

Ice-age endurance: DNA evidence of a white spruce refugium in Alaska

Lynn L. Anderson*, Feng Sheng Hu**§, David M. Nelson*, Rémy J. Petit[¶], and Ken N. Paige*^{||}

*Program in Ecology and Evolutionary Biology and Departments of [†]Plant Biology, [‡]Geology, and [§]Animal Biology, University of Illinois, Urbana, IL 61801; and [¶]Institut National de la Recherche Agronomique, Unité Mixte de Recherche Biodiversité, Gènes et Ecosystèmes, 69 Route d'Arcachon, F-33612 Cestas Cedex, France

Communicated by H. E. Wright, University of Minnesota, Minneapolis, MN, June 26, 2006 (received for review February 19, 2006)

Paleorecords offer key information for evaluating model simulations of species migration in response to forecast climatic change. However, their utility can be greatly compromised by the existence of glacial refugia that are undetectable in fossil records (cryptic refugia). Despite several decades of investigation, it remains controversial whether Beringia, the largely unglaciated area extending from northeastern Siberia to the Yukon Territory, harbored small populations of certain boreal tree species during the last glaciation. Here, we present genetic evidence for the existence of a glacial refuge in Alaska that helps to resolve this long-standing controversy. We sequenced chloroplast DNA (cpDNA) of white spruce (*Picea glauca*), a dominant boreal tree species, in 24 forest stands across northwestern North America. The majority of cpDNA haplotypes are unique, and haplotype diversity is relatively high in Alaska, arguing against the possibility that this species migrated into the region from areas south of the Laurentide Ice Sheet after the end of the last glaciation. Thus, white spruce apparently survived long glacial episodes under climatic extremes in a heterogeneous landscape matrix. These results suggest that estimated rates of tree migration from fossil records may be too high and that the ability of trees to track anthropogenic warming may be more limited than previously thought.

Beringia | *Picea glauca* | refugia | chloroplast | climate change

In his classic book published in 1937, Eric Hultén (1) postulated that the ice-free area of Beringia served as a northern refugium for arctic and boreal biota during Quaternary glaciations. Numerous paleoecological studies (2–7) have since attempted to determine whether the dominant tree species of the North American boreal biome survived the last glacial maximum (LGM) (25,000–17,000 years ago) in eastern Beringia (Alaska and adjacent Canada). Results of pollen and macrofossil analyses yielded ambiguous evidence for the spruce species, with some studies arguing for this hypothesis (2–4) and others against it (5, 6). For example, trace amounts of spruce pollen are not uncommon in the lake sediments of the last glacial period, and their spatial patterns appear to be consistent with the idea that eastern Beringia was a glacial refugium (2). However, whether they represent local trees remains “a hypothesis that invites testing” (2), and spruce macrofossils have never been found in the LGM sediments from Beringia (5). These interpretations have profoundly different implications for rates of tree migration and resilience to climatic forcing. Evidence that is independent of fossil pollen records is necessary to resolve this controversy because of the inherent limitations of palynological analyses for identifying species ranges (2, 7, 8).

Genetic analysis of modern species offers an effective approach for detecting cryptic glacial refugia (9, 10). We conducted a chloroplast DNA (cpDNA) survey of white spruce foliar samples from 24 forest stands across northwestern North America along a putative migration route (Fig. 1a; Table 1, which is published as supporting information on the PNAS web site; and ref. 6). These stands are arrayed along a broad southeast–northwest transect that allows us to use the geographic patterns of cpDNA to infer the general locations of glacial refugia during the LGM (6). Our results

provide several lines of evidence for the existence of a glacial refuge of white spruce in Alaska, supporting Hultén’s hypothesis (1) and offering important information for interpreting paleorecords of tree responses to climatic change.

Results and Discussion

To identify potential cpDNA markers of glacial refugia, we sequenced the most variable noncoding portions of the white spruce chloroplast genome (11) on a subset of our samples. This initial screening revealed three regions with polymorphisms: the *trnF/L* and *trnL/T* intergenic spacers and the *ndhK/C* pseudogene (11). These three variable regions are linked, as indicated by the circular structure of the chloroplast genome. Therefore, we combined them to derive composite haplotypes for inferring glacial refugia (12).

The number of haplotypes reaches a plateau with the cpDNA sequences of 163 white spruce trees from 12 non-Alaskan populations (Fig. 1b), indicating that we have likely uncovered the existing haplotype variation in that area. The number of haplotypes approaches, but does not reach, a plateau with the same number of trees from 12 Alaskan populations (Fig. 1b), suggesting that we have captured the vast majority of haplotype diversity in that area. These inferences are confirmed by plotting the number of detected haplotypes against the number of investigated populations (Fig. 1c). The 326 sampled individuals together yielded a total of 17 haplotypes that are either one or two single point mutations (both transitions and transversions) away from haplotype I (Fig. 2 and Table 1). Haplotype I is ancestral, as indicated by its sequence similarity to black and Norway spruce (*Picea mariana* and *Picea abies*, respectively), two sister species.

No phylogeographic structure exists on the basis of haplotype-identity permutations because all others derive from haplotype I or II, both of which occur in all 24 populations (Fig. 2). This result reflects the extremely slow mutation rate of the chloroplast genome in conifers (estimated to be 5.3×10^{-10} mutations per gene per generation for sequenced regions) (14). The slow mutation rate, along with the fact that spruce cpDNA is paternally inherited through widely dispersed pollen (15, 16), could have resulted in minimal genetic differentiation and obliterated the cpDNA signature of refuge locations in modern samples. Nevertheless, we found significant patterns in the geographic distribution of the cpDNA haplotypes that unambiguously differentiate the two regions (i.e., within Alaska versus outside Alaska), unveiling several lines of evidence for a glacial refuge of white spruce in Alaska.

Assuming that a LGM spruce refuge was absent in eastern Beringia and that spruce immigrated from the southeastern United States after the end of the last glaciation, we expected the Alaskan

Conflict of interest statement: No conflicts declared.

Abbreviations: cpDNA, chloroplast DNA; LGM, last glacial maximum.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ112004–DQ112006, DQ112008–DQ112016, DQ112018, and DQ315942–DQ315947).

§To whom correspondence should be addressed at: University of Illinois, 265 Morrill Hall, 505 South Goodwin Avenue, Urbana, IL 61801. E-mail: fshu@life.uiuc.edu.

© 2006 by The National Academy of Sciences of the USA

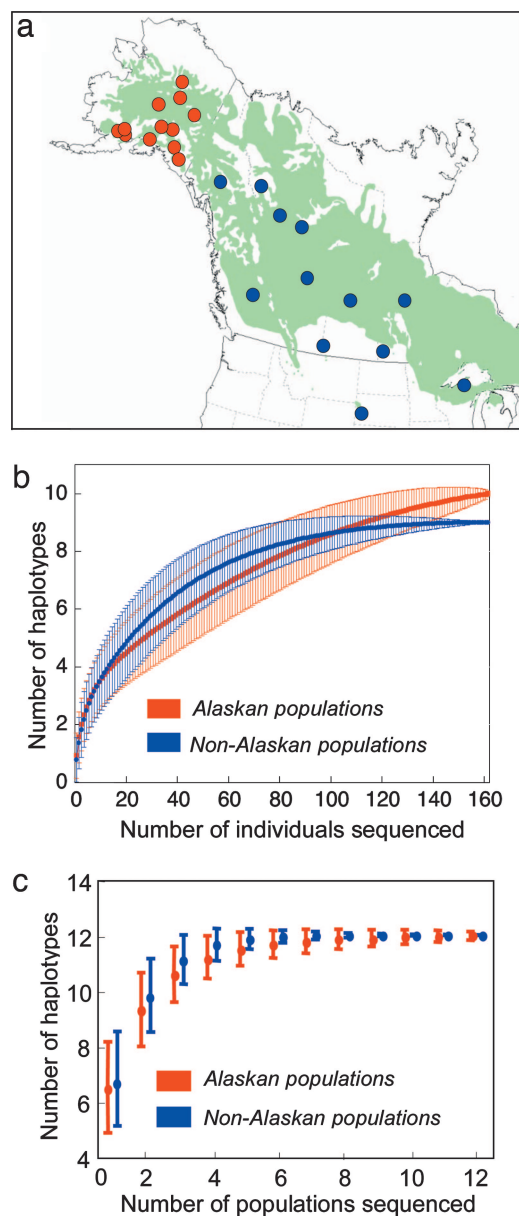


Fig. 1. Sampling locations and haplotype diversity of white spruce. (a) Locations of Alaskan (red) and non-Alaskan (blue) populations sampled for this study (all 12 non-Alaskan populations are outside the boundary of Berlingia). The modern white spruce range is indicated in green. (b) Means and 95% confidence intervals of white spruce cpDNA haplotype diversity in relation to the number of sequenced individuals based on rarefaction resampling (using the statistical program EstimateS). (c) Means and 95% confidence intervals of white spruce cpDNA haplotype diversity in relation to the number of sequenced populations based on rarefaction resampling.

haplotypes to form a subset of those in the southeastern portion of our transect. Contrary to this expectation, the majority of the Alaskan haplotypes (7 of 12: Ic, Ie, If, Ig, IId, IJe, and IIJf) are unique to that region (Figs. 2 and 3). This result offers strong evidence that white spruce trees, although of low densities as suggested by only trace amounts of pollen in lake sediments (2), indeed survived the LGM in that region and expanded locally during the Holocene. We also found five haplotypes that were unique to the southeastern portion of the transect (Figs. 2 and 3a), which likely derive from the southern refuge well documented in the paleorecord (6). Because the LGM northern limit of this species' range does not overlap with

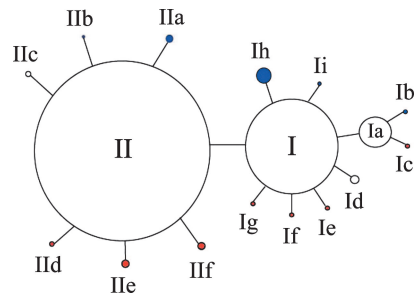


Fig. 2. Haplotype network assembled by using the method of Templeton *et al.* (13). Each haplotype is represented by a number–letter combination; red haplotypes occur only in Alaskan populations, blue haplotypes occur only in non-Alaskan populations, and white haplotypes occur in both regions. Circle sizes are proportional to haplotype frequencies in all 24 populations, with the largest circle representing the most abundant haplotype.

its southern limit today (6), a major loss of haplotype diversity must have occurred on the “trailing edge” (18) during the range displacement resulting from climatic warming at the end of the last glaciation. The haplotypes in the southeastern populations therefore represent only a fraction of the suite in the southern refugium of white spruce. Given the extremely slow mutation rate for the sequenced regions (19), the polymorphisms detected in our cpDNA

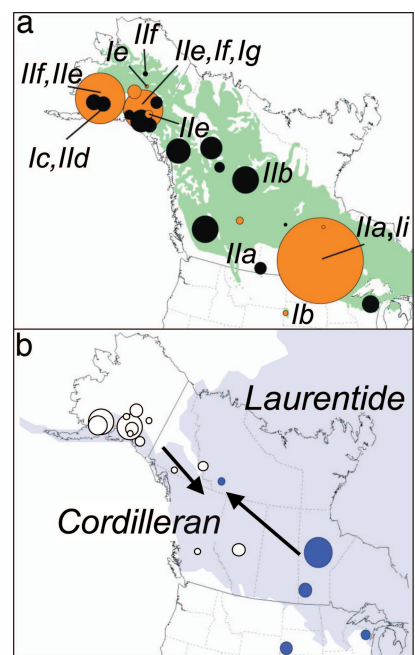


Fig. 3. Geographic patterns of white spruce cpDNA haplotypes in northwestern North America. (a) Modern white spruce range (green), haplotype diversity (circles), and locations of unique haplotypes (labeled). Intrapopulation haplotype diversity was calculated by using the computer program Rarefac to account for differences in population sample sizes. Orange and black circles indicate diversity values above and below the mean (1.97) of all 24 populations, respectively. Circle sizes are proportional to deviations from the mean, with the largest orange circle representing the highest diversity, the largest black circle representing the lowest diversity, and the smallest circles having values closest to the mean. Haplotypes unique to each region are indicated by the number-letter combinations. (b) Frequency patterns of haplotypes 1a (white circles) and 1h (blue circles) suggest bidirectional mixing of white spruce cpDNA. Circle sizes are proportional to intrapopulation haplotype frequencies, with the largest circle representing the highest frequency. Light-blue shading indicates the extent of the Laurentide and Cordilleran Ice Sheets during the LGM (17).

of trees to keep pace with anthropogenic climatic warming is more limited than previously inferred from fossil pollen records.

Methods

We determined our sample density in each of the two regions on the basis of a statistical program (EstimateS) that can identify the sample size for detecting all or nearly all of the haplotypes (38). The greater total area of sampling and the lower population density per unit area outside Alaska than inside Alaska should skew the results toward a higher total diversity value outside Alaska, according to the principle underlying the species–area relationship (34). Thus, our estimated higher diversity in the Alaskan population should be conservative and robust.

The locations of our study populations were selected partially for site accessibility and sample availability. We did not choose sites to maximize the potential of identifying the specific locations of glacial refuge populations, such as ice-free areas of eastern Beringia with the earliest establishment of spruce on the basis of ^{14}C -dated pollen records (2). This limitation, along with the fact that cpDNA is paternally inherited through pollen dispersal, means that our cpDNA data reflect the overall genetic patterns of each large region rather than of each forest stand where samples were taken. At each of the 24 forest stands along our sampling transect (Table 1), needle fascicles were collected from an average of 14 trees. Fascicles were generally obtained between ground level and 4 m above ground, and individual trees were ≥ 100 m apart within each population. The samples were stored with dry ice in the field and during shipment and then at -70°C in the laboratory before DNA extraction.

We extracted DNA from the needles of 326 individual trees by using a Plant Mini Kit (Qiagen, Valencia, CA). Sequence data as well as banding patterns from chloroplast and random amplified polymorphic DNA primers verified species identity (35, 36), ruling out the possibility that some samples originated from *P. sitchensis*, *P. mariana*, or *Picea engelmannii* (results not shown). Universal primers were used for the amplification of the *trnF*/L and *trnL*/T introns (37). Primers were designed for the *ndhK*/C pseudogene by using Primer v3 and sequence data from *Pinus thunbergii* (11).

Primer sequences for the *ndhK*/C region included the following: forward, GGAGAATGATGAAGAAGTTG; reverse, ATTCGATGGAACACGCTACC. Each PCR mixture contained 20 ng of DNA, $1\times$ PCR buffer (Invitrogen, Carlsbad, CA), 1.6 mM MgCl_2 , a 0.2 μM concentration of each primer, a 0.1 mM concentration of each dNTP, 0.2 unit of *Taq*DNA polymerase (Invitrogen), and sterile water to a volume of 25 μL . The samples were then subjected to the following PCR profile: step 1, 3 min at 94°C ; step 2, 30 s at 94°C ; step 3, 45 s at 56°C ; step 4, 1 min at 72°C ; step 5, a final extension 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 sequenced with the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and analyzed with an ABI PRISM 377 and 3730 at the W. M. Keck Center for Comparative and Functional Genomics (University of Illinois, Urbana, IL).

We used Sequencher 4.5 (Gene Codes, Ann Arbor, MI) to edit and combine DNA sequences. The haplotype network was created by using the method of Templeton *et al.* (13). The program DnaSP (39) was used to estimate K_{ST} and corresponding P values, Permut (40) was used to test for phylogeographic significance, and the program SAMOVA (41) was used to identify genetically differentiated groups of proximal populations. See Table 2, which is published as supporting information on the PNAS web site, for GenBank accession numbers of the cpDNA haplotypes.

We thank J. E. Beever, S. R. Downie, J. Germano-Presby, C. A. Phillips, and K. C. Rowe for advice and discussion; M. C. Berg, B. F. Clegg, A. Henderson, G. Clarke, M. L. Farmer, and P. D. Henne for comments; F. S. Chapin and two anonymous reviewers for constructive comments that helped improve the manuscript; 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. This work was supported by a Packard Fellowship in Science and Engineering (to F.S.H.), National Science Foundation grants (to F.S.H., K.N.P., and L.L.A.), a PEO Scholar Award (to L.L.A.), and a Sigma Xi Award (to L.L.A.).

- Hultén, E. (1937) *Outline of the History of Arctic and Boreal Biota During the Quaternary Period* (Lehre J. Cramer, New York).
- Brubaker, L. B., Anderson, P. M., Edwards, M. E. & Lozhkin, A. V. (2005) *J. Biogeogr.* **32**, 833–848.
- Colinvaux, P. A. (1964) *Ecol. Monogr.* **34**, 297–329.
- Hopkins, D. M. (1972) *Inter Nord* **12**, 121–150.
- Hopkins, D. M., Smith, P. A. & Matthews, J. V., Jr. (1981) *Quat. Res.* **15**, 217–249.
- Ritchie, J. C. & MacDonald, G. M. (1986) *J. Biogeogr.* **13**, 527–540.
- Davis, M. B., Schwartz, M. W. & Woods, K. (1991) *J. Biogeogr.* **18**, 653–668.
- Bennett, K. D., Tzedakis, P. C. & Willis, K. J. (1991) *J. Biogeogr.* **18**, 103–115.
- Stewart, J. R. & Lister, A. M. (2001) *Trends Ecol. Evol.* **16**, 608–612.
- Petit, R. J., Aguinalde, I., de Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., *et al.* (2003) *Science* **300**, 1563–1565.
- Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T. & Sugiura, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9794–9798.
- Gamache, I., Jaramillo-Correa, J. P., Payette, S. & Bousquet, J. (2003) *Mol. Ecol.* **12**, 891–901.
- Templeton, A. R., Crandall, K. A. & Sing, C. F. (1992) *Genetics* **132**, 619–633.
- Graur, D. & Li, W.-H. (1999) *Fundamentals of Molecular Evolution* (Sinauer, Sunderland, MA).
- Sutton, B. C. S., Flanagan, D. J., Gawley, J. R., Newton, C. H., Lester, D. T. & El-kassaby, Y. A. (1991) *Theor. Appl. Genet.* **82**, 242–248.
- Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D. & Vendramin, G. G. (2005) *Mol. Ecol.* **14**, 689–701.
- Ehlers, J. & Gibbard, P. L., eds. (2004) *Quaternary Glaciations: Extent and Chronology. Part II: North America* (Elsevier, Boston).
- Hampe, A. & Petit, R. J. (2005) *Ecol. Lett.* **8**, 461–467.
- Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9054–9058.
- Ibrahim, K. M., Nichols, R. A. & Hewitt, G. M. (1996) *Heredity* **77**, 282–291.
- Hudson, R. R., Boos, D. D. & Kaplan, N. L. (1992) *Mol. Biol. Evol.* **9**, 138–151.
- Comes, H. P. & Kadereit, J. W. (1998) *Trends Plant Sci.* **3**, 432–438.
- Hewitt, G. M. (1996) *Biol. J. Linn. Soc. London* **58**, 247–276.
- Bialozyt, R., Ziegenhagen, B. & Petit, R. J. (2006) *J. Evol. Biol.* **19**, 12–20.
- Sutton, B. C. S., Pritchard, S. C., Gawley, J. R., Newton, C. H. & Kiss, G. K. (1994) *Can. J. For. Res.* **24**, 278–285.
- Malcolm, J. R., Markham, A., Neilson, R. P. & Garaci, M. (2002) *J. Biogeogr.* **29**, 835–849.
- Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebe, T. J. C., Fromentin, J. M., Hoegh-Guldberg, O. & Bairlein, F. (2002) *Nature* **416**, 389–395.
- Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T. & Prentice, I. C. (2005) *Proc. Natl. Acad. Sci. USA* **102**, 8245–8250.
- Chapin, F. S., Callaghan, T. V., Bergeron, Y., Fukuda, M., Johnstone, J. F., Juday, G. & Zimov, S. A. (2004) *Ambio* **33**, 361–365.
- McAvaney, B. J., Covey, C., Jousaume, S., Kattsov, V., Kitoh, A., Ogana, W., Pitman, A. J., Weaver, A. J., Wood, R. A., Zhao, Z.-C., *et al.* (2001) in *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, New York), pp. 471–523.
- Fastie, C. L. (1995) *Ecology* **76**, 1899–1916.
- Clark, J. S., Lewis, M. & Horvath, L. (2001) *Am. Nat.* **157**, 537–554.
- McLachlan, J. S., Clark, J. S. & Manos, P. S. (2005) *Ecology* **86**, 2088–2098.
- MacArthur, R. H. & Wilson, E. O. (1963) *Evolution* (Lawrence, Kans.) **17**, 373–387.
- Germano, J. & Klein, A. S. (1999) *Theor. Appl. Genet.* **99**, 37–49.
- Khasa, P. D. & Dancik, B. P. (1996) *Theor. Appl. Genet.* **92**, 46–52.
- Demesure, B., Sodji, N. & Petit, R. J. (1995) *Mol. Ecol.* **4**, 129–131.
- Colwell, C. K. (2005) EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples (Univ. of Connecticut, Storrs, CT), Version 7.5.
- Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2006) DnaSP (Universitat de Barcelona, Barcelona), Version 4.10.8.
- Petit, R. J. (2005) Permut (Institut National de la Recherche Agronomique, Bordeaux, France), Version 1.0.
- Dupanloup, I. (2002) SAMOVA (Univ. of Bern, Bern, Switzerland) Version 1.0.