rust could also influence the genetic diversity of WWP populations. For example, Kim et al. (2003) showed that WWP stands under higher disease pressure had lower genetic diversity than stands with low disease pressure in northern Idaho. While genetic diversity could be reduced locally under high blister rust pressure, the collection sites (e.g., provenances) included in this study have experienced relatively low levels of disease-related mortality, which should have only limited impact on landscape-level genetic diversity.

The present AFLP marker data demonstrate that populations from the southern Cascades, Siskiyou, and Sierra Mountains exhibited a higher level of genetic diversity than northern populations. Furthermore, southern populations have more genetic structure than northern populations despite having shorter geographic distances among populations and small geographic areas. Most of these observations can be attributed to changes in geographic distribution caused by past changes in climate and glaciation followed by an expansion of WWP distribution at higher latitudes (Mehes et al. 2009). However, large F_{ST} values between *DeCr* and *NaCr* are highly unusual for a conifer species. A parallel study was conducted to (1) better understand the processes involved in genetic discontinuity across the southern Cascade crest, the northern Rocky Mountains, and the Pacific Northwest; and (2) determine why some putatively neutral markers from this study are distributed in a manner analogous to quantitative traits (Richardson et al. 2009). Studies using DNA sequencing (nuclear and organelle genes) to examine more natural WWP populations are necessary to better assess evolutionary relationships among southern and northern populations and determine the effective population sizes of WWP. In addition, it is also important to evaluate how populations are responding under climate changes. Currently, a modeling approach is available to predict current and future distribution of forest trees using climate variables (Rehfeldt et al. 2006). Modeling of past climate and considerations of phylogeography with glacial refugia and adaptation are needed to support hypothetical population structure and responses to climate change and ecosystem reconstruction (Carstens and Richards 2007; Waltari et al. 2007). The results from the present study can provide a basis for identifying and selecting WWP populations for species management, conservation, and restoration. Furthermore,

continued studies can identify the influence of climate in shaping these populations and develop strategies to manage WWP under climate-change scenarios.

Acknowledgments This work was supported by the USDA Forest Service—Rocky Mountain Research Station, Forest and Woodland Ecosystems Program in Moscow, ID, USA, Research Joint Venture Agreements (01-JV-11222062-275 and 07-JV-11221662-078), and the new faculty research program 2009 of Kookmin University in Korea. We thank Drs. John King and Mike Carlson for providing sample collections of western white pine for Canadian populations. We also thank Dr. Seok-Woo Lee, Dr. Bum-Yong Kang, and Jonalea R. Tonn for their comments on an earlier version of this manuscript and Jane E. Stewart and John W. Hanna for their technical assistance. Use of trade names does not constitute endorsement by the USDA Forest Service.

References

Bingham RT (1983) Blister rust resistant western white pine for the Inland Empire: the story of the first 25 years of the research and development program. US Dept. Agric. For. Serv. Gen Tech Rep INT-146 Intermountain Forest and Range Experiment Station, Ogden, UT, 45 p

Bingham RT, Hoff RJ, Steinhoff RJ, United States, Forest Service (1974) Genetics of western white pine. US Dept. Agric. For. Serv. Res. Pap. WO-12. USDA Forest Service, Washington, D.C, 28 p

Bonin A, Bellemain B, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

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

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. In: Silvertown J, Antonovics J (eds) Integrating ecology and evolution in a spatial context. Blackwell Publishing, Williston, pp 319–339

Campbell RK, Sugano AI (1989) Seed zones and breeding zones for white pine in the Cascade Range of Washington and Oregon. USDA For. Serv. Res. Pap. PNW-RP-407

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

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620

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

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package). Distributed by the author (http://evolution.gs.washington.edu/phylip.html)

Franklin JF, Dyrness CT (1973) Natural vegetation of Oregon and Washington. USDA For. Serv. Gen Tech Rep PNW-8 417

Griffin JR, Critchfield WB (1972) The distribution of forest trees in California. USDA Forest Service Res. Paper PSW-82. Pacific Southwest Forest and Range Experiment Station, Berkeley, California

Grigg LD, Whitlock C (1998) Late-glacial vegetation and climate change in Western Oregon. Quatern Res 49:287–298

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467

Hamrick JL, Godt MJ (1996) Conservation genetics of endemic plant species. In: Avise JC, Hamrick JL (eds) Conservation genetics: case histories from nature. Chapman & Hall, New York, pp 281–333

Hebda RJ, Haggarty JC (1997) Brooks Peninsula: an ice age refugium on Vancouver Island. Occasional Paper 5. British Columbia Ministry of Environment, Lands and Parks: Victoria, British Columbia

Johansen AD, Latta RG (2003) Mitochondrial haplotype distribution, seed dispersal and patterns of post glacial expansion of ponderosa pine. Mol Ecol 12:293–298

Kim MS, Brunsfeld SJ, McDonald GI, Klopfenstein NB (2003) Effect of white pine blister rust (*Cronartium ribicola*) and rust-resistance breeding on genetic variation in western white pine (*Pinus monticola*). Theor Appl Genet 106:1004–1010

Liston A, Parker-Defeniks M, Syring JV, Willyard A, Cronn R (2007) Interspecific phylogenetic analysis enhances intraspecific phylogeographical inference: a case study in *Pinus lambertiana*. Mol Ecol 16:3926–3937

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecol 3:91–99

MacDonald GM, Cwynar LC, Whitlock C (1998) The late Quaternary dynamics of pines in northern North America. In: Richardson DM (ed) Ecology and biogeography of *Pinus*. Cambridge University Press, New York, pp 122–136

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209

McDonald G, Zambino P, Sniezko R (2004) Breeding rust-resistant five-needle pines in the western United States: lessons from the past and a look to the future. In: Sniezko RA, Samman S, Schlarbaum SE, Kriebel HB (eds) Breeding and Genetic Resources of Five-Needle Pines: Growth, Adaptability, and Pest Resistance. Proceedings RMRS-P-31. US Dept. Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, pp 28–50

Mehes MS, Nkongolo KK, Michael P (2007) Genetic analysis of *Pinus strobes* and *Pinus monticola* populations from Canada using ISSR and RAPD markers: development of genome-specific SCAR markers. Pl Syst Evol 267:47–63

Mehes M, Nkongolo KK, Michael P (2009) Assessing genetic diversity and structure of fragmented populations of eastern white pine (*Pinus strobes*) and western white pine (*P. monticola*) for conservation management. J Plant Ecol 2:143–151

Miller MP (1997) Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Distributed by the author (http://www.marksgeneticsoftware.net/tfpga.htm)

Mohr JA, Whitlock C, Skinner CN (2000) Postglacial vegetation and fire history, eastern Klamath Mountains, California, USA. Holocene 10:587–601

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Neuenschwander LF, Byler JW, Harvey AE, McDonald GI, Ortiz DS, Osborne HL, Snyder GC, Zack, A (1999) White Pine in the American West: A Vanishing Species – Can We Save it? Gen. Tech. Report RMRS-GTR-35. US Dept. Agriculture, Forest Service, Rocky Mountain Research Station and University of Idaho, Fort Collins, CO, 20 p.

Newton AC, Allnutt TR, Gillies ACM, Lowe AJ, Ennos RA (1999) Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends Ecol Evol 14:140–145

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood JA et al (2001) Terrestrial ecoregions of the world: a new map of life on earth. Bioscience 51:933–938

Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R et al (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565

Pinedo-Cancino V, Sheen P, Tarazona-Santos E, Oswald WE, Jeri C, Vittor A et al (2006) Limited diversity of Anopheles darlingi in the Peruvian Amazon regions of Iquitos. Am J Trop Med Hyg 75:238–245

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rehfeldt GE, Hoff RJ, Steinhoff RJ (1984) Geographic patterns of genetic variation in *Pinus monticola*. Bot Gaz 145(2):229–239

Rehfeldt GE, Crookston NL, Warwell MV, Evans JS (2006) Empirical analyses of plant-climate relationships for the western United States. Int J Plant Sci 167:1123–1150

Remington DL, Whetten RW, Liu BH, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98:1279–1292

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*). Mol Ecol 11:215–227

Richardson BA, Rehfeldt GE, Kim M-S (2009) Congruent climate-related genecological responses from molecular markers and quantitative traits for western white pine (*Pinus monticola*). Int J Plant Sci 170:1120–1131

SAS Institute Inc (2000) SAS OnlineDoc version 8. SAS Institute Inc., Cary NC

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Syst Evol 206:353–373

SPSS Inc (1998) SPSS Base 8.0 for windows user's guide. SPSS Inc, Chicago IL

Steinhoff RJ, Joyce DJ, Fins L (1983) Isozyme variation in *Pinus monticola*. Can J For Res 13:1122–1132

Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am Nat 166:581–591

Thompson JN (2005) Coevolution: the geographic mosaic of coevolutionary arms races. Curr Biol 15:992–994

Vekemans X (2002) Distributed by the author. Universitie Libre de Bruxelles, Belguim, Laboratorie de Genetique et Ecologie Vegetale

Waltari E, Hijmans RJ, Peterson AT, Nyari AS, Perkins SL, Guralnick RP (2007) Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS ONE 2: e563. doi:10.1371/journal.pone.0000563

Walter R, Epperson BK (2001) Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. Mol Ecol 10:103–111

Wellner CA (1962) Silvics of western white pine. US Dept Agriculture Forest Service Misc Pub 26 Intermountain Forest and Range Experiment Station, Ogden, Utah

Yeh FC, Boyle T, Yang RC, Ye Zh, Mao X (1999) POPGENE. University of Alberta, Edmonton, Canada

Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. Mol Ecol 6:907–913