Phylogeographic History of White Spruce During the Last Glacial Maximum: Uncovering Cryptic Refugia

ARTICLE *in* THE JOURNAL OF HEREDITY · OCTOBER 2010

Impact Factor: 1.97 · DOI: 10.1093/jhered/esq110 · Source: PubMed

3 AUTHORS, INCLUDING:

Lynn L Anderson-Carpenter
University of Michigan
5 PUBLICATIONS **207** CITATIONS

SEE PROFILE

Journal of Heredity 2011:102(2):207–216
doi:10.1093/jhered/esq110
Advance Access publication October 18, 2010

Phylogeographic History of White Spruce During the Last Glacial Maximum: Uncovering Cryptic Refugia

Lynn L. Anderson, Feng Sheng Hu, and Ken N. Paige

From the Program in Ecology, Evolution and Conservation Biology, School of Integrative Biology, University of Illinois, Urbana, IL. Lynn L. Anderson is now at the Department of Ecology and Evolutionary Biology, University of Michigan, 2019 Kraus Natural Science Building, 830 North University, Ann Arbor, MI 48109.

Address correspondence to Ken N. Paige, Department of Animal Biology, School of Integrative Biology, University of Illinois, 515 Morrill Hall, 505 S, Goodwin Avenue, Urbana, IL 61801, or e-mail: k-paige@illinois.edu.

Abstract

Although a recent study of white spruce using chloroplast DNA uncovered the presence of a glacial refuge in Alaska, chloroplast failed to provide information on the number or specific localities of refugia. Recent studies have demonstrated the utility of nuclear microsatellites to refine insights into postglacial histories. The greater relative rate of mutation may allow finer scale resolution of historic dynamics, including the number, location, and sizes of refugia. Genetic data were acquired from screening 6 microsatellite loci on approximately 14 trees from each of 22 populations located across the central and western boreal forests of Canada and Alaska. Our studies combining microsatellites with Bayesian analyses of population structure in white spruce support the phylogeographic patterns uncovered using chloroplast, separating Alaskan from non-Alaskan regions. Results also support the idea that north-central Alaska served as a glacial refugium during the last glacial maximum. Additionally, the relationship between the degree of genetic differentiation and geographic distance indicated that gene flow played a more important role in structuring non-Alaskan populations, whereas drift played a more important role in structuring Alaskan populations (R_{ST}'s for non-Alaskan populations 0.029 ± 0.007 and 0.083 ± 0.012 for Alaskan populations). Microsatellites also substantiate the bidirectional patterns of gene flow previously uncovered using chloroplast DNA but indicate much greater movement and mixing. Results from our Bayesian analyses also suggest the existence of additional cryptic refugia. However, the locations have been obscured by high gene flow (R_{ST} averaging 0.057 ± 0.004).

Key words: *last glacial maximum, microsatellites, refugia, white spruce*

The contemporary geographic distribution of boreal trees in North America is a result of their expansion and migration from glacial refugia. Pollen and macrofossil records provide evidence that many boreal species existed south of the Laurentide and Cordilleran ice sheets during the last glacial maximum (LGM ~20 000 years before present [yBP]) (Bryant and Holloway 1985; Ritchie and MacDonald 1986). A recent synthesis of pollen data, however, suggests that glacial refugia for a number of boreal tree species may have existed in the ice-free areas of eastern Beringia (Alaska and adjacent Canada) (Brubaker et al. 2005). Recent molecular evidence supports this hypothesis and suggests that the recolonization of some North American tree species was facilitated by cryptic northern refugia, influencing patterns and rates of migration (McLachlan et al. 2005). For example, chloroplast DNA analyses of white spruce uncovered not only a northern refugium in eastern Beringia during the LGM but also the potential for bidirectional gene flow from both northern and southern refugia during recolonization (Anderson et al. 2006).

Chloroplast DNA data, however, did not provide information on specific refuge localities within eastern Beringia due to the coarse spatial resolution of pollen-inherited chloroplast DNA markers and the fact that the genetic variation detected almost certainly predates the last glacial period, given the extremely slow mutation rate of chloroplast DNA (estimated to be 5.3×10^{-9} mutations per gene per generation) (Wolfe et al. 1987; Anderson et al. 2006). Therefore, the differential distributions of chloroplast haplotypes within Alaska likely reflect mutation, sorting, and genetic drift through repeated glacial events during the Pleistocene (Anderson et al. 2006), perhaps obscuring details of the more recent history of the LGM. In addition, cytoplasmic inheritance patterns represent but a single gene genealogy and fail to capture all the historical events populations have experienced (Nordborg 2003; Heuertz et al. 2004). Alternatively, nuclear

markers allow for recombination thereby integrating multiple genealogical processes. Recent studies have demonstrated the utility of nuclear markers (in particular, microsatellites) in gaining insights into postglacial histories (see e.g., Heuertz et al. 2004; Lee et al. 2006). In particular, the greater relative rate of mutation may allow finer scale resolution of the historic dynamics of populations, including the number, location, and population sizes of refugia.

In this study, using a suite of microsatellite markers in combination with a Bayesian approach, we assess the phylogeographic history of white spruce populations across northwestern North America in an attempt to refine our previous results and gain new information on population dynamics following the LGM. Specifically, we address the following questions: 1) do microsatellites help in refining the location(s) of Alaskan refugia?, 2) do phylogenetic relationships uncovered with microsatellites parallel those uncovered using chloroplast DNA?, 3) is there nuclear evidence for bidirectional mixing of white spruce through pollen dispersal, seed dispersal, and/or tree migration as previously found for chloroplast DNA?, 4) what were the roles of gene flow and genetic drift in structuring the patterns we see today?

Materials and Methods

Genetic analyses were conducted on 22 populations located across the central and western boreal forests of Canada and Alaska (Figure 1). Needles for DNA extractions were collected from an average of 14 trees from each of the forest stands (a sample size of 14 is sufficient for estimates of allele frequencies but will likely miss some low frequency alleles, leading to an underestimate of genetic diversity and larger estimates of population structure; Table 1). Needles were generally obtained between ground level and 4 m above-ground, and individual trees were $\geq$100 m apart within each population to minimize genetic relatedness within a stand (Wright 1955). The samples were stored at −70 °C in the laboratory after field collection and prior to DNA extraction.

We extracted DNA from the needles of 290 individuals using a Qiagen Plant Mini Kit. DNA quantities were measured using either a Hoefer DNA flourimeter or a Nanodrop (Nanodrop Technologies). Six loci were amplified using the polymerase chain reaction (PCR) and published microsatellite primers (Hodgetts et al. 2001). Each PCR mixture contained 20 ng of DNA, 1× PCR buffer (Invitrogen), 1.6 mM $MgCl_2$, 0.2 μM of each primer, 0.1 mM of each dNTP, 0.2 units of *Taq* DNA polymerase (Invitrogen), and sterile water to a volume of 25 μl. The samples were then subjected to the following PCR profile: 1) 3 min at 94 °C, 2) 30 s at 94 °C, 3) 45 s at 59 °C, 4) 1 min at 72 °C, and 5) a final extension step of 10 min at 72 °C. Steps 2–4 were repeated 29 times. Five microliters of each sample was run on a 1% agarose gel for 1.5 h at 60 V to verify amplification. The remaining PCR product was analyzed with an ABI PRISM 3730 at the University of

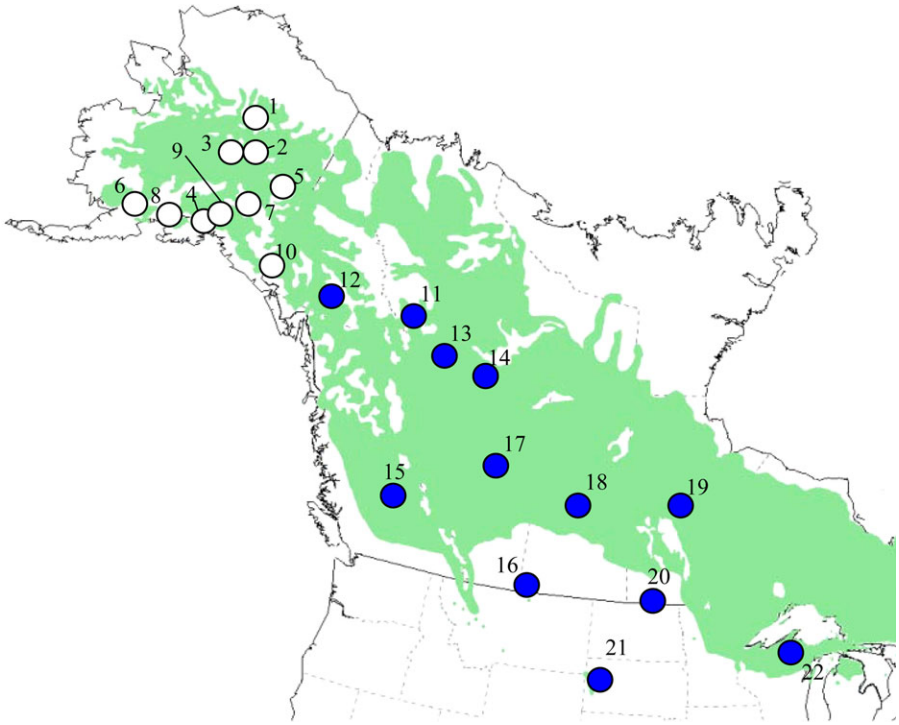

Figure 1. Location of 22 populations of white spruce sampled for studies of genetic differentiation using microsatellites. Extent of contemporary spruce forest shaded in light gray (after http://climchange.cr.usgs.gov/data/atlas/little). Sampled Alaskan populations are indicated in white, whereas non-Alaskan sampled populations are indicated in black.

Table 1 Population number (corresponding to Figure 1), sample locations, number of samples per population (#), coordinates of populations, allelic richness, population heterozygosity (hetero.), and inbreeding coefficients for white spruce

Pop. number	Sample location	#	Latitude (N), longitude (W)	Allelic richness	Population hetero.	Inbreeding coefficient (F_{IS})
1	Dalton Hwy, AK	08	65°34′, −148°59′	21.72	0.651	0.015
2	Fairbanks, AK	14	65°05′, −147°43′	29.70	0.700	0.135*
3	Denali Park, AK	14	63°09′, −149°24′	26.32	0.676	0.179*
4	Anchorage, AK	17	61°47′, −148°26′	25.87	0.685	0.212*
5	Tok, AK	14	63°20′, −142°58′	29.01	0.735	0.116
6	Sunshine Lake, AK	15	59°29′, −159°24′	24.13	0.640	0.278*
7	Carrot Lake, AK	12	62°02′, −146°48′	32.74	0.780	0.243*
8	Hole Lake, AK	15	59°22′, −159°52′	21.99	0.732	0.293*
9	Hudson Lake, AK	15	61°54′, −145°41′	33.17	0.739	0.255*
10	Slana River, AK	10	62°42′, −143°58′	30.11	0.758	0.196*
11	Mackenzie Mtns, YT	12	61°03′, −118°31′	27.35	0.739	0.174*
12	White Horse, YT	14	60°48′, −135°11′	31.87	0.738	0.289*
13	Ft. Liard, NT	14	60°00′, −122°56′	29.87	0.768	0.153*
14	Enterprise, NT	15	60°33′, −116°07′	32.94	0.747	0.253*
15	Quesnel, BC	15	53°09′, −122°21′	32.32	0.799	0.400*
16	Cypress Hills, AB	14	49°40′, −109°30′	29.10	0.746	0.378*
17	Lesser Slave Lake, AB	16	55°16′, −114°42′	33.38	0.751	0.243*
18	Prince Albert, SK	14	53°33′, −105°51′	35.42	0.765	0.246*
19	Upper Manitoba	07	55°42′, −97°47′	31.72	0.796	0.237*
20	Riding Mountain, MB	15	52°26′, −101°54′	28.92	0.730	0.089
21	Black Hills, SD	16	44°09′, −103°38′	31.83	0.773	0.217*
22	Shawano, WI	11	44°50′, −88°50′	34.01	0.757	0.216*

*significance at the 0.05 level. Sampled Alaskan populations, 1–10, are shaded in gray.

Illinois Keck Center for Functional Genomics. Genemapper (Applied Biosystems) was used to assess band sizes for all individuals.

Departures from Hardy–Weinberg equilibrium were tested using the Hardy–Weinberg probability test in Genepop 3.4 (Raymond and Rousset 1995). Exact *P* values were computed using the complete enumeration method for loci with fewer than 4 alleles (Louis and Dempster 1987) and the Markov Chain method (dememorization 1000, batches 100, iterations per batch 1000) for loci with more than 4 alleles (Guo and Thompson 1992).

Allelic variation, heterozygosity estimates, and F_{ST} values for all 6 microsatellite loci were calculated using SPAGeDi (version 1.2b) and GENEPOP (version 4.0.10). We also assessed the potential for null alleles (the lack of PCR amplification due to a mutation within the primer binding site), which can be identified by an excess of homozygosity within a locus, using MICRO-CHECKER (version 2.2.3; Oosterhout et al. 2004).

To verify whether data derived from microsatellite markers reflect broadscale phylogeographic patterns uncovered using chloroplast markers, we conducted a Bayesian analysis of population structure clustering groups of individuals (BAPS, Corander et al. 2003). This was followed by an admixture analysis (in BAPS) directed at identifying "ancestral" alleles across populations to assess the number and potential location of Alaskan refugia. The power of Bayesian approaches in determining patterns of population structure is that information is combined from several loci into a single probability model, as opposed to the simple averaging used in other analyses (Corander et al. 2003; Heuertz et al. 2004). The method of Corander et al. (2003) requires multilocus genotypes at unlinked genetic markers. Therefore, to verify the independence of our microsatellite loci, we analyzed linkage disequilibrium for all pairs of loci in each sampled population with exact tests with GENEPOP (http://genepop.curtin.edu.au/genepop_op2.html) (Raymond and Rousset 1995) and applied a sequential Bonferroni correction (Rice 1989) to discard chance correlations between loci (Heuertz et al. 2004). We also incorporated R_{ST} analyses to add additional information concerning population structure (i.e., relative importance of drift vs. gene flow and degree of differentiation; Hutchinson and Templeton 1999) in Alaska and non-Alaskan regions, regions previously delineated with chloroplast DNA (Anderson et al. 2006). R_{ST} values were quantified in SPAGeDi (version 1.2b) using a nested analysis of variance approach (Michalakis and Excoffier 1996). This approach was developed for loci undergoing a stepwise mutation process and hence a stepwise mutation model (Shriver et al. 1993; Slatkin 1995; Michalakis and Excoffier 1996). Support for use of the stepwise mutation model comes from the fact that alleles in this study were typically found to be one mutational step apart (although from a phylogenetic standpoint it is hard to tell whether the differences occurred in sequence). Values theoretically range from 0 (no differentiation among populations) to 1 (fixation of different alleles in populations).

Glacial refuge populations of trees in Alaska were likely small and isolated, as suggested by trace amounts of pollen in lake sediments (Ritchie and MacDonald 1986; Brubaker

et al. 2005; Anderson et al. 2006). We used several genetic measures gathered from microsatellite data to evaluate signatures of small population size. These measures compared Alaskan populations with non-Alaskan populations for allelic richness, levels of heterozygosity, and inbreeding coefficients (F_{IS}). Allelic richness, defined as the average number of alleles per population, is often lost during a population bottleneck. Richness estimates were obtained for each population using FSTAT (http://www2.unil.ch/popgen/softwares/fstat.htm), and results were adjusted for unequal sample sizes using Rarefac (http://www.pierroton.inra.fr/genetics/labo/Software/Rarefac/index.html). Richness differences between Alaskan and non-Alaskan regions were then calculated after 500 permutations in FSTAT. A decrease in overall allele number for each locus was tested using a multivariate analysis of variance for allelic richness followed by Bonferroni corrections using SPSS (version 11). In addition to richness, population heterozygosity and inbreeding coefficients (F_{IS}) were estimated in FSTAT (version 2.9.9.2) and Bottleneck (version 1.2.02), respectively. Heterozygosity is generally lost well after allelic richness, indicating a relatively long and severe genetic bottleneck (Hedrick et al. 1986; Luikart and Cornuet 1998). Inbreeding coefficients are indicators of the level of inbreeding due to small population size (Wright 1965; Luikart and Cornuet 1998; Luo et al. 2005).

It is important to point out that the issue of homoplasy has been raised as a possible concern in the use of microsatellites. Homoplasy occurs when different copies of a locus are identical in state (size), but not by descent, due to factors such as convergence, parallelism, or reversion (Estoup et al. 2002). A recent review of data gathered from both simulation and experimental studies suggests that homoplasy does not represent a significant problem for many types of population genetic analyzes undertaken by molecular ecologists (e.g., see Angers et al. 2000; Estoup et al. 2002; Queloz et al. 2010). For example, Rousset (1996) showed that there is virtually no effect of homoplasy on the parameters F_{ST} (see also Queloz et al. 2010) and F_{IS}. Homoplasy, however, may result in the reduction of number of alleles per population, the proportion of heterozygous individuals, and gene diversity (given identical size but differences in sequence), resulting in a decreased ability to detect differentiation among populations. Nonetheless, the large amount of variability at microsatellite loci often compensates for their homoplasious evolution (Estoup et al. 2002). Additionally, for high mutation rate loci, such as microsatellites, F_{ST}/R_{ST} values are likely underestimated thereby decreasing the ability to detect population structure (e.g., see Hedrick 1999), thus results reported here are conservative.

Results

The microsatellite loci used in this study generated 7–59 alleles and levels of heterozygosity from 0.255 to 0.745 (see Table 2). These ranges suggest differential mutation rates among loci. Specifically, Primers 8, 24, 87, and 144 had a greater number of alleles, many of which were locally distributed, suggesting a comparatively high mutation rate. In contrast, Primers 6 and 25 had few alleles, which were generally widespread throughout most of the areas sampled suggesting a slower rate of mutation (Table 2). As expected, the more conserved loci used in this study (Primers 6 and 25) were generally in Hardy–Weinberg equilibrium (Supplementary Appendix 1), whereas more variable and faster mutating loci (Primers 8, 24, 87, 144) were often not in equilibrium.

Table 2 Number of detected alleles, observed (H_o) and expected heterozygosity (H_e), average F_{ST}'s, and variation in allele size (in base pairs) across each of the 6 loci (indicated by primer number) for all 22 white spruce populations

Locus	No of alleles	H_o	H_e	F_{ST}	Range in allele size
Primer 6	7	0.734	0.7760	0.0366	58–64
Primer 8	24	0.745	0.9179	0.0618	97–129
Primer 24	43	0.359	0.9351	0.0396	100–144
Primer 25	16	0.255	0.2777	0.0312	45–73
Primer 87	59	0.666	0.9567	0.0239	54–138
Primer 144	22	0.666	0.7559	0.0225	67–95

Among a total of 330 tests for linkage disequilibrium between pairs of loci, 6 were significant ($P \leq 0.05$) and associated with only 2 of 22 populations, 3 with Mackenzie Mountain in northwest Canada and 3 with Tok in central Alaska. Therefore, we concluded that analyzed loci were sufficiently independent for the application of Bayesian methods for the analysis of population structure. The Bayesian analysis of population structure resulted in 3 genetically distinct clusters based on allele frequency differences and unique alleles, Dalton Highway in north central Alaska (the northern-most population sampled in Alaska), Riding Mountain in southern Manitoba (located close to the southeastern end of the transect), and the remaining 20 populations ranging from Wisconsin and continuing northwestwardly into southern and central Alaska (see Figures 1 and 2). The admixture analysis shows a general decline in Dalton Highway alleles in individuals from the western portion of the transect to the southeastern end of the transect in Canada (Figure 2). Dalton Highway alleles were also present in a portion of the individuals sampled in all populations south of Dalton Highway, in Alaska. Similarly, Riding Mountain alleles tend to show a general decline from southeastern Canada into southern and central Alaska. Few Riding Mountain alleles are found in the Dalton Highway population and vice versa (see Figure 2). No evidence for the presence of null alleles was found in either the Dalton Highway population or the Riding Mountain, but there was evidence for the potential presence of null alleles in some or most populations for 4 of the 6 loci based on our analysis using MICRO-CHECKER; detected as an excess in homozygosity (see Supplementary Appendix 2).

Regional comparisons indicate significantly fewer alleles in Alaska relative to areas outside of Alaska (27.5 ± 1.03 vs. 31.6 ± 0.94, degrees of freedom [df] = 1, 20, $P = 0.008$) and significantly lower levels of heterozygosity (0.71 ± 0.011

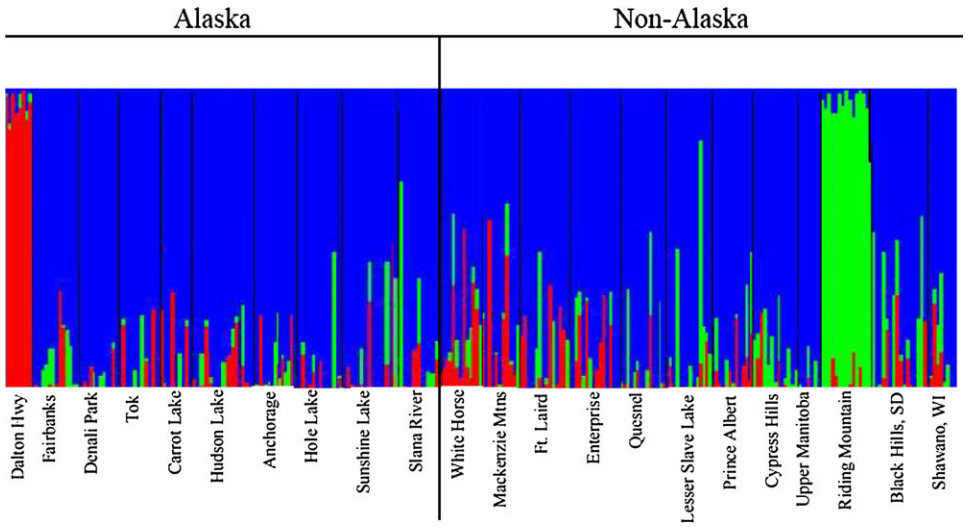

Figure 2. Bayesian analysis of population structure showing genetic differentiation among 3 regions, Dalton Highway in red, Riding Mountain in green, and the remaining 20 populations in blue. Proportions of ancestry in each of the 3 clusters are shown. Log marginal likelihood of optimal partition = −8321.9.

vs. 0.76 ± 0.010, df = 1, 20, $P = 0.004$; Table 1). Inbreeding coefficients were on average lower in Alaska than outside of Alaska but nonsignificant (0.192 ± 0.027 vs. 0.241 ± 0.025, df = 1, 20, $P = 0.198$; Table 1). However, most individual populations were significantly inbred (i.e., had high inbreeding coefficients) with the exception of 3: 2 populations in northern Alaska (Tok and Dalton Highway) and 1 outside of Alaska (Riding Mountain, Manitoba) (Table 1).

Genetic differentiation as measured by R_{ST} was, for the most part, low among populations, averaging 0.057 ± 0.004 (range: 0.000–0.298) for each population compared with all others (Table 3). However, one population in Alaska, Dalton Highway, had significantly higher R_{ST} values compared with all other Alaskan and non-Alaskan populations (averaging 0.156 ± 0.014 for all pairwise comparisons, $P < 0.001$). There was also significantly greater gene flow among non-Alaskan populations than among Alaskan populations (R_{ST}'s for non-Alaskan populations 0.029 ± 0.007 and 0.083 ± 0.012 for Alaskan populations, $P < 0.0001$).

Among 171 alleles from 6 microsatellite loci, 68 (39.8%) were localized (i.e., in a single population or a cluster of neighboring populations), 65 (38%) were found everywhere with no predictable pattern, and 38 (22.2%) showed bi-directional behavior. Overall, 14 microsatellite alleles moved from Alaskan populations southeastward and 24 alleles moved from non-Alaskan populations northwestward, fitting the predicted pattern of bi-directional movement uncovered using chloroplast DNA (Figure 3).

Discussion

Evidence for Ice Age Refugia

A Bayesian analysis of population structure using microsatellite markers resulted in 3 genetically distinct regions; Dalton Highway in north-central Alaska, Riding Mountain in Manitoba, and one large deme consisting of the remaining 20 populations ranging from southeastern Canada to southern and central Alaska (Figure 2). For the Dalton Highway population, we see a predominant one-way pattern of gene flow out of Dalton Highway southward into Alaska and southeastward into Canada with little gene flow (as compared with all other populations except Riding Mountain, see below) into the Dalton Highway region from other populations (Figure 2). Additionally, R_{ST} values were significantly different ($P < 0.001$) for all pairwise population comparisons except Tok ($P = 0.161$) indicating comparatively less gene flow between Dalton Highway and the majority of populations (averaging 0.156; with 4 of 21 comparisons <0.10 and 17 of 21 comparisons >0.10; 7 of these latter comparisons were >0.20, see Table 3). The range of R_{ST} values indicate "low-moderate" genetic differentiation (range for all plant species is ca. 0.01 to 0.60, Hamrick JL, personal communication). These results suggest that the north-central region of Alaska, Dalton Highway, served as a refugium during the LGM, narrowing the region from the whole of Alaska based on chloroplast markers (Anderson et al. 2006) down to a more localized northern region of interest. Interestingly, the notion that north-central Alaska may serve as a glacial refuge was first proposed by Hopkins (1967).

Similarly, Riding Mountain located in the southeast portion of the transect appears to represent a genetically unique early extension of the northern-most region of the well-established southern refugium (which extended from northern Kansas through Missouri, Iowa, and Illinois; Dort et al. 1985; Ritchie and MacDonald 1986) from which spruce expanded and migrated north-northwestward as the ice sheets receded (Riding Mountain was covered by the southern edge of the Cordilleran and Laurentide ice sheets

Table 3 Pairwise population comparisons for calculated R_{ST} values using SpaGeDi (version 1.2b)

	Fair	Anch	Denali	Dalt	Tok	HudLk	CartLk	Slana	Sunsh	Hole	MacMtn	Ft.Liard	Ent	WhtHrs	PAlb	PGQu	LSL	UM	RM	CH	SD	Wisc
Fair	0.0000	0.0335	0.0532	0.2140	0.2035	−0.0282	−0.0217	0.0389	0.0892	−0.0024	0.1324	−0.0255	0.0303	0.0442	−0.0241	−0.0068	0.0313	0.0019	0.0129	0.0103	−0.0085	0.0937
Anch		0.0000	−0.0361	0.0446	0.0804	−0.0036	0.0208	0.0315	0.0656	0.0130	0.0337	0.0018	0.0007	−0.0115	0.0194	0.0638	−0.0286	−0.0331	−0.0108	0.0669	0.0176	−0.0072
Denali			0.0000	0.0240	0.1137	0.0161	0.0236	0.0812	0.0977	0.0391	0.0793	0.0170	0.0152	−0.0246	0.0223	0.0975	−0.0186	−0.0192	0.0139	0.0646	0.0440	−0.0323
Dalt				0.0000	0.1282	0.1965	0.1416	0.2080	0.2697	0.2362	0.1542	0.1952	0.1410	0.1078	0.1890	0.2488	0.0616	0.1259	0.1050	0.2675	0.2253	−0.0037
Tok					0.0000	0.1638	0.2126	0.1476	0.0893	0.1379	−0.0394	0.1805	0.0823	0.1103	0.2308	0.2118	0.0194	0.0328	0.1155	0.2983	0.1225	0.0677
HudLk						0.0000	−0.0028	0.0379	0.0671	−0.0286	0.0891	−0.0382	−0.0098	0.0080	−0.0164	0.0018	−0.0101	−0.0225	−0.0063	0.0275	−0.0262	0.0552
CartLk							0.0000	0.0432	0.1276	0.0453	0.1589	−0.0060	0.0633	0.0455	−0.0411	0.0288	0.0502	0.0075	0.0111	0.0009	0.0414	0.0717
Slana								0.0000	0.1965	0.0850	0.0723	0.0334	0.0745	0.1347	0.0616	0.0266	0.0349	0.0485	−0.0398	0.1583	0.0798	0.1450
Sunsh									0.0000	0.0281	0.0647	0.1052	0.0289	0.0352	0.1434	0.1528	0.0066	−0.0282	0.1172	0.2028	0.0045	0.0573
Hole										0.0000	0.0669	−0.0176	−0.0363	−0.0033	0.0248	0.0191	−0.0244	−0.0254	0.0247	0.0643	−0.0432	0.0447
MacMtn											0.0000	0.1071	0.0213	0.0776	0.1711	0.1285	−0.0233	−0.0045	0.0457	0.2491	0.0599	0.0560
Ft.Liard												0.0000	−0.0090	0.0165	−0.0294	−0.0163	−0.0018	0.0063	−0.0113	0.0047	−0.0093	0.0588
Ent													0.0000	−0.0241	0.0363	0.0243	−0.0422	−0.0243	0.0204	0.0628	−0.0242	−0.0018
WtHrs														0.0000	0.0355	0.0880	−0.0300	−0.0279	0.0477	0.0566	0.0028	−0.0486
PAlb															0.0000	0.0073	0.0378	0.0291	0.0104	−0.0237	0.0291	0.0722
PGQu																0.0000	0.0446	0.0620	0.0081	0.0427	0.0247	0.1307
LSL																	0.0000	−0.0564	−0.0006	0.0761	−0.0187	−0.0253
UM																		0.0000	−0.0005	0.1052	−0.0399	−0.0159
RM																			0.0000	0.0688	0.0300	0.0600
CH																				0.0000	0.0703	0.0931
SD																					0.0000	0.0486
Wisc																						0.0000

The negative R_{ST} values noted in some cells should be interpreted as no genetic differentiation between the 2 populations and reflects the imprecision of the algorithm used by the software to estimate this value.

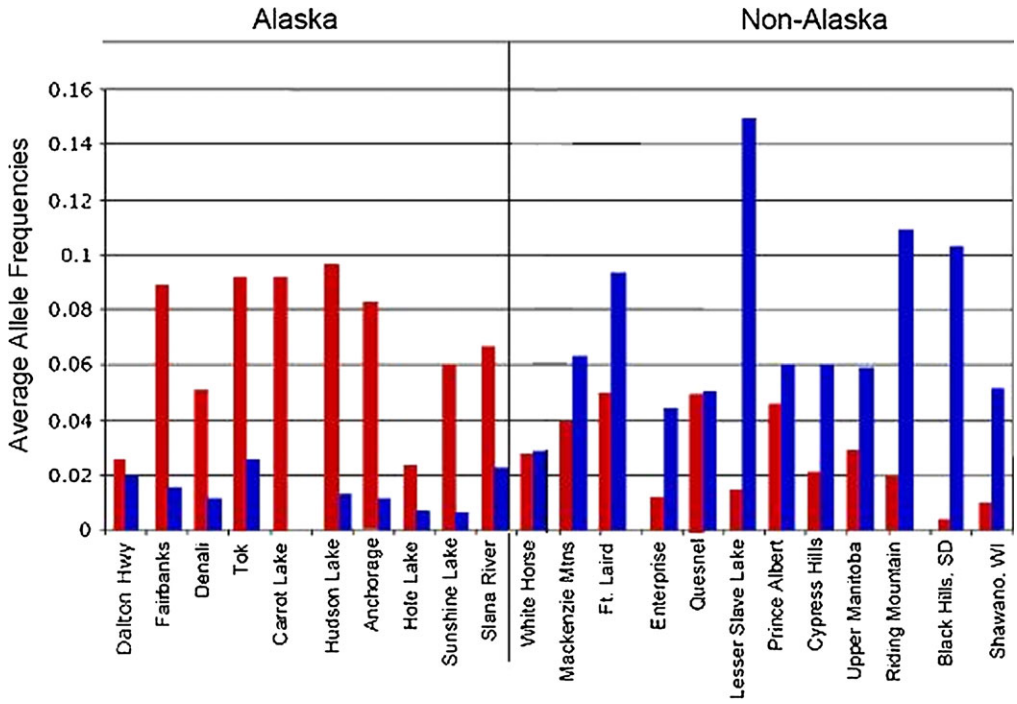

Figure 3. Average frequency distribution of Alaskan (red) and non-Alaskan alleles (blue) that support bidirectional genetic exchange. Black lines indicate the division between Alaskan and non-Alaskan populations. Populations are arranged from northwest to southeast as one moves from Alasaka to non-Alaskan populations.

during the LGM). We also note a predominate one-way pattern of gene flow out of Riding Mountain indicating that it served as a source of unique genotypes/alleles that have contributed to the genetic structuring of white spruce as it expanded following the LGM. Although there are no known barriers to gene flow between Riding Mountain and other southeastern populations or between Dalton Highway and other Alaskan populations, we surmise that gene flow may have been limited by the earlier expansion of trees in these refugial populations, acting to minimize the influx and establishment of new genotypic variants (see Figure 2).

The third genetically distinct group, comprised 20 populations ranging from southeastern Canada to southern and central Alaska, also suggests that white spruce expanded from Alaskan and southern refugia. Frequencies of nuclear microsatellite alleles clearly indicate bidirectional gene flow (Figure 3), paralleling patterns uncovered using chloroplast DNA (Anderson et al. 2006). Fourteen microsatellite alleles had an average high concentration in Alaska, appearing to have dispersed southeastward and gradually trailing off in frequency. In contrast, 24 microsatellite alleles were concentrated in the southeastern end of the transect, appearing to have moved northwestward and gradually decreasing in frequency (Figure 3). This pattern suggests bidirectional mixing of white spruce through gene flow, with no allele exchange in opposition to the chloroplast DNA (Anderson et al. 2006). Two other patterns were also detected in nuclear alleles, one was an even dispersal likely explained by common/possibly ancestral alleles found throughout Alaska and non-Alaskan populations, and the other a localized pattern, potentially explained by the accumulation of new mutations within sampled areas that have not had time to disperse (averaging 2.13 ± 0.15 unique alleles per population). Overall, the bidirectional patterns of microsatellites indicate much greater movement and mixing than those of chloroplast-DNA, with northern nuclear alleles filtering to the southernmost populations and vice versa (Anderson et al. 2006). The explanation for the lack of greater penetrance of chloroplast haplotypes may be that by chance these increasingly rare haplotypes (in comparison with common haplotypes shared across populations) were no longer carried to adjacent populations or alternatively selection may have played a role such that chloroplast were lost with changing climatic conditions.

In addition, the lack of Hardy–Weinberg equilibrium and the potential presence of null alleles (an excess in homozygosity) can best be explained by the bidirectional flow of alleles discussed above rather than technical problems in the amplification of alleles. The excess in homozygosity observed across populations is likely due to a Wahlund effect, which is caused by the merging of subpopulations with different allele frequencies reducing overall heterozygosity. Consistent with this argument is the fact that the 2 refugial regions, Dalton Highway and Riding Mountain, show no signs of excess homozygosity/null alleles and there is little gene flow into these regions, as noted above (see Figure 2). Both the Dalton Highway and Riding Mountain populations were also predominantly in Hardy–Weinberg equilibrium

with the exception of locus 6 in Riding Mountain. Alternatively, observed patterns may be due to within-population patterns of family structure rather than the merging of subpopulations leading to the high levels of inbreeding observed (Table 1).

One might ask whether Dalton Highway and Riding Mountain served as the sole northern and southern refugial regions, respectively. Based on microsatellite and chloroplast data, we would argue that they were not the sole regions from which trees migrated (although they did contribute some genetic material, as described above and clearly shown in Figure 2). Support for this statement comes from the fact that Dalton Highway and Riding Mountain are genetically differentiated from the remaining populations not only in terms of microsatellite variation but also in terms of chloroplast variation, that is, the 10 unique haplotypes within these 20 populations cannot be explained by migration out of the 2 refugia uncovered in this study (Dalton Highway and Riding Mountain), given that they are not all shared with those of Dalton Highway or Riding Mountain. More important, mutational input predates the LGM, estimated to be 5×10^{-10} mutations per locus per generation; a rate far too slow to account for so many unique chloroplast haplotypes following the LGM. Although data suggest that there were additional refugia at the height of the LGM, the precise locations are likely obscured by the bidirectional patterns of gene flow described above for microsatellite data that resulted in the genetic similarity of these 20 populations. Furthermore, there are areas that have not yet been sampled in our investigation that likely harbor refugial populations. Zazula et al. (2006) recently uncovered macrofossils within the unglaciated Yukon Territory in far northwest Canada from a time period somewhere between 26 000 and 24 500 ^{14}C yBP, indicating the existence of an additional refugial population.

Population Size Structure and the Relative Importance of Gene Flow Versus Drift

Within Alaskan populations, regional comparisons of allelic richness, levels of heterozygosity, and inbreeding support the presence of small-localized refugia during the LGM. Allelic richness and levels of heterozygosity were significantly lower in Alaskan relative to non-Alaskan populations, indicative of small population size, perhaps the result of bottleneck effects during the last ice age. Inbreeding coefficients also indicate most Alaskan populations endured bottlenecks, although they may be somewhat inflated given potential Wahlund effects. The levels of inbreeding observed here have been observed in other species of *Picea* as well using microsatellite markers. For example, inbreeding coefficients in Sitka spruce have been reported to range from an F of -0.38 to an F as high as 0.54 with a mean of 0.23 (A′Hara and Cottrell 2009), similar to the range reported here. However, as A′Hara and Cottrell (2009) point out these inflated levels may possibly be due to the effects of null alleles or a Wahlund effect. Pollen records do, however, show trace amounts of spruce pollen during the last glacial period, suggesting small sizes of refugial populations (Ritchie and MacDonald 1986; Brubaker et al. 2005). Although the majority of Alaskan populations show potential genetic signatures of small size, the Dalton Highway population appears to have maintained sufficiently large size given its low inbreeding coefficient relative to all other populations in Alaska (0.015 for Dalton Highway, vs. an average of 0.224 for all other Alaskan populations; Table 1).

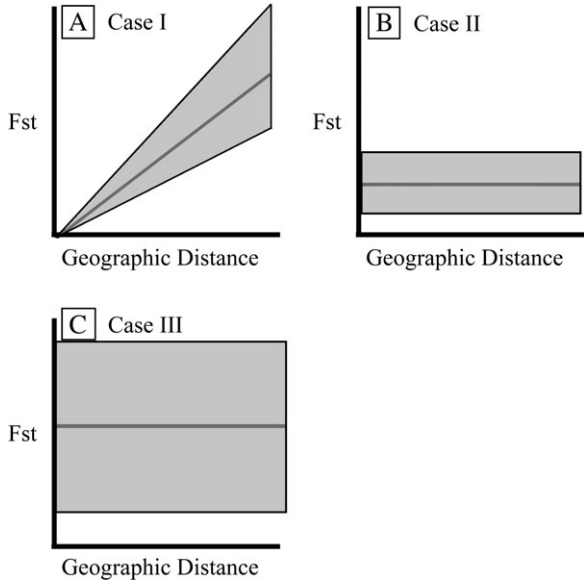

Figure 4. Graphical representations of hypothesized relationships between genetic distance (defined by F_{ST}) and geographic distances, reproduced from Hutchinson and Templeton (Hutchinson and Templeton 1999). (**A**) Case I: gene flow and drift in regional equilibrium. (**B**) Case II: gene flow much more influential than genetic drift, lack of regional equilibrium. (**C**) Case III: genetic drift more influential than gene flow, lack of regional equilibrium.

Regional comparisons of the relationship between genetic and geographic distance also support the view that populations in Alaska were predominantly small and isolated during the LGM. When populations are in genetic equilibrium, there is a positive relationship between genetic and geographic distance with an increasing variance with geographic distance (Case I of Hutchinson and Templeton (1999), Figure 4A). Hutchinson and Templeton (1999) proposed that deviation from this expected relationship can give valuable insights into the relative influences of genetic drift and gene flow on regional population structure, given that most natural populations are probably not in equilibrium (McCauley 1993). If gene flow is strong relative to drift a pattern reflecting panmixia would be apparent, that is, the homogenizing effects of gene flow would spread genetic variants throughout the region without regard to the extent of geographic separation. This would show a pattern similar to that of Case II from Hutchison and Templeton (1999) shown in Figure 4B. Alternatively, if populations are small and isolated, then drift would become relatively more

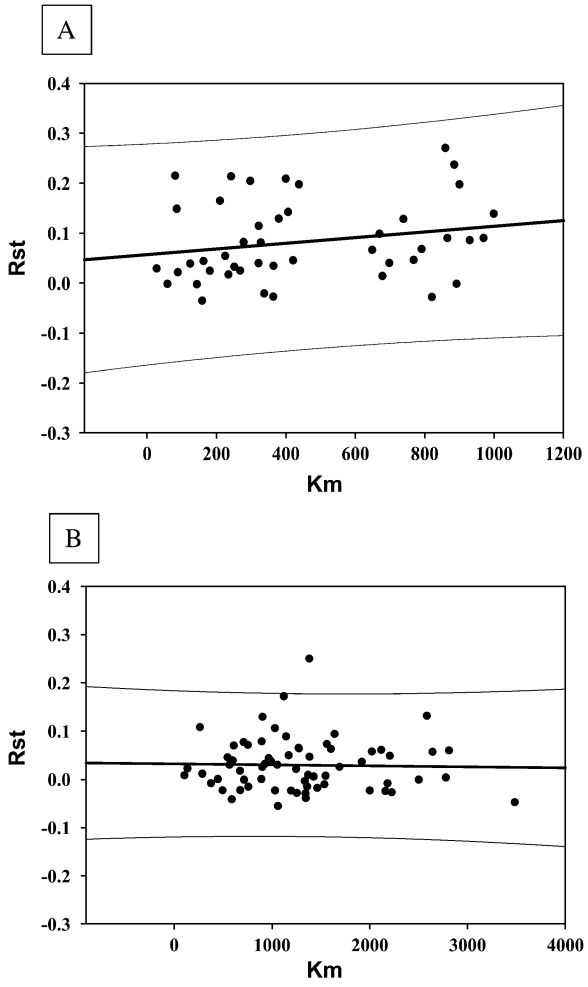

Figure 5. R_{ST} values for Alaska (**A**) and non-Alaska (**B**) regions plotted against increasing distance between pairwise populations; 99% confidence intervals are shown. Note greater variance for Alaska versus non-Alaskan populations.

influential than gene flow. This would show a pattern similar to that of Case III from Hutchison and Templeton (1999) shown in Figure 4C. Comparisons between R_{ST} and geographic distance among Alaskan populations revealed a greater degree of isolation where populations were influenced moreso by drift independent of geographic distance, creating a much wider degree of scatter (or variance in estimates of divergence) (Figure 5A; $R_{ST} = 0.083 \pm 0.013$), most similar to Case III of Hutchinson and Templeton . In contrast, comparisons between pairwise R_{ST} values and geographic distance within non-Alaskan populations supports the invasion from a homogeneous source population (the southern refuge) where gene flow is more influential than drift (Figure 5B; $R_{ST} = 0.029 \pm 0.007$). This parallels Case II of Hutchison and Templeton (1999) (Figure 4B) where there is no relationship between genetic and geographic distance and little variance in estimates of divergence relative to Alaskan populations.

Concluding Remarks

Genetic relationships derived from microsatellite markers support the broadscale phylogeographic patterns uncovered using chloroplast markers (Anderson et al. 2006), indicating that microsatellite markers are useful in investigating millennial-scale vegetational events. Others as well have found utility in using microsatellites for uncovering historic pattern. For example, microsatellite variation within the common ash (*Fraxinus excelsior*) and white oak (*Quercus glauca*) coincides with organelle, pollen, and macrofossil data that indicate separation into multiple refugia during the LGM (Heuertz et al. 2004; Lee et al. 2006). In particular, the use of microsatellites in uncovering phylogeographic pattern may be most useful in plant species where genetic variation is difficult to detect within their highly conserved plastid DNA (Wolfe et al. 1987; Heuertz et al. 2004).

Results from this study also indicate that microsatellites can reveal aspects of a plant's glacial history that data from more conserved plastid markers cannot. Although chloroplast DNA revealed the presence of a northern refugium during the LGM, chloroplast revealed little about the number or historic sizes of glacial refugia or the roles of gene flow and drift in structuring the patterns we see today. Microsatellites indicated that there were multiple refugia and a mix of historically large and small populations that contributed significantly to recolonization of the boreal forest after the end of the last glaciation. Particularly, noteworthy is the Dalton Highway population located in northern Alaska, clearly identified as one of the primary contributors to the spread and expansion of white spruce following the LGM. In addition, the relationship between the degree of genetic differentiation and geographic distance between the 2 phylogeographic regions indicates that gene flow plays a more important role in structuring non-Alaskan populations while drift plays a more important role in structuring Alaskan populations. Microsatellite markers also substantiate the bidirectional patterns of gene flow previously uncovered using chloroplast markers but indicate much greater movement and mixing. These results contrast with existing paleoecological records that revealed an absence of significant pollen or macrofossils from Alaskan lake sediments (perhaps because Alaskan LGM populations were too small to be reliably detected by pollen analysis) and a unidirectional northward migration from southern refugia.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

National Science Foundation grants to K.N.P., F.S.H., and L.L.A.; PEO Scholar Award to L.L.A.; Sigma Xi Award to L.L.A.; Packard Fellowship in Science and Engineering to F.S.H.

Acknowledgments

We thank J.E. Beever, S.R. Downie, J.S. McLachlan, C.A. Phillips for advice and discussion, and D.D. Biesboer, D. Bishop, J. Buchanan-Mappin, P. Camill, M.J. Carroll, K.M. Cavey, B.E. Cook, G. Fewless, E.C. Grimm, A. Henderson, B. Jaquish, C.L. Lehman, P. LePage, D. Sharp, and K. Wolfe for assistance with sample collections.

References

A′Hara SW, Cottrell JE. 2009. Development of a set of highly polymorphic genomic microsatellites (gSSRs) in Sitka spruce (*Picea sitchensis*(Bong.) Carr.). Mol Breed. 23:349–355.

Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN. 2006. Ice-age endurance: DNA evidence for a white spruce refugium in Alaska. Proc Natl Acad Sci. 103:2447–12450.

Angers B, Estoup A, Jarne P. 2000. Microsatellite size homoplasy, SSCP, and population structure: a case study in the freshwater snail *Bulinus truncatus*. Mol Biol Evol. 17:1926–1932.

Brubaker LB, Anderson PM, Edwards ME, Lozhkin AV. 2005. Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data. J Biogeogr. 32:833–848.

Bryant VM, Holloway RG, editors. 1985. Pollen records of late-quaternary North American sediments, American Association of Stratigraphic Palynologists Foundation Contributions Series, Dallas, Texas.

Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. Genetics. 163:367–374.

Dort W, Johnson WC, Fredlund GG, Rogers RA, Martin LD, Stewart JD, Wells PV. 1985. Evidence for an open conifer woodland in the Central Great Plains during the late Wisconsin glacial maximum. Lethbridge (Canada): Canadian Quaternary Association. p. 23.

Estoup A, Jarne P, Cornuet J-M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol. 11:1591–1604.

Guo SW, Thompson EA. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics. 48:361–372.

Hedrick PW. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution. 53:313–318.

Hedrick PW, Brussard PF, Allendorf FW, Beardmore JA, Orzack S. 1986. Protein variation, fitness, and captive propagation. Zoo Biol. 5:91–99.

Heuertz M, Hausman J-F, Hardy OJ, Vendramin GG, Frascaria-LaCoste N, Vekemans X. 2004. Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior*, L.). Evolution. 58:976–988.

Hodgetts RB, Aleksiuk MA, Brown A, Clarke C, Macdonald E, Nadeem S, Khasa D. 2001. Development of microsatellite markers for white spruce (*Picea glauca*) and related species. Theor Appl Genet. 102:1252–1258.

Hopkins DM. 1967. The Cenozoic history of Beringia—a synthesis. In: Hopkins DM, editor. The Bering land Bridge. Stanford (CA): Stanford University Press. p. 451–484.

Hutchison DW, Templeton AR. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution. 53:1898–1914.

Lee Y, Hwang S-Y, Ho K-C, Lin T-P. 2006. Source populations of Quercus glauca in the Last Glacial Age in Taiwan revealed by nuclear microsatellite markers. J Hered. 97:261–269.

Louis EJ, Dempster ER. 1987. An exact test for Hardy–Weinberg and multiple alleles. Biometrics. 43:805–811.

Luikart G, Cornuet J-M. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv Biol. 12:228–237.

Luo J, Want Y, Korpelainen H, Li C. 2005. Allozyme variation in natural populations of *Picea asperata*. Silva Fennica. 39:167–176.

McCauley DE. 1993. Genetic consequences of extinction and recolonization in fragmented habitats. In: Kareiva PM, Kingsolver JG, Huey RB, editors. Biotic interactions and global change. Sunderland (MA): Sinauer. p. 217–233.

McLachlan JS, Clark JS, Manos PS. 2005. Molecular indicators of tree migration capacity under rapid climate change. Ecology. 86:2088–2098.

Michalakis Y, Excoffier L. 1996. A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics. 142:1061–1064.

Nordborg M. 2003. Coalescent theory. Hoboken (NJ): John Wiley & Sons.

Oosterhout CV, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 4:535–538.

Queloz V, Duo A, Sieber TN, Grunig CR. 2010. Microsatellite size homoplasies and null alleles do not affect species diagnosis and population genetic analysis in a fungal species complex. Mol Ecol Resour. 10:348–367.

Raymond M, Rousset F. 1995. GENEPOP (version1.2): population genetics software for exact tests and ecumenicism. J Hered. 86:248–249.

Rice WR. 1989. Analyzing tables of statistical tests. Evolution. 43:223–225.

Ritchie JC, MacDonald GM. 1986. The patterns of post-glacial spread of white spruce. J Biogeogr. 13:527–540.

Rousset F. 1996. Equilbrium values of measures of population subdivision for stepwise mutation processes. Genetics. 142:1357–1362.

Shriver MD, Jin L, Chakraborty R, Boerwinkle E. 1993. VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics. 134:983–993.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics. 139:457–462.

Wolfe KH, Li W-H, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial chloroplast and nuclear DNA. Proc Natl Acad Sci. 84:9054–9058.

Wright JW. 1955. Species crossability in spruce in relation to distribution and taxonomy. Forest Sci. 1:319–349.

Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. 19:395–420.

Zazula GD, Telka AM, Harington CR, Schweger CE, Mathewes RW. 2006. New spruce (*Picea* spp.) macrofossils from Yukon territory: implications for late pleistocene refugia in eastern Beringia. Arctic. 59:391–400.

Received August 31, 2009; Revised September 16, 2010; Accepted September 16, 2010

Corresponding Editor: James L. Hamrick