Genetic Composition and Diploid Hybrid Speciation of a High Mountain Pine, *Pinus densata*, Native to the Tibetan Plateau

Xiao-Ru Wang,*,† Alfred E. Szmidt‡ and Outi Savolainen§

**Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 100093 Beijing, China,* ‡*Department of Biology, Graduate School of Science, Kyushu University, Fukuoka 812-8581, Japan,* §*Department of Biology, Oulu University, 90570 Oulu, Finland and* †*National Institute for Working Life, 907 13, Umeå, Sweden*

Manuscript received February 12, 2001
Accepted for publication June 22, 2001

ABSTRACT

Pinus densata has been suggested to have originated from hybridization events involving *P. tabulaeformis* and *P. yunnanensis*. In this study, allozyme differentiation at 12 loci was studied in 14 populations of *P. tabulaeformis*, *P. densata*, and *P. yunnanensis* from China. The observed genetic composition of *P. densata* supported the hybrid hypothesis and showed varying degrees of contribution from *P. yunnanensis* and *P. tabulaeformis* among its populations. These data, together with previous chloroplast DNA results, indicated different evolutionary histories among *P. densata* populations. To examine the possibility of ongoing hybridization among the three species, we analyzed patterns of linkage disequilibria between allozyme loci in ovule, pollen, and zygote pools. None of these tests suggested that there is significant ongoing gene exchange, implying that populations of *P. densata* have a stabilized hybrid nature. The normal fertility and high fecundity of *P. densata* indicate that this hybrid is maintained through sexual reproduction. *P. densata* represents an example of diploid hybrid speciation in an extreme ecological habitat that is both spatially and ecologically separated from that of its parents.

IN plants, hybridization and introgression are known to be important evolutionary forces (Anderson and Stebbins 1954; Stebbins 1969; Grant 1981; Arnold 1992). Gene flow between species may lead to uni- or bi-directional gene exchange, may eliminate species boundaries after secondary contact, may generate unstable tension zones, or may initiate stable speciation events (Barton and Hewitt 1985; Paige *et al.* 1991; Whittemore and Schaal 1991; Rieseberg 1997). Hybrid speciation in plants through polyploidy is much more common than homoploid hybrid speciation, because the duplication in chromosome number provides the hybrid with an effective means of reproductive isolation from its parental species. Hybrid speciation without an increase in ploidy must, in contrast, involve other isolating factors such as ecological and spatial isolation and/or some degree of chromosomal or genetic incompatibility (Anderson 1948; McCarthy *et al.* 1995; Rieseberg and Carney 1998; Buerkle *et al.* 2000).

The occurrence of several crossable sympatric species from the genus Pinus in Asia has led to suggestions that some species arose as a result of hybridization (Wu 1956; Mirov 1967). One example is *Pinus densata*, which is distributed in southwestern China and the Tibetan plateau. It has been suggested, following analysis of its morphological characters, that *P. densata* arose as the product of natural hybridization between *P. tabulaeformis* and *P. yunnanensis* (Wu 1956; Mirov 1967). This view is supported by morphological data (Wang 1961), allozyme analysis (Wang *et al.* 1990), and the distribution of chloroplast DNA (cpDNA) polymorphism (Wang and Szmidt 1990, 1994). Like most pine species, *P. densata* is diploid ($2N = 24$) and has a karyotype similar to *P. tabulaeformis* (Yang 1987). Thus, the evolution of *P. densata* does not involve polyploidy.

Hybridization and introgression have been reported for several other conifer species complexes where hybridization goes on in the narrow hybrid zone of sympatry of the parental taxa (*e.g.*, Szmidt *et al.* 1988; Ernst *et al.* 1990; Sutton *et al.* 1991; Wagner *et al.* 1991; Sigurgeirsson 1992). In such a situation the hybrid zone is maintained mainly by current gene flow and the evolutionary fate of the hybrids remains to be determined. In contrast, *P. densata* appears as a well-developed and stabilized species that has a large distribution range and occupies a unique habitat at high elevations that appears inaccessible to other pines in the region (Wu 1956; Guan 1981). The geographic distribution of the three species forms a succession, with *P. tabulaeformis* in the north, *P. densata* in the middle, and *P. yunnanensis* in the south (Figure 1). *P. densata* overlaps, in latitude, only slightly with *P. tabulaeformis* and *P. yunnanensis* at the margins. The degree of overlap is further diminished when we consider the distribution of the three species with respect to altitude. *P. densata* occurs at higher elevations (2700–4200 m above sea level) than both *P. tabulaeformis* (0–2600 m) and *P. yunnanensis*

Corresponding author: Xiao-Ru Wang, National Institute for Working Life, P.O. Box 7654, S-907 13 Umeå, Sweden.
E-mail: xiao-ru.wang@niwl.se

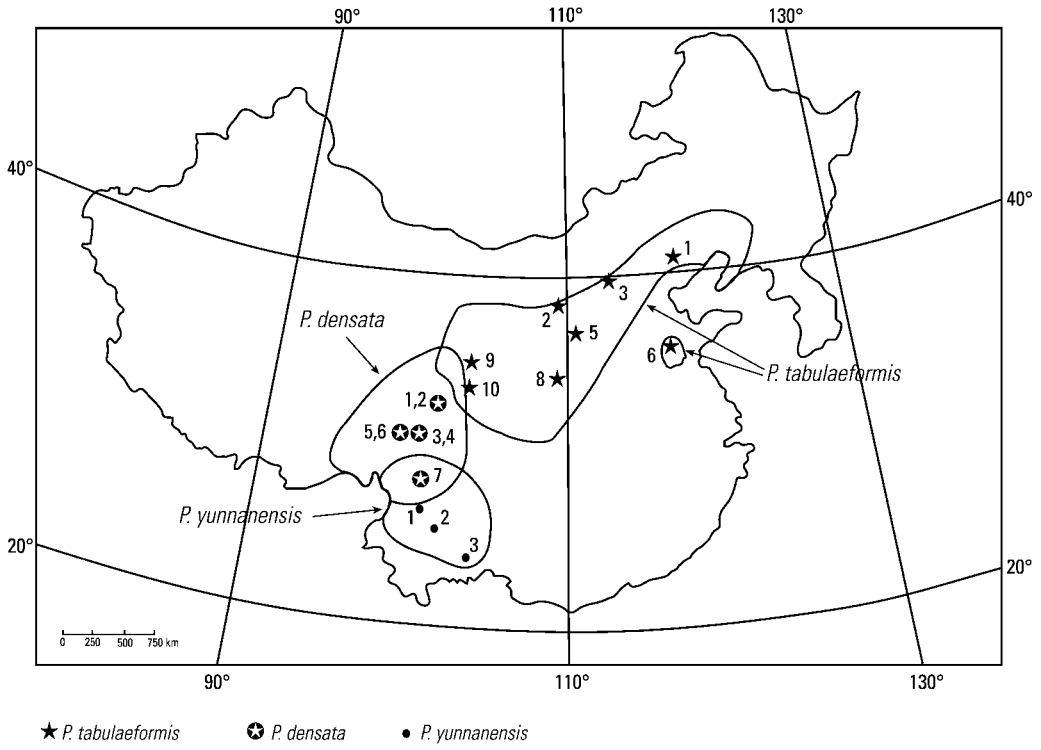

FIGURE 1.—Geographic distribution of the sampled populations of *P. tabulaeformis*, *P. densata*, and *P. yunnanensis*.

(600–3100 m) (WU 1956; GUAN 1981). However, extensive pollen flow is often observed in conifers (ENNOS 1994). Thus, the possibility of current gene exchange with the putative parents remains to be examined. In addition, the genetic composition of the hybrid populations, the mating pattern, and population genetic structure also require quantification.

Detailed studies on population genetic structure are important for understanding the effects of introgression on the maintenance of species differences, hybrid population dynamics, and the potential differences in the evolutionary biology of organellar *vs.* nuclear genes (ASMUSSEN *et al.* 1987; RIESEBERG and CARNEY 1998). Estimation of the degree of gametic disequilibrium can provide an indirect reference for the significance of current gene flow in the maintenance of the hybrid zone. However, despite much theoretical work in this field, there have been very few studies on associations between loci in outcrossing conifer populations (MUONA and SZMIDT 1985). Inferences about linkage disequilibrium rely on measures of gametic frequencies. Unless gametic frequencies are directly observable, they are inferred from genotypic frequencies under the assumption that the union of gametes is random. Thus, the analysis of linkage disequilibrium in diploid samples is subject to large uncertainties, due to the fact that double heterozygotes cannot be apportioned to coupling and repulsion phases unless the sample size is sufficiently large (BROWN 1975; WEIR and COCKERHAM 1979). Conifers are interesting materials for detailed studies on population genetic structure. Simultaneous analysis of the haploid megagametophyte and diploid embryo in a seed allows the female and male gametic contributions to the zygote to be separated. Thus, it is possible to study separately the ovule and pollen pools that make up the zygotes. This approach can also be used in studying multilocus associations in conifer populations by measuring the disequilibria separately in haploid ovules and pollen.

In our previous analysis, only one population of each of the three species, *P. tabulaeformis*, *P. densata*, and *P. yunnanensis*, was analyzed for allozyme variation (WANG *et al.* 1990), and no other report has been published concerning the population structure of these three pines. The study presented here was designed to provide a detailed assessment of genetic structure in these three taxa at the allozyme level. We examined the level and distribution of genetic variation within and among populations, the mating pattern, gametic pool composition, and linkage disequilibrium in populations of all three species. On the basis of this information we discuss the mechanisms that led to the speciation of *P. densata*. In particular, four questions are addressed. Do the new allozyme data conform to earlier evidence indicating that *P. densata* has a hybrid origin? Is there significant ongoing hybridization among species in this complex? What is the relative gene contribution from the putative parental taxa to the investigated populations of *P. densata*? Finally, do different *P. densata* populations have the same origin and a similar evolutionary history?

MATERIALS AND METHODS

Seed material: Bulked seed samples were obtained from eight natural populations of *P. tabulaeformis* (Pt-1, -2, -3, -5, -6, -8, -9, and -10), three populations of *P. densata* (Pd-1, -2,

and -7) and three populations of *P. yunnanensis* (Py-1, -2, and -3) from the natural range of each species in China. The geographic distribution of these populations is presented in Figure 1. Two populations of *P. tabulaeformis* (Pt-9 and Pt-10) and one population of *P. yunnanensis* (Py-1) originated from the region of close sympatry with *P. densata* (Figure 1). The remaining populations of *P. tabulaeformis* and *P. yunnanensis* were located outside the range of *P. densata*. Of the three populations of *P. densata*, one population (Pd-7) was located in the region of sympatry with *P. yunnanensis*, while the remaining two populations originated from the north-central part of its distribution, which is outside the ranges of both putative parental species (Figure 1). The exact number of trees included in these collections is unknown, but appears to be >50. The megagametophyte and embryo from each seed were separated and analyzed simultaneously for all the enzyme systems used in this study. From 100 to 195 megagametophytes and their corresponding embryos were analyzed from each population.

Allozyme analyses: For allozyme electrophoresis, seeds were germinated for ~14 days until a 3-mm radicle emerged from the seed coat. Enzymes were extracted as previously described (Wang *et al.* 1990), and allozymes were separated in 12% starch gels. Seven enzyme systems, encoded by 13 loci, were scored as described previously (Wang *et al.* 1990). One locus, *Pgm*-2, was excluded from further analysis since the reproducibility of the results associated with it from some populations was poor. All the remaining 12 loci were assessed simultaneously in each megagametophyte and the corresponding embryo. The separate female and male gamete contributions to the zygote were deduced through comparison of allozyme patterns in the haploid megagametophyte and diploid embryo.

Statistical analysis: *Diversity measurements:* Allozyme frequencies, expected (H_e, Nei 1978) and observed (H_o) heterozygosity, and gene diversity statistics were calculated using version 1.7 of the BIOSYS program (Swofford and Selander 1981). A locus was considered polymorphic if the frequency of the most common allele did not exceed 0.95. All calculations were made separately for ovules, pollen, and zygotes. Differences in allele frequencies between ovule and pollen pools of each population were tested (Fisher's exact test) by the Markov chain method, involving 1000 dememorizations with 100 batches and 2000 iterations, using the GENEPOP ver. 3.1 program (Raymond and Rousset 1995). If there were allelic differences between the ovule and pollen pool, the expected heterozygosity was calculated as $H_e' = 1 - \Sigma P_{if}P_{im}$, where P_{if} and P_{im} are the allelic frequencies for ovule and pollen, respectively.

The fit of genotypic frequencies to the Hardy-Weinberg (H-W) expectation was tested using the exact test (Haldane 1954; Guo and Thompson 1992). For this, the complete enumeration method (Louis and Dempster 1987) was applied for loci with up to four alleles and the Markov chain method for loci with more than four alleles, again involving 1000 dememorizations with 100 batches and 2000 iterations, using the GENEPOP ver. 3.1 program (Raymond and Rousset 1995). In addition, single-locus and mean fixation indices (FI) were computed for polymorphic loci as described by Curie-Cohen (1982).

Population differentiation was analyzed for polymorphic loci in ovules, pollen, and zygotes by *F*-statistics (Weir and Cockerham 1984). F_{st} (θ) values and the probability test (exact *G*-test, Goudet *et al.* 1996) were estimated using 1000 randomizations with the program FSTAT ver. 2.8 (Goudet 1995), not assuming random mating within samples. Population relationships were inferred from the allozyme data applying the UPGMA clustering method on the basis of Nei's (1978) unbiased genetic distance. The UPGMA tree was constructed using TFPGA program ver. 1.3 (Miller 1997). To generate confidence estimates in the constructed tree, the bootstrap procedure with 1000 permutations was employed.

Linkage disequilibria: To examine the possibility of ongoing hybridization among the three species, we analyzed patterns of linkage disequilibria between loci in ovule, pollen, and zygote pools. Measures of linkage disequilibrium were calculated as described by Weir (1996). The probability test (Fisher's exact test) for each within-population locus-pair contingency table was performed using the Markov chain method, with 1000 dememorizations, 100 batches, and 2000 iterations per batch. An association between loci was considered significantly different from zero if the exact test gave a probability <0.05. The multiple testing was performed using the GENEPOP ver. 3.1 computer program (Raymond and Rousset 1995).

Allozyme admixture: The relative contributions from *P. tabulaeformis* and *P. yunnanensis* to the investigated populations were estimated using a least-squares procedure developed by Roberts and Hiorns (1965) and Elston (1971). The calculated index of gene migration (m) is considered an estimate of the proportion of genes in a population derived from the reference taxon. The standard error of m was calculated as described by Draper and Smith (1966) and Wheeler and Guries (1987). All calculations were made as the proportion of genes derived from *P. yunnanensis* and *P. tabulaeformis*, respectively. The frequencies for all alleles scored were included in the analysis.

RESULTS

Allele frequency: Twelve loci were inferred for ovules, pollen, and zygotes in each of the 14 investigated populations. Of these 12 loci, 9 were polymorphic (frequency limit for most common allele: 0.95) in at least 1 population. Allozyme frequencies in zygotes for each of the 3 populations of *P. densata* are given in Table 1 and were compared to the mean allozyme frequencies found for *P. tabulaeformis* and *P. yunnanensis*. Allozyme frequencies for individual populations of *P. tabulaeformis* and *P. yunnanensis* can be obtained upon request from the senior author. The most polymorphic loci included *Pgm*-1, *Fes*, *Lap*-2, *Got*-3, and *Sdh*-1. Little polymorphism was found at the *Mdh*-1, *Mdh*-2, and *Gdh*-1 loci. Distinct differences were found at the *Lap*-2 locus, which was highly variable in *P. yunnanensis* and *P. densata* but nearly monomorphic in *P. tabulaeformis*. The *Sdh*-1 locus, however, was more variable in *P. tabulaeformis* and *P. densata* than in *P. yunnanensis*. The *Got*-3 locus was highly polymorphic in all three species, but for different alleles (Table 1). At all the loci that varied markedly between *P. tabulaeformis* and *P. yunnanensis*, *P. densata* nearly always showed an intermediate allozyme composition, with Pd-7 tending to be more similar to *P. yunnanensis*. Alleles restricted to *P. tabulaeformis* and *P. yunnanensis* were detected at several loci, *e.g.*, *Pgm*-1 and *SDh*-1, and they were also often found in *P. densata*. At other loci, most of the unique alleles for *P. tabulaeformis* and *P. yunnanensis* had very low frequencies (<0.05) and they were not always detected in *P. densata*. Among all the 12 loci examined, only two alleles specific to *P. densata* were detected, at *Got*-3 and *SDh*-2, with frequencies ranging from 0.004 to 0.075.

TABLE 1

Mean allele frequencies at 12 loci in populations of *P. tabulaeformis* and *P. yunnanensis* and in each of the *P. densata* populations

Loci	*P. tabulaeformis* (No. pop. = 8; N = 959)	*P. densata* Pd-1 (N = 195)	*P. densata* Pd-2 (N = 133)	*P. densata* Pd-7 (N = 147)	*P. yunnanensis* (No. pop. = 3; N = 334)
Pgm-1					
1	0.230	0.049	0.045	0.031	0.000
2	0.706	0.884	0.823	0.677	0.865
3	0.054	0.067	0.132	0.265	0.131
4	0.010	0.000	0.000	0.003	0.000
5	0.001	0.000	0.000	0.024	0.004
MDh-1					
1	0.014	0.033	0.041	0.024	0.024
2	0.984	0.964	0.940	0.963	0.976
3	0.003	0.003	0.019	0.014	0.000
MDh-2					
1	0.977	0.962	1.000	0.990	0.999
2	0.022	0.038	0.000	0.003	0.001
3	0.001	0.000	0.000	0.007	0.000
Fes					
1	0.206	0.443	0.244	0.133	0.188
2	0.632	0.461	0.605	0.779	0.554
3	0.100	0.096	0.135	0.065	0.051
4	0.042	0.000	0.004	0.000	0.002
5	0.018	0.000	0.000	0.000	0.008
6	0.003	0.000	0.011	0.024	0.197
Lap-1					
1	0.930	0.885	0.906	0.959	0.985
2	0.039	0.000	0.004	0.000	0.002
3	0.026	0.115	0.090	0.037	0.012
4	0.003	0.000	0.000	0.003	0.001
5	0.002	0.000	0.000	0.000	0.000
Lap-2					
1	0.027	0.003	0.008	0.000	0.000
2	0.948	0.863	0.799	0.539	0.731
3	0.018	0.131	0.193	0.461	0.269
4	0.007	0.003	0.000	0.000	0.000
Got-1					
1	0.107	0.079	0.085	0.191	0.144
2	0.884	0.897	0.867	0.799	0.853
3	0.009	0.023	0.048	0.010	0.003
Got-2					
1	0.007	0.000	0.000	0.000	0.084
2	0.955	1.000	1.000	1.000	0.854
3	0.038	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.060
5	0.000	0.000	0.000	0.000	0.001
Got-3					
1	0.265	0.605	0.627	0.824	0.783
2	0.634	0.379	0.354	0.176	0.207
3	0.099	0.011	0.015	0.000	0.002
4	0.000	0.005	0.004	0.000	0.000
5	0.000	0.000	0.000	0.000	0.006
6	0.001	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.001
GDh					
1	0.995	1.000	1.000	1.000	1.000
2	0.004	0.000	0.000	0.000	0.000
3	0.001	0.000	0.000	0.000	0.000

(*continued*)

TABLE 1

(Continued)

Loci	*P. tabulaeformis* (No. pop. = 8; N = 959)	*P. densata* Pd-1 (N = 195)	*P. densata* Pd-2 (N = 133)	*P. densata* Pd-7 (N = 147)	*P. yunnanensis* (No. pop. = 3; N = 334)
SDh-1					
1	0.002	0.000	0.000	0.000	0.000
2	0.828	0.361	0.355	0.010	0.000
3	0.125	0.637	0.645	0.990	0.976
4	0.029	0.003	0.000	0.000	0.021
5	0.014	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.003
7	0.003	0.000	0.000	0.000	0.000
SDh-2					
1	0.015	0.077	0.040	0.090	0.025
2	0.980	0.923	0.881	0.903	0.975
3	0.005	0.000	0.004	0.000	0.000
4	0.000	0.000	0.075	0.007	0.000

N = sample size.

Allele frequency differences between the two gametic pools were very minor (data not shown). If allelic differences between ovule and pollen pools were present, the expected heterozygosity would be higher than the H-W expectation. The effect of allelic differences on estimates of expected heterozygosity is shown in Table 2. The expected heterozygosity (H_e') based on the allele frequencies in the ovules and pollen was calculated for each locus, and the mean value for each population is listed in Table 2. The difference in H_e' compared to H_e based on allele frequencies in zygotes is negligible. Thus, in the following H-W test and FI estimates, allele frequencies in the zygotic population were used.

Diversity measures and population differentiation: A summary of genetic variability measures at 12 loci in the investigated populations is given in Table 2. The average proportion of polymorphic loci ranged from 50% in *P. yunnanensis* to 67% in *P. densata*. The average

TABLE 2

Gene diversity measures in ovule, pollen, and zygote pools and estimates of departures from H-W equilibrium in zygote populations

Population	Sample size per locus	P	h_f (ovule)	h_m (pollen)	H_e' (zygote)	H_e (zygote)	H_o (zygote)	FI
Pt-1	115.3 ± 0.4	50.0	0.232	0.189	0.215	0.212	0.197	0.043
Pt-2	102.8 ± 0.2	50.0	0.181	0.151	0.168	0.167	0.162	0.034
Pt-3	109.3 ± 0.4	58.3	0.191	0.185	0.190	0.188	0.175	0.040
Pt-5	144.1 ± 2.8	41.7	0.175	0.181	0.178	0.178	0.167	0.020
Pt-6	110.0 ± 0.0	50.0	0.193	0.187	0.192	0.190	0.177	0.035
Pt-8	108.5 ± 0.2	50.0	0.196	0.198	0.199	0.197	0.190	0.018
Pt-9	119.5 ± 0.2	58.3	0.226	0.204	0.218	0.216	0.213	0.020
Pt-10	136.4 ± 3.3	66.7	0.255	0.238	0.247	0.246	0.248	−0.019
Average Pt	946.0 ± 3.8	58.3	0.214	0.198	0.206	0.205	0.193	
Pd-1	191.4 ± 2.6	66.7	0.237	0.204	0.224	0.222	0.191	0.097
Pd-2	128.7 ± 2.5	75.0	0.253	0.219	0.242	0.239	0.229	0.038
Pd-7	144.8 ± 0.8	50.0	0.204	0.179	0.196	0.194	0.165	0.069
Average Pd	464.9 ± 3.9	66.7	0.246	0.216	0.234	0.232	0.194	
Py-1	114.8 ± 0.3	58.3	0.193	0.178	0.187	0.186	0.184	0.014
Py-2	111.3 ± 0.3	50.0	0.197	0.190	0.194	0.193	0.179	0.021
Py-3	105.6 ± 0.3	50.0	0.182	0.172	0.181	0.178	0.170	0.007
Average Py	331.8 ± 0.6	50.0	0.195	0.183	0.190	0.190	0.178	

P, percentage of polymorphic loci (0.95 criterion); h_f, h_m, gene diversity in ovule and pollen pools, respectively; H_e and H_o, expected and observed heterozygosity; H_e', expected heterozygosity calculated from ovule and pollen allele frequencies; FI, mean fixation index over polymorphic loci.

TABLE 3

F_{st} (WEIR and COCKERHAM 1984) values among the investigated populations of *P. tabulaeformis*, *P. densata*, and *P. yunnanensis* and the significance test

	P. tabulaeformis			*P. densata*			*P. yunnanensis*		
Locus	Ovule	Pollen	Zygote	Ovule	Pollen	Zygote	Ovule	Pollen	Zygote
Pgm-1	0.027*	0.013*	0.017*	0.051*	0.080*	0.064*	0.042*	0.025*	0.037*
MDh-1	0.043*	0.014*	0.014*	0.009	−0.001	0.001	0.004	0.001	0.003
MDh-2	0.007	−0.003	0.002*	0.005	0.038*	0.020*	—	0.000	0.000
Fes	0.026*	0.043*	0.034*	0.111*	0.080*	0.095*	0.033*	0.064*	0.050*
Lap-1	0.024*	0.021*	0.019*	0.027*	0.004	0.016*	0.022*	0.006	0.015*
Lap-2	0.019*	0.013*	0.016*	0.109*	0.194*	0.147*	0.001	−0.003	0.004
Got-1	0.083*	0.046*	0.063*	0.048*	0.007	0.024*	0.033*	−0.007	0.009*
Got-2	0.011*	0.006*	0.007*	—	—	—	0.022*	0.027*	0.009*
Got-3	0.040*	0.027*	0.033*	0.079*	0.030*	0.055*	0.034*	−0.006	0.012*
GDh	0.003	0.001	0.002*	—	—	—	—	—	—
SDh-1	0.013*	0.006*	0.009*	0.180*	0.203*	0.191*	−0.003	0.009	0.005
SDh-2	0.018*	0.009*	0.011*	0.019*	0.007*	0.011*	0.078*	0.014	0.047*
Average	0.032*	0.024*	0.026*	0.083*	0.097*	0.086*	0.027*	0.022*	0.023*

* $P < 0.05$.

observed heterozygosity in zygotes was 0.193, 0.194, and 0.178 for *P. tabulaeformis*, *P. densata*, and *P. yunnanensis*, respectively, and the corresponding numbers for the expected heterozygosity in zygotes were 0.205, 0.232, and 0.190 (Table 2). Comparison between the ovule and pollen pools revealed very similar gene diversity but slightly higher values in the ovules of each species (Table 2). In both the ovule and pollen, as well as in the zygote pools, *P. densata* had the highest gene diversity among the three pines.

Exact tests for departure from H-W equilibrium in the zygotes revealed that, among the eight populations of *P. tabulaeformis*, two populations (Pt-2 and Pt-5) showed significant deviation from H-W expectations (Table 2). All the three populations of *P. densata* and one population of *P. yunnanensis* (Py-1) showed deviation from H-W expectations. However, although statistically significant all the FI values were small and close to zero. The average FI for each population ranged between −0.019 and 0.043 in *P. tabulaeformis*, between 0.038 and 0.097 in *P. densata*, and between 0.007 and 0.021 in *P. yunnanensis*, at the seed stage (Table 2). All populations, except one (Pt-10), had positive fixation indices, indicating slight homozygote excess.

Single locus estimates of population differentiation (F_{st}) are presented in Table 3. In *P. tabulaeformis* all 12 polymorphic loci gave significant F_{st} values among the zygote populations. In *P. densata* 9 of the 10 polymorphic loci, and in *P. yunnanensis* 7 out of 11 loci, were significantly differentiated among the zygote populations. Mean F_{st} values in zygotes ranged from 0.023 (*P. yunnanensis*) to 0.086 (*P. densata*). Within each species, the average F_{st} values for ovule, pollen, and zygote populations were very similar, although slightly higher F_{st} values were found for ovules in *P. tabulaeformis* and *P. yunnanensis*. In *P. densata*, however, the pollen pools gave the highest F_{st} (0.097). Compared to *P. tabulaeformis* and *P. yunnanensis*, *P. densata* had noticeably higher F_{st} values for all the ovule, pollen, and zygote populations.

Linkage disequilibrium test: Results of the linkage disequilibria tests in ovules, pollen, and zygotes are summarized in Table 4. The numbers of significant loci associations among all the possible pair-wise loci combinations were low and similar in all three species. For example, in *P. tabulaeformis*, 9 and 1 out of 66 loci combinations were found in close association in zygote populations of Pt-1 and Pt-6, respectively. In the pollen pool

TABLE 4

Linkage disequilibrium in the investigated populations, showing the number of significant ($P < 0.05$) loci associations among all the possible pairwise combinations of loci

	Linkage disequilibrium		
Population	Ovule	Pollen	Zygote
Pt-1	5/55	3/55	9/66
Pt-2	1/36	3/36	3/36
Pt-3	1/45	1/36	1/45
Pt-5	4/43	3/36	2/43
Pt-6	3/45	0/55	1/66
Pt-8	6/44	2/55	3/55
Pt-9	3/45	3/55	2/55
Pt-10	3/66	3/55	3/66
Pd-1	5/45	1/45	5/45
Pd-2	5/36	2/36	2/36
Pd-7	2/45	0/36	3/45
Py-1	3/45	1/28	0/45
Py-2	2/36	1/44	3/54
Py-3	4/21	2/36	3/36

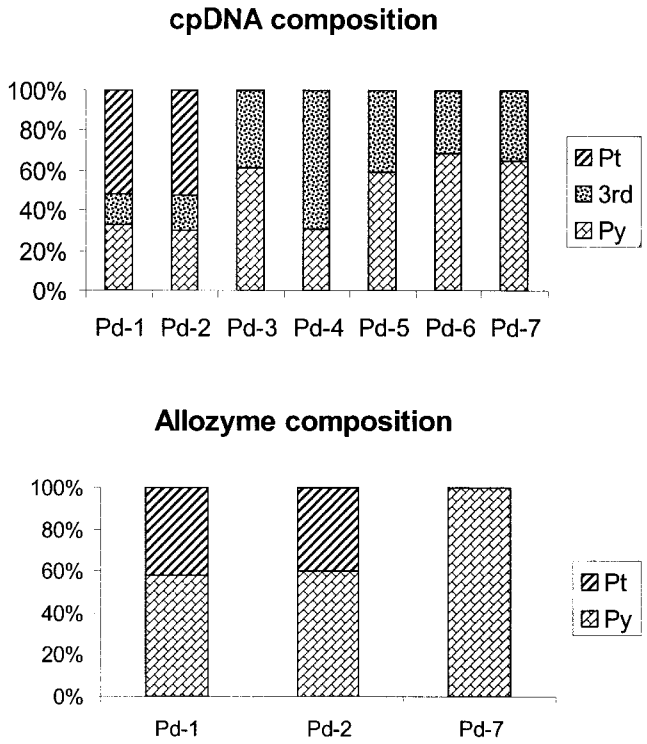

FIGURE 2.—Comparison of allozyme and cpDNA composition in *P. densata* populations.

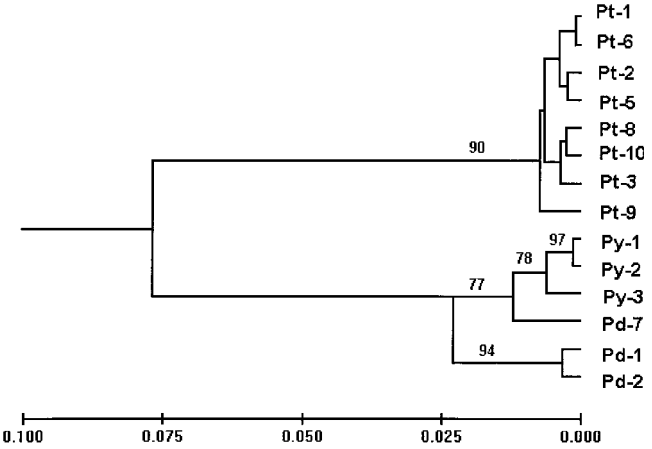

FIGURE 3.—UPGMA tree for the 14 populations of *P. tabulaeformis*, *P. densata*, and *P. yunnanensis* based on NEI's (1978) unbiased distance. Bootstrap support for main nodes is indicated on the corresponding branch.

of Pt-6, all 55 possible loci combinations were found to be in random association. Similarly, in the pollen pool of Pd-7 and zygote pool of Py-1, all loci were found in random association. Compared to *P. tabulaeformis* and *P. yunnanensis*, *P. densata* did not show a noticeably higher degree of disequilibrium in ovule, pollen, and zygote pools.

Continuing admixture of different gene pools could generate linkage disequilibrium due to different allelic frequencies between the parental populations. The power for detecting linkage disequilibrium is highest between polymorphic loci (*e.g.*, *Fes*, *Pgm*-1, *Sdh*-1, and *Got*-3). If continuing hybridization were the cause of linkage disequilibrium, we would expect that *P. densata* would have more disequilibrium than the parental populations. In addition, we would also expect to find that those loci that were highly differentiated between the parental populations (such as *Sdh*-1 and *Got*-3) would be more often involved in significant disequilibria than those that were equally polymorphic in all species, without differentiation between the supposed parental species (such as *Pgm*-1 and *Fes*, Table 1). In fact, this is not the case. Among the significant disequilibria of Pd-1 and Pd-2, these two groups of loci are equally frequently involved in the few significant pairs of disequilibria.

Gene admixture analysis and population clustering pattern: The allozyme composition in each population was analyzed for the respective gene admixtures from reference species of *P. yunnanensis* and *P. tabulaeformis*. Of the three populations of *P. densata*, populations Pd-1 and Pd-2 had similar amounts of apparent allozyme mixture from *P. yunnanensis* (57.8 and 60.4%, respectively, Figure 2), while population Pd-7 showed a very strong affinity to *P. yunnanensis* (100%, Figure 2). The sampled populations of the two reference species, *P. tabulaeformis* and *P. yunnanensis*, showed nearly no admixture from each other and appeared to be nearly pure representatives of their respective species (data not shown).

The UPGMA tree based on NEI's (1978) unbiased genetic distance for all populations revealed three distinct groups of populations with strong bootstrap support (Figure 3). The eight populations of *P. tabulaeformis* showed little differentiation and clustered as one group. Another major cluster was composed of the three *P. yunnanensis* populations and one population of *P. densata* (Pd-7). The third group comprised the other two populations of *P. densata* (Pd-1 and Pd-2) from the north-central part of its distribution.

DISCUSSION

Genetic structure of *P. densata* populations: In general, we observed similar patterns of gene diversity and population differentiation in ovule, pollen, and zygote pools of all three species. All the populations had fixation index very close to zero, suggesting they can be regarded as random mating populations. The partitioning of the gene diversity revealed that most of this diversity occurred within populations of each species and only 2.3–8.6% (F_{st}) of it resided among populations of each species. This range of F_{st} values is within the commonly observed degree of population differentiation for wind-pollinated conifers with continuous distributions (see HAMRICK 1983; EL-KASSABY 1991). Compared to *P. tabulaeformis* and *P. yunnanensis*, *P. densata*

had noticeably higher diversity measures and population differentiation. The present results confirmed the previous report, which also showed the intermediate allozyme frequencies and high genetic diversity in *P. densata*, as well as the slight homozygote excess in all populations of the three species (Wang *et al.* 1990).

Gene admixture analysis based on allozyme frequencies revealed that *P. tabulaeformis* and *P. yunnanensis* provided nearly equal contributions to the allozyme composition of populations Pd-1 and Pd-2. Population Pd-7, however, appeared to be much more strongly related to *P. yunnanensis*. On the UPGMA tree, Pd-7 grouped with the *P. yunnanensis* populations, although farther from them than the distance they spanned. These findings agree well with the results from a previous cpDNA analysis (Wang and Szmidt 1994). About 50% of the *P. tabulaeformis* and 30% of the *P. yunnanensis* cpDNA haplotypes were found in Pd-1 and Pd-2, whereas 65% of the *P. yunnanensis* and no *P. tabulaeformis* haplotypes were found in Pd-7 (Figure 2). However, in all three *P. densata* populations a third novel cpDNA haplotype was detected, which does not appear to have originated from any extant pines in Asia (Wang and Szmidt 1994). Thus, we cannot undertake further allozyme admixture analysis by reference to an additional taxon. Nevertheless, from both allozyme and cpDNA data, it seems that if there were a third species involved in the origin of *P. densata*, its allozyme composition was probably more similar to that of *P. yunnanensis* than *P. tabulaeformis*. Alleles characteristic of *P. densata* were rare and occurred at very low frequencies. It is possible that these alleles were contributed by the putative third species. The admixture of genetic material in *P. densata* populations is concordant with its observed higher gene diversity and F_{st} values as compared to *P. tabulaeformis* and *P. yunnanensis*. In addition, the topography of the region occupied by *P. densata* is very complex and presents considerable geographic barriers. Therefore, gene flow among populations from different regions is likely to have been limited for *P. densata*.

The heterogeneity in allozyme and cpDNA composition among *P. densata* populations suggests that different populations of *P. densata* have had different evolutionary histories. It seems that *P. yunnanensis* and another species were involved in the origin of *P. densata*, and *P. tabulaeformis* was involved locally. Thus, *P. densata* might be multiply derived or certain populations might have been in contact with one or the other parental species in the past, leading to the observed differences in allele and haplotype frequencies among its populations. Population Pd-7 showed very similar allozyme composition to *P. yunnanensis*. This might be suggestive of introgression with its parents in the marginal populations. Given the huge and complex geographic region occupied by *P. densata* it is not surprising to discover different factors involved in the evolution of different populations.

Speciation and maintenance of *P. densata*: Homoploid speciation in plants involves different processes and mechanisms from polyploid speciation (Grant 1981; McCarthy *et al.* 1995; Rieseberg 1997; Rieseberg *et al.* 1999; Buerkle *et al.* 2000). In a genetic simulation study, Buerkle *et al.* (2000) have shown that ecological and spatial isolation are required to achieve reproductive isolation, and thus the stabilization, of homoploid hybrid derivatives. The opportunity for ecological isolation can arise as a result of adaptation to extreme habitats that are not accessible to parental species. When a new ecological niche becomes available, speciation may occur as selection operates on the recombinant genotypes to form the new coadapted systems that come to characterize the new species (Carson 1975). Thus, many examples of homoploid hybrid speciation are associated with range extension, particularly in novel or variable environments (Anderson 1948; Lewontin and Birch 1966; Arnold 1997; Rieseberg 1997).

P. densata represents another example of diploid hybrid speciation that involves isolation and adaptation in an extreme habitat. *P. densata* differs from many other plant hybrids in its stability as a well-recognized species. It occupies a huge territory at high elevations that is not accessible to the putative parents and other pines in the region. In this environment, the putative parental species cannot grow. However, *P. densata* forms extensive, pure forests and regenerates well. In contrast to hybrid zones that are maintained by intensive current gene flow, the origin of *P. densata* is hypothesized to be ancient and to be related to the uplift of the Tibetan plateau, which would date back to at least 20 mya (Guan 1981; Ruddiman and Kutzbach 1991; Harrison *et al.* 1992; Ruddiman 1998). Significant increases in altitude of the Tibetan plateau are thought to have occurred ~10–8 mya (Harrison *et al.* 1992; Zhisheng *et al.* 2001). Drastic geographic and climatic changes in that period could have either brought distant species together or separated sympatric species and thus altered the flora (Florin 1963; Frenzel 1968; Guan 1990). This would provide an opportunity for gene exchange among species that were otherwise allopatric. The chronology of *P. densata* development, however, cannot be precisely defined in these geographic events. Nevertheless, the uplift of the plateau clearly created a new territory and opportunity for hybrids to develop.

Estimates of the degree of nonrandom associations among gene loci lend further support to the hypothesis that populations of *P. densata* have a stabilized hybrid nature. If significant gametic disequilibrium is detected among unlinked neutral markers, like allozymes, current gene flow would be expected to be one of the factors maintaining the hybrid genetic structure (Hedrick *et al.* 1978; Weir and Cockerham 1979; Lewontin 1988). In the absence of selection in random mating populations, the decay of linkage for random unlinked loci from the baseline of disequilibrium brought about

by gene exchange would take a few generations (Weir 1996). In this study, we analyzed both ovule and pollen pools. They can be regarded as representing different time events, since pollen represents the most current gene pool while the ovules represent the mother trees and therefore were generated some time earlier. We detected no significant disequilibrium in either gene pool, even in the populations close to sympatry with other species. The suspected hybrid had no more disequilibrium than the parental species, and little linkage disequilibrium was found between polymorphic loci, independent of whether they were highly differentiated between the parents or not. This would indicate that there is no significant ongoing gene exchange among the complex of species and that the hybrid populations are at an advanced stage of stabilization. The normal fertility and high seed production of *P. densata* indicate that this hybrid is maintained through normal sexual reproduction. *P. densata* may have become genetically stabilized through selection for certain hybrid genotypes that show increased fertility and viability in the species' ecologically extreme habitat.

The high proportion of mixed allozyme and cpDNA composition in *P. densata* populations does not indicate strong selection acting for or against a particular mating direction. Artificial hybridization experiments have shown that genetic barriers among the Asian pines in the subgenus Pinus are weak. Most of the species can cross with each other and produce viable seeds (Little and Righter 1965; Kormutak and Lanakova 1988). Furthermore, *P. tabulaeformis* and *P. yunnanensis* have been crossed successfully (Institute of Forest Genetics at Placerville, California, unpublished data). If there was no fertility difference between mating directions, both putative parental species could have acted as pollen donors in the initial hybridization event(s), and random mating within the population would maintain the cpDNA admixture in the initial hybrid population, even in the advanced generations, without producing the "capture" of a particular type as has been observed in some other plant hybrids (see Rieseberg and Soltis 1991). Other factors such as long generation time, large population size, and repeated hybridization in the past would also help to maintain the mixed genetic composition in *P. densata*.

It should be pointed out that our sampling in the close-sympatric regions is not extensive, due to the complex and inaccessible terrain. Detailed sampling of *P. densata* from areas to both the north and south linking to the *P. tabulaeformis* and *P. yunnanensis* populations would be needed to accurately assess the gene flow in the species complex. In addition, artificial hybridization involving *P. densata* with either of the putative parents would be of great value in clarifying several aspects related to this speciation event, such as the feasibility of backcrossing, the viability and fertility of the backcross progeny, the genetic structure of these progeny populations, and the potential differences in the evolutionary biology of organellar *vs.* nuclear genes. Furthermore, information on maternally inherited mtDNA markers, in combination with the paternal cpDNA and biparental nuclear genome data, would shed new light on the evolution of *P. densata*.

This study was supported by grants from the Chinese Academy of Sciences, the National Natural Science Foundation of China (NSFC-30070058), the Swedish Council for Forestry and Agricultural Research, and the Swedish International Development Agency.

LITERATURE CITED

Anderson, E., 1948 Hybridization of the habitat. Evolution **2:** 1–9.
Anderson, E., and G. L. Stebbins, 1954 Hybridization as an evolutionary stimulus. Evolution **8:** 378–388.
Arnold, M. L., 1992 Natural hybridization as an evolutionary process. Annu. Rev. Ecol. Syst. **23:** 237–261.
Arnold, M. L., 1997 *Natural Hybridization and Evolution*. Oxford University Press, New York.
Asmussen, M. A., J. Arnold and J. C. Avise, 1987 Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. Genetics **115:** 755–768.
Barton, N. H., and G. M. Hewitt, 1985 Analysis of hybrid zones. Annu. Rev. Ecol. Syst. **16:** 113–148.
Brown, A. H. D., 1975 Sample sizes required to detect linkage disequilibrium between two or three loci. Theor. Popul. Biol. **8:** 184–201.
Buerkle, C. A., R. J. Morris, M. A. Asmussen and L. H. Rieseberg, 2000 The likelihood of homoploid hybrid speciation. Heredity **84:** 441–451.
Carson, H. L., 1975 The genetics of speciation at the diploid level. Am. Nat. **109:** 83–92.
Curie-Cohen, M., 1982 Estimates of inbreeding in a natural population: a comparison of sampling properties. Genetics **100:** 339–358.
Draper, N., and H. Smith, 1966 *Applied Regression Analysis*. John Wiley & Sons, New York.
El-Kassaby, Y. A., 1991 Genetic variation within and among conifer populations: review and evaluation of methods, pp. 61–76 in *Biochemical Markers in the Population Genetics of Forest Trees*, edited by S. Fineschi, M. E. Malvolti, F. Cannata and H. H. Hattemer. SPB Academic Publishing bv, The Hague.
Elston, R. C., 1971 The estimation of admixture in racial hybrids. Ann. Hum. Genet. **35:** 9–17.
Ennos, R. A., 1994 Estimating the relative rates of pollen and seed migration among plant populations. Heredity **72:** 250–259.
Ernst, S. G., J. W. Hanover and D. E. Keathley, 1990 Assessment of natural interspecific hybridization of blue and Engelmann spruce in southwestern Colorado. Can. J. Bot. **68:** 1489–1496.
Florin, R., 1963 The distribution of conifer and taxad genera in time and space. Acta Horti Bergiani **20:** 122–312.
Frenzel, B., 1968 The Pleistocene vegetation of northern Eurasia. Science **161:** 637–649.
Goudet, J., 1995 FSTAT (Version 1.2): a computer program to calculate F-statistics. J. Hered. **86:** 485–486.
Goudet, J., M. Raymond, T. Demeeus and F. Rousset, 1996 Testing differentiation in diploid populations. Genetics **144:** 1933–1940.
Grant, V., 1981 *Plant Speciation*. Columbia University Press, New York.
Guan, C.-T., 1981 Fundamental features of the distribution of *Coniferae* in Sichuan. Acta Phytotaxonom. Sinica **11:** 393–407. (in Chinese).
Guan, Z., 1990 *The Geography of Conifers in Sichuan*. Sichuan Science and Technology Press, Chengdu, China. (in Chinese).
Guo, S. W., and E. A. Thompson, 1992 Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics **48:** 361–372.
Haldane, J. B. S., 1954 An exact test for randomness of mating. J. Genet. **52:** 631–635.
Hamrick, J. L., 1983 The distribution of genetic variation within and among natural plant populations, pp. 335–348 in *Genetics*

and Conservation, edited by C. M. Schonewald-Cox, S. M. Chambers, B. MacBryde and W. L. Thomas. Benjiamin/Cumming Publishing, Menlo Park, CA.

Harrison, T. M., P. Copeland, W. S. F. Kidd and A. Yin, 1992 Raising Tibet. Science **255:** 1663–1670.

Hedrick, P., S. Jain and L. Holden, 1978 Multilocus systems in evolution. Evol. Biol. **11:** 101–184.

Kormutak, A., and M. Lanakova, 1988 Biochemistry of reproductive organs and hybridological relationships of selected species of pines (*Pinus* sp.), pp 1–119 in *Acta Dendrobiologica*. VEDA Publishing House of the Slovak Academy of Sciences, Bratislava, Slovakia.

Lewontin, R. C., 1988 On measures of gametic disequilibrium. Genetics **120:** 849–852.

Lewontin, R. C., and L. C. Birch, 1966 Hybridization as a source of variation for adaptation to new environments. Evolution **20:** 315–336.

Little, Jr., E. L., and F. I. Righter, 1965 Botanical descriptions of forty artificial pine hybrids. U.S. Dep. Agric. U.S. For. Serv. Tech. Bull. **1345:** 1–47.

Louis, E. J., and E. R. Dempster, 1987 An exact test for Hardy-Weinberg and multiple alleles. Biometrics **43:** 805–811.

McCarthy, E. M., M. A. Asmussen and W. W. Anderson, 1995 A theoretical assessment of recombinational speciation. Heredity **74:** 502–509.

Miller, M. P., 1997 Tools for population genetics analysis (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Mirov, N. T., 1967 *The Genus Pinus*. The Ronald Press Company, New York.

Muona, O., and A. E. Szmidt, 1985 A multilocus study of natural populations of *Pinus sylvestris*. Lect. Notes Biomath. **60:** 226–240.

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

Paige, K. N., W. C. Capman and P. Jennetten, 1991 Mitochondrial inheritance patterns across a cottonwood hybrid zone: cytonuclear disequilibria and hybrid zone dynamics. Evolution **45:** 1360–1369.

Raymond, M., and F. Rousset, 1995 GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. **86:** 248–249.

Rieseberg, L. H., 1997 Hybrid origins of plant species. Annu. Rev. Ecol. Syst. **28:** 359–389.

Rieseberg, L. H., and S. E. Carney, 1998 Plant hybridization. New Phytol. **140:** 599–624.

Rieseberg, L. H., and D. E. Soltis, 1991 Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants **5:** 65–84.

Rieseberg, L. H., J. Whitton and K. Gardner, 1999 Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics **152:** 713–727.

Roberts, D. F., and R. W. Hiorns, 1965 Methods of analysis of the genetic composition of a hybrid population. Hum. Biol. **37:** 38–43.

Ruddiman, W., 1998 Early uplift in Tibet? Nature **394:** 723–725.

Ruddiman, W. F., and J. E. Kutzbach, 1991 Plateau uplift and climatic change. Sci. Am. **3:** 42–50.

Sigurgeirsson, A., 1992 Insights into the evolution of Picea inferred from chloroplast DNA. Ph.D. Thesis, Swedish University of Agricultural Sciences, Faculty of Forestry, Department of Forest Genetics and Plant Physiology, Umeå, Sweden. ISBN 91-576-4617-1.

Stebbins, G. L., 1969 The significance of hybridization for plant taxonomy and evolution. Taxon **18:** 26–35.

Sutton, B. C. S., D. J. Flanagan, J. R. Gawley, C. H. Newton, D. T. Lester *et al.*, 1991 Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. Theor. Appl. Genet. **82:** 242–248.

Swofford, D. L., and R. B. Selander, 1981 BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. **72:** 281–283.

Szmidt, A. E., Y. A. El-Kassaby, A. Sigurgeirsson, T. Alden, D. Lindgren *et al.*, 1988 Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. Theor. Appl. Genet. **76:** 841–845.

Wagner, D. B., Z.-X. Sun, D. R. Govindaraju and B. P. Dancik, 1991 Spatial patterns of chloroplast DNA and cone morphology variation within populations of a *Pinus banksiana-Pinus contorta* sympatric region. Am. Nat. **138:** 156–170.

Wang, C. W., 1961 *The Forests of China*. Maria Moors Cabot Foundation Publ., Harvard University, Cambridge, MA.

Wang, X.-R., and A. E. Szmidt, 1990 Evolutionary analysis of *Pinus densata* (Masters), a putative Tertiary hybrid. 2. A study using species-specific chloroplast DNA markers. Theor. Appl. Genet. **80:** 641–647.

Wang, X.-R., and A. E. Szmidt, 1994 Hybridization and chloroplast DNA variation in a *Pinus species* complex from Asia. Evolution **48:** 1020–1031.

Wang, X.-R., A. E. Szmidt, A. Lewandowski and Z.-R. Wang, 1990 Evolutionary analysis of *Pinus densata* (Masters), a putative Tertiary hybrid. 1. Allozyme variation. Theor. Appl. Genet. **80:** 635–640.

Weir, B. S., 1996 *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.

Weir, B. S., and C. C. Cockerham, 1979 Estimation of linkage disequilibrium in randomly mating populations. Heredity **42:** 105–111.

Weir, B. S., and C. C. Cockerham, 1984 Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358–1370.

Wheeler, N. C., and R. P. Guries, 1987 A quantitative measure of introgression between lodgepole and jack pines. Can. J. Bot. **65:** 1876–1885.

Whittemore, A. T., and B. A. Schaal, 1991 Interspecific gene flow in oaks. Proc. Natl. Acad. Sci. USA **88:** 2540–2544.

Wu, C. L., 1956 The taxonomic revision and phytogeographical study of Chinese pines. Acta Phytotaxonom. Sinica **5:** 131–163. (in Chinese).

Yang, X.-Y., 1987 Nuclear type analysis of *Pinus densata*. J. Northwest. Coll. For. **2:** 51–54. (in Chinese).

Zhisheng, A., J. E. Kutzbach, W. L. Prell and S. C. Porter, 2001 Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since late Miocene times. Nature **411:** 62–66.

Communicating editor: D. Charlesworth