

**LOW LEVELS OF POPULATION GENETIC STRUCTURE IN
PINUS CONTORTA (PINACEAE) ACROSS A GEOGRAPHIC
 MOSAIC OF CO-EVOLUTION¹**

THOMAS L. PARCHMAN^{2,4}, CRAIG W. BENKMAN³, BRITTANY JENKINS²,
 AND C. ALEX BUERKLE²

²Department of Botany, Dept. 3165, 1000 E. University Ave., University of Wyoming, Laramie, Wyoming 82071 USA;
 and ³Department of Zoology and Physiology, Dept. 3166, 1000 E. University Ave., University of Wyoming, Laramie,
 Wyoming 82071 USA

- *Premise of the study:* Population genetic analyses provide information on the population context in which evolutionary processes operate and are important for understanding the evolution of geographically variable traits. Earlier studies showed that cone structure of lodgepole pine in the Rocky Mountains diverged among populations because of geographic variation in co-evolutionary interactions involving mammalian and avian seed predators. Analyses of population genetic variation are needed to determine whether this divergence has arisen despite extensive gene flow and whether populations to the east and west of the Rocky Mountains have evolved convergent phenotypes independently.
- *Methods:* We investigated genetic structuring across 22 stands of lodgepole pine in the central Rocky Mountains and in isolated peripheral populations that experience different seed predators and exhibit parallel divergence in cone traits using a set of nine simple sequence repeats and 235 AFLP loci.
- *Key results:* Our analyses reveal high levels of genetic diversity within and low genetic differentiation among populations. Nonetheless, geographic and genetic distances were correlated, and isolated populations to the east and west of the Rocky Mountains had higher levels of differentiation than did populations in the central part of the range.
- *Conclusions:* These data indicate not only that adaptive divergence of cone traits across a geographic mosaic of coevolution has occurred despite minimal genetic differentiation, but also that isolated populations to the east and west of the Rocky Mountains have evolved distinctive cones independently and in parallel. The population structure quantified here will inform future research aimed at detecting genetic variants associated with divergent adaptive traits.

Key words: AFLP; genetic structure; geographic variation; Pinaceae; *Pinus contorta*; SSR; Rocky Mountains.

The extent of gene flow among populations shapes the potential for independent population histories and adaptive evolution. Thus, knowledge of population genetic structure across the geographical range of a species provides information on the population context in which evolutionary processes operate (Avice, 2004; Hedrick, 2005). When combined with an understanding of natural selection and trait variation, studies of molecular marker variation can thus improve our understanding of how adaptive geographic variation arises and is maintained in plant populations. Accounting for population structure is also a prerequisite for research on the genetic basis of phenotypic traits in natural populations. The geographic distribution of genetic variation affects the interpretation of traits that vary geographically by allowing an assessment of the extent to which divergence has been aided or impeded by gene flow. Similarly, knowledge of population genetic structure is used for genetic mapping in natural populations to distinguish true from spurious

associations between genotype and phenotype (i.e., association genetics; Aranzana et al., 2005; Yu et al., 2006).

As a result of large population sizes, wind dispersal, and high levels of outcrossing, conifers typically harbor large amounts of genetic variation within populations and have low levels of genetic differentiation among populations (Hamrick et al., 1992; Newton et al., 1999; Savolainen and Pyhajarvi, 2007a). Despite this typical lack of neutral genetic differentiation among populations, many studies indicate that conifers are adapted to local environments and exhibit substantial geographic variation in adaptive traits (Morgenstern, 1996; Savolainen and Pyhajarvi, 2007a). Provenance trials demonstrate pronounced variation in adaptive phenotypes (Wheeler and Guries, 1982a; Xie and Ying, 1995), and geographic differentiation in morphology is often far greater than neutral genetic divergence (Wheeler, 1982b; Yang et al., 1996). High levels of phenotypic variation across populations that have limited neutral genetic differentiation indicate that conifers are readily capable of local adaptation, even in the face of extensive gene flow (Petit et al., 2004).

Lodgepole pine (*Pinus contorta*) is a dominant and widespread species occurring in the mountainous regions of western North America (Critchfield, 1980). Of its four recognized subspecies, the Rocky Mountain subspecies (*P. contorta* subsp. *latifolia*) has the most extensive range, growing throughout the Rocky Mountains and neighboring regions from the Yukon to Colorado (Fig. 1; Critchfield, 1980). Previous studies of Rocky Mountain lodgepole pine have highlighted substantial adaptive

¹Manuscript received 25 September 2010; revision accepted 3 January 2011.

The authors thank Z. Gompert, A. Siepielski, and two anonymous reviewers for valuable comments on the analyses and manuscript. This research was supported by an NSF DBI award (0701757) to C.A.B., and an NSF DEB award (0344503) to C.W.B. The authors thank the Nevada Genomics Center at the University of Nevada, Reno for facilitating rapid genotyping of AFLP markers.

⁴Author for correspondence (e-mail: tparchma@uwyo.edu)

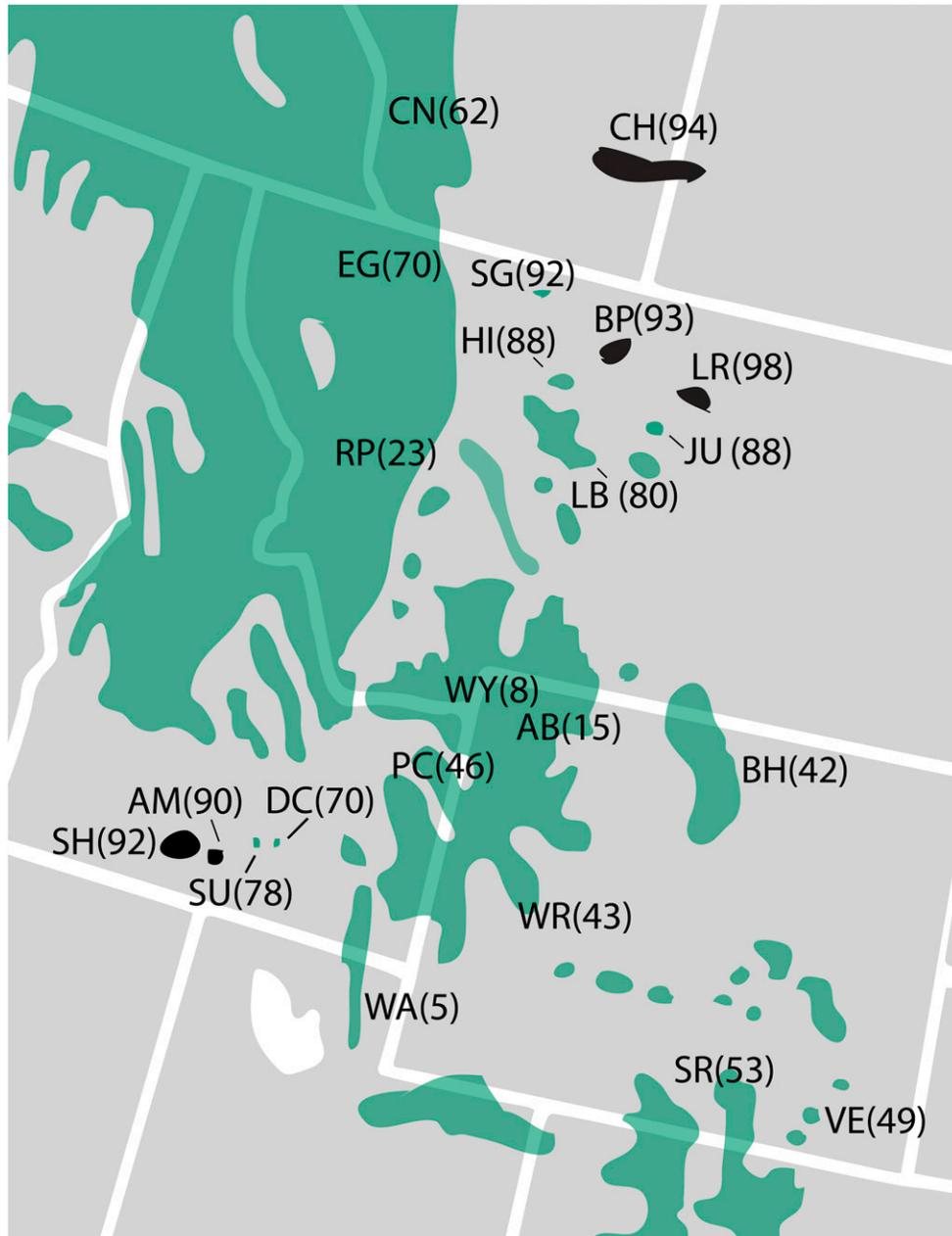


Fig. 1. The distribution of lodgepole pine in the study area. Abbreviations for locations of sampled populations are defined in Table 2, and the percentage of serotinous trees in each stand is in parentheses. Isolated populations found in previous studies to have diverged in cone morphology and serotiny (South Hills, ID; Albion Mts., ID; Bears Paw Mts., MT; Little Rocky Mts., MT; and Cypress Hills, Alberta) (Benkman, 1999; Benkman et al., 2001; Benkman and Siepielski, 2004) are filled in black.

divergence in cone structure, and found that cone traits have evolved in response to variation in the distribution of seed predators (Benkman, 1999; Benkman et al., 2001, 2003; Benkman and Siepielski, 2004; Siepielski and Benkman, 2005). Throughout much of the pine's range, red squirrels (*Tamiasciurus hudsonicus*) feed extensively on lodgepole pine seeds and are dominant and preemptive competitors for these seeds (Smith, 1970; Benkman, 1999; Benkman et al., 2001, 2003). In the portion of the geographic range in which red squirrels are present, lodgepole pine has evolved cone traits that defend seeds against squirrels, including a reduction in the ratio of seed mass to cone

mass (Smith, 1970; Benkman et al., 2001). However, in isolated ranges east and west of the Rocky Mountains where red squirrels are absent, lodgepole pine has evolved a higher ratio of seed mass to cone mass in response to relaxation of selection by red squirrels (Benkman et al., 2001). In addition, in these regions without squirrels, lodgepole pine has evolved larger cones with thicker distal scales in response to increases in seed predation and selection exerted by red crossbills (*Loxia curvirostra*) and crossbills have reciprocally evolved deeper bills (Benkman, 1999; Benkman et al., 2001, 2003). Divergence resulting from predator-prey arms races has occurred in a parallel

and replicated fashion in isolated ranges lacking squirrels to the east and west of the Rocky Mountains (CH and SH/AM in Fig. 1), leading to a geographic mosaic of coevolution (Thompson, 2005) that has driven geographic divergence in cone morphology (Benkman et al., 2001, 2003).

Populations of lodgepole pine across this region exhibit substantial geographic variation in the percentage of serotinous trees in a stand (Lotan, 1975; Tinker et al., 1994; Benkman and Siepielski, 2004). Individual trees largely produce either serotinous or nonserotinous cones (Teich, 1970; Lotan, 1975), and most stands of lodgepole pine contain a mix of such trees. Serotinous cones remain attached to branches and can hold viable seeds for decades (Critchfield, 1980), until fire heats and breaks resinous bonds between the cone scales and causes the release of seeds to the forest floor. Serotiny is favored by fire (Enright et al., 1998), and the percentage of serotinous trees within stands of several species of pines is positively correlated with the frequency of stand-replacing fires (Givnish, 1981; Muir and Lotan, 1985; Gauthier et al., 1996). However, predation by red squirrels reduces the accumulation of seeds in a canopy seed bank and thereby removes the benefit of serotiny (Benkman and Siepielski, 2004). The presence of red squirrels therefore should select against serotinous individuals. In areas with red squirrels, stands of lodgepole pine are generally less than 50% serotinous, whereas stands in ranges lacking squirrels are more than 90% serotinous (Fig. 1; Benkman and Siepielski, 2004). Thus, diverse selection pressures from the distribution of seed predators, fire, and perhaps other factors have given rise to substantial geographic variability in this trait.

An understanding of population genetic structuring across this region is needed to determine whether the phenotypic differentiation described has arisen despite extensive gene flow and whether populations to the east and west of the Rocky Mountains have evolved convergent phenotypes independently. Previous studies of population genetic structuring of lodgepole pine in the Rocky Mountain region have focused mostly on the part of the range north of Montana (Wheeler and Guries, 1982a; Marshall et al., 2002; Godbout et al., 2008) and have included only one of the squirrel-less populations described (Cypress Hills). Much of the current distribution of lodgepole pine is thought to have resulted from expansion and northward migration of populations over the last 12 000 years following the last glacial retreat (MacDonald and Cwynar, 1985, 1991; Marshall et al., 2002; Godbout et al., 2008). Reflecting this migrational history, results of studies of genetic variation in *P. c. latifolia* based on allozymes (Wheeler and Guries, 1982a; Dancik and Yeh, 1983; Yeh et al., 1985; Epperson and Allard, 1989), mtDNA (Dong and Wagner, 1994; Godbout et al., 2008), and cpDNA SSR loci (Marshall et al., 2002) have shown high diversity within populations but minor, if any, differentiation among populations. However, these studies have used individual organellar markers or small sets of allozymes and have focused on the northern part of the range (mostly in Canada), where occupancy is recent (last 12 000 years). The portion of the range south of Canada (and where geographic variation in cone traits is evident) is characterized by a more fragmented distribution (Fig. 1) that has been less altered by glacial activity than the northern range (Dong and Wagner, 1994). In addition, some populations along the eastern and western periphery of the geographic range are especially isolated from the central Rocky Mountains (Fig. 1) and undergo divergent selection from different seed predators (Benkman et al., 2001, 2003; Benkman and Siepielski, 2004), and the level of population structure characterizing this region is unknown.

Here we surveyed genetic variation at nine nuclear simple sequence repeat (SSR) loci and 235 AFLP markers across 22 lodgepole pine populations that span the central and southern portion of the Rocky Mountains, including isolated populations occurring to the east and west that have diverged in cone structure and serotiny frequency. We used this combination of marker types because nuclear SSRs are typically highly variable and informative and because we could readily assay large numbers of AFLP markers and thereby base analyses of population structure on a larger portion of the genome. A population genetic survey based on large numbers of polymorphic markers should provide a context for understanding the adaptive phenotypic divergence in cone traits. A second motivation for this research is that a thorough understanding of population structure is a prerequisite for association genetics, in which future studies will map variation in cone serotiny and defense traits to variable genetic loci.

MATERIALS AND METHODS

Genetic resources and molecular markers—We sampled needles from a minimum of 18 individual trees from each of 22 populations ranging from southern Wyoming to southern Alberta (Fig. 1). Dessicated needles (50 mg) served as a source of DNA, which was extracted using a modification of the cetyltrimethyl ammonium bromide (CTAB) method (Doyle, 1991). We quantified the amount of extracted DNA using a Nanodrop spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA) and verified the presence of high molecular weight genomic DNA by electrophoresis in 1.5% agarose gels and staining and visualization with ethidium bromide.

We used the polymerase chain reaction (PCR) to amplify a set of nine previously developed nuclear SSR loci (Liewlaksaneeyanawin et al., 2004) and determined the genotypes at these loci for 423 individuals from the 22 populations (Table 1). Each reaction consisted of 50–100 ng total genomic DNA, 2 pmol of each primer, 0.5 mM each of dATP, dCTP, dGTP, and dTTP, 1× PCR buffer, and 0.4 units of *Taq* polymerase. All PCR amplifications were performed with the following conditions: 94°C for 5 min, followed by 32 cycles of 94°C for 1 min, annealing temperature for 1 min (which varied, see Liewlaksaneeyanawin et al., 2004), and 72°C for 1 min, followed by a final extension step of 72°C for 3 min. Forward primers were labeled with various fluorescent dyes, and fragments were detected via capillary electrophoresis using an ABI 3130 genetic analyzer (ABI, Foster City, California, USA). We assessed the presence of fragments representing different alleles using the program Genemapper (ABI).

We generated AFLP markers based on a modified version of the Vos et al. (1995) protocol. Preselective and selective primers were modified and made more specific to accommodate the large genome size of lodgepole pine, following Remington et al. (1999). Restriction digestion and adaptor-ligation were carried out simultaneously on 0.5 µg of genomic DNA using the restriction

TABLE 1. Characteristics of SSR loci used in the study. Locus names correspond to those in Liewlaksaneeyanawin et al. (2004). Expected heterozygosity (H_e) is the estimated fraction of individuals that would be heterozygous for a given locus under Hardy–Weinberg equilibrium with data pooled across all populations.

Locus	H_e	H_o	No. of alleles	Average allele size	Maximum allele size
<i>Lop1</i>	0.331	0.344	14	156	174
<i>Lop5</i>	0.91	0.838	28	180	211
<i>PiTX2123</i>	0.57	0.575	6	200	206
<i>PiTX3011</i>	0.956	0.72	48	182	226
<i>PiTX3034</i>	0.838	0.637	20	209	233
<i>PiTX3049</i>	0.901	0.604	30	315	340
<i>PiTX3052</i>	0.598	0.592	14	244	267
<i>PiTX3107</i>	0.797	0.509	21	167	199
<i>PiTX3127</i>	0.651	0.604	13	178	199

Note: H_o = observed heterozygosity

endonucleases *EcoRI* and *MseI* (NEB, Ipswich, Massachusetts). AFLP adaptor pairs were attached to digested fragments using *T4* DNA ligase (NEB). We performed restriction and ligation reactions simultaneously in 11 μ L volumes that were incubated for 18 h at 38°C. After incubation, we diluted these reactions with 170 μ L 0.1 \times Tris-EDTA (TE) buffer. Preselective and selective primers were based on primer core sequences *EcoRI* 5'-GACTGCGTACCA-ATTC-3' and *MseI* 5'-GATGAGTCTGAGTAA-3' (*EcoRI* and *MseI* hereafter). Preselective amplification reactions contained 4 μ L of the diluted restriction-ligation products, 15 μ L PCR core mix (385 μ L H₂O, 68 μ L 10 \times PCR buffer, 41 μ L 25 mM MgCl₂, and 6.8 μ L 40 mM dNTPs for a total volume of 500 μ L), and 1 μ L preselective primers, which each consisted of the adaptor primer sequences with one additional nucleotide at the 3' ends (*EcoRI-A* and *MseI-C*). Preselective PCR conditions included 20 PCR cycles (94°C for 30 s, 56°C for 1 min, 72°C for 2 min), and a final extension at 60°C for 30 min. We diluted preselective amplification products with 170 μ L 0.1 \times Tris-EDTA (TE). Selective amplification reactions contained 3 μ L of diluted preselective amplification product, 15 μ L standard PCR core mix, 1 μ L of selective *MseI* primer, and 1 μ L of the fluorescently labeled *EcoRI* selective primer. Both *EcoRI* and *MseI* selective amplification primers had three extra nucleotides at the 3' ends to produce an appropriate number of amplified fragments.

We used three selective primer combinations (*EcoRI*-ACT and *MseI*-CTA; *EcoRI*-AAC and *MseI*-CTT; *EcoRI*-AAT and *MseI*-CAT) to generate AFLP fragments. We combined 1 μ L of each selective amplification product with 8.75 μ L formamide and 0.45 μ L GeneScan 500 ROX labeled size standard (ABI) and resolved products on an ABI 3730 capillary sequencer. The presence or absence of AFLP fragments in each lane file was analyzed using the program Genemapper (ABI). We scored fragments generated by each selective primer combination as dominant marker loci (i.e., as present or absent). We limited analyses to fragment sizes between 70 and 350 bp, only scored unambiguously discernible fragments, and genotyped a total of 536 individuals spanning the 22 populations (compared to 423 individuals with SSR data).

Data analysis—For the SSR loci, we calculated descriptive statistics for each population using microsatellite analyzer (MSA version 4.05; Dieringer and Schlötterer, 2003), which included observed and expected heterozygosities (Nei, 1987) and the mean number of alleles per locus. Prior to further analyses, we used the program GenePop (Raymond and Rousset, 1995) to check for violations of Hardy-Weinberg equilibrium for each SSR locus across all populations. To quantify genetic differentiation among populations based on the nine SSR loci, we used the program SMOGD (Crawford, 2010) to calculate G_{ST} , G_{ST}' and Jost's D for each population, and used MSA to calculate θ as in Weir and Cockerham (1984). G_{ST}' and Jost's D are not dependent on heterozygosity and are potentially more appropriate measures of genetic differentiation given loci with high diversity and numerous allelic states (Hedrick, 2005; Jost, 2008). We calculated basic descriptive statistics for the AFLP data, including heterozygosity and the proportion of polymorphic loci, for each population using the program AFLP-SURV (Vekemans, 2002). We analyzed overall population structure with the method of Lynch and Milligan (1994) and the Bayesian method of Holsinger et al. (2002), and estimated pairwise values of F_{ST} between different populations by the Bayesian method of Zhivotovsky (1999) using AFLP-SURV (Vekemans, 2002). Because of major differences in the characteristics of SSRs and AFLPs and the processes affecting their diversity, we analyze and interpret the AFLP and SSR data separately except where noted otherwise.

We used the program STRUCTURE (version 2.2; Pritchard et al., 2000; Falush et al., 2003, 2007) to assess whether the sampled genotypes were consistent with a single or multiple (K) populations, each in Hardy-Weinberg equilibrium. For these analyses, we used the combined data set of nine SSR and 235 AFLP loci. Log-likelihoods from Markov chain Monte Carlo (MCMC) sampling provide the basis for evaluating the number of clusters that best fit the data. We also used STRUCTURE to assign genotypic proportions of each individual to the K clusters based on admixture coefficients. We used the admixture model and ran simulations for 12 replicates for each value of K ranging from one to 15. We used a burn-in of 50000 steps and ran the MCMC for 150000 steps for each round of simulations. The number of populations (K) that was most likely given the data were determined using the method of Pritchard et al. (2000). We also calculated the change in model likelihoods between successive values of K (ΔK) and used the method of Evanno et al. (2005) to infer K .

We obtained 1000 bootstrapped matrices of estimates of Nei's D for pairs of populations for the SSR data using the program FSTAT (Goudet, 1995). Similarly, we created 1000 bootstrapped matrices of Nei's D for the AFLP data set using AFLP-SURV. Neighbor-joining trees were then constructed separately for the SSR and AFLP data sets based on the 1000 bootstrapped distance matrices

using the neighbor routine in the program PHYLIP (version 3.6; Felsenstein, 2005). We used PHYLIP to construct consensus trees based on the 1000 trees generated as described, and the number of trees out of 1000 sharing nodes were used as bootstrap support values. We used correlations and Mantel tests to examine the relationship between genetic differentiation and geographic distance for both the SSR and AFLP data. Geographic coordinates of all populations were recorded and pairwise geographic distances between all sets of populations were obtained using the point-distance tool in the program ArcMap version 9.3 (RockWare, Golden, Colorado, USA). We compared geographic distances to estimates of $\theta/(1-\theta)$ for the SSR data, and $F_{ST}/(1-F_{ST})$ for the AFLP data because these adjusted values are expected to vary linearly with geographic distances and are more appropriate for comparison (Rousset, 1997).

RESULTS

All nine of the nuclear SSR loci were polymorphic and highly heterozygous, with the number of alleles per locus ranging from 6 to 48 (Table 1). Pooled across all populations, some loci deviated from Hardy-Weinberg equilibrium; however, such departures were uncommon for individual SSR loci within populations and no tests were globally significant across populations. The mean number of alleles per locus was similar across populations, with no populations having significantly lower allelic diversity than the average (Table 2). Of the 235 AFLP loci we scored, 96% were also polymorphic. As indicated by estimates of heterozygosity, genetic diversity within populations was generally high and similar across populations for both the SSR and AFLP data sets (Table 2). As expected, per locus heterozygosity estimates were higher for the multiallelic SSR data (ranging from 0.543 to 0.708) than for the dominant, biallelic AFLP data (ranging from 0.264 to 0.306; Table 2).

Clustering analyses run in STRUCTURE based on the combined SSR and AFLP data sets found the highest log-likelihood support for three genotypic clusters in the population genetic data ($K = 3$; Fig. 2A), and the ΔK method of Evanno et al. (2005) resulted in support for the same value of K (Fig. 2B). Nonetheless, overall F_{ST} among the clusters in the $K = 3$ model was very low (0.013), and assignment of individuals to these three genotypic clusters was not consistent with population or geographic origin. Nearly all individuals had partial assignments to all three clusters, and the credible intervals on assignments (admixture coefficients) were very large (i.e., many ranged from 0 to 1). These analyses suggest an absence of pronounced population structure.

Estimates of genetic differentiation among populations were low for both the SSR and AFLP data sets. However, differences in the nature of these markers led to different information content and values of differentiation estimates, and there was substantial variation among loci and populations in the level of differentiation (Figs. 3, 4). Population pairwise and overall estimates of genetic differentiation for the SSR data were low, and many were statistically indistinguishable from zero (θ , mean = 0.016, range = 0.00–0.058; G_{ST}' mean = 0.111, range = 0.01–0.29; and Jost's D , mean = 0.104, range = 0.01–0.26; Appendix S1, see Supplemental Data with online version of this article). Estimates of G_{ST}' and Jost's D were considerably higher than θ and G_{ST} , perhaps suggesting that estimates based on heterozygosity (θ and G_{ST}) do not adequately measure differentiation for highly diverse SSR loci (Hedrick, 2005; Jost, 2008). Nonetheless, the two standardized estimates of genetic differentiation among populations for the SSR data (G_{ST}' and Jost's D) were strongly correlated with θ and result in similar relative rankings for population differentiation. Analyses of the AFLP

TABLE 2. The number of individuals genotyped in each population for nine SSR loci (n), expected (H_e) and observed heterozygosities (H_o), and the mean number of SSR alleles per locus. Following is the number of individuals genotyped across 235 AFLP loci (N), expected heterozygosity, and the percentage of polymorphic loci (PLP). Abbreviations after location names correspond to the labels in Fig. 1.

Population, State/Province (Abbreviation)	SSR			Mean no. alleles	AFLP		
	N	H_e	H_o		n	H_e	PLP
Absaroka Mts., WY (AB)	21	0.702	0.616	8.78	42	0.281	84.3
Albion Mts., ID (AM)	14	0.709	0.648	6.67	36	0.268	83
Bears Paw Mts., MT (BP)	24	0.723	0.638	11.33	24	0.288	87.7
Big Horn Mts., WY (BH)	14	0.712	0.618	7	19	0.277	82.1
Crow's Nest Pass, AB (CN)	21	0.716	0.566	9.56	25	0.28	82.1
Cypress Hills, AB (CH)	20	0.746	0.609	9.11	20	0.285	87.2
Deep Creek Range, ID (DC)	23	0.695	0.624	7.78	22	0.265	84.3
East Glacier, MT (EG)	19	0.746	0.652	8.56	19	0.294	85.5
Highwood Mts., MT (HI)	21	0.732	0.622	9.11	19	0.3	85.1
Judith Mts., MT (JU)	20	0.744	0.577	8.89	20	0.306	88.5
Little Belt Mts., MT (LB)	25	0.731	0.6	9	17	0.305	85.5
Little Rocky Mts., MT (LR)	21	0.74	0.655	9	24	0.277	89.4
Pine Creek Pass, ID (PC)	17	0.706	0.543	8.11	17	0.293	83
Roger's Pass, MT (RP)	16	0.753	0.708	9.11	23	0.291	87.7
Snowy Range, WY (SR)	16	0.707	0.594	7.33	17	0.282	81.3
South Hills, ID (SH)	24	0.652	0.57	8.11	56	0.264	81.7
Sublett Range, ID (SU)	21	0.668	0.565	7.44	29	0.264	79.6
Sweetgrass Hills, MT (SG)	10	0.685	0.607	6.56	9	0.303	84.7
Wasatch Mts., UT (WA)	18	0.695	0.577	8	19	0.288	85.5
West Yellowstone, WY (WY)	19	0.754	0.631	9.44	18	0.269	77.4
Wind River Mts., WY (WR)	20	0.719	0.654	8.56	43	0.267	82.1
Vedauwoo, WY (VE)	19	0.671	0.555	8	19	0.29	83

Notes: AB = Alberta, Canada; ID = Idaho, MT = Montana, UT = Utah, WY = Wyoming, USA

data based on the method of Lynch and Milligan (1994) yielded an overall F_{ST} of 0.01, and the Bayesian method of Holsinger et al. (2002) led to F_{ST} of 0.02. Pairwise F_{ST} estimates from the AFLP data for the 22 populations were low, ranging from 0 to 0.07, but many were statistically distinguishable from zero (online Appendix S2). Although overall estimates of genetic differentiation were low, estimates for individual loci varied substantially for both the SSR and AFLP data sets, with some loci being characterized by far higher levels of genetic differentiation (Fig. 4). A wide range of differentiation estimates also characterized the pairwise comparisons between populations (Fig. 3).

Despite the low average level of genetic differentiation, there was a tendency for geographically proximate populations to be

less differentiated than were more distant populations, consistent with a pattern of isolation by distance (Fig. 5). Geographic and genetic distances were correlated for both the SSR ($\rho = 0.28$; Mantel test, $P < 0.001$) and AFLP data sets ($\rho = 0.31$, Mantel test, $P < 0.001$). In addition, neighbor-joining dendrograms constructed from pairwise estimates of Nei's D between populations generally clustered geographically proximate populations together (Fig. 6). Pairwise comparisons of populations at the opposite extremes of the sampled geographic distribution exhibited among the highest differentiation estimates (e.g., CH vs. SH, CH vs. SU; Fig. 5). For example, the average differentiation was higher for comparisons between isolated populations at the northeastern (CH, BP, LR, SG) and southwestern margins of the range (SH, AM, DC, SB) (mean $G_{ST}' = 0.17$)

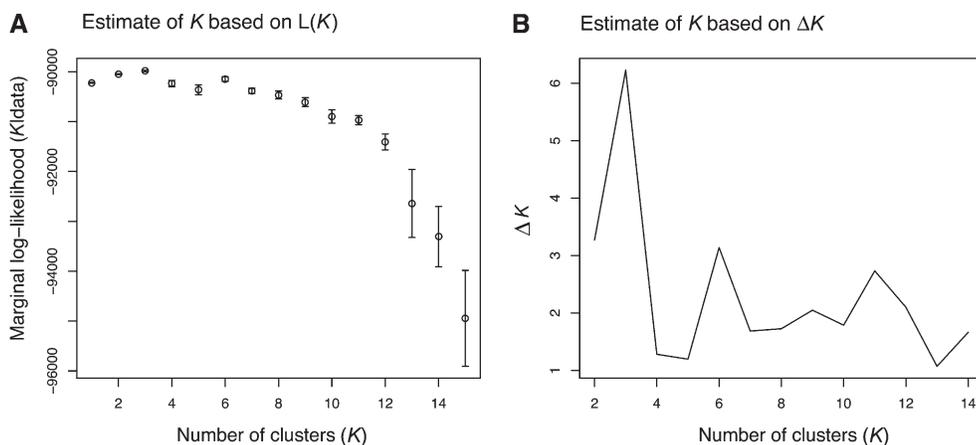


Fig. 2. Results from analyses in STRUCTURE estimating the $\ln(K|data)$. (A) Mean estimates of $\ln(K|data)$ for K ranging from one to 15 and their standard errors for 12 replicate MCMC runs at each K . (B) ΔK , as estimated by the technique of Evanno et al. (2005) across all tested values of K .

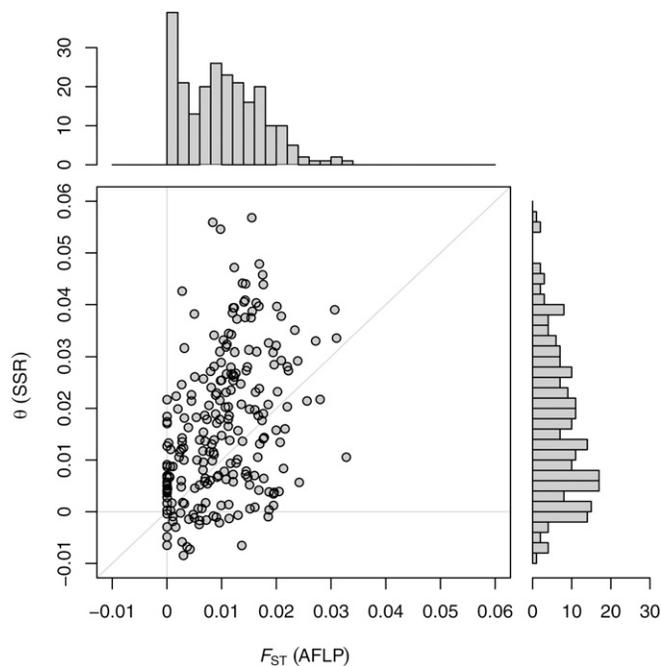


Fig. 3. Estimates of F_{ST} from the AFLP data and θ for the SSR data between all pairs of populations. Frequency distributions for estimates of F_{ST} from the AFLP data and θ for the SSR data for all pairs of populations are in the marginal histograms.

than for comparisons among populations in the central part of the range (mean $G_{ST}' = 0.06$). However, some populations in close proximity were also characterized by higher differentiation estimates (e.g., BP vs. LR, SH vs. WA; Fig. 5), indicating exceptions in the relationship between genetic and geographic distance. These higher pairwise differentiation estimates of geographically proximal populations often involved isolated peripheral populations (e.g., South Hills, Idaho and Cypress Hills, Alberta; Fig. 5).

DISCUSSION

Previous studies of substantial adaptive divergence in cone traits across a geographic mosaic coevolution motivated our assessment of genetic variation across lodgepole pine populations in the central Rocky Mountain region and in isolated peripheral populations. Consistent with studies based predominantly in the northern part of the distribution (Wheeler and Guries, 1982a; Wagner et al., 1987; Marshall et al., 2002; Godbout et al., 2008), our data revealed high levels of genetic variation within and low levels of differentiation among populations. The subtle levels of genetic differentiation indicate that substantial divergence in cone morphological traits has evolved despite ongoing gene flow and shared recent ancestry of populations. However, estimates of genetic differentiation spanned a wide range, varied across marker types and loci, and indicated that genetic variation is not entirely homogeneous across the sampled geographic region. The correlation of geographic and genetic distances between populations was positive, and the isolated peripheral populations on either side of the Rocky Mountains exhibited relatively higher levels of differentiation from other populations. Below we discuss this variation in

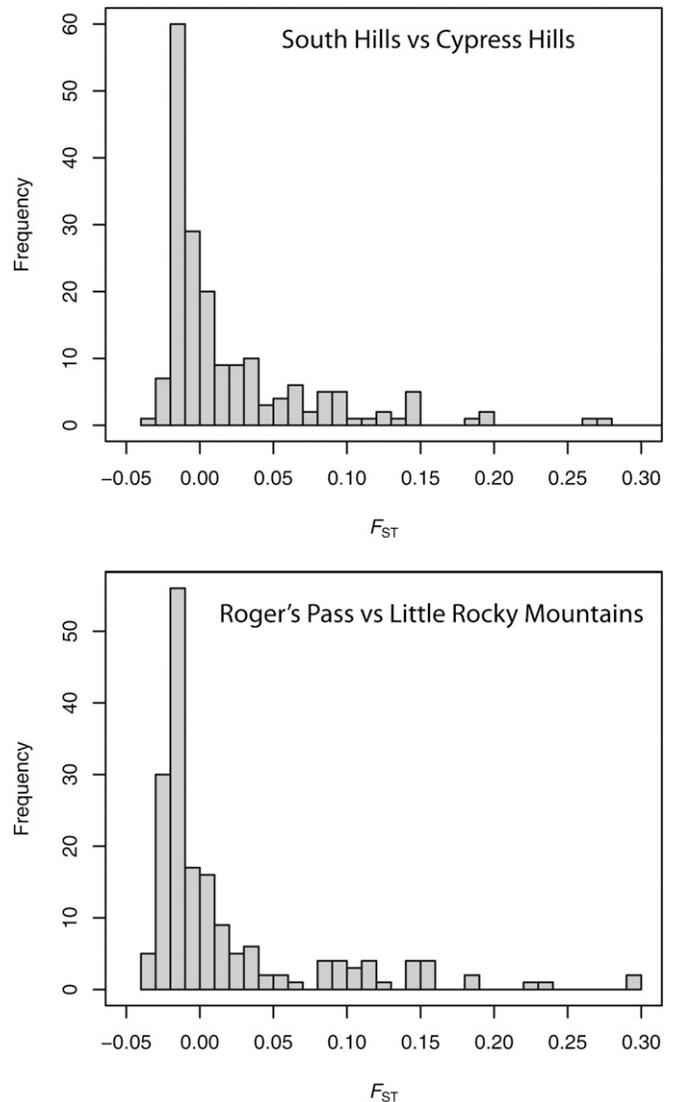


Fig. 4. Example frequency distributions of per locus F_{ST} estimates calculated for the AFLP data for two pairs of populations. Most loci are consistent with minimal population genetic differentiation, but a small percentage yield F_{ST} of 0.15 or greater. These two comparisons were chosen for illustrative purposes, but are representative of the pattern common to all comparisons.

light of substantial phenotypic divergence in cone traits and its implications for future research aimed at detecting the genetic architecture of such traits.

Population genetic structure in lodgepole pine—The SSR and AFLP data sets indicate high levels of genetic diversity within and subtle levels of differentiation between lodgepole populations in our study region. Overall estimates of population genetic differentiation for the AFLPs and the SSRs were low ($F_{ST} = 0.01$ – 0.02 , $\theta = 0.016$), indicating that the majority of genetic variation is partitioned within populations. Despite utilizing a large number of polymorphic markers, analyses conducted using the Bayesian clustering method of Pritchard et al. (2000, with extensions from Falush et al., 2003) similarly produced little evidence of clear or pronounced population

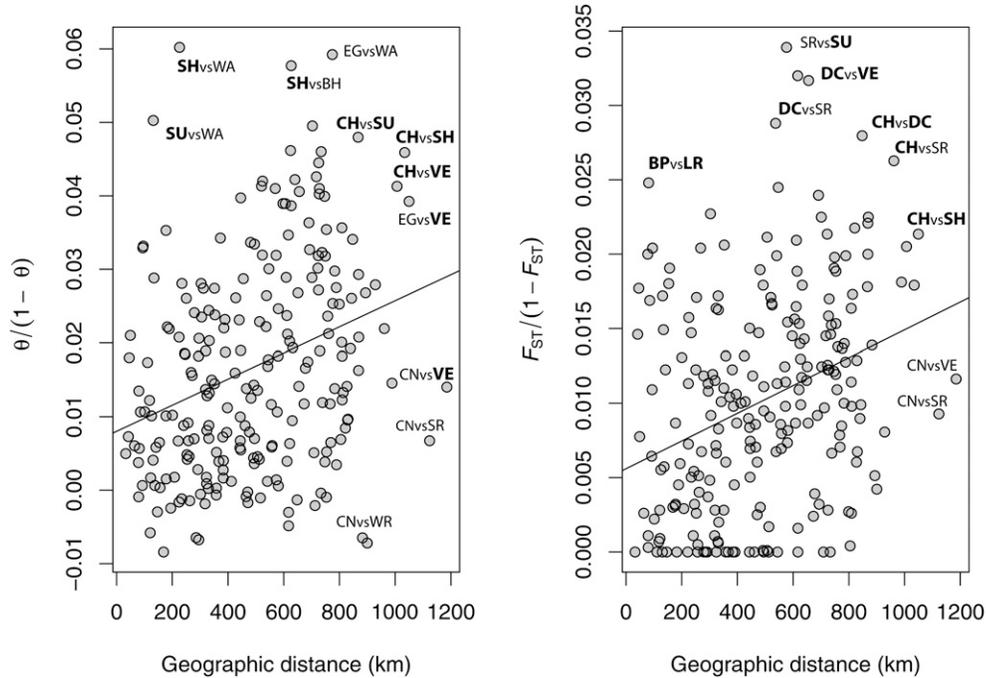


Fig. 5. Plots of geographic and genetic distances between all pairs of populations sampled and the lines of best fit. Genetic differentiation for the SSR data are based on θ and differentiation for the AFLP data are based on F_{ST} . Several pairwise comparisons exhibiting noteworthy patterns of variation are labeled with abbreviations that match those in Fig. 1. Populations indicated in bold face are relatively isolated and occur at the periphery of the distribution.

structure (Fig. 2). Given the average low levels of differentiation estimates ($F_{ST} < 0.02$), it is not surprising that STRUCTURE did not detect evidence of true genotypic clustering, as analyses of simulated data sets failed to detect the true K when overall F_{ST} was less than 0.02 (Latch et al., 2006). These results are similar to those documented in previous population genetic surveys of lodgepole pine based on the northern part of the distribution and smaller sets of markers (Wheeler and

Guries, 1982a; Yeh et al., 1985; Dong and Wagner, 1994; Marshall et al., 2002) and are consistent with other population genetic surveys of conifers, where per locus estimates of F_{ST} are typically less than 0.1 (Hamrick et al., 1992). High levels of outcrossing, wind dispersal, and large population sizes predispose these populations to harbor large amounts of genetic diversity and to exhibit low levels of geographic differentiation and structure.

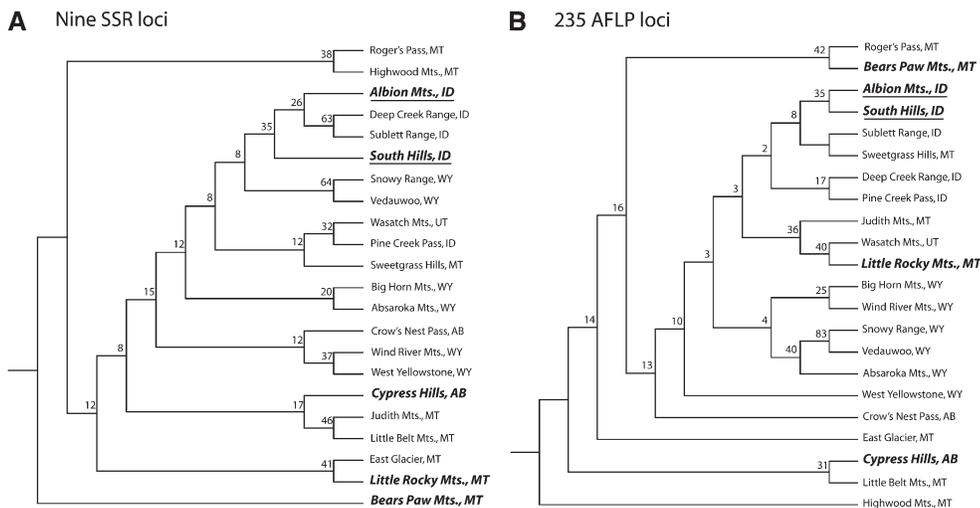


Fig. 6. Neighbor-joining trees based on pairwise estimates of Nei's (1972) genetic distance between all pairs of 22 populations for the (A) nine SSR loci, and (B) 235 AFLP loci. Values at the branches indicate the percentage of trees constructed from 1000 bootstrapped distance matrices that shared that node. Isolated populations that have diverged in cone morphology from populations in the central part of the range are indicated in bold, italicized font. Those west of the Rocky Mountains are also underlined.

Despite the lack of pronounced range-wide population structure, there was substantial variation in estimated differentiation among populations (Fig. 3) and among individual loci (Fig. 4) and marker types (Fig. 3). Pairwise population differentiation estimates spanned a wide range (Fig. 3), and those comparisons exhibiting higher levels of differentiation often involved peripheral populations (Fig. 5). Estimates of genetic differentiation based on heterozygosity (θ and F_{ST}) can underestimate levels of differentiation for markers such as SSRs that have high allelic diversity (Hedrick, 2005; Jost, 2008). Our estimates of differentiation for the SSR data based on Jost's D and G_{ST}' are substantially higher than those based on θ , suggesting these corrected estimates reveal higher levels of differentiation than those based on heterozygosity (θ and F_{ST}). With highly diverse SSR loci, the presence of rare alleles at low frequencies in different populations could contribute to this pattern. Because our sample sizes per population are small relative to the number of different alleles in populations, our estimates of Jost's D and G_{ST}' could be biased upwards. The strong correlation and similar rankings of Jost's D and G_{ST}' with θ suggest similar interpretations despite differences in the absolute numerical values of these estimates. Nonetheless, the nine SSR loci do give rise to slightly higher estimates of differentiation (based on θ) than the 235 AFLP markers (F_{ST} ; Fig. 3). Overall, the differentiation estimates and the large number of pairwise comparisons that were statistically distinguishable from zero indicate that there is subtle genetic divergence across this geographic region.

Although the average estimates of differentiation across all loci are low, some loci exhibit higher levels of differentiation (Fig. 4), which could reflect the influence of natural selection on these or closely linked loci. A common approach for detecting evidence of selection in such data sets involves simulating neutral distributions of F_{ST} to detect markers that have higher than expected F_{ST} values (Beaumont and Balding, 2004; Beaumont, 2005). Comparison of neutral distributions of F_{ST} (based on simulations of the coalescent) with estimates for the SSR and AFLP loci (conducted with the programs FDist and DFDist; Beaumont and Balding, 2004) indicated that a small number (<3% for all comparisons) of these markers fall outside of the simulated neutral distribution and thus could be influenced by selection. Because the loci detected as possibly experiencing selection were different for separate independent pairwise population comparisons and exhibited no relationship with geographic or phenotypic patterns, we do not present these results in detail here. In addition, we have concerns about the utility of AFLPs for such single-locus analyses, due to their dominant nature, the possibility of AFLP products arising from endosymbiont DNA (e.g., Gompert et al., 2010), and uncertainty about their molecular evolution (what proportion of variation corresponds to unrecognized length variation rather than the presence or absence of products?). Nonetheless, the large variance in differentiation estimates highlights that genetic differentiation is heterogeneous across the genome (Fig. 4), and suggests that genome scans based on higher quality marker or sequence data (SSRs or haplotypes) should be effective in detecting genetic regions involved in adaptive evolution (Beaumont, 2005; Nosil et al., 2009).

Pairwise geographic and genetic distances between populations for both sets of markers were correlated, indicating a pattern of isolation by distance (Fig. 5). This pattern was also evident in neighbor-joining phenograms where geographically proximate populations typically clustered together (Fig. 6). A similar relationship between genetic and geographic distances

has been documented in previous research on lodgepole pine (Marshall et al., 2002) and other conifers in western North America (Krutovsky et al., 2009) and is likely influenced by a recent northward migration and expansion occurring over the past 10 000–12 000 years following the last glacial retreat (Wheeler and Guries, 1982a; MacDonald and Cwynar, 1991; Godbout et al., 2008). Although populations have not been isolated for long enough to accumulate substantial divergence at neutral markers, divergence, and expansion has apparently resulted in an association between geographical and genetic distance.

Whereas previous population genetic surveys of lodgepole pine focused mostly on the northern part of the range (Wheeler and Guries, 1982a; Yeh et al., 1985; Dong and Wagner, 1994; Marshall et al., 2002; Godbout et al., 2008), we sampled a portion of the range where lodgepole pine has been present for longer time periods and is characterized by a more fragmented distribution (Fig. 1). During the last glacial maximum, ice covered much of the current distribution north of the Canadian border, but lodgepole pine was present to the south in areas of Idaho, Montana, and Wyoming (Baker, 1976; Mehringer et al., 1977; Beiswenger, 1991). Perhaps surprisingly, the level of genetic differentiation among the populations in our study is not more pronounced than to the north. Nonetheless, the isolated peripheral populations have slightly higher levels of genetic differentiation from other populations than the populations in the core of the range (Fig. 5). Many of these peripheral populations are isolated from the central Rocky Mountain region by large expanses of grassland or sagebrush (Fig. 1, Benkman, 1999; Edelaar and Benkman, 2006), which likely reduces gene flow. Previous studies have not included these populations, with the exception of the Cypress Hills, which have higher genetic divergence than other populations in the Rocky Mountain region based on mtDNA (Godbout et al., 2008) and allozymes (Wheeler and Guries, 1982a; Dancik and Yeh, 1983). During a warming period that ended approximately 7000 years ago, lodgepole pine likely had a wider distribution that extended to the Cypress Hills and the isolated ranges in northern Montana, after which time these forests became isolated by large expanses of grassland (Anderson et al., 1989). In southeastern Idaho, lodgepole pine is known to have been in the Albion Mountains for at least 12 000 years (Davis et al., 1986). These forests may have been connected to those in the Rocky Mountains as recently as 10 000 years ago (Wells, 1983; Beiswenger, 1991) although they have been isolated since. That genetic differentiation of these populations is not greater is potentially the result of the relatively recent isolation of lodgepole pine in these mountain ranges rather than high levels of contemporary gene flow.

The population genetic context for adaptive variation within *Pinus contorta*—Together with previous studies of morphological variation, our results indicate that the substantial phenotypic divergence in cone traits across these populations (Benkman, 1999; Benkman et al., 2001, 2003; Benkman and Siepielski, 2004; Edelaar and Benkman, 2006) has occurred within a background of limited neutral genetic divergence. This suggests that adaptive divergence has arisen despite recent shared ancestry and has been relatively unimpeded by gene flow and agrees with other studies examining the relationship between phenotypic variation and genetic structure in lodgepole pine (Wheeler, 1982b; Xie and Ying, 1995; Yang et al., 1996). The ability of forest trees in general to evolve rapidly in response to changing biotic and abiotic conditions has been widely recognized (Petit et al., 2004; Savolainen and Pyhajarvi, 2007a;

Savolainen et al., 2007b) and is likely enabled by large amounts of genetic variation resulting from high levels of outcrossing, high levels of gene flow, and large effective population sizes.

Despite subtle genetic differentiation among populations, there are patterns in these data that further bear on our understanding of geographic variation in cone traits. In particular, populations to the east and west of the Rocky Mountains that exhibit parallel divergence in cone traits are more genetically similar to populations in the middle of the range than they are to each other (Fig. 6). Trees in these populations, including the South Hills, Idaho and the Cypress Hills, Alberta have evolved larger cones with thicker distal scales as a result of coevolution with crossbills (Benkman, 1999; Benkman et al., 2001, 2003). These populations also exhibit the most extreme levels of cone serotiny, where the frequency of trees with serotinous cones approaches 100% (Fig. 1). This is in stark contrast to populations in the central part of the range, where the frequency of serotiny is consistently less than 50% (Benkman and Siepielski, 2004). Phenotypic plasticity is unlikely to contribute substantially to the divergence in cone traits among populations because cone traits in pines generally have high heritabilities (broad sense heritabilities of 0.6–0.8, Matziris, 1998; Sivacioglu and Ayan, 2010) and serotiny is thought to be under strong genetic control (Teich, 1970; Critchfield, 1980). Patterns of genetic variation presented here along with the likely biogeographical history of lodgepole pine in this region suggest that cone traits diverged in a parallel and independent fashion in the isolated ranges to the east and west of the Rocky Mountains (Benkman, 1999; Benkman et al., 2001, 2003).

Finally, an understanding of population structure and geographic variation in phenotypic traits is crucial for investigating the genetic control and architecture of traits in natural populations (González-Martínez et al., 2006; Zhao et al., 2007). If not accounted for, population structure can lead to spurious associations between genetic and phenotypic variation, particularly when populations differ in genotypic and phenotypic composition (Aranzana et al., 2005; Yu et al., 2006). The subtle levels of population differentiation across our study region mean that markers associated with phenotypic variation should be distinguishable from low levels of background neutral variation. Nonetheless, knowledge of this subtle genetic structure should still provide important information for accounting for population structure in association studies. Recent transcriptome sequencing for lodgepole pine (Parchman et al., 2010) and high-throughput sequencing assays (e.g., Hohenlohe et al., 2010) will facilitate future association and population genomic approaches involving these populations. The population genetic analyses presented here will provide a source of information for the planning and execution of such studies.

LITERATURE CITED

- ANDERSON, T. W., R. W. MATHEWES, AND C. E. SCHWEGER. 1989. Holocene climatic trends in Canada with special reference to the Hypsithermal interval. In R. J. Fulton [ed.], *Quaternary geology of Canada and Greenland*, 522–528. Geological Society of America, Boulder, Colorado, USA and Geological Survey of Canada, Ottawa, Ontario, Canada.
- ARANZANA, M. J., S. KIM, K. Y. ZHAO, E. BAKKER, M. HORTON, K. JAKOB, C. LISTER, ET AL. 2005. Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genetics* 1: e60.
- AVISE, J. C. 2004. *Molecular markers, natural history, and evolution*. Sinauer, Sunderland, Massachusetts, USA.
- BAKER, R. G. 1976. Late Quaternary history of the Yellowstone Lake Basin, Wyoming. Technical Report No. 729-E, Geological Survey Professional Paper. U.S. Government Printing Office, Washington, D.C., USA.
- BEAUMONT, M. A. 2005. Adaptation and speciation: What can F_{ST} tell us? *Trends in Ecology & Evolution* 20: 435–440.
- BEAUMONT, M. A., AND D. J. BALDING. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology* 13: 969–980.
- BEISWENGER, J. M. 1991. Late Quaternary vegetational history of Grays Lake, Idaho. *Ecological Monographs* 61: 165–182.
- BENKMAN, C. W. 1999. The selection mosaic and diversifying coevolution between crossbills and lodgepole pine. *American Naturalist* 154: S75–S91.
- BENKMAN, C. W., W. C. HOLIMON, AND J. W. SMITH. 2001. The influence of a competitor on the geographic mosaic of coevolution between crossbills and lodgepole pine. *Evolution* 55: 282–294.
- BENKMAN, C. W., T. L. PARCHMAN, A. FAVIS, AND A. M. SIEPIELSKI. 2003. Reciprocal selection causes a coevolutionary arms race between crossbills and lodgepole pine. *American Naturalist* 162: 182–194.
- BENKMAN, C. W., AND A. M. SIEPIELSKI. 2004. A keystone selective agent? Pine squirrels and the frequency of serotiny in lodgepole pine. *Ecology* 85: 2082–2087.
- CRAWFORD, N. G. 2010. SMOGD: Software for the measurement of genetic diversity. *Molecular Ecology Resources* 10: 556–557.
- CRITCHFIELD, W. B. 1980. Genetics of lodgepole pine. U. S. Forest Service Research Paper WO-37. U.S. Department of Agriculture, Forest Service, Washington, D.C., USA.
- DANCIK, B. P., AND F. C. YEH. 1983. Allozyme variability and evolution of lodgepole pine (*Pinus contorta* var. *latifolia*) and Jack pine (*Pinus banksiana*) in Alberta. *Canadian Journal of Genetics and Cytology* 25: 57–64.
- DAVIS, O. K., J. C. SHEPPARD, AND S. ROBERTSON. 1986. Contrasting climatic histories for the Snake River Plain, Idaho, resulting from multiple thermal maxima. *Quaternary Research* 26: 321–339.
- DIERINGER, D. AND C. SCHLÖTTERER. 2003. Microsatellite analyzer (MSA): A platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3: 167–169.
- DONG, J. S., AND D. B. WAGNER. 1994. Paternally inherited chloroplast polymorphism in *Pinus*—Estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics* 136: 1187–1194.
- DOYLE, J. 1991. DNA protocols for plants: A CTAB total DNA isolation. In G. M. Hewitt and A. Johnston [eds.], *Molecular techniques in taxonomy*, 283–293. Springer, Berlin, Germany.
- EDELAAR, P., AND C. W. BENKMAN. 2006. Replicated population divergence caused by localised coevolution? A test of three hypotheses in the red crossbill-lodgepole pine system. *Journal of Evolutionary Biology* 19: 1651–1659.
- ENRIGHT, N. J., R. MARSULA, B. B. LAMONT, AND C. WISSEL. 1998. The ecological significance of canopy seed storage in fire-prone environments: A model for non-sprouting shrubs. *Journal of Ecology* 86: 946–959.
- EPPELSON, B. K., AND R. W. ALLARD. 1989. Spatial auto-correlation analysis of the distribution of genotypes within populations of lodgepole pine. *Genetics* 121: 369–377.
- EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620.
- FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7: 574–578.
- FELSENSTEIN, J. 2005. PHYLIP (phylogeny inference package), version 3.6. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle, Washington, USA.
- GAUTHIER, S., Y. BERGERON, AND J. P. SIMON. 1996. Effects of fire regime on the serotiny level of Jack pine. *Journal of Ecology* 84: 539–548.

- GIVNISH, T. J. 1981. Serotiny, geography, and fire in the pine barrens of New Jersey. *Evolution* 35: 101–123.
- GODBOUT, J., A. FAZEKAS, C. H. NEWTON, F. C. YEH, AND J. BOUSQUET. 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.
- GOMPERT, Z., M. L. FORISTER, J. A. FORDYCE, C. C. NICE, R. WILLIAMSON, AND C. A. BUERKLE. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butterflies. *Molecular Ecology* 19: 2455–2473.
- GONZÁLEZ-MARTÍNEZ, S. C., K. V. KRUTOVSKY, AND D. B. NEALE. 2006. Forest-tree population genomics and adaptive evolution. *New Phytologist* 170: 227–238.
- GOUDET, J. 1995. FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity* 86: 485–486.
- HAMRICK, P. L., M. J. W. GODT, AND M. SHERMAN BROYLES. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95–124.
- HEDRICK, P. 2005. Genetics of populations, 3rd ed. Jones and Bartlett, Sudbury, Massachusetts, USA.
- HOHENLOHE, P. A., S. BASSHAM, P. D. ETTER, N. STIFFLER, E. A. JOHNSON, AND W. A. CRESKO. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* 6: e1000862.
- HOLSINGER, K. E., P. O. LEWIS, AND D. K. DEY. 2002. A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology* 11: 1157–1164.
- JOST, L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015–4026.
- KRUTOVSKY, K. V., J. B. S. CLAIR, R. SAICH, V. D. HIPKINS, AND D. B. NEALE. 2009. Estimation of population structure in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii*] using allozyme and microsatellite markers. *Tree Genetics & Genomes* 5: 641–658.
- LATCH, E. K., G. DHARMARAJAN, J. C. GLAUBITZ, AND O. E. RHODES. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* 7: 295–302.
- LIEWLAKSANEEYANAWIN, C., C. E. RITLAND, Y. A. EL-KASSABY, AND K. RITLAND. 2004. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* 109: 361–369.
- LOTAN, J. E. 1975. The role of cone serotiny in lodgepole pine forests. In D. M. Baumgartner [ed.], Symposium Proceedings, Management of Lodgepole Pine Ecosystems, 471–495, Washington State University, Pullman, Washington, USA.
- LYNCH, M., AND B. G. MILLIGAN. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3: 91–99.
- MACDONALD, G. M., AND L. C. CWYNAR. 1985. A fossil pollen based reconstruction of the late quaternary history of lodgepole pine (*Pinus contorta* ssp. *latifolia*) in the western interior of Canada. *Canadian Journal of Forest Research* 15: 1039–1044.
- MACDONALD, G. M., AND L. C. CWYNAR. 1991. Postglacial population growth rates of *Pinus contorta* ssp. *latifolia* in western Canada. *Journal of Ecology* 79: 417–429.
- MARSHALL, H. D., C. NEWTON, AND K. RITLAND. 2002. Chloroplast phylogeography and evolution of highly polymorphic microsatellites in lodgepole pine (*Pinus contorta*). *Theoretical and Applied Genetics* 104: 367–378.
- MATZIRIS, D. 1998. Genetic variation in cone and seed characteristics in a clonal seed orchard of Aleppo pine grown in Greece. *Silvae Genetica* 47: 37–41.
- MEHRINGER, P. J., S. R. ARNO, AND K. L. PETERSEN. 1977. Post glacial history of Lost Trail Pass Bog, Bitterroot Mountains, Montana. *Montana Arctic and Alpine Research* 9: 345–368.
- MORGENSTERN, E. K. 1996. Geographic variation in forest trees. UBC Press, Vancouver, British Columbia, Canada.
- MUIR, P. S., AND J. E. LOTAN. 1985. Disturbance history and serotiny of *Pinus contorta* in western Montana. *Ecology* 66: 1658–1668.
- NEI, M. 1972. Genetic distance between populations. *American Naturalist* 106: 283–292.
- NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.
- NEWTON, A. C., T. R. ALLNUTT, A. C. M. GILLIES, A. J. LOWE, AND R. A. ENNOS. 1999. Molecular phylogeography, intraspecific variation and the conservation of tree species. *Trends in Ecology & Evolution* 14: 140–145.
- NOSIL, P., D. J. FUNK, AND D. ORTIZ-BARRIENTOS. 2009. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 18: 375–402.
- PARCHMAN, T. L., K. S. GEIST, J. A. GRAHNEN, C. W. BENKMAN, AND C. A. BUERKLE. 2010. Transcriptome sequencing in an ecologically important tree species: Assembly, annotation, and marker discovery. *BMC Genomics* 11: 180.
- PETIT, R. J., R. BIALOZYT, P. GARNIER-GERE, AND A. HAMPE. 2004. Ecology and genetics of tree invasions: From recent introductions to Quaternary migrations. *Forest Ecology and Management* 197: 117–137.
- PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- RAYMOND, M., AND F. ROUSSET. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 82: 648–649.
- REMINGTON, D. L., R. W. WHETTEN, B. H. LIU, AND D. M. O'MALLEY. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theoretical and Applied Genetics* 98: 1279–1292.
- ROUSSET, F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* 145: 1219–1228.
- SAVOLAINEN, O., AND T. PYHAJARVI. 2007a. Genomic diversity in forest trees. *Current Opinion in Plant Biology* 10: 162–167.
- SAVOLAINEN, O., T. PYHAJARVI, AND T. KNURR. 2007b. Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics* 38: 595–619.
- SIEPIELSKI, A. M., AND C. W. BENKMAN. 2005. A role for habitat area in the geographic mosaic of coevolution between red crossbills and lodgepole pine. *Journal of Evolutionary Biology* 18: 1042–1049.
- SIVACIOGLU, A., AND S. AYAN. 2010. Variation in cone and seed characteristics in a clonal seed orchard of Anatolian black pine [*Pinus nigra* Arnold subsp. *pallasiana* (Lamb.) Holmboe]. *Journal of Environmental Biology* 31: 119–123.
- SMITH, C. C. 1970. Coevolution of pine squirrels (*Tamiasciurus*) and conifers. *Ecological Monographs* 40: 349–371.
- TEICH, A. H. 1970. Cone serotiny and inbreeding in natural populations of *Pinus banksiana* and *Pinus contorta*. *Canadian Journal of Botany* 48: 1805–1809.
- THOMPSON, J. N. 2005. The geographic mosaic of coevolution. University of Chicago Press, Chicago, Illinois, USA.
- TINKER, D. B., W. H. ROMME, W. W. HARGROVE, R. H. GARDNER, AND M. G. TURNER. 1994. Landscape-scale heterogeneity in lodgepole pine serotiny. *Canadian Journal of Forest Research* 24: 897–903.
- VEKEMANS, X. 2002. AFLP-SURV version 1.0. Distributed by the author, Laboratoire de Génétique et Ecologie Végétales, Université Libre de Bruxelles, Brussels, Belgium.
- VOS, P., R. HOGERS, M. BLEEKER, M. REIJNS, T. VANDELEE, M. HORNES, A. FRUITERS, ET AL. 1995. AFLP—A new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.
- WAGNER, D. B., G. R. FURNIER, M. A. SAGHAIMAROO, S. M. WILLIAMS, B. P. DANCIC, AND R. W. ALLARD. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proceedings of the National Academy of Sciences, USA* 84: 2097–2100.
- WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating *F*-statistics for the analysis of population-structure. *Evolution* 38: 1358–1370.
- WELLS, P. V. 1983. Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs* 53: 341–382.

- WHEELER, N. C., AND R. P. GURIES. 1982a. Biogeography of lodgepole pine. *Canadian Journal of Botany* 60: 1805–1814.
- WHEELER, N. C., AND R. P. GURIES. 1982b. Population structure, genetic diversity, and morphological variation in *Pinus contorta* Dougl. *Canadian Journal of Forest Research* 12: 595–606.
- XIE, C. Y., AND C. C. YING. 1995. Genetic architecture and adaptive landscape of interior lodgepole pine (*Pinus contorta* ssp. *latifolia*) in Canada. *Canadian Journal of Forest Research* 25: 2010–2021.
- YANG, R. C., F. C. YEH, AND A. D. YANCHUK. 1996. A comparison of isozyme and quantitative genetic variation in *Pinus contorta* ssp. *latifolia* by F_{ST} . *Genetics* 142: 1045–1052.
- YEH, F. C., W. M. CHELIAK, B. P. DANCİK, K. ILLINGWORTH, D. C. TRUST, AND B. A. PRYHITKA. 1985. Population differentiation in lodgepole pine, *Pinus contorta* spp. *Latifolia*—A discriminant analysis of allozyme variation. *Canadian Journal of Genetics and Cytology* 27: 210–218.
- YU, J., W. PRESSOIR, W. BRIGGS, I. BI, M. YAMASAKI, J. DOEBLEY, M. McMULLEN, ET AL. 2005. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203–208.
- ZHAO, K., M. J. ARANZANA, S. KIM, C. LISTER, C. SHINDO, C. TANG, C. TOOMAJIAN, ET AL. 2007. An *Arabidopsis* example of association mapping in structured samples. *PLoS Genetics* 3: e4.
- ZHIVOTOVSKY, L. A. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* 8: 907–913.