Heredity (2003) 91, 510–518

www.nature.com/hdy

Genetic diversity of nuclear and mitochondrial genomes in *Pinus parviflora* Sieb. & Zucc. (Pinaceae) populations

N Tani[1], K Maruyama[2,4], N Tomaru[3], K Uchida[2], M Araki[2], Y Tsumura[1], H Yoshimaru[1] and K Ohba[2]
[1]*Department of Forest Genetics, Forestry and Forest Products Research Institute, Matsunosato, Tsukuba, Ibaraki 305-8687, Japan;* [2]*Institute of Agriculture and Forestry, University of Tsukuba, Tennodai, Tsukuba, Ibaraki 305-8572, Japan;* [3]*Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan*

Genetic diversities of the nuclear and mitochondrial genomes in *Pinus parviflora* were studied in 16 populations, which were distributed across most of the species' range in Japan. Six mitochondrial DNA haplotypes were identified among the 16 populations. The intrapopulation diversity of allozymes was similar to that of other endemic woody species ($H_S = 0.259$). Although *P. parviflora* is distributed in discrete populations, differentiation between these was very low ($G_{ST} = 0.044$). In addition, the extent of genetic differentiation between two varieties (var. *pentaphylla* and var. *parviflora*) was extremely low ($G_{VT} = 0.001$). Intrapopulation diversity of mitochondrial DNA was also very low ($H_S = 0.098$), but population differentiation was high ($G_{ST} = 0.863$). Moreover, the distribution of haplotypes reflected the taxonomic differences between *P. parviflora* var. *pentaphylla* and var. *parviflora*. The populations of var. *pentaphylla* and var. *parviflora* contained different haplotypes. Differing modes of inheritance may account for the differences in nuclear and mitochondrial genetic diversity.
Heredity (2003) **91,** 510–518. doi:10.1038/sj.hdy.6800349

Keywords: allozyme; mitochondrial DNA; RFLP; genetic differentiation; maternal inheritance; phylogeography

Introduction

Pinus parviflora Sieb. & Zucc. (subsection *Strobi*) is distributed as discrete populations in montane and subalpine zones of Japan. The species occupies the northern part of the range of the *P. parviflora* complex, which extends from Vietnam to Hokkaido Island in Japan, along the east coast of Eurasia (Mirov, 1967). Two varieties of *P. parviflora*, var. *pentaphylla* (Mayr) Henry and var. *parviflora*, are generally recognized (Yagashira, 1964; Ohwi, 1978). Although many taxonomic studies of the two varieties have been conducted (e.g. Mayr, 1890; Wilson, 1916; Miyabe and Kudo, 1921; Makino and Nemoto, 1931), the taxonomic differences have not been rigorously defined. In addition, natural hybridization events between var. *pentaphylla* and *P. pumila* were observed, and organelle capture events were detected at some var. *pentaphylla* populations using molecular markers (Watano *et al*, 1995, 1996; Senjo *et al*, 1999). Thus, subsection *Strobi* has many taxonomic problems, especially, in eastern Asia (Mirov, 1967).

Although, in Pine species, nuclear genomes are inherited biparentally and chloroplast genomes paternally, mitochondrial genomes are inherited maternally (Neale and Sederoff, 1989; Wagner *et al*, 1991). In previous studies of some Pinaceae species, high levels of population differentiation for mitochondrial DNA (mtDNA), compared with nuclear-encoded allozymes, have been observed. The large mtDNA genetic differentiation was inferred to be mainly caused by limited seed dispersal (due to maternal inheritance) and reduced effective population size (eg, Dong and Wagner, 1993; Strauss *et al*, 1993; Tsumura and Suyama, 1998). Therefore, mtDNA polymorphisms would be an efficient tool to address taxonomic problems at lower taxonomic levels (Palmer, 1992).

In this paper, we focused on the genetic diversity of nuclear-encoded allozymes and mtDNA to determine the influence of such factors on the genetic composition of this species, on allelic (allozyme) and haplotype (mtDNA) geographical distributions to infer phylogeography and population history of the species, and to resolve the taxonomical controversy related to these two varieties.

Materials and methods

Collection of samples

Current-year needles were collected from 467 trees, representing 11 natural populations of *P. parviflora* var. *pentaphylla* and 190 trees from five natural populations of var. *parviflora* (Figure 1). The sampled populations were distributed across most of the natural ranges of the two varieties in Japan. The needles were individually collected with more than 20 m between sampled trees, to avoid sampling of half-siblings. No attention was paid to the age or size of the trees. Samples were stored at −80°C, prior to analysis of allozymes and DNA isolation.

Correspondence: N Tani, Department of Forest Genetics, Forestry and Forest Products Research Institute, Matsunosato, Tsukuba, Ibaraki 305-8687, Japan. E-mail: ntani@ffpri.affrc.go.jp
[4]*Current address: Nippon Paper Industries, Yuraku-cho, Chiyoda-ku, Tokyo 100-0006, Japan.*
Received 14 January 2002; accepted 17 June 2003

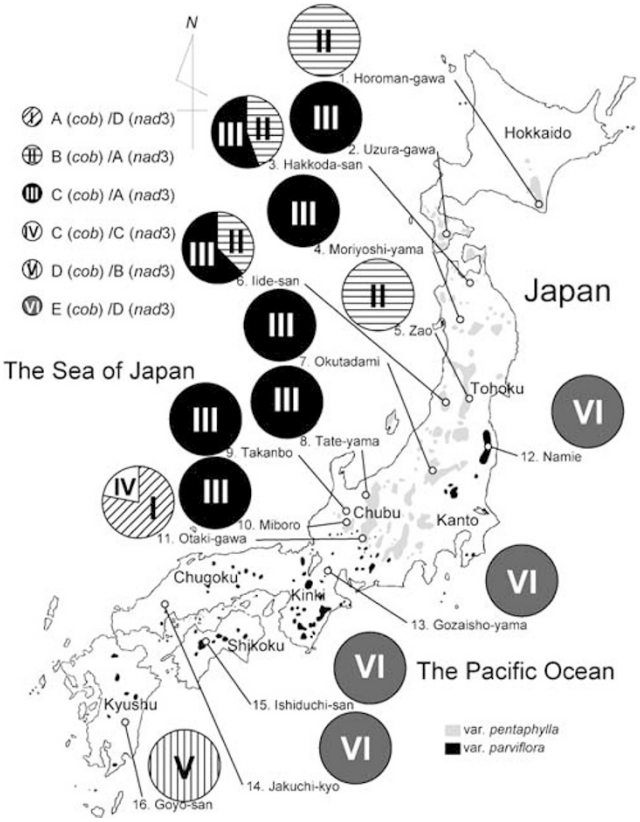

Figure 1 Distribution of *Pinus parviflora* var. *pentaphylla* and var. *parviflora*, and the geographical distribution of mtDNA haplotypes in the 16 populations of *P. parviflora*. Circles indicate the locations of the populations surveyed and show the haplotype frequencies.

Electrophoresis of allozymes

Allozyme analysis was carried out following the procedure described by Tsumura and Ohba (1993). In all, 10 enzymes were stained consistently, shikimate dehydrogenase (*Shd*-1, *Shd-2*; EC 1.1.1.25), glycerate-2-dehydrogenase (*G2d*; EC 1.1.1.29), malate dehydrogenase (*Mdh-1*; EC 1.1.1.37), 6-phosphogluconate dehydrogenase (*6pg*; EC 1.1.1.44), glutamate dehydrogenase (*Gdh*; EC 1.4.1.2), diaphorase (*Dia*; EC 1.6.4.3), glutamate oxaloacetate transaminase (*Got-1*; EC 2.6.1.1), phosphoglucomutase (*Pgm-2*; EC 2.7.5.1), alanine aminopeptidase (*Aap-2*; EC 3.4.11.1), and triosephosphate isomerase (*Tpi*; EC 5.3.1.1). The genotypes at all loci were deduced using known subunit structures and cellular compartmentalization of the enzymes (Weeden and Wendel, 1989).

DNA extraction and RFLP analysis

Total genomic DNA was extracted following a modification of the CTAB method (Murray and Thompson, 1980). Primers for the PCR-amplified probes were designed using a complete nucleotide sequence of mtDNA from *Marchantia polymorpha* (Oda *et al*, 1992). The primer sequences for the analysis of mitochondrial genetic diversity in all populations were: *cob* gene (forward 5′ CCC CGA GCA ATC TTA GTT AT 3′, reverse 5′ GGA GAA ATT TGT CAA ATA GT 3′) and *nad*3 gene (forward 5′ TCC CAC TTG GTG TTC CTT TT 3′, reverse 5′ ATT TAG ATC TGC CCC TTT TT 3′). The probes were labeled with digoxygenin-dUTPs (DIG) for PCR amplification (Roche Diagnostics). In the preliminary screening, total genomic DNA was digested using 18 restriction enzymes, then electophoretically separated on 0.8% agarose gel and hybridized with two PCR-amplified gene probes. In all, 47 (var. *parviflora*) and 123 (var. *pentaphylla*) individuals from 16 natural populations of *P. parviflora* (*c.* 10 trees per population) out of 657 samples were subsequently analyzed using two restriction enzyme/probe combinations (*Xba*I/*cob* and *Sac*I/*nad3*) that detected polymorphisms in the preliminary screening. Length mutations between *P. pumila* and *P. parviflora* var. *pentaphylla* were detected using agarose electrophoresis. An intron between exons B and C of the mitochondrial gene encoding subunit 1 of NADH dehydrogenase was amplified by PCR, using primers designed by Demesure *et al* (1995). This region can be used to classify the mitochondrial type of the two species (Watano *et al*, 1996; Senjo *et al*, 1999).

Statistical analysis

The number of alleles per locus, the effective number of alleles per locus, the proportion of polymorphic loci (95% criteria), and the observed and expected heterozygosities were calculated using the allozyme data. The geographical distribution of rare alleles (defined as alleles with an average frequency lower than 0.05) was also investigated. For both the allozyme and the mitochondrial data, gene diversity statistics were estimated following the methods of Nei (1973) and Nei and Chesser (1983), using FSTAT ver. 2.9.1 (Goudet, 1995). In addition, the fixation index, θ (Wright, 1965), and its standard deviation were estimated, following the method of Weir (1990). The genetic differentiation of nuclear and mitochondrial genomes for the two varieties of *P. parviflora* was estimated using the equation:

$$G_{VT} = (H_T - H_V)/H_T$$

where H_T is the gene diversity of the whole population, H_V is the gene diversity of all the populations of both varieties, and G_{VT} represents a coefficient of gene differentiation between the two varieties.

The number of migrants per generation (N_em) was estimated on the basis of nuclear and mitochondria genomes, following two equations: $N_em = 1/4(1/G_{ST(N)} - 1)(s-1)/s$ (for the nuclear genome; Takahata and Nei, 1984) and $N_{eo}m_e = 1/2(1/G_{ST(M)} - 1)(s-1)/s$ (for the mitochondrial genome; Birky *et al*, 1989), where $G_{ST(N)}$ and $G_{ST(M)}$ are the coefficients of gene differentiation among the populations, for nuclear and mitochondrial genomes, respectively, and the variable s is the number of populations. Since the actual number of populations of the species was extremely large, we assume that $(s-1)/s$ provides an unbiased estimate. Mutation rates were not considered when estimating the values of N_em.

A phenogram was constructed to demonstrate the mitochondrial relationships between populations. The neighbor-joining (NJ) method was used (Saitou and Nei, 1987), based on a D_A distance matrix (Nei *et al*, 1983), calculated from the haplotype frequencies. Bootstrap estimates, generated by the DISPAN computer program (Ota, 1993), were used to test the reliability of the trees. Nei *et al* (1983) noted that the D_A distance measure is accurate in demonstrating the true topology of an

evolutionary tree, when it is reconstructed from allele frequency data.

Results

Allozyme variation

Eleven putative loci, with a total of 37 alleles, were determined by analyzing the 10 enzyme systems. The averages of the statistics of genetic variation within the population of var. *pentaphylla* were slightly higher than those of var. *parviflora* (Table 1).

Fixation indices (F_{IS}) were calculated for loci where no null alleles were detected and where H_e exceeded 0.1. The mean value of F_{IS} for the populations was 0.115, with values of F_{IS} for individual populations ranging from 0.003 (Takanbo-yama) to 0.204 (Goyo-san). F_{IS} values of three populations (Tate-yama, Miboro, and Goyo-san) deviated significantly (0.01) from expectations based on the Hardy–Weinberg principle (Table 1).

The gene diversity in the combined populations (H_T) and the average gene diversity within individual populations (H_S) for allozyme variation were 0.272 and 0.259, respectively. Therefore, the coefficient of gene differentiation among populations (G_{ST}) was 0.044. G_{ST} values for each locus ranged from 0.006 (*Tpi*) to 0.123 (*Got-1*; Table 2). The average gene diversities within the var. *pentaphylla* and var. *parviflora* populations were 0.283 and 0.251, respectively. The coefficient of gene differentiation between the two varieties (G_{VT}) was 0.001 (Table 2). The coefficients of gene differentiation between the two varieties (G_{VT}) for each locus ranged from 0.000 (*Shd-1*, *6Pg*, *Gdh*, *Dia*, and *Aap-2*) to 0.013 (*Got-1*).

Although genetic variation within individual populations was high, the rare alleles (18 out of 37) were unevenly distributed between two regions: northern Tohoku and southern Chubu. The average number of rare alleles was 6.1 in the north (sites 1, 2, 3, 4, 5, 6, 7, and 12), but was only 2.1 in the south (sites 8, 9, 10, 11, 13, 14, 15, and 16). Although 18 rare alleles were identified in both regions, nine more were only in the north. Furthermore, alleles *Shd-2-a*, *6Pg-e*, and *Tpi-c* were only found in the Zao population. In contrast, no rare alleles were detected in the Gozaisho-yama and Goyo-san populations (Table 1).

Mitochondrial DNA variation

We surveyed polymorphisms at the two mtDNA regions (*cob*, *nad*3) using southern hybridization for the preliminary screening. Of the 36 enzyme/probe combinations, 15 showed polymorphisms; the *Xba*I/*cob* and *Sac*I/*nad*3 combinations were selected because they displayed the clearest autoradiographs (Figure 2). Five haplotypes were detected with *Xba*I/*cob* and four with *Sac*I/*nad*3. Since both pairs of the restriction enzymes and probes simultaneously assessed each haplotype, we combined the results of the *Xba*I/*cob* and *Sac*I/*nad*3 analyses. The combination of two haplotypes was defined as the mitochondrial type. As a result, six mitochondrial types were recognized (Table 3).

Mitochondrial variation within populations was only observed at Hakkoda-san, Iide-san, and Otaki-gawa, where there were multiple haplotypes (Table 3, Figure 1). Individuals that had acquired *P. pumila* mitochondria were identified by a 2.2 kbp PCR fragment of the intron between exons B and C of the mitochondrial gene encoding subunit 1 of NADH dehydrogenase which had previously been confirmed as originating from *P. pumila* mitochondria (Watano *et al*, 1996; Senjo *et al*, 1999); 2.2 kbp PCR fragments were amplified from all individuals with mitochondrial Type II for the *cob*/*nad*3 combination. PCR fragments of 2.6 kbp were amplified for all the other individuals of var. *pentaphylla* with mitochondrial Type III (Table 3). Mitochondrial Type III was sympatrically distributed with mitochondrial Type II, captured from *P. pumila*, in two var. *pentaphylla* populations (Hakkoda-san and Iide-san). Furthermore, in the Horoman-gawa and Zao populations, *P. pumila* type mitochondria had completely replaced the var. *pentaphylla* type (Figure 1).

The *pentaphylla* and *parviflora* varieties contained different mitochondrial types (Figure 1). Although the trees in Otaki-gawa population belong to var. *pentaphylla*, the population was located very close to the border of the two varieties. Although two mitochondrial types observed in Otaki-gawa population (demonstrated by the combination of the two RFLPs using *cob* and *nad*3 gene probes) were unique, one (mitochondrial Type IV) possessed a unique *Sac*I/*nad*3 haplotype (C) and the C *Xba*I/*cob* haplotype found in var. *pentaphylla*. The other (mitochondrial Type I) also possessed a unique *Xba*I/*cob* haplotype (A) and the D *Sac*I/*nad*3 haplotype observed in var. *parviflora* (Table 3).

Restriction fragments associated with *cob* and *nad*3 genes were highly variable in this species ($H_T = 0.708$). The average gene diversity within populations (H_S) was 0.092. Therefore, the coefficient of gene differentiation among populations (G_{ST}) was 0.870. The gene differentiation between the two varieties of mtDNA was 0.401. In contrast to the allozyme variation, the gene differentiation among populations and between the two varieties of mtDNA was considerably larger (Table 2).

The NJ phenetic tree (Figure 3) shows five main groups (bootstrap value $\geqslant 75\%$). The first group included four populations; the Horoman-gawa and Zao populations displayed no mtDNA polymorphisms, and only contained mitochondrial type II, while the Iide-san and Hakkoda-san populations contained both mitochondrial types II and III. The second group of the tree contained six populations, all with mitochondrial type III. The third group of the tree contained only one population, which had two unique mitochondrial types (I and IV). Four populations were present in the fourth group, all of which possessed mitochondrial type VI. The fifth group contained only the Goyo-san population, which had the unique mitochondrial type V.

Discussion

Population genetic aspects of mtDNA and nuclear-encoded allozymes

In contrast to the genetic diversity of allozymes, mtDNA was strongly differentiated between populations. The mode of mtDNA inheritance, in pine species, is thought to be maternal (Neale and Sederoff, 1989; Wagner *et al*, 1991). The high population differentiation for mtDNA in *P. parviflora* is probably due to this mode of inheritance and to the species' limited seed-dispersal ability. This agrees with previous studies of the population genetics

Table 1 Genetic variation (standard errors in parentheses) and fixation index of nuclear-encoded allozymes within the 16 Japanese *Pinus parviflora* populations

Variety	*Population number and name*	*Sample size*	*Proportion of polymorphic loci 95% criterion (P_1)*	*Average number*			*Average heterozygosity*		*Fixation index*[b]
				Number of rare alleles[a]	*Alleles per locus (N_a)*	*Effective alleles per locus (N_e)*	*Observed (H_o)*	*Expected (H_e)*	*(F_{IS})*
Pentaphylla	1 Horoman-gawa	52	54.5	4	2.09 (0.28)	1.46 (0.19)	0.194 (0.065)	0.228 (0.073)	0.143
	2 Uzura-gawa	48	63.6	7	2.36 (0.28)	1.55 (0.19)	0.252 (0.065)	0.269 (0.074)	0.067
	3 Hakkoda-san	34	63.6	2	1.91 (0.21)	1.55 (0.19)	0.244 (0.070)	0.272 (0.075)	0.108
	4 Moriyoshi-yama	44	63.6	6	2.27 (0.24)	1.50 (0.19)	0.246 (0.068)	0.251 (0.071)	0.026
	5 Zao	44	81.8	9	2.55 (0.21)	1.71 (0.21)	0.303 (0.058)	0.339 (0.068)	0.113
	6 Iide-san	44	54.5	9	2.55 (0.34)	1.54 (0.20)	0.238 (0.067)	0.261 (0.075)	0.095
	7 Okutadami	50	63.6	6	2.27 (0.24)	1.50 (0.17)	0.227 (0.053)	0.262 (0.067)	0.142
	8 Tate-yama	43	63.6	5	2.00 (0.19)	1.57 (0.14)	0.258 (0.067)	0.306 (0.071)	0.162**[c]
	9 Takanbo-yama	32	63.6	5	2.00 (0.27)	1.52 (0.18)	0.265 (0.072)	0.266 (0.072)	0.003
	10 Miboro	44	72.7	5	2.00 (0.19)	1.54 (0.17)	0.241 (0.053)	0.284 (0.069)	0.156**
	11 Ohtaki-gawa	32	54.5	5	2.00 (0.27)	1.39 (0.13)	0.188 (0.049)	0.228 (0.065)	0.186
	Average for var. *pentaphylla*		63.6	5.7	2.18 (0.25)	1.53 (0.18)	0.241 (0.062)	0.270 (0.071)	0.109
Parviflora	12 Namie	50	63.6	6	2.27 (0.36)	1.57 (0.19)	0.235 (0.066)	0.274 (0.078)	0.144
	13 Gozaisho-yama	31	54.5	0	1.73 (0.24)	1.34 (0.12)	0.179 (0.061)	0.199 (0.065)	0.106
	14 Jakuchi-kyo	37	54.5	2	1.91 (0.28)	1.47 (0.18)	0.218 (0.065)	0.234 (0.075)	0.074
	15 Ishiduchi-san	38	54.5	3	2.00 (0.23)	1.44 (0.17)	0.204 (0.062)	0.227 (0.070)	0.108
	16 Goyo-san	34	54.5	0	1.73 (0.24)	1.49 (0.17)	0.198 (0.063)	0.248 (0.077)	0.204**
	Average for var. *parviflora*		56.3	2.2	1.92 (0.27)	1.46 (0.17)	0.207 (0.063)	0.236 (0.073)	0.127
	Average		61.33	4.6	2.10 (0.25)	1.51 (0.17)	0.231 (0.063)	0.259 (0.072)	0.115

[a]The number of rare alleles was defined as alleles with an average frequency lower than 0.05.
[b]The statistical significance of the deviation from Hardy–Weinberg expectations was tested by the χ^2 test (Li and Horvitz, 1953).
[c]The levels of significance: $*P<0.05$; $**P<0.01$.

Table 2 Gene diversity statistics for nuclear-encoded allozymes and mtDNA among the 16 Japanese populations and two variaties of *Pinus parviflora*

Locus	*Gene diversity in all populations* (H_T)	*Gene diversity within each two variaties* (H_V)	*Gene diversity within populations* (H_S)	*Gene differentiation between the two variaties* (G_{VT})[a]	*Gene differentiation between populations* (G_{ST})[a]	*Relationship between alleles of different populations* (θ)[b]
Sdh-1	0.066	0.066	0.063	0.000	0.050	0.051 (0.022)
Sdh-2	0.057	0.057	0.055	0.008	0.037	0.048 (0.018)
G2d	0.468	0.467	0.456	0.002	0.024	0.025 (0.013)
Mdh-1	0.488	0.468	0.459	0.003	0.056	0.064 (0.032)
6pg	0.288	0.287	0.272	0.000	0.053	0.064 (0.029)
Gdh	0.331	0.331	0.314	0.000	0.052	0.053 (0.020)
Dia	0.660	0.660	0.639	0.000	0.032	0.029 (0.014)
Got-1	0.079	0.078	0.068	0.013	0.123	0.196 (0.108)
Pgm-2	0.035	0.035	0.031	0.007	0.111	0.191 (0.112)
Aap-2	0.502	0.502	0.482	0.000	0.040	0.040 (0.012)
Tpi	0.016	0.016	0.016	0.003	0.006	0.007 (0.003)
Average	0.272	0.271	0.259	0.001	0.044	0.046 (0.008)
mtDNA	0.708	0.425	0.092	0.401	0.870	0.831 (0.091)

[a]Nei's G_{ST} (G_{VT} is the same value as G_{ST}) unbiased for sample size and population number (Nei and Chesser, 1983).
[b]θ is estimated. The standard deviation over alleles or haplotypes, derived by jack-knifing, is noted in parentheses (Weir and Cockerham, 1984).

Figure 2 Southern hybridization patterns of *Xba*I fragments with the *cob* probe (**a**), *Sac*I fragments with the *nad*3 probe (**b**). Numbers indicate the size of the fragments detected. Lane M refers to the molecular size marker (lambda/*Hind*III digest DNA).

of conifers (Dong and Wagner, 1993; Strauss *et al*, 1993; Latta and Mitton, 1997; Aagaard *et al*, 1998; Tsumura and Suyama, 1998; Wu *et al*, 1998). These results are in accordance with the theory that DNA spatial structure variation in organelles should be greater than the nuclear variation (Birky *et al*, 1989; Petit *et al*, 1993).

Under migration-drift equilibrium in a finite island model (Wright, 1943), the parameters determining genetic differentiation are the effective population size, the migration rate, the number of subpopulations, and the mutation rate. When mutation rates of allozymes and mtDNA are similar, reductions in both the effective population size and migration rate will enhance the population differentiation of mtDNA relative to nuclear markers. This is because of the influence of maternal inheritance on migration rate and is a consequence of the effective halving of the number of genes in the haploid genome compared with the diploid genome (Birky, 1988; Petit *et al*, 1993).

We also estimated the number of migrants exchanged per generation of the species in a finite island model, under migration-drift equilibrium, with negligible mutation rates. The estimate for the nuclear genome (N_em) affected by both pollen flow and seed dispersal, based on allozyme data, was 5.4. The estimate for the mitochondrial genome ($N_{eo}m_e$), affected solely by seed dispersal, based on RFLP, was 0.075. These estimates indicate that pollen flow is a more effective method of gene exchange between populations than seed dispersal; the estimate for N_em being about 73 times higher than that for $N_{eo}m_e$. Thus, allozyme and mtDNA variations can be successfully used to measure gene exchange by both pollen flow and seed dispersal.

Phylogeographic pattern and a consideration of population history

The phenetic tree derived from mtDNA variation divided the 16 populations into five groups, based on a high bootstrap value (Figure 3). The clustering may be linked to the recolonization history of the species following the last glaciation, as discussed in previous studies of organelle genome diversity (Dumolin-Lapègue

Table 3 Frequency of each haplotype of mitochondrial DNA in the 16 Japanese populations of *P. parviflora* var. *pentaphylla* and var. *parviflora*

Population number and name	*Sample size*	*XbaI/cob*					*SacI/nad3*				*cob/nad3 combination*						*nad1 exon B-C*[a]	
		A (6.3 kbp)	*B (6.8 kbp)*	*C (8.8 kbp)*	*D (9.4 kbp)*	*E (16.7 kbp)*	*A (5.2 kbp)*	*B (8.2 kbp)*	*C (7.9/13.6 kbp)*	*D (12.8 kbp)*	*I (A/D)*	*II (B/A)*	*III (C/A)*	*IV (C/C)*	*V (D/B)*	*VI (E/D)*	*P. pumila type (2.2 kbp)*	*Var. pentaphylla type (2.6 kbp)*
1 Horoman-gawa	11	—	1.000	—	—	—	1.000	—	—	—	—	1.000	—	—	—	—	1.000	—
2 Uzura-gawa	9	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
3 Hakkoda-san	16	—	0.500	0.500	—	—	1.000	—	—	—	—	0.500	0.500	—	—	—	0.500	0.500
4 Moriyoshi-yama	9	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
5 Zao	9	—	1.000	—	—	—	1.000	—	—	—	—	1.000	—	—	—	—	1.000	—
6 Iide-san	15	—	0.600	0.400	—	—	1.000	—	—	—	—	0.600	0.400	—	—	—	0.600	0.400
7 Okutadami	9	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
8 Tate-yama	12	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
9 Takanbo	9	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
10 Miboro	10	—	—	1.000	—	—	1.000	—	—	—	—	—	1.000	—	—	—	—	1.000
11 Otaki-gawa	14	0.786	—	0.214	—	—	—	—	0.214	0.786	0.786	—	—	0.214	—	—	—	1.000
12 Namie	9	—	—	—	—	1.000	—	—	—	1.000	—	—	—	—	—	1.000	—	—
13 Gozaisho-yama	9	—	—	—	—	1.000	—	—	—	1.000	—	—	—	—	—	1.000	—	—
14 Jakuchi-kyo	10	—	—	—	—	1.000	—	—	—	1.000	—	—	—	—	—	1.000	—	—
15 Ishiduchi-san	9	—	—	—	—	1.000	—	—	—	1.000	—	—	—	—	—	1.000	—	—
16 Goyo-san	10	—	—	—	1.000	—	—	1.000			—	—	—	—	1.000	—	—	—

[a]The length variation of an intron between exons B and C of the mitochondrial gene encoding subunit 1 of NADH dehydrogenase was used to differentiate between mitochondrial types of *P. pumila* and *P. parviflora* var. *pentaphylla* (Watano *et al*, 1996).

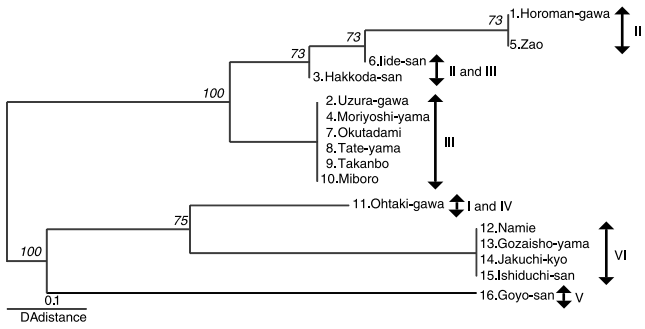

Figure 3 Neighbor-joining phenogram, using mtDNA D_A distance, for the 16 Japanese *P. parviflora* populations. Numbers in italics represent bootstrap estimates, based on 1000 replicates. Roman numerals indicate mtDNA types (Figure 1).

et al, 1998; Mitton *et al*, 2000). According to maps reconstructing the vegetation during the last ice age (Tsukada, 1974, 1982), boreal conifer forest, mostly containing no *P. parviflora*, was distributed from Kanto and Niigata to the western part of Hokkaido, and across highlands such as the Chubu mountains. Therefore, the present populations in the Chubu highlands, Tohoku and Hokkaido must have established after the ice age. The typical var. *parviflora* mitochondrial Type, VI (E/D), was present in four populations, extending from the western part of Honshu and Shikoku Island to the Namie population (located in the Tohoku district on the Pacific coast). According to Tsukada's vegetation map of the last ice age, Tohoku district was mainly covered with boreal conifer forest (Tsukada, 1982). Since the Namie population has been suggested to have established during global warming following the last glaciation, *P. parviflora* with mitochondrial Type VI (E/D) may have recolonized the Pacific coast from refuges in the western part of Japan as far as the southern Tohoku region on the Pacific coast. The typical var. *pentaphylla* mitochondrial Type III (C/A) is distributed within the area that was covered with the boreal conifer forest during the last glaciation. Although the location of var. *pentaphylla* refuges has not been confirmed, individuals with mitochondrial Type III (C/A) probably colonized along the central mountain ridge in Honshu Island from the lower altitudes in the Chubu district to the north.

We detected evidence of reductions in population size in the postglacial age. Smaller number of rare alleles were detected in populations in western Chubu than in the Tohoku populations. In particular, the Gozaisho-yama and Goyo-san populations contained no rare alleles (Table 1). A simulation of the 'bottleneck effect' has demonstrated that the loss of alleles is profoundly affected by the bottleneck size, and the loss of alleles in the first generation is great (Nei *et al*, 1975). Since the number of alleles is affected by sample size, it is necessary to consider this attribute for each population. The number of rare alleles detected in populations in western Chubu is so small that the bottleneck effect may have completely eliminated low-frequency alleles. However, these populations have maintained high genetic variation. This could be because the loss of alleles was faster than the loss of heterozygosity when the bottleneck effect occurred (Barrett and Kohn, 1991; Luikart *et al*, 1998). Although var. *pentaphylla* populations appear to have retained a larger number of rare alleles than populations of var. *parviflora*, this may simply be a result of the recent reduction in population size. This would account for the fact that the Namie population possessed a large number of rare alleles despite consisting of var. *parviflora* (Table 1).

Genetic differentiation between the two *P. parviflora* varieties

When the extent of genetic differentiation between the two varieties was investigated using 11 allozyme loci, the derived G_{VT} value (the coefficient of genetic differentiation between the two varieties) was found to be only 0.001 (Table 3). However, it has been reported previously that there are morphological differences between the two varieties, including differences in the morphology of winter buds, the length of leaves, the presence of lines of stomata on the abaxial side of the leaf, the size and shape of the cones, and the length of the wings of cones (eg Makino and Nemoto, 1931; Isii, 1941; Hayashi, 1960). However, some of the morphological differences between the two varieties may be the result of environmental conditions, since they form a gradient among populations along a geographical axis. For example, continuous exposure to daylight turns lines of stomata on leaves bright white, while on shaded and older leaves, they remain pale (Iwata and Kusaka, 1952). Furthermore, there may be no linkage between the 11 allozyme loci examined in this study and loci governing morphological traits. If this is the case, the allozyme data may not reflect genuine morphological differences between the varieties. However, mtDNA variation did correspond to the distribution of the two varieties. This observation suggests that genetic differentiation in mtDNA between the two varieties is maintained by limited seed dispersal.

We found two unique haplotypes in the Otaki-gawa population, which occurs between the ranges of the two varieties. These were related to either var. *pentaphylla* or var. *parviflora* (Figure 1). Our observation could not be solely explained by organelle capture through hybridization, because we observed that all individuals in the Otaki-gawa population contained mitochondria with unique haplotypes. Several authors have predicted the origin of novel genetic variants in hybrid zones (Morgan and Strobeck, 1979; Golding and Strobeck, 1983). Recently, 'rare allele phenomenon', namely, novel allele in hybrid populations, was observed in a mammalian hybrid zone (Bradley *et al*, 1993; Hoffman and Brown, 1995). Dong and Wagner (1993) found a unique mitochondrial haplotype in which, *P. banksiana* population had received *P. contorta* mtDNA through past hybridization. Wang and Szmidt (1994) also detected a novel chloroplast DNA haplotype in *P. densata* populations which had been formed by a putative tertiary hybridization event between *P. tabulaeformis* and *P. yunnanensis*. At least three hypotheses as to how novel RFLP haplotypes arise have been suggested. Firstly, mutation of specific restriction enzyme sites may yield the unique haplotype A from E of *cob*, and C from A of *nad*3. Secondly, it may be caused by leakage of paternal mtDNA followed by intermolecular recombination between maternal and paternal typical mitochondria types for var. *parviflora* and var. *pentaphylla*. Thirdly, it was

possible that an origin of the unique haplotypes of the population was traced back to a different refugium. However, more extensive studies are required in the area where the ranges overlap, in order to elucidate the relationship between the two varieties. An investigation of genetic variation within individual is also required to test whether heteroplasmy or multiple haplotypes arise within individual tree or not in future analysis.

Comparisons between nuclear-encoded allozymes and mtDNA variations demonstrated that pollen flow produced genetic uniformity among populations. Although mtDNA rearrangement did not appear to be of value for determining the phylogenetic relationship between mitochondrial types, mtDNA clustering data may shed light on the phylogeography and population history of *P. parviflora*, supporting the current taxonomy of a northern part of the *P. parviflora* complex.

Acknowledgements

We thank Dr T Sato for supplying the plant material from the Takanbo population, and Mr M Takahashi for his help in sampling the four populations in the Hokkaido and Tohoku districts. We also thank Dr X-R Wang and anonymous reviewers for their helpful comments in the latest version of the manuscript. This research was partly supported by a grant No. 07456066, from the Ministry of Education, Culture, Sports, Science and Technology.

References

Aagaard JE, Krutovskii KV, Strauss SH (1998). RAPD markers of mitochondrial origin exhibit lower population diversity and higher differentiation than RAPDs of nuclear origin in Douglas fir. *Mol Ecol* **7**: 801–812.

Barrett SH, Kohn JR (1991). Genetic, evolutionary consequences of small population size in plants: implication for conservation. In: Falk DA, Holsinger KE (eds) *Genetics and Conservation of Rare Plants*. Oxford University Press: New York, USA. pp 3–30.

Birky CW (1988). Evolution and variation in plant chloroplast and mitochondrial genomes. In: Gottlieb LD, Jain SK (eds) *Plant Evolutionary Biology*. Chapman & Hall: New York, USA. pp 25–53.

Birky CW, Fuerst P, Maruyama T (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* **121**: 613–627.

Bradley RD, Bull JJ, Johnson AD, Hills DM (1993). Origin of a novel allele in a mammalian hybrid zone. *Proc Natl Acad Sci USA* **90**: 8939–8941.

Demesure B, Sodzi N, Petit RJ (1995). A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Mol Ecol* **4**: 129–131.

Dong J, Wagner DB (1993). Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theor Appl Genet* **86**: 573–578.

Dumolin-Lapègue S, Pemonge M-H, Petit RJ (1998). Association between chloroplast and mitochondrial lineages in oaks. *Mol Biol Evol* **15**: 1321–1331.

Golding GB, Strobeck C (1983). Increased number of alleles found in hybrid populations due to intragenic recombination. *Evolution* **37**: 17–29.

Goudet J (1995). FSTAT. Vers. 1.2. A computer program to calculate *F*-statistics. *J Hered* **86**: 485–486.

Hayashi Y (1960). *Taxonomical, Phytogeographical Study of Japanese Conifers*. Norin-Shuppan: Tokyo, Japan. (in Japanese).

Hoffman SM, Brown WM (1995). The molecular mechanism underlying the 'rare allele phenomenon' in a subspecific hybrid zone of the California field mouse, *Peromyscus californicus*. *J Mol Evol* **41**: 1165–1169.

Isii S (1941). On the various forms of *Pinus pumila* and other northern Japanese soft pines with special reference to their distribution (V). *J Jpn For Soc* **23**: 107–114 (in Japanese).

Iwata T, Kusaka M (1952). *Coniferae Japonicae Illustrated*. Sangyo Tosho: Tokyo, Japan. (in Japanese).

Latta RG, Mitton JB (1997). A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). *Genetics* **146**: 1153–1163.

Li CC, Horvitz DG (1953). Some methods of estimating the inbreeding coefficient. *Am J Hum Genet* **5**: 107–117.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998). Usefulness of molecular markers for detecting population bottlenecks via monitering genetic change. *Mol Ecol* **7**: 963–974.

Makino T, Nemoto K (1931). *Nippon-shokubutsu-soran*. 2nd edn. Shunyodo: Tokyo. (in Japanese).

Mayr H (1890). *Monographie der abietineen der Japanischen Reiches*. (in German).

Mirov NT (1967). *The Genus Pinus*. The Ronald Press Company: New York.

Mitton JB, Kreiser BR, Latta RG (2000). Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondria DNA. *Mol Ecol* **9**: 91–97.

Miyabe K, Kudo Y (1921). *Woody Species Illustrated in Hokkaido I*. Hokkaido University Press: Sapporo. (in Japanese).

Morgan K, Strobeck C (1979). Is intergenic recombination a factor in the maintenance of genetic variation in natural populations? *Nature* **277**: 383–384.

Murray MG, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**: 4321–4325.

Neale DB, Sederoff RR (1989). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theor Appl Genet* **77**: 212–216.

Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* **70**: 3321–3323.

Nei M, Chesser RK (1983). Estimation of fixation indices and gene diversities. *Ann Hum Genet* **47**: 253–259.

Nei M, Maruyama T, Chakraborty R (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

Nei M, Tajima F, Tateno Y (1983). Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol* **19**: 153–170.

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992). Complete nucleotide sequence of the mitochondrial DNA from a liverwort, *Marchantia polymorpha*. *Plant Mol Biol Rep* **10**: 105–163.

Ohwi J (1978). *Flora of Japan*. Shibundo: Tokyo. (in Japanese).

Ota T (1993). *DISPAN. A computer program for the genetic distance and phylogenetic analysis*. Institute of Molecular Evolutionary Genetics, The Pennsylvania State University.

Palmer JD (1992). Mitochondrial DNA in plant systematics: applications and limitations. In: Solitis PS, Solitis DE, Doyle JJ (eds) *Molecular Systematics of Plants*. Chapman & Hall: New York, USA. pp 36–49.

Petit RJ, Kremer A, Wagner DB (1993). Finite island model for organelle and nuclear genes in plants. *Heredity* **71**: 630–641.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Senjo M, Kimura K, Watano Y, Ueda K, Shimizu T (1999). Extensive mitochondrial introgression from *Pinus pumila* to

P. parviflora var. *pentaphylla* (Pinaceae). *J Plant Res* **112**: 97–105.

Strauss SH, Hong Y-P, Hipkins VD (1993). High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *muricata* and *attenuata*. *Theor Appl Genet* **86**: 605–611.

Takahata N, Nei M (1984). F_{ST} and G_{ST} statistics in the infinite island model. *Genetics* **107**: 501–504.

Tsukada M (1974). *Paleoecology II. Synthesis*. Kyoritsu: Tokyo, Japan (in Japanese).

Tsukada M (1982). *Cryptomeria japonica*, glacial refugia and late-glacial and postglacial migration. *Ecology* **63**: 1091–1105.

Tsumura Y, Ohba K (1993). Genetic structure of geographical marginal populations of *Cryptomeria japonica*. *Can J For Res* **23**: 859–863.

Tsumura Y, Suyama Y (1998). Differentiation of mitochondrial DNA polymorphisms in populations of five Japanese *Abies* species. *Evolution* **52**: 1031–1042.

Wagner DB, Dong J, Carlson MR, Yanchuk AD (1991). Paternal leakage of mitochondrial DNA in *Pinus*. *Theor Appl Genet* **91**: 1222–1236.

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

Watano Y, Imazu M, Shimizu T (1995). Chloroplast DNA typing by PCR-SSCP in the *Pinus pumila–P. parviflora* var. *pentaphylla* complex (Pinaceae). *J Plant Res* **108**: 493–499.

Watano Y, Imazu M, Shimizu T (1996). Spatial distribution of cpDNA and mtDNA haplotypes in a hybrid zone between *Pinus pumila* and *P. parviflora* var. *pentaphylla (Pinaceae)*. *J Plant Res* **109**: 403–408.

Weeden CR, Wendel JF (1989). Visualization, interpretation of plant isozymes. In: Soltis DE, Soltis PS (eds) *Isozymes in Plant Biology*. Dioscorides Press: Portland, OR, USA. pp 5–45.

Weir BS (1990). *Genetic Data Analysis: Methods for Discrete Population Genetics Data*. Sinauer: Sunderland, MA, USA.

Wilson EH (1916). *The conifers and taxa of Japan*. Publications of the Arnold Arboretum, no. 8. Cambridge University Press: Cambridge, UK.

Wright S (1943). Isolation by distance. *Genetics* **28**: 114–138.

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

Wu J, krutovskii KV, Strauss SH (1998). Abundant mitochondrial genome diversity, population differentiation and convergent evolution in pines. *Genetics* **150**: 1605–1614.

Yagashira K (1964). *Dendrology-Coniferae Illustrated*. Asakura Shoten: Tokyo. (in Japanese).