High population differentiation and unusual haplotype structure in a shade-intolerant pioneer tree species, *Zanthoxylum ailanthoides* (Rutaceae) revealed by analysis of DNA polymorphism at four nuclear loci

K. KAMIYA, E. MORITSUKA, T. YOSHIDA, T. YAHARA and H. TACHIDA
Department of Biology, Faculty of Sciences, Kyushu University, 4-2-1 Ropponmatsu, Chuo-ku, Fukuoka 810-8560, Japan

Abstract

Differences in demographic history, life-history traits, and breeding systems affect nucleotide variation patterns. It is expected that shade-intolerant pioneer tree species have different patterns of genetic polymorphism and population structure than climax species. We studied patterns of nucleotide polymorphism at four putative starch pathway loci (agpSA, agpSB, agpL, and GBSSI) in *Zanthoxylum ailanthoides*, a shade-intolerant pioneer tree species that occupies forest gaps in warm-temperate forests of East Asia. Genetic diversity was lower within each population than among populations, and differentiation among populations was high across the loci \(F_{ST} = 0.32–0.64\), as expected from the insect-pollinated breeding system and the metapopulation structure of this pioneer species. Numbers of haplotypes were smaller than those expected from the observed numbers of segregating sites. Single haplotypes accounted for more than 47% of all the sampled genes at the respective loci. These variation patterns were incompatible with neutral predictions for populations of a finite island model. Complex population dynamics, such as bottleneck and/or admixture, in the history of this pioneer tree species might have resulted in the observed patterns of genetic variation and population structure, which are different from those of climax wind-pollinated tree species, such as conifers. In contrast to the other loci investigated in this study, agpL showed nearly no variation in *Z. ailanthoides* (one singleton only), but there was some extent of variation in a closely related species, *Zanthoxylum schinifolium*. This suggests possibly a recent selective sweep at or near the locus in *Z. ailanthoides*.

Keywords: nucleotide variation, pioneer tree, population structure, *Zanthoxylum*

Received 3 November 2007; revision accepted 25 February 2008

Introduction

The levels and patterns of genetic diversity in genomes are shaped by demographic processes and natural selection. Previous isozyme studies of genetic variation reveal that life history and breeding systems, in addition to demography, considerably affect the pattern of genetic polymorphism and structure in natural populations (e.g. Hamrick & Godt 1989, 1996). However, it is still uncertain to what extent those findings from the isozyme studies hold true for DNA polymorphisms. For example, although isozyme studies suggest that wind-pollinated and widely distributed coniferous species have higher levels of heterozygosity than angiosperms at the species level, lower levels of DNA polymorphisms generally have been found in several coniferous species such as *Pinus* (Dvornyk et al. 2002) and *Cryptomeria* (Kado et al. 2003) than in herbaceous plants (*Arabidopsis*, *Zea*, and rice; Wright & Gaut 2005) and woody angiosperms (*Populus*; Ingvarsson 2005). Moreover, multilocus analyses of DNA polymorphisms have successfully discriminated between demographic processes that affect all genes similarly in the genome and natural selection that influences a particular gene and adjacent regions in human (Stajich & Hahn 2005), model plants, and important feeding crops (Nordborg et al. 2005; Wright & Gaut 2005; Arunyawat et al. 2007; Moeller et al. 2007). Yet, such analyses remain scarce for nonmodel organisms, such as woody species,
except for the economically important coniferous tree species as mentioned above. Therefore, before we can generalize about the DNA polymorphism pattern in plants, it is important to accumulate DNA polymorphism data from a wide variety of plant species that differ in their distributions, ecological requirements, and mating systems.

It is known that pioneer trees species differ from climax trees in seed morphologies, ecological requirements, germination patterns, and growth rates (Dalling & Hubbell 2002). For example, pioneer species require light for germination, and their offspring cannot perpetuate under canopy shade. Therefore, regeneration and maintenance of populations in such species largely depend on the dynamics of the forest gap. So, one might expect that a pioneer species can extend its range of distribution by some means, but cannot establish large continuous populations. As a result, many pioneer species show patchy distributions and seem to have experienced frequent founder events in a given forest. These characteristics of pioneer species affect the levels and patterns of genetic diversity and population differentiation.

In this study, we examine the nucleotide sequence variation at four putative starch pathway loci in five populations of a species, Japanese prickly ash, Zanthoxylum ailanthoides Sieb. et Zucc. (Rutaceae), well known for its typical characteristics of pioneer trees. Zanthoxylum ailanthoides is distributed throughout the east coast of Asia from the Philippines to the main island of Japan and is often found in warm-temperate evergreen oak forests below 1500 m (Horikawa 1972). This species is essentially dioecious and consequently is an outbreeder. It is a medium-sized tree, reaching 15 m tall, and its density is usually low in the forests (Miura et al. 2001). The pollen is dispersed by bees and syrphid flies (T. Kawasaki, personal communication), and seeds are dispersed by several species of birds (Zosterops) (Shimoda 1998). Zanthoxylum ailanthoides has a number of characteristics of typical pioneer species, including shade intolerance (Yamamoto 1992), rarity of recruitment under closed canopies (Miura et al. 2001), formation of soil seed banks (Numata 1961; Naka & Yoda 1984), and faster growth rates in height (Watanabe 1979). The objective of the present study is to determine whether the genetic diversity pattern in this species is consistent with the predictions from its historical demography, habitat characteristics, life-history traits, and migration. More specifically, we try to answer the following questions. (i) How much nucleotide variation do these populations have and is there any differentiation between populations? (ii) Can we find any characteristic haplotype structure expected from the population dynamics of pioneer species? (iii) Is there any indication of selection on those putative starch pathway genes, which are considered to be involved in seed production and thus could be targets of selection in plants forming soil seed banks? This is the first study to examine DNA variation in shade-intolerant pioneer tree species based on multilocus nuclear sequence data.

Materials and methods

Materials

Forty-four individuals of Zanthoxylum ailanthoides were collected from five populations in Kyushu Island, Japan: Hirado (HR), Ikitsuki (IK), Motoooka (MT), Ozasa (OZ), and Shiiba (SB) (Fig. 1). In addition, six individuals of Zanthoxylum schinifolium, the closest relative of Z. ailanthoides as shown in a phylogenetic study using chloroplast matK sequences (K. Kamiya, E. Moritsuka, T. Yoshida, T. Yahara & H. Tachida unpublished data), were sampled from HR and SB. We note that the MT population is relatively young (1 to 4 years old), having grown from seeds in soil after artificial clearing of underbrush for the construction of a new campus of Kyushu University (Normile 2004). Other samples were potentially reproductive trees taken from natural or mature secondary forests dominated by evergreen trees of Fagaceae (Quercus glauca, Castanopsis sieboldii), Lauraceae (Machilus thunbergii, Neolitsea sericea) and Theaceae (Camellia japonica, Eurya japonica) mixed with deciduous pioneers such as Mallotus japonicus and Clerodendrum trichotomum. The OZ population is located in a suburb of a fairly big city (Fukuoka) and other populations are in rural areas.
Table 1 Primers used in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5′–3′)</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>agpSAF</td>
<td>AGACATTAATACTGTGCAAT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>agpSAR</td>
<td>AGTAAAAGCCTCACAAGT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggSBF</td>
<td>AGTGAATATCCTGCACT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggSBR</td>
<td>AGTAAAGCCTCACAAGT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggSB4F</td>
<td>TTTGAGGAGGCACTTA</td>
<td>sequencing</td>
</tr>
<tr>
<td>aggSB4R</td>
<td>AGCAGACGTCACTAGAGG</td>
<td>sequencing</td>
</tr>
<tr>
<td>aggLF</td>
<td>TCAACAGTGGCATAAACGAT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggLF1</td>
<td>AGCCTGAGAACAATGAAA</td>
<td>sequencing</td>
</tr>
<tr>
<td>aggLR1–1R</td>
<td>GTTTTCCTCAGTTTCCAG</td>
<td>sequencing</td>
</tr>
<tr>
<td>aggLR1–2R</td>
<td>CTTGATGTCAGACCTT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggLR2–2F</td>
<td>CCGAGGTTGATTTTATG</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>aggLR2–1F</td>
<td>AGAGCTCTCTCAATGAGG</td>
<td>sequencing</td>
</tr>
<tr>
<td>aggLR1R</td>
<td>ATATCCACTCCACCATCA</td>
<td>sequencing</td>
</tr>
<tr>
<td>GBSSIf156</td>
<td>GCCTGCGTGCAATAGAT</td>
<td>PCR, sequencing</td>
</tr>
<tr>
<td>GBSSIf152</td>
<td>TTAGAGAAGAACACTCTTG</td>
<td>sequencing</td>
</tr>
<tr>
<td>GBSSIf627</td>
<td>ATTCCACCTGCTGATACAT</td>
<td>sequencing</td>
</tr>
<tr>
<td>GBSSIr642</td>
<td>ACCTCAAGCCTTTATCTCC</td>
<td>PCR, sequencing</td>
</tr>
</tbody>
</table>

Total DNA was extracted from fresh or frozen leaf tissues using the DNeasy Plant Mini Kit (QIAGEN). The leaves were ground into fine powder with liquid nitrogen. Some DNA samples were further treated with phenol–chloroform extraction to remove remaining phenol and other impurities.

DNA sequencing

We selected four genes homologous to known starch pathway genes. Because those genes are considered to be important for seed production (Whitt et al. 2002), we thought that they might be possible targets of selection in plants forming soil seed banks. ADP-glucose pyrophosphorylase catalyses the formation of ADP-glucose from glucose-1-phosphate and ATP. This reaction is an essential step at the beginning of starch synthesis (Dickinson & Preiss 1969; Kim et al. 1998). The enzyme consists of two subunits; we studied the genes that encode ADP-glucose pyrophosphorylase small and large subunits (agpS and agpL, respectively). These genes affect the ripening of several fruits: tomato (Park & Chung 1998), watermelon (Kim et al. 1998), and citrus (Kim et al. 2001). The enzyme encoded by the GBSSI gene produces amyllose. The agpS and agpL loci were amplified by polymerase chain reaction (PCR) using primers designed from published complete cDNA sequences in Citrus unshiu (Kim et al. 2001). The PCR amplified two fragments with different sizes, suggesting at least two copies of the agpS (henceforth called agpSA and agpSB, respectively). After both fragments were sequenced independently, we designed additional primers that can amplify each locus specifically. The primers designed at exons 3 and 14 of agpL (agpLF and agpL0R2) yielded a PCR fragment of over 4 kb, and initially we determined the sequences of a single accession of Z. ailanthoides and Z. schinifolium. Based on these sequences, two additional primers (agpLr1–2R and agpLr2–2F) were designed. For other samples, an upstream region extending from exons 3–6 (agpLr1) and a downstream region from exons 10–13 (agpLr2) were amplified using the newly designed primers, sequenced, and analysed separately. For the GBSSI, primers were obtained from Evans et al. (2000) for initial PCR, and new primers specific for Zanthoxylum were designed. All the newly designed primers in this study are shown in Table 1. PCR was performed with 25 μL of reaction mixture containing 10 ng genomic DNA, 1× PCR buffer with 2.5 mM MgCl₂, 0.2 mM dNTPs, 0.1 μM of each primer, and 1 U of Ex-Taq DNA polymerase (TaKaRa). The cycling profile consisted of a primary denaturing of 2 min at 95 °C, followed by 30 cycles of 45 s at 95 °C, 45 s at 50 °C, and 1 min at 72 °C, then a final extension of 10 min at 72 °C. The PCR products were directly sequenced after purification using either the MiniElute PCR Purification kit (QIAGEN), or ExoSAP treatment, which degrades surplus primers and dNTPs by exonuclease-I and shrimp alkaline phosphatase. When more than one heterozygous site was encountered from the direct sequencing, the PCR products were cloned into a pGEM-T vector (Promega). Then, more than six clones were sequenced to determine two possible alleles in a sample. DNA sequencing was performed using an ABI BigDye Terminator Cycle version 3.1 sequencing kit and an ABI PRISM 3100 Genetic Analyser (PerkinElmer) following the manufacturer’s instructions. The sequences obtained in this study have been deposited in DNA Data Bank of Japan (DDBJ) under accession nos AB383173–AB383664.

Data analysis

The obtained sequences were aligned manually using a sequence alignment editor, se-al version 2.0a11 (Rambaut 1996). All indels found in the aligned sequences were excluded from the following analysis. Standard measures of DNA polymorphism, including π (Nei & Li 1979), and nucleotide divergence, K (Jukes & Cantor 1969), were calculated at total, nonsynonymous, and silent sites separately for each locus. Average within-population diversity (πw) was weighted by sample size (Wright et al. 2003). The minimum number of recombination events within each locus, Rw, was estimated using the method of Hudson & Kaplan (1985).

Neutrality test statistics, including Tajima’s D (1989) as well as Fu & Li’s D and F (1993), were estimated. The HKA test (Hudson et al. 1987) was used to assess the contrast between patterns of polymorphism and the divergence at different loci. To test the hypothesis that gene variations are distributed under the standard neutral model, the distribution of these statistics across 10 000 data sets was generated.
by coalescent simulation. All of these analyses were carried out with either DNAsP version 4.10.3 (Rozas et al. 2003) or the HKA program (developed by Jody Hey, available at www.lifesci.rutgers.edu/~heylab/HeylabSoftware.htm).

The $F_{ST}$ statistics (Hudson et al. 1992a) were used to measure the degree of genetic differentiation between the sampled populations. Significance levels of the nearest-neighbour statistic ($S_{nn}$; Hudson 2000) and a weighted measure of the ratio of the average pairwise differences within populations to the total average differences ($K_{ST}^*$; Hudson et al. 1992b) were computed by 1000 permutation tests using DNAsP. Haplotype networks were constructed by using TCS version 1.21 software (Clement et al. 2000).

To estimate the probability, under the finite island model, that the haplotype diversity is less than or equal to the observed value, we used the MS program developed by Hudson (2002) to generate sequence variation under the model. We wrote a C program to compute the haplotype diversity from the outputs and counted the times when the simulated value was less than or equal to the observed value. We employed the migration parameter ($4N_{m}$), estimated by $(1 - F_{ST}/F_{ST})$ using the average $F_{ST}$ from the data, the sample configuration of our data set, and a subpopulation number of 10 or 20 as inputs to MS. For each parameter set, 10,000 data sets were generated by MS to estimate the probability.

Results

DNA polymorphism

Combining the four nuclear loci, we analysed 4375 bp, involving 1657 nonsynonymous and 2718 silent sites (Table 2). Among 88 sequences analysed in Zanthoxylum ailanthoides, we found 53 polymorphic sites, of which only five sites were singletons. Table 3 shows a summary of nucleotide diversity in Z. ailanthoides and Zanthoxylum schinifolium, as well as the divergence between the species at each locus. Levels of genetic variation varied among the loci in both Z. ailanthoides and Z. schinifolium. Table 4 shows nucleotide diversities within two populations (MT and OZ) for which we sampled at least 15 individuals. In Z. ailanthoides, within-population diversities were generally lower than between- and total-population ones. Surprisingly, there was almost no variation in the two segments of agpL in Z. ailanthoides, and only one singleton was found in agpL1. In contrast, there was some variation at this locus in Z. schinifolium. The polymorphism levels were about twofold higher in Z. ailanthoides than in Z. schinifolium when the least variable locus, agpL, was excluded. The levels of nonsynonymous polymorphisms were approximately a magnitude lower than those for the silent sites, consistent with strong purifying selection at the studied loci (data not shown).

Tests of neutrality

We tested the obtained data for compatibility with the standard neutral model. As shown in Table 3, in all the studied loci except for agpLr1, positive Tajima’s $D$ values were obtained in Z. ailanthoides, but estimates were statistically significant only at agpSA. Similar results were obtained for Fu & Li’s $D$ and Fu & Li’s $F$-tests, with significant values found only at either one of the loci (i.e. agpSA or GBSSI). In Z. ailanthoides, positive mean Tajima’s $D$ and Fu & Li’s $D$ values were obtained, and both were shown to be statistically significant ($P < 0.01$) by the HKA program. No tests of neutrality were significant in Z. schinifolium, suggesting that this species does fit the neutral equilibrium model, although more extensive sampling is necessary to confirm this inference.

The HKA test was used to assess whether or not the levels of intraspecific polymorphism relative to those of interspecific divergence are constant between different loci (Hudson et al. 1987). Significant departure from the neutral equilibrium model was detected in all comparisons.
The multilocus HKA test was applied to the data of all the loci and then to the data excluding agpL. Although the latter test revealed a nonsignificant result ($P < 0.178$), the former test was statistically significant ($P < 0.00056$), indicating that the very low level of polymorphism at agpL is the result of selection rather than a lower mutation rate.

### Haplotype structure

The nucleotide variation pattern was highly structured at the agpSA, agpSB, and GBSSI loci (Figs 2 and 3). At each of the three loci, the most frequent haplotype accounted for more than 47% of all the sampled genes (i.e. haplotypes 10, 2, and 1 for agpSA, agpSB, and GBSSI, respectively). Relatively few haplotypes appear only once in the samples (four of 10 at agpSA, four of eight at agpSB, and zero of six at GBSSI), consistent with the positive values of Fu & Li’s $D$ statistics.

Moreover, three divergent haplogroups differed by the numbers of nucleotide substitutions at these three loci (A, B, and C in Fig. 2). The minimum number of recombination events, $R_M$, was one at agpSA and zero at the other three loci. Because of the small $R_M$, the networks showed no ambiguous connections (i.e. loops). Intragenic linkage disequilibrium was very strong and most pairs of sites within a gene showed $D' = 1$ (data not shown).

We computed the probability of haplotype diversity being less than or equal to the observed value under the finite island model by using $ms$, and the results are shown in Table 6. The migration parameter $4Nm$ was estimated from the average of $F_{ST}$ (Table 7). The observed values of...
The haplotype diversities were 0.697, 0.586, and 0.489 for agpSA, agpSB, and GBSSI, respectively. As the number of subpopulations increases or the population migration rate $N_m$ decreases, the probability becomes larger. However, the probability at each locus was generally low, and the joint probability of observing such low haplotype diversity at the three loci was very low ($P < 0.0005$, if we assume $N_m = 1.0$ and $d = 20$).

The distribution of haplotypes appeared to be different among populations. The major haplotype was different between two populations where we analysed sufficient numbers of sequences (MT and OZ populations) at agpSA and agpSB (Fig. 2).

**Population differentiation**

The present data show that the haplotype structure does not correspond to the sampling locations. Most of the
haplotypes were shared by at least two different populations, and major haplotypes were found in all of the studied populations, except for SB population, in which we studied only six sequences (Fig. 2). Table 7 shows genetic differentiation among populations. The estimates of $F_{ST}$ among the five populations were high (0.318, 0.322, and 0.644 at agpSA, agpSB, and GBSSI, respectively). The nearest-neighbour statistic ($S_{nn}$) of Hudson (2000) and that of $K_{ST}*$ (Hudson et al. 1992b) were highly significant in all the cases ($P < 0.001$). Estimates of pairwise $F_{ST}$ were also high at most of the loci and in population pairs, but there was considerable variation among the loci. This variation might be due to a large variance of estimates of $F_{ST}$ in addition to the limited number of our samples for some populations.

### Discussion

**Genetic diversity and population differentiation**

In *Zanthoxylum ailanthoides*, we found low levels of within-population variation and moderate levels of species-wide variation at all the studied loci. The within-population silent diversity averaged across the three loci excluding agpL was 0.0053 in the MT population and 0.0063 in the OZ population (Table 4). These values are comparable to the
estimates at BpMADS2 in a European population of *Betula pendula* (\(\pi_S = 0.0043\), Järvinen et al. 2003) and at Fal1 in Scots pine (\(\pi_S = 0.0043\), Dvornyk et al. 2002) but lower than the average in *Populus tremula* (\(\pi_S = 0.0145\), Ingvarsson 2005).

We consider the average number of pairwise differences between populations as a measure of genetic variation at the species-wide level (Wakeley 2000; Ramos-Ortins et al. 2004). The species-wide level of silent polymorphism in *Z. ailanthoides* averaged across the three loci was 0.0106. This value is larger than those in coniferous species, such as Scots pine (Dvornyk et al. 2002), *Cryptomeria japonica* (Kado et al. 2003), Douglas fir (Krutovsky & Neale 2005); and broad-leaf tree species *B. pendula* (Järvinen et al. 2003) and *Dunnia sinensis* (Chiang et al. 2002). However, a wind-pollinated tree species, *P. tremula*, has larger species-wide genetic variation than *Z. ailanthoides* (Ingvarsson 2005). The greater among-population variation than the within-population variation is in agreement with the predictions of subdivided population models (Wakeley 2000).

The genetic differentiation in *Z. ailanthoides*, an insect-pollinated species, was higher than that in wind-pollinated species, such as *Pinus sylvestris* (Dvornyk et al. 2002), *B. pendula* (Järvinen et al. 2003), and *C. japonica* (Kado et al. 2003). A moderate level of differentiation was observed in out-breeding and wind-pollinated species such as *P. tremula* (Ingvarsson 2005) and *Picea abies* (Heuertz et al. 2006). Indeed, the level of differentiation was higher in *Z. ailanthoides* than in those species. It is argued that the level of population differentiation is affected by the breeding system and life-history characteristics (Hamrick & Godt 1989). The level of population differentiation in *Z. ailanthoides* is consistent with the prediction that early succession species and insect-pollinated species are genetically more differentiated than late-succession species and wind-pollinated species.

**Causes for unusual haplotype structures**

We found an excess of high-frequency polymorphisms in *Z. ailanthoides* as indicated by positive values of Tajima’s *D* in the total population except at agpL (Table 2). This was observed consistently across the loci studied and led to the low haplotype number and low haplotype diversity. However, the high estimates of \(F_{ST}\) indicate that the populations of *Z. ailanthoides* are geographically structured. Consequently, we need to consider the population structure when we interpret the polymorphism data in this species. For this purpose, we ran ms to compute the probability of observing the haplotype diversity being less than or equal to the observed value under the island model using the migration parameter estimated from the data. We found that the observed haplotype diversity is low compared to the expectation under the island model. Although the result was not significant at individual loci, the joint probability of observing such low haplotype diversity at all three loci was very small. Therefore, we cannot explain the polymorphism data by the island model with a constant subpopulation size.

The low haplotype number and low haplotype diversity might be the result of the population dynamics of this pioneer tree species, which experienced a number of founder events in the past. In temperate broad-leaf forest, *Z. ailanthoides* needs large gaps created by typhoons to regenerate and maintain its population (Miura et al. 2001), but such large gaps rarely exist in many gap-dynamic forests (Naka & Yoda 1984). The gaps that *Z. ailanthoides* occupies seem to be isolated from each other in old-growth evergreen broad-leaf forests (Fujita et al. 2003). Usually, newly created gaps occur by chance and are often small and scattered spatially. Therefore, only a small number of individuals can occupy a newly established small forest gap, and thus rare variants may be lost easily. We showed that MT and OZ populations had only a few haplotypes, presumably as a result of repeated founder events in the past. The repeated founder events reduce genetic variation within a local population, but the species-wide variation would not decrease (Savolainen et al. 2000; see also Galtier et al. 2000). Founder events caused by gap dynamics in the course of a population history are consistent with the observed pattern of genetic variation in *Z. ailanthoides*.

Recently, De & Durrett (2007) showed that decreases of the number of low-frequency-derived alleles and linkage disequilibria are more pronounced under the stepping-stone model than under the island model. Because such effects are expected to cause a decrease of the haplotype diversity, the stepping-stone model may also explain the low haplotype diversity found in our data. In order to examine more complex models of population structure like the stepping-stone model and those involving effects of gap dynamics, we need data that are more comprehensive.

Another notable aspect of the polymorphism data is that the ratios of silent nucleotide divergences between the distinct haplogroups in *Z. ailanthoides* to those between *Z. ailanthoides* and *Z. schinifolium* at the three polymorphic loci were in a narrow range (0.398–0.479). Indeed, the haplotype networks of the three nuclear loci consistently show three distinct major haplotypes with several rare ones (Fig. 3). Although we did not carry out a formal test, the narrow range seems difficult to explain just by assuming the island model with frequent bottlenecks. Simulated gene trees from a pair of populations of constant size that exchange an average of one-half of a gene every generation showed that the time of the most recent common ancestor (MRCA) and the number of major clades varied widely among the trees (Harpending et al. 1998). Even in the case of random mating, the variance in the time of MRCA is known to be very large (Tajima 1983). This leads us to consider the possibility of a recent admixture of a few long-term isolated populations. Fossil data indicated...
that the Japanese Archipelago was almost entirely covered by coniferous forests during the last glacial maximum (Tsukada 1983). Small populations of various broadleaf forest species were scattered in the full-glacial temperate forests mainly along the coastal belt (Tsukada 1985). Thus, the population size reduction and isolation may have been caused by recent events of glaciations and postglacial recolonization. We admit that our data, which covers only a part of the distribution of this species, are not sufficient to reveal this complex demographic scenario or estimate the time scales. However, we can at least say that no simple scenario can explain the current polymorphism data. To understand the patterns of polymorphisms of *Z. ailanthoides*, further studies, with greater numbers of loci and sampling locations, are necessary. We are currently developing microsatellite markers for this species to pursue this issue.

Although the breeding system, life history, and historical demography of *Z. ailanthoides* would most consistently explain the patterns of genetic variation and population structure observed at agpSA, agpSB, and GBSSI, we cannot definitely rule out selective explanations. The presence of several major haplotypes may be associated with some selective scenarios, such as balanced polymorphism or incomplete hitch-hiking. Whitt *et al.* (2002) studied the nucleotide variation of domesticated maize and its closest wild relative, *Zea mays ssp. parviglumis*, at six loci involved in starch metabolism. Their results showed that these starch-related loci exhibited less than half the diversity of those at the other 20 random maize loci at both nonsynonymous and silent sites (0.0008 vs. 0.0038 at the nonsynonymous sites, and 0.0052 vs. 0.0122 at the silent sites, according to Table 1 in Whitt *et al.* 2002). They found that at least three of the six loci in the starch pathway are targets of selection, and concluded that the reduction of diversity in the starch pathway should motivate a paradigm shift for maize breeding. Although those authors found evidence of natural selection at brittles2, which is homologous to agpS, we found no evidence of selection at this putative gene orthologue in *Z. ailanthoides*. Our finding of almost no variation at aggL is suggestive of a signature of natural selection (see Table 5), but maize shows no evidence of selection at this locus. This suggests that, not surprisingly, different targets of natural selection have acted in *Z. ailanthoides* and maize. If the patterns of genetic variation observed in the starch pathway loci of *Z. ailanthoides* are due to natural selection, different patterns will be found at loci involved in different pathways. A comparative study of genetic variation at such loci using large population samples would allow the testing of the selective hypothesis in the starch pathway loci.

### Acknowledgements

This study is partially supported by the Kyushu University Interdisciplinary Programs and Projects (P & P) and Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (nos. 16370101, 19370099). We thank M. Kanzaki for providing his unpublished data; K. Odagiri and Y. Obshima for support of fieldwork. We also thank the referees for many helpful comments.

### References


Tsukada M (1985) Map of vegetation during the last glacial maximum in Japan. Quaternary Research, 23, 369–381.


© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd