Genetic Structure and Evolutionary History of a Diploid Hybrid Pine *Pinus densata* Inferred from the Nucleotide Variation at Seven Gene Loci

Xiao-Fei Ma, Alfred E. Szmidt,† and Xiao-Ru Wang**

*State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China; and †Department of Biology, Graduate School of Science, Kyushu University, Fukuoka, Japan

Although homoploid hybridization is increasingly recognized as an important phenomenon in plant evolution, its evolutionary genetic mechanisms are poorly documented and understood. *Pinus densata,* a pine native to the Tibetan Plateau, represents a good example of a homoploid hybrid speciation facilitated by adaptation to extreme environment and ecological isolation from the parents. Its ecologically and reproductively stabilized nature offers excellent opportunity for studying genetic processes associated with hybrid speciation. In this study, we investigated the levels and patterns of nucleotide variation in *P. densata* and its putative parents. Haplotype composition, gene genealogies, and the levels and patterns of nucleotide variation gave further support to the hybrid nature of *P. densata*. Allelic history, as revealed by our data, suggests the ancient nature of the hybrid preceding elevation of the Tibetan Plateau. We detected more deviations from neutrality in *P. densata* than in the parental species. Thus, at least some of the evolutionary forces that have shaped the genetic variation in *P. densata* are likely to be different from those acting upon parental species. We speculate that when populations of *P. densata* invaded new territories, they had elevated rates of response to selection in order to develop traits that help them to survive and adapt in the new environments.

Introduction

Hybridization has long been considered as an important evolutionary force in plant evolution (Anderson 1948; Stebbins 1950; Lewontin and Birch 1966; Grant 1981; Arnold 1997; Rieseberg 1997). It can quickly create evolutionary novelties that promote adaptive evolution and speciation (Arnold and Hodges 1995; Turelli, Barton, and Coyne 2001; Arnold, Bouck, and Cornman 2003; Rieseberg et al. 2003). Historically, polyploid speciation is thought to be the main form of hybrid speciation (Grant 1981; D. E. Soltis and P. S. Soltis 1995) while the role of homoploid hybridization has not been fully examined due to the difficulties in its detection that the most parsimonious conclusion was that it was not as prevalent (and therefore important) as polyploidy hybrid speciation. Only recently research employing molecular approaches has revealed several cases of homoploid hybrid speciation suggesting this form of speciation may be more common than previously thought (Wolfe, Xiang, and Kephart 1998; Gross and Rieseberg 2005; Howarth and Baum 2005). However, the genetic mechanisms underlying homoploid hybrid speciation in plants are poorly understood. A few investigations targeted traits that are distinct in hybrids and are thought to be important in species-specific adaptations (Rieseberg, Archer, and Wayne 1999; Rieseberg et al. 2003; Lexer, Lai, and Rieseberg 2004). These studies suggested that transgressive segregation appears to be sufficient to explain the origin of adaptations in hybrids. Another important mechanism underlying homoploid hybrid speciation is chromosomal rearrangement relative to parental species. Different models of chromosomal speciation have been proposed that facilitate reproductive isolation in the process of speciation (see review by Rieseberg 2001 and references therein). A different approach to understand the mechanisms of hybrid speciation is to study the patterns of nucleotide variation at randomly selected gene loci to infer the selective forces and demographic events involved in the history of speciation (Avise 1989; Kliman et al. 2000).

Pinus densata is a conifer tree with large distribution in the southeastern part of the Tibetan Plateau. It forms pure forest at high elevations (2700–4200 m asl) and regenerates well (Wu 1956; Guan 1981). The origin of *P. densata* has been investigated with allozyme, chloroplast (cp) DNA, and mitochondrial (mt) DNA markers (Wang and Szmidt 1994; Wang, Szmidt, and Savolainen 2001; Song et al. 2002, 2003). All results indicate that *P. densata* originated from hybridization between *Pinus tabuliformis* and *Pinus yunnanensis* without alteration in the ploidy level (Wang, Szmidt, and Savolainen 2001; Song et al. 2002, 2003; Liu et al. 2003). *Pinus tabuliformis* is widely distributed from northern to central China, and *P. yunnanensis* has a relatively limited range in southwestern China (Wu 1956). The geographic distribution of the three pines forms a succession, with *P. tabuliformis* in the north, *P. densata* in the middle, and *P. yunnanensis* in the south. The high elevation habitat occupied by *P. densata* is inaccessible to any other pine species growing in the region (Wu 1956; Guan 1981). Patterns of variation in allozymes, cp, and mt DNA show that individual populations of *P. densata* have very diverse genetic compositions, with varying degrees of genomic contribution from each parental species (Wang and Szmidt 1994; Wang, Szmidt, and Savolainen 2001; Song et al. 2002, 2003). In addition, populations of *P. densata* from different parts of the plateau show reciprocal parentage. These results suggest that populations of *P. densata* have unique evolutionary histories and most likely independent origins.

Habitat divergence plays a crucial role in plant speciation. It is even more important in homoploid hybrid speciation to facilitate reproductive isolation from the parents. Theory indicates that this process is most likely to be driven by ecological divergence (McCarthy, Asmussen, and Anserson 1995; Buerkle et al. 2000; Barton 2001). Indeed, all well-documented examples of homoploid hybrid plant species occur in habitats that are different from those of

Key words: genealogy, hybrid speciation, nucleotide polymorphism, *Pinus densata*, population heterogeneity, selection.

E-mail: xiao-ru.wang@niwl.se.

Mol. Biol. Evol. 23(4):807–816. 2006
doi:10.1093/molbev/msj100
Advance Access publication January 30, 2006

Published by Oxford University Press 2006.

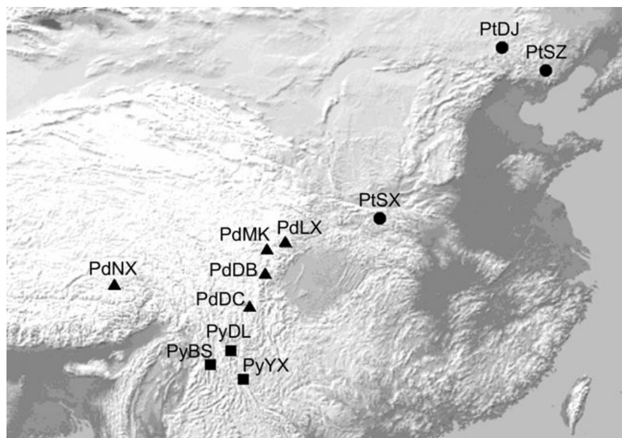

FIG. 1.—Geographic distribution of the sampled populations. ● *Pinus tabuliformis,* ▲ *Pinus densata,* ■ *Pinus yunnanensis.*

their parental species (Gross and Rieseberg 2005). *Pinus densata* represents a good example of a homoploid hybrid speciation facilitated by adaptation to extreme environment and ecological isolation from the parents. Its ecologically and reproductively stabilized nature offers excellent opportunities for studying genetic processes associated with the hybrid speciation and evolution.

To date, there are very few studies of polymorphism and divergence across several nuclear gene loci in conifers (Garcia-Gil, Mikkonen, and Savolainen 2003; Kado et al. 2003; Brown et al. 2004; Neale and Savolainen 2004; Bouillé and Bousquet 2005). Furthermore, there is particularly little information about the levels and patterns of nucleotide variation in plant species that evolved through interspecific hybridization. Studies on the levels and patterns of nucleotide variation in nuclear genes can help to better understand the evolutionary forces that have shaped genetic variation in *P. densata* and its parents and permit a genome-wide assessment of speciation, which cannot be revealed by uniparentally inherited cp and mt DNA markers.

In this study, we surveyed the nucleotide polymorphism and haplotype structure over seven loci in *P. densata* and its two putative parents. Our main objectives were (1) to determine the levels and patterns of DNA polymorphism in the investigated species complex, (2) to characterize the population heterogeneity and allelic coalescence history in *P. densata* and its parental species, and (3) to better understand the history of the hybrid speciation of *P. densata*.

Materials and Methods

Population Sampling

Seeds from three, three, and five populations were sampled for *P. tabuliformis*, *P. yunnanensis*, and *P. densata*, respectively (fig. 1). For *P. densata*, seeds were collected from 14 to 17 individual trees in each population. One seed from each tree was germinated on a Petri dish and a haploid megagametophyte from each seed was used for genomic DNA isolation. For *P. tabuliformis* and *P. yunnanensis*, composite seed samples were collected from more than 100 trees per population. Megagametophytes from 13 to 16 seeds per population were used for DNA extraction. These megagametophytes were regarded as random gamete samples from each population. A total of 164 haploid DNA samples were analyzed. The geographic locations of the populations are shown in figure 1. Two populations of *P. tabuliformis* (PtDJ and PtSZ) were from the northern limit of its distribution, the other one (PtSX) was from the central part of its distribution. The three populations of *P. yunnanensis* (PyBS, PyDL, and PyYX) were from the Yunnan province. Of the five populations of *P. densata*, four (PdLX, PdMK, PdDB, and PdDC) were located along the eastern edge of the Tibetan Plateau, the other one (PdNX) was further into the center of the plateau. These populations were selected based on the previous investigations, which showed a diverse genetic composition of individual populations and verified their hybrid nature (Wang and Szmidt 1994; Wang, Szmidt, and Savolainen 2001; Song et al. 2002, 2003).

Sampled Loci and Sequencing

In a preliminary survey, about 80 gene loci were screened for amplification in the three species investigated in this study. Only the loci that were represented by a single polymerase chain reaction (PCR) band were selected. The PCR products were cloned into a pGEM T-easy vector (Promega Inc., Madison, Wisc.) and 7–8 clones were sequenced for each locus to examine whether they consisted of a single sequence. Finally, seven loci (*ARA*, *DEH*, *PHO*, *POD*, *PtIFG2009*, *PtIFG8744*, and *PtIFG8887*) were selected for sequence analysis. The putative function and structure of these loci and the PCR primers are described in Supplementary Table 1.

DNA of the haploid megagametophytes was used to amplify the seven loci. PCR was performed in a volume of 20 μl consisting of 0.2 μM of each primer, 2.0 mM of $MgCl_2$, 200 μM of each dNTP, and 1.0 U of Ex Taq DNA polymerase (Takara Biotechnology, Dalian, China), using a PTC100 (MJ Research Inc., Waltham, Mass.) thermal cycler programmed for an initial denaturation at 94°C for 3 min followed by 34 cycles of 30 s at 94°C, annealing at a specific temperature (as listed in Supplementary Table 1) for 30 s and extension for 30 s at 72°C, and a final extension of 10 min at 72°C. Amplification products were first examined through electrophoresis in 1.0% agarose gel. The products were then cut from the gel and purified using a GFX PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech, Little Chalfont, Buckinghamshire, UK). The purified products were sequenced directly using a BigDye Terminator Cycle Sequencing Ready Reaction Kit v3.0 (Applied Biosystems, Foster City, Calif.). Unique haplotype sequences for each locus are deposited in the GenBank with the accession numbers DQ232904–DQ233201.

The genus *Pinus* is divided into two subgenera, *Strobus* and *Pinus*. *Pinus tabuliformis*, *P. yunnanensis*, and *P. densata* all belong to the subgenus *Pinus*. To estimate the divergence between the two subgenera, two haploid DNA samples from each of *Pinus armandii* and *Pinus koraiensis*, both from the subgenus *Strobus*, were sequenced for the seven loci included in this study. These sequences were used as outgroup.

Table 1
Nucleotide Diversity over the Seven Loci in 11 Populations of the Three Pine Species

Population	Total Length 3,040 bp				Silent			Nonsilent				Coalescence
	n	S	π	θ_w	S	π	θ_w	S	π	θ_w	$N_e \times 10^5$	Time (MYA)
PtDJ	13	71	0.0075	0.0076	64	0.0108	0.0110	7	0.0019	0.0020	5.63	28.2
PtSX	15	95	0.0092	0.0099	81	0.0125	0.0136	14	0.0038	0.0039	6.96	34.8
PtSZ	15	96	0.0083	0.0099	87	0.0118	0.0143	9	0.0025	0.0025	7.34	36.7
P. tabuliformis	43	136	0.0085	0.0107	121	0.0119	0.0153	15	0.0030	0.0031	7.85	39.3
PyBS	10	53	0.0071	0.0064	46	0.0100	0.0089	7	0.0024	0.0022	4.56	22.8
PyDL	10	54	0.0064	0.0064	47	0.0087	0.0089	7	0.0026	0.0022	4.55	22.8
PyYX	9	47	0.0065	0.0059	41	0.0093	0.0083	6	0.0020	0.0020	4.25	21.3
P. yunnanensis	29	63	0.0067	0.0055	55	0.0095	0.0077	8	0.0023	0.0018	3.95	19.7
PdDB	12	61	0.0079	0.0069	54	0.0114	0.0098	7	0.0022	0.0021	5.01	25.0
PdDC	13	55	0.0070	0.0060	47	0.0099	0.0083	8	0.0022	0.0023	4.25	21.3
PdLX	14	90	0.0089	0.0095	82	0.0126	0.0138	8	0.0027	0.0023	7.09	35.4
PdMK	13	98	0.0092	0.0107	85	0.0128	0.0150	13	0.0033	0.0038	7.69	38.5
PdNX	14	56	0.0065	0.0060	49	0.0096	0.0084	7	0.0015	0.0020	4.33	21.6
P. densata	66	141	0.0086	0.0101	123	0.0122	0.0143	18	0.0028	0.0034	7.32	36.6

NOTE.—n, S, π, and θ_w refer to haploid sample size, number of segregating sites, nucleotide diversity (Nei and Li 1979), and Watterson's parameter (Watterson 1975), respectively. Estimates of the effective population size N_e and allelic coalescence time in each population and species are based on $\theta_{ws.}$

Data Analysis

One hundred sixty-four haploid DNA samples were sequenced for each locus except for *DEH* and *PtIFG8887* loci, for which due to the failure of PCR amplification in some samples only 163 and 140 samples were sequenced, respectively. Sequences were aligned with ClustalX (Thompson et al. 1997) and further manually adjusted using BioEdit program (Hall 1999). The total aligned sequence length over the seven loci was 3,040 bp, which included 1,463 bp of exons, 610 bp of introns, and 977 bp of the 3′UTR regions. Nucleotide polymorphism measured by θ_w (Watterson 1975) and diversity measured by π (Nei and Li 1979), intragenic minimum recombination events (Rm) (Hudson and Kaplan 1985) and haplotype diversity (H_d) (Nei and Li 1979), at each locus and in each population and species, were estimated using DnaSP v4.0 program (J. Rozas and R. Rozas 1999). The ratio of replacement (π_a) and synonymous (π_s) polymorphism (π_a/π_s) was calculated for each locus in each population. Population differentiation was estimated by F_{st} (Hudson, Slatkin, and Maddison 1992) as implemented in the ProSeq v2.9 program (Filatov 2002) with 1,000 permutations to obtain the significance estimates of F_{st}. Gaps were excluded in all analyses.

The measures of linkage disequilibrium (LD) D' and r^2 among informative sites were calculated using DnaSP v4.0. The statistical significance of LD was determined by Fisher's exact test with and without Bonferroni correction.

Each locus was tested for departure from neutrality by Tajima's (1989) *D* and Fu and Li's (1993) D^* and F^* statistics using DnaSP v4.0. Except for *DEH* locus at which sequencing failed for the two pines of subgenus *Strobus*, the Fu and Li statistics were performed with a *P. armandii* sequence as outgroup. The use of the subgenus *Strobus* sequence as outgroup is to improve the power of the test as it is less related to the three pines investigated in this study. Under neutral model, the level of polymorphism within species correlates with the degree of divergence between species across loci. The HKA test (Hudson, Kreitman, and Aguade 1987) was developed to test this prediction. When applied to multilocus data, the HKA test assesses the overall fit of the data to the neutral model that assumes the same ratios of polymorphism and divergence at each locus. The HKA test was performed for *P. densata* versus *P. tabuliformis*, *P. densata* versus *P. yunnanensis*, and *P. tabuliformis* versus *P. yunnanensis* as implemented in Jody Heys' multilocus HKA program (http://lifesci.rutgers.edu/~heylab/HeylabSoftware.htm). The simulations were run 1,000 times.

Gene/allele genealogies of each locus were constructed by coalescent simulations using the Median-Joining model as implemented in the Network v4.0 program (Bandelt, Forster, and Röhl 1999). We treated the segregating sites as independent evolutionary events. All indels were excluded from the analysis. Sequences of the subgenus *Strobus* were used as outgroup in the network construction of each locus except for the *DEH* locus.

Results

Nucleotide Polymorphism

Measures of nucleotide polymorphism at each locus and in each population and species are presented in Supplementary Table 2. The average estimates over the seven loci for each population and species (calculated using pooled population data for each species) are presented in table 1. At species level, the nucleotide polymorphism over all loci (θ_w) was similar in *P. tabuliformis* (0.0107) and *P. densata* (0.0101), which was nearly twofold higher than that in *P. yunnanensis* (0.0055). Silent polymorphism in the three species was four- to fivefold higher than the replacement polymorphism (table 1). Within each species, the levels of polymorphism (π and θ_w) among loci differed by four- to sixfold. The *POD* locus was the most polymorphic in *P. tabuliformis* and *P. densata*, while the *PtIFG8744* locus was the most polymorphic in *P. yunnanensis*. *PHO* was the least polymorphic locus in all the three pines (Supplementary Table 2).

The ratios of replacement (π_a) and synonymous (π_s) polymorphism at five of the seven loci are listed in Supplementary Table 3. Two loci (*PtIFG8887* and *PtIFG8744*) were excluded from this analysis due to their very short

Table 2
Population Differentiation (F_{st}) in the Three Pine Species

Species	Locus							Multilocus
	ARA	*DEH*	*PHO*	*POD*	*PtIFG2009*	*PtIFG8744*	*PtIFG8887*	
P. tabuliformis	0.0594	0.0326	0.0452	0.1251**	0.0937**	0.2117**	−0.0188	0.0860***
P. yunnanensis	−0.0440	−0.0551	0.0708	−0.0381	0.0237	0.1720*	−0.0744	0.0310
P. densata	0.0718	0.1417**	0.2048**	0.0894*	0.1170**	0.1056*	0.0560	0.1047***

NOTE.—Probability obtained by permutation test with 1,000 replicates: *0.01 ≤ P < 0.05; **0.001 ≤ P < 0.01; ***P < 0.001.

coding regions. Most of the values for π_a/π_s at *ARA, DEH, PHO*, and *PtIFG2009* (except for the *PHO* locus in population PyYX) were smaller than 1 in the investigated populations. At the *POD* locus, however, three populations (PtSX, PdLX, and PdMK) had π_a/π_s greater than 1.

Minimum numbers of recombination events in the three pine species at the seven loci are listed in Supplementary Table 2. Among the seven loci, recombination was detected at five (*ARA*, *DEH*, *POD*, *PtIFG2009*, and *PtIFG8744*), four (*ARA*, *DEH*, *PtIFH2009*, and *PtIFG8744*) and three (*ARA, PtIFG2009,* and *PtIFG8744*) loci in *P. tabuliformis*, *P. densata*, and *P. yunnanensis*, respectively. No recombination was detected at the *PHO* and *PtIFG8887* loci in any of the three pines. The lack of recombination at *PHO* and *PtIFG8887* correlates with the relatively low haplotype diversity (H_d) at these loci.

Segregating sites at each locus were used to compute LD in each population and species. Among 1,367, 1,360, and 466 pairwise intragenic comparisons over the seven loci in *P. tabuliformis*, *P. densata*, and *P. yunnanensis*, 16%, 18%, and 33% comparisons were significant by Fisher's exact test and 4%, 8%, and 18% were still significant after Bonferroni correction (Weir 1996), respectively. *Pinus yunnanensis* showed the highest number of significantly associated sites (two- to fourfold higher than that in the other two pines). No intergenic LD was detected in any of the three pines.

There was significant population differentiation in *P. tabuliformis* and *P. densata*, with multilocus F_{st} values of 0.086 and 0.105, respectively (table 2). Population differentiation in *P. yunnanensis* was low (0.031) and not significant.

Tests of Selective Neutrality

Tajima's *D*, Fu and Li's *D** and *F**, and HKA tests were used to detect the departure from the neutral model of molecular evolution at each locus. The Tajima's *D* and Fu and Li's *F** values at the seven loci in each population are given in table 3. Fu and Li's *D** test gave a similar trend as *F** and thus is not shown. In *P. tabuliformis*, no significant *D* and *F** values were detected at any locus in any population. In *P. densata*, however, significant departures from neutrality were detected by both Tajima's *D* and Fu and Li's *F**. Populations PdMK and PdDB had positive significant *D* and *F** values at the *ARA* and *PtIFG8744* loci, respectively, and population PdNX had positive significant *F** and positive but not significant *D* at the *PtIFG2009* locus. In *P. yunnanensis*, population PyBS had positive significant *D* and *F** at the *PtIFG8744* locus while populations PyDL and PyYX had positive significant *F** but not significant *D* at the *ARA* locus. In addition, in two populations of *P. densata* (PdNX and PdMK) both *D* and *F** or only *F** were significantly negative at the *DEH* and *POD* loci (table 3).

We performed HKA tests for *P. densata* versus *P. tabuliformis*, *P. densata* versus *P. yunnanensis*, and *P. tabuliformis* versus *P. yunnanensis* over the seven loci. The results showed that none of the pairwise species comparisons gave significant chi-square statistic at any locus.

Effective Population Size

For neutral alleles in mutation-drift equilibrium, synonymous polymorphism θ_{ws} in diploid genome is equal to $4N_e\mu_g$, where N_e is the effective population size and μ_g the mutation rate per generation (Tajima 1989). We used *P. armandii* of subgenus *Strobus* to calculate the average divergence between the subgenera *Pinus* and *Strobus* at six of the seven loci, excluding *DEH* locus due to the failure in amplification in *P. armandii*. According to the fossil records, pines diversified into the two subgenera during the early Cretaceous ca. 130 MYA (Miller 1977). The estimated average divergence at silent sites (*K*s) between *P. armandii* and *P. tabuliformis* and *P. yunnanensis* was 0.0507, which can be treated as the accumulation of silent mutations over 130 Myr. Thus, the mutation rate per year (μ_y) was estimated to be $Ks/2T = 1.95 \times 10^{-10}$ for *P. tabuliformis*, *P. yunnanensis*, and *P. densata*. Assuming pines take 25 years to reach full seed production, the mutation rate per generation (μ_g) can be estimated as $25\mu_y = 4.875 \times 10^{-9}$. N_e can then be estimated as $\theta_{ws}/4\mu_g$. A summary of the estimated effective population sizes for individual populations of the three pines is listed in table 1. *Pinus tabuliformis* and *P. densata* had similar and large effective population sizes (7.85×10^5 and 7.32×10^5), which were approximately twofold larger than that of *P. yunnanensis* (3.95×10^5).

Genealogy of Each Locus

Genealogies of the haplotypes observed at each of the seven loci were constructed by coalescent simulations using the Median-Joining network (fig. 2). Based on the topology and the frequency of the haplotypes, the genealogies of the seven loci can be grouped into three classes. The first class includes simple and relatively small networks with short branches such as those obtained for the *PHO* and *PtIFG8887* loci. The *PHO* locus harbored 12 haplotypes among the 164 samples of the three investigated pine species, with four haplotypes (H1, H2, H3, and H9) occurring at high frequencies. Excluding *P. densata*, the H1 haplotype was the main haplotype specific to *P. tabuliformis*

Table 3
Neutrality Test at Each Locus as Measured by Tajima's *D* and Fu and Li's *F**

Population	*ARA*		*DEH*		*PHO*		*POD*		*PtIFG2009*		*PtIFG8744*		*PtIFG8887*	
	D	*F**	*D*	*F**	*D*	*F**	*D*	*F**	*D*	*F**	*D*	*F**	*D*	*F**
P. tabuliformis, PtDJ	0.8985	0.6297	−0.6255	−0.0191	0.6120	0.9730	−0.9002	−2.0149	−0.0459	0.1494	−1.0993	−0.1048	0.7531	1.5679
PtSX	0.1048	−0.6191	0.4071	0.1133	−0.5019	−1.0934	−0.3444	0.7101	−0.3488	−0.3688	0.6824	−0.3749	−0.3970	−0.1744
PtSZ	−0.6973	−0.8748	0.5785	1.3655	−0.8267	−0.5166	−1.2065	−1.9661	−0.2748	−0.8817	−1.2502	−0.7999	−0.9910	−0.9531
P. yunnanensis, PyBS	1.3800	0.9483	1.1926	1.0757	−0.0269	0.8495	−0.5796	0.7397	−0.7984	−1.0389	2.2110*	2.0632*	0.6789	0.4792
PyDL	1.8240	1.9292*	1.4128	1.0740	0.8943	1.0922	−1.4512	−1.8697	−0.2604	−0.5212	−1.5431	−1.8059	−0.4451	1.0542
PyYX	1.6949	1.8845*	1.5328	1.5440	−0.7638	−0.5704	0.0345	0.9294	1.4021	1.6052	−0.2920	1.2381	0.4680	−0.0376
P. densata, PdDB	0.9421	0.5113	−0.5218	0.0343	−0.5826	−1.7547	1.3106	0.6241	0.1523	0.4827	2.0274*	1.9603*	0.2777	0.5779
PdDC	0.0964	0.1389	1.3734	1.1844	0.6685	0.5860	0.5913	0.5426	0.7520	0.6654	1.9800	1.4022	0.4323	−0.1081
PdLX	0.7726	0.9304	−0.3286	−0.6682	−0.0981	−0.4671	0.0545	0.3369	−0.8276	−0.7213	−0.3442	−0.5216	−0.4479	0.3166
PdMK	2.0186*	2.0038*	−0.4024	−0.8447	−0.1712	−1.6823	−1.7239	−3.3028*	−0.4175	−0.3806	−1.4542	−2.0811	−0.6757	0.3489
PdNX	0.2925	−0.3281	−1.9099*	−2.5436*	nc	nc	0.0033	−0.3534	1.8939	2.0848*	1.6449	1.5283	0.7681	1.5195

Note.—nc: not calculated due to the lack of polymorphism. *$0.01 < P < 0.05$

and the H3 and H9 haplotypes were specific to *P. yunnanensis*. When *P. densata* was included, all these haplotypes become shared with *P. densata*. At the *PtIFG8887* locus, 21 haplotypes were detected among which the H2 haplotype dominated in all the three pines. Thus, this haplotype could be regarded as the ancestral haplotype with many of the minor haplotypes derived from it. A smaller cluster surrounding the H10 haplotype was specific to *P. yunnanensis* and *P. densata*. The second class of genealogy included more complex networks composed of two to three main haplotypes with relatively deep coalescence shared by all the three species. This class included the *ARA*, *PtIFG8744*, *DEH*, and *POD* loci. The *ARA* locus harbored 22 haplotypes of which three (H4, H9, and H5) accounted for 75% of the total number of haplotypes. These three haplotypes were shared by all the three pines but the H9 haplotype was more common in *P. yunnanensis* and *P. densata*. Six mutation steps separated the haplotypes H9 and H4, and three mutation steps separated haplotypes H9 and H5. Similar to the *PtIFG8887* locus, the topology of the *PtIFG8744* locus had two clusters, one including the haplotypes H1 and H2, which were shared by all the three species, and another cluster surrounding the H11 haplotype, which was shared by only *P. yunnanensis* and *P. densata*. Six mutation steps separated the two clusters. At the *DEH* locus, 27 haplotypes were found among which 13 were unique to *P. densata*. The three common haplotypes (H1, H4, and H5) were shared by all the three pine species. A few intermediate frequency haplotypes were largely species-specific, like H2 and H3 to *P. tabuliformis*, H15 to *P. densata*, and H13, H11, and H14 to *P. yunnanensis*. At the *POD* locus, 25 haplotypes were detected of which two (H1 and H8) were the most frequent. The H1 haplotype was shared by all the three species, but H8 haplotype was most abundant in *P. densata* (65%). The cluster composed of H8, H21, H23, and H24 haplotypes was nearly specific to *P. densata*. The last class of genealogy, which included the *PtIFG2009* locus, was particularly large and reticulated. A total of 48 haplotypes were found at this locus and the haplotypes were spread with no clear structure. The high reticulation of the haplotypes at this locus suggests more recombination events over the evolutionary history.

In general, despite different types of genealogies among the seven loci, some common features can be noticed: (1) most of the *P. densata* haplotypes coalesced to *P. tabuliformis* and *P. yunnanensis*, (2) at five of the seven loci, *P. densata* harbored the largest number of haplotypes, (3) many haplotypes specific to either *P. tabuliformis* or *P. yunnanensis* were shared with *P. densata*, and (4) among the 175 haplotypes detected at the seven loci in the three investigated pines, 52 (30%) were unique to *P. densata*, 57 (33%) to *P. tabuliformis*, and 11 (6%) to *P. yunnanensis*.

Discussion

Levels and Patterns of Nucleotide Polymorphism

Results from a few available studies on the patterns and levels of nucleotide polymorphism in conifers are

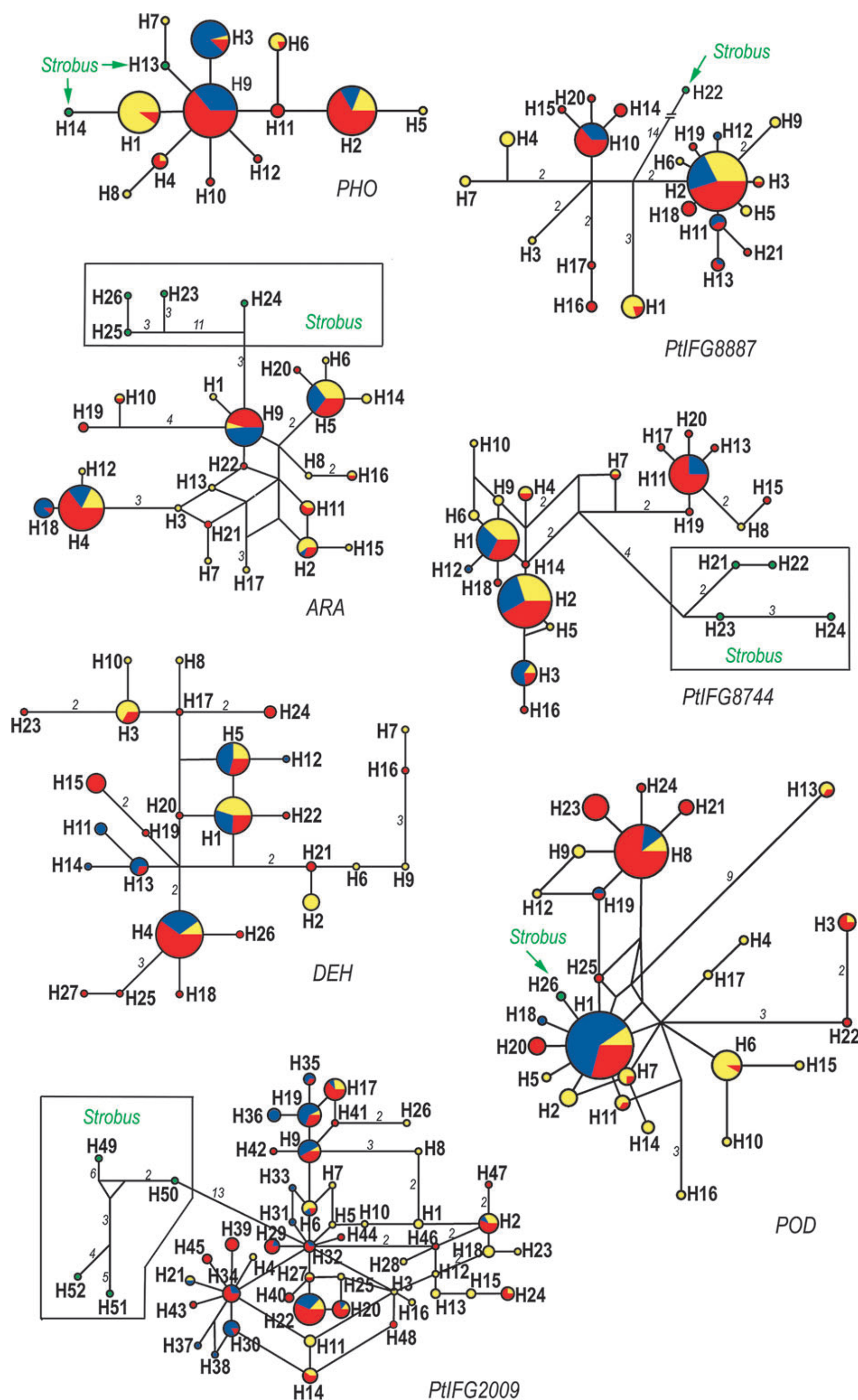

Fig. 2.—Gene genealogies of the seven loci. Colors in the pie chart indicate the haplotype origin; yellow for *Pinus tabuliformis*, blue for *Pinus yunnanensis*, and red for *Pinus densata*. Outgroup sequences from the subgenus *Strobus* are indicated as *Strobus*. The size of the pie is proportional to the haplotype frequency found in the three pines. Branch lengths longer than one mutation step are marked on each branch.

summarized in table 4. Very low levels of polymorphism (θ_w or π) were observed in *Pinus sylvestris* (0.0004–0.0014) and *Cryptomeria japonica* (0.0002–0.0038), intermediate levels were observed in *Pinus taeda* (0.0041), *P. yunnanensis* (0.0055), *Picea abies*, *Picea glauca*, and *Picea mariana* (0.0066–0.0081). On the other hand, high levels were found in *P. tabuliformis* (0.0107), *P. densata* (0.0101), and *Pseudotsuga menziesii* (0.0085). The level of nucleotide polymorphism in *P. tabuliformis* and *P. densata* was about two- and sevenfold higher than that of *P. taeda* and *P. sylvestris*, respectively. Among the three pine species analyzed in this study, *P. yunnanensis*

Table 4
Levels of Nuclear Gene Sequence Polymorphism (θ_w, or π When Indicated) in Outcrossing Conifers

Species	No. Loci	All Site θ_w	θ_{ws}	θ_{wns}	$N_e \times 10^5$	$\mu_y \times 10^{-10}$	Reference
Pinus tabuliformis	7	0.0107	0.0153	0.0031	7.85	1.95	This study
Pinus yunnanensis	7	0.0055	0.0077	0.0018	3.95	1.94	This study
Pinus densata	7	0.0101	0.0143	0.0034	7.32	1.99	This study
Pinus sylvestris	1	0.0014 (π)	0.0049 (π_s)	0.0003 (π_{ns})	3.29	1.49	Dvornyk et al. (2002)
	2	0.0004–0.0010 (π)	0.0013–0.0020 (π_s)	0.0002–0.0003 (π_{ns})	—[a]	—[a]	Garcia-Gil, Mikkonen, and Savolainen (2003)
Pinus taeda	19	0.0041	0.0066	0.0011	5.6	1.17	Brown et al. (2004)
	44	0.0046–0.0049	0.0059–0.0070	0.0018–0.0019	—[a]	—[a]	Neale and Savolainen (2004)
Cryptomeria japonica	7	0.0002–0.0038	0.0038 (π_s)	0–0.0015	—[a]	10x of *P. sylvestris*	Kado et al. (2003)
Pseudotsuga menziesii	12	0.0085	0.0126	0.0046	—[a]	—[a]	Neale and Savolainen (2004)
Picea abies, Picea glauca, Picea mariana	3	0.0066–0.0081 (π)	—[a]	—[a]	0.96–1.82	2.23–3.24	Bouillé and Bousquet (2005)

[a] Not reported.

harbored about half of the polymorphism found in the other two pines. The level of polymorphism in *P. tabuliformis* and *P. densata* was comparable to that reported in *Populus tremula* (θ_w = 0.0167, π_s = 0.0160, π_{ns} = 0.0059) (Ingvarsson 2005), a diecious, wind-pollinated leaf tree, and in outcrossing maize (θ_w = 0.0096, θ_{ws} = 0.0173, θ_{wns} = 0.0039) (Tenaillon et al. 2001). Furthermore, our results show that within each species the levels of polymorphism among loci can differ by as much as four- to sixfold. Substantial locus-to-locus variation in the levels of polymorphism was also reported in other studies on conifers (Kado et al. 2003; Brown et al. 2004; Bouillé and Bousquet 2005). Thus, many more loci should be sampled to obtain a representative genome-wide estimate of the polymorphism in individual species of conifers.

The level of polymorphism at neutral or near-neutral loci is determined by effective population size, mutation rate, recombination, and demography (Hedrick 1980; Kimura 1983). The high polymorphism detected in *P. tabuliformis* is consistent with its larger N_e and suggests less drastic population size reduction in the Pleistocene glaciations as compared to *P. sylvestris* and *P. taeda*. Indeed, much of the region currently occupied by *P. sylvestris* has been covered by Pleistocene ice while the regions occupied by *P. tabuliformis* were ice-free (Prentice, Bartlein, and Webb 1991; Lindsey 2002). The high polymorphism observed in *P. densata* is rather unusual for the homoploid hybrid plant species. Among the case studies on homoploid hybrid speciation in plants, the hybrids often show lower (or roughly equivalent) genetic diversity than those found in parental species (Gallez and Gottlieb 1982; Maki and Murata 2001; Schwarzbach and Rieseberg 2002; Welch and Rieseberg 2002; Gross, Schwarzbach, and Rieseberg 2003). The low levels of genetic diversity in hybrids are explained by restricted distribution, small population sizes, and recent and single origins. *Pinus densata*, on the other hand, differs from the other hybrids in that it has an advanced evolutionary history with multiple origins, and its populations from different geographic regions show reciprocal parentage and varying degrees of genomic contribution from parental species (Wang and Szmidt 1994; Wang, Szmidt, and Savolainen 2001; Song et al. 2002, 2003). The complex evolutionary history and genetic composition of *P. densata* should have contributed to its present considerable polymorphism.

Very low levels of intragenic LD (after Bonferroni correction) were detected at the seven loci in *P. tabuliformis* (4%) and *P. densata* (8%). These results add further evidence for the high recombination rate expected for outcrossing species with large N_e. The higher LD (18%) in *P. yunnanensis* is consistent with its smaller N_e estimated in the present study. Substantial interspecific gene flow can also cause LD. However, our results did not show increased LD in *P. densata*, suggesting its populations are in drift-mutation equilibrium and there is no on-going hybridization with the parental species in the sampled populations. This conclusion is in accordance with the allozyme data which also did not detect increased LD in *P. densata* (Wang, Szmidt, and Savolainen 2001).

Signature of Selection in *P. densata*

Compared to *P. tabuliformis* and *P. yunnanensis*, we detected more departures from neutrality in *P. densata*. Significant positive Tajima's D and/or Fu and Li's F^* were observed at the *ARA*, *PtIFG8744*, and *PtIFG2009* loci in populations PdMK, PdDB, and PdNX, respectively. Significant positive D and F^* would suggest either balancing selection, or a reduction in population size in the recent past. Whereas demographic processes are likely to affect all loci in a similar manner, the effects of selection are usually restricted to a specific locus. The significant positive D and F^* at the *ARA* locus in PdMK contrast the negative D and F^* observed at all the other six loci in this population, which suggests against the reduction in population size. In fact, PdMK had the largest N_e in *P. densata* (table 1). Thus, balancing selection seems the more likely reason for the observed significant departure from the neutral model. As for the *PtIFG8744* and *PtIFG2009* loci in populations PdDB and PdNX, without other supporting evidence it is difficult to distinguish between the possibilities

of selection and demographic factors. Interestingly, significant positive F^* at *ARA* locus was also observed in populations PyDL and PyYX and significant positive D and F^* at *PtIFG8744* locus in population PyBS of *P. yunnanensis*. This parallel pattern in departure from neutrality between *P. yunnanensis* and *P. densata* at these two loci may suggest certain functional significance. More information from the full gene sequences is needed to clarify the reasons for the non-neutral variation at these two loci.

Significant negative D and/or F^* were observed at the *DEH* and *POD* loci in populations PdNX and PdMK, respectively. Significant negative D and F^* may result from either some form of selective pressure or from population expansion. The ratio of π_a/π_s at the *POD* locus in the PdMK population was 1.626, suggesting selective advantage of the replacements. Thus, the observed significant negative F^* in this population could be the result of positive selection. The *POD* gene encodes a peroxidase. Peroxidases are a family of enzymes that catalyze oxidation–reduction reactions and play important roles in detoxification, cell protection, and defense responses in plants. The functional significance of the observed departure from neutrality at this locus needs to be verified with full-length gene sequences. Together with the observation at the *ARA* locus in the PdMK population, we see an example of different evolutionary forces shaping the genetic variation at different loci in the same population. In contrast to the *POD* locus, π_a/π_s ratio at the *DEH* locus in the PdNX population was 0.141, which is consistent with the action of purifying selection. Taking into account the absence of negative D and F^* values across other loci to signify expansion of the PdNX population, some form of negative selection could be the cause for the observed deviation from neutrality at this locus. More detailed studies are needed to better explain the observed locus-specific deviations from neutrality.

Based on the observed deviations, which may represent signatures of selection in *P. densata,* it is tempting to speculate that when populations of *P. densata* invaded new territories they had elevated rates of response to selection in order to develop traits that help them survive and adapt in the new environments. Populations with different genetic composition that have occupied different regions of the Tibetan Plateau might have experienced varying degrees and types of selective forces and demographic processes. The effects of diverse evolutionary forces on populations could be preserved by limited gene flow among populations. Indeed, *P. densata* showed the highest F_{st} (0.105) among the three pines. This result is in agreement with the F_{st} estimated from allozyme loci which also revealed high differentiation (0.086) in *P. densata* (Wang, Szmidt, and Savolainen 2001). The F_{st} estimates for *P. tabuliformis*, however, varied between allozyme (0.026) and DNA sequence (0.086) data. This difference was likely caused by the sampling of very distant populations used in the DNA study. The level of population differentiation in *P. densata* is much higher than those commonly found in outcrossing, wind-pollinated conifers ($F_{st} < 0.05$) (Hamrick and Godt 1996). This high differentiation could be related to two factors, the diverse genetic composition among populations and the complex geography of the southeastern Tibetan Plateau. The topology of the region occupied by *P. densata* is characterized by high mountains and deep valleys. Therefore, gene flow among populations from different regions is likely to have been limited in *P. densata*.

Hybrid Nature and Multiple Origins of *P. densata*

Haplotype distribution at individual loci showed that most of the major haplotypes are shared by the three pines. At some loci (e.g., *PHO*, *PtIFG8887*) species-specific haplotypes were found for *P. tabuliformis* and *P. yunnanensis* when *P. densata* was excluded. Inclusion of *P. densata*, however, resulted in sharing of all these formerly species-specific haplotypes. This pattern gives further support to the hybrid nature of *P. densata* as most of its haplotypes appeared to be inherited or derived from the two parental species. The fixation time of neutral polymorphisms in a population depends on the population size (Kimura and Ohta 1969; Tajima 1983; Hudson 1990). The average coalescence time of two randomly selected alleles at a gene locus is estimated by $2N_e$ generations (Tajima 1983). Assuming a generation time of 25 years in pines, the estimates of N_e in table 1 would translate into 19.7–39.3 MYA of allele coalescence time in the three pine species. Alleles in *P. yunnanensis* had shorter coalescence history (19.7 MYA) than those in *P. tabuliformis* (39.3 MYA). The coalescence history in *P. densata* varied among populations (21.3–38.5 MYA) but within the boundary set by the two parental species (table 1). This supports multiple origins of this species. The five populations of *P. densata* differ in the levels of gene admixture from each of the parental species (Wang, Szmidt, and Savolainen 2001; Song et al. 2003). Thus, it is not surprising to see much variation in allelic history among its populations. Our present estimates of the allele coalescence time suggest that the PdMK and PdLX populations have been much influenced by *P. tabuliformis*, while populations PdDB, PdDC, and PdNX by *P. yunnanensis*. This suggestion is in agreement with the previous allozyme and mt and cpDNA results, which showed ca. 50% of *P. tabuliformis* and 30% of *P. yunnanensis* cpDNA components in the PdMK population (Pd-1 and 2 in Wang, Szmidt, and Savolainen 2001), 65% of *P. yunnanensis* cpDNA component in PdDC (Pd-7 in Wang, Szmidt, and Savolainen 2001), and dominant *P. yunnanensis* mt and cpDNA in PdDB (Song et al. 2003). It should be mentioned, however, that due to the observed deviations from neutrality for some loci in *P. densata* and *P. yunnanensis*, the allelic coalescence estimates may be biased to some extent.

The estimated allele coalescence time in the three pines also indicates the ancient nature of the allelic polymorphism. When there is much-shared polymorphism among related species, the coalescence of many alleles would precede species divergence (Bouillé and Bousquet 2005). Hybridization between divergent species would make this phenomenon more pronounced. The speciation of *P. densata* is suggested to be related to its adaptation to a unique ecological niche at high elevations on the Tibetan Plateau, where neither of the two parental pines can normally grow (Wang and Szmidt 1994; Wang, Szmidt, and Savolainen 2001). The uplift of the Tibetan Plateau dates back at least 20 MYA (Ruddiman and Kutzbach 1991; Harrison et al. 1992; Ruddiman 1998). However,

significant increases in altitude of the plateau are thought to have occurred only about 8–10 MYA followed by continued development of the plateau toward the north and east (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). The hybrid nature of *P. densata* suggests gene exchange between *P. tabuliformis* and *P. yunnanensis* occurred before they became separated by the uplift of the plateau. Although we cannot date precisely the origin of each population in these geographic events, it is evident that the extant populations of *P. densata* from different parts of the plateau differ in their origin, and their allele coalescence history much exceeds the establishment of the hybrid in its final unique territory on the Tibetan Plateau. Further surveys of candidate genes associated with important adaptive traits in *P. densata* could provide better insights into the functional significance of its allelic diversity and evolution.

Supplementary Material

Supplementary Tables 1–3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This study was supported by grants from the Natural Science Foundation of China (NSFC 30325006 and 30121003). A.E.S. was supported by grant 17405032 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Literature Cited

Anderson, E. 1948. Hybridization of the habitat. Evolution **2**:1–9.

Arnold, M. L. 1997. Natural hybridization and evolution. Oxford University Press, New York.

Arnold, M. L., A. C. Bouck, and R. S. Cornman. 2003. Verne Grant and Louisiana Irises: is there anything new under the sun? New Phytol. **161**:143–149.

Arnold, M. L., and S. A. Hodges. 1995. Are natural hybrids fit or unfit relative to their parents. Trends Ecol. Evol. **10**:67–71.

Avise, J. C. 1989. Gene trees and organismal histories—a phylogenetic approach to population biology. Evolution **43**:1192–1208.

Bandelt, H. J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16**:37–48.

Barton, N. H. 2001. The role of hybridization in evolution. Mol. Ecol. **10**:551–568.

Bouillé, M., and J. Bousquet. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Am. J. Bot. **92**:63–73.

Brown, G. R., G. P. Gill, R. J. Kuntz, C. H. Langley, and D. B. Neale. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. USA **101**: 15255–15260.

Buerkle, C. A., R. J. Morris, M. A. Asmussen, and L. H. Rieseberg. 2000. The likelihood of homoploid hybrid speciation. Heredity **84**:441–451.

Dvornyk, V., A. Sirviö, M. Mikkonen, and O. Savolainen. 2002. Low nucleotide diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. Mol. Biol. Evol. **19**:179–188.

Filatov, D. A. 2002. ProSeq: a software for preparation and evolutionary analysis of DNA sequence data sets. Mol. Ecol. Notes **2**:621–624.

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.

Fu, Y.-X., and W.-H. Li. 1993. Statistical tests of neutrality of mutations. Genetics **133**:693–709.

Gallez, G. P., and L. D. Gottlieb. 1982. Genetic evidence for the hybrid origin of the diploid plant *Stephanomeria diegenesis*. Evolution **36**:1158–1167.

Garcia-Gil, M. R., M. Mikkonen, and O. Savolainen. 2003. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. Mol. Ecol. **12**:1195–1206.

Grant, V. 1981. Plant speciation. Columbia University Press, New York.

Gross, B. L., and L. H. Rieseberg. 2005. The ecological genetics of homoploid hybrid speciation. J. Hered. **96**:241–252.

Gross, B. L., A. E. Schwarzbach, and L. H. Rieseberg. 2003. Origin(s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). Am. J. Bot. **90**:1708–1719.

Guan, C.-T. 1981. Fundamental features of the distribution of Coniferae in Sichuan. Acta Phytotaxonom. Sinica **11**: 393–407 [in Chinese].

Hall, T. A. 1999. BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. Phil. Trans. R. Soc. Lond. B **351**:1292–1298.

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

Hedrick, P. W. 1980. Hitch-hiking: a comparison of linkage and partial selfing. Genetics **94**:791–808.

Howarth, D. G., and D. A. Baum. 2005. Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (Goodeniaceae) in the Hawaiian islands. Evolution **59**:948–961.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Pp. 1–44 *in* D. Futuyma and J. Antonivics, eds. Oxford surveys in evolutionary biology. Oxford University Press, Oxford, UK.

Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111**:147–164.

Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics **116**:153–159.

Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics **132**:583–589.

Ingvarsson, P. K. 2005. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae). Genetics **169**:945–953.

Kado, T., H. Yoshimaru, Y. Tsumura, and H. Tachida. 2003. DNA variation in a conifer, *Cryptomeria japonica* (Cupressaceae sensu lato). Genetics **164**:1547–1559.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

Kimura, M., and T. Ohta. 1969. The average number of generations until fixation of a mutant gene in a finite population. Genetics **61**:763–771.

Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics **156**:1913–1931.

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

Lexer, C., Z. Lai, and L. H. Rieseberg. 2004. Candidate gene polymorphisms associated with salt tolerance in wild sunflower hybrids: implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. New Phytol. **161**:225–233.

Lindsey, R. 2002. The migrating boreal forest. Earth Observatory. (http://earthobservatory.nasa.gov/Study/BorealMigration/boreal_migration.html).

Liu, Z.-L., D. Zhang, D.-Y. Hong, and X.-R. Wang. 2003. Chromosomal localization of 5S and 18S-5.8S-25S ribosomal DNA sites in five Asian pines using fluorescence *in situ* hybridization. Theor. Appl. Genet. **106**:198–204.

Maki, M., and J. Murata. 2001. Allozyme analysis of the hybrid origin of *Arisaema ehimense* (Araceae). Heredity **86**:87–93.

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

Miller, C. N. 1977. Mesozoic conifers. Bot. Rev. **43**:217–280.

Neale, D. B., and O. Savolainen. 2004. Association genetics of complex traits in conifers. Trends Plant Sci. **9**:325–330.

Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76**:5269–5273.

Prentice, I. C., P. J. Bartlein, and T. Webb. 1991. Vegetation and climate change in North America since the last glacial maximum. Ecology **72**:2038–2056.

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

———. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. **16**:351–358.

Rieseberg, L. H., M. A. Archer, and R. K. Wayne. 1999. Transgressive segregation, adaptation and speciation. Heredity **83**:363–372.

Rieseberg, L. H., O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingstone, T. Nakazato, J. L. Durphy, A. E. Schwarzbach, L. A. Donovan, and C. Lexer. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science **301**:1211–1216.

Rozas, J., and R. Rozas. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15**:174–175.

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. **264**:66–75.

Schwarzbach, A. E., and L. H. Rieseberg. 2002. Likely multiple origins of a diploid hybrid sunflower species. Mol. Ecol. **11**:1703–1715.

Soltis, D. E., and P. S. Soltis. 1995. The dynamic nature of polyploid genomes. Proc. Natl. Acad. Sci. USA **92**:8089–8091.

Song, B.-H., X.-Q. Wang, X.-R. Wang, K.-Y. Ding, and D.-Y. Hong. 2003. Cytoplasmic composition in *Pinus densata* and population establishment of the diploid hybrid pine. Mol. Ecol. **12**:2995–3001.

Song, B.-H., X.-Q. Wang, X.-R. Wang, L.-J. Sun, D.-Y. Hong, and P.-H. Peng. 2002. Maternal lineages of *Pinus densata*, a diploid hybrid. Mol. Ecol. **11**:1057–1063.

Stebbins, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics **105**:437–460.

———. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**: 585–595.

Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley, and B. S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays* L.). Proc. Natl. Acad. Sci. USA **98**:9161–9166.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends Ecol. Evol. **16**:330–343.

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, and O. Savolainen. 2001. Genetic composition and diploid hybrid speciation of a high mountain pine, *Pinus densata*, native to the Tibetan plateau. Genetics **159**:337–346.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7**:256–276.

Weir, B. S. 1996. Genetic data analysis II. Sinauer Associates, Inc., Sunderland, Mass.

Welch, M. E., and L. H. Rieseberg. 2002. Patterns of genetic variation suggest a single, ancient origin for the diploid hybrid species *Helianthus paradoxus*. Evolution **56**:2126–2137.

Wolfe, A. D., Q.-Y. Xiang, and S. R. Kephart. 1998. Diploid hybrid speciation in *Penstemon* (Scrophulariaceae). Proc. Natl. Acad. Sci. USA **95**:5112–5115.

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

Pekka Pamilo, Associate Editor

Accepted January 25, 2006