Intra- and interspecific allozyme variability in Eurasian *Larix* Mill. species

VLADIMIR L. SEMERIKOV†§, LEONID F. SEMERIKOV† & MARTIN LASCOUX*‡

†*Institute of Plant and Animal Ecology, 620144 Ekaterinburg, Russia and ‡Department of Genetics, Uppsala University, Box 7003, 75007 Uppsala, Sweden*

The variation at 15 allozyme loci was examined in 32 populations of *Larix* taxa across their range in northern Eurasia. Fifteen populations of *L. sibirica* Ledeb., six populations of *L. gmelinii* Rupr., three populations of *L. olgensis* A. Henry, two populations of *L. decidua* Mill., and one population each of *L. cajanderi* Mayr, *L. kaempferi* (Lamb.) Carr. (= *L. leptolepis* (Sieb. et Zucc.) Endl.), *L. kamtschatica* (Rupr.) Carr., *L. czekanovskii* Szaf., *L. amurensis* Kolesn. and *L. ochotensis* Kolesn. were analysed. Interpopulational genetic differentiation within taxa was estimated with Wright's F_{ST} and varied from 2.6 to 8.1% according to taxon. Highly significant allelic differentiation was detected with Fisher's exact test. Genetic distances among *L. sibirica* populations were associated primarily with longitude. Allozyme frequencies changed gradually from the Urals eastward to Siberia, and eastern populations of *L. sibirica* are genetically more similar to *L. olgensis* than to western *L. sibirica*. Far-eastern populations of *Larix* species, morphologically similar to *L. gmelinii* and defined as *L. cajanderi*, *L. amurensis* and *L. ochotensis*, appear to be genetically close to both *L. gmelinii* and to *L. olgensis*, suggesting that they originated through introgressive hybridization between these two species.

Keywords: allozyme polymorphism, *Larix*, population structure.

Introduction

The present distribution of genetic variation in larch in Eurasia may, in large part, reflect the way the species spread and interacted over the postglacial recolonization period. Unlike northern Europe which was entirely covered by ice at the time of the last glacial maximum (18 kya), northern Siberia and the Russian far east were only locally glaciated (Bennett, 1997). This was recently confirmed indirectly by molecular studies in Arctic lemmings (*Dicrostonyx*); using mitochondrial DNA, Fedorov *et al.* (in press) revealed a complex phylogeographical pattern reflecting the past subdivisions of the present range by ice sheets. They also showed that the formerly glaciated Verchoyansky Mountains, on the eastern shore of the Lena river, acted as a strong barrier to gene flow. In comparison, postglacial recolonization of Siberia by plant species is still poorly understood. Glacial refuges and routes of recolonization have been partly documented for a

few forest species: *Abies sibirica* Ledeb., *Picea obovata* Ledeb., *Pinus sibirica* Du Tour. and *Pinus sylvestris* L. All of these species were absent from the northern part of the current range during the last glaciation. In contrast, four *Larix* fossils from northern Siberia were radiocarbon-dated to 13, 15, 16 and 21 kya, respectively (Kremenetski, 1994), thus suggesting the presence of *Larix* species during the last glaciation, or at least shortly before or after its maximum. On the other hand, palaeobotanical data indicate a radical transformation of plant cover during the late Pleistocene and Holocene in Siberia (Neishtadt, 1957; Bojarskaja & Malajeva, 1967; Belova, 1985).

Introgression has played an important part in plant evolution as, for instance, in the recolonization of Europe by oaks after the last glaciation (Dumolin-Lapégue, 1998). Larch species can easily be hybridized, and morphological and genetic studies indicate that hybridization is widespread in natural populations. The natural hybrid complex between *L. sibirica* Ledeb. and *L. gmelinii* Rupr. occurs along a long belt running from the Taymir Peninsula down to Mongolia (Kruklis & Milyutin,

*Correspondence. E-mail: martin.lascoux@genetik.uu.se
§Present address: Department of Genetics, Uppsala University, Box 7003, 75007 Uppsala, Sweden.

1977). Another possible case of larch introgressive hybridization can be found in the Russian far east (Dylis, 1961; Bobrov, 1972), where hybrid origins are suspected for many of the larch species described previously (*L. amurensis* Dylis, *L. maritima* Sukacz., *L. komarovii* Kolesn., *L. middendorfii* Kolesn., *L. ochotensis* Kolesn. and *L. lubarskii* Sukacz.). Because morphological differences between larch taxa are small, the taxonomy of *Larix* cannot be resolved on the basis of morphological traits only and there is currently no common view on the taxonomic status of some geographical varieties, and hence on the total number of species (Bobrov, 1972). Unsurprisingly, there is little congruence among phylogenetic trees based on morphological traits (Sukaczew, 1924; Kolesnikov, 1946; Dylis, 1947, 1961; Lepage & Basinger, 1995) and phylogenies based on molecular data (Kisanuki *et al.*, 1995; Qian *et al.*, 1995). Introgression and the population genetic structure of Eurasian larch species remain poorly known (but see Semerikov & Matveev, 1995; Shigapov *et al.*, 1998, on *L. sibirica*; Maier, 1992 on *L. decidua* Mill.).

In the present study, allozyme diversity within and between Eurasian *Larix* species of the section Pauciseriales will be examined. More specifically, we will (i) compare genetic distances within and between species at allozyme loci with the differentiation pattern based on morphological traits, and (ii) relate genetic divergence within and between species to what is known of *Larix* fossil history and the extent of glaciation in eastern Siberia. If *Larix* species survived at high latitudes during the last glaciation, as suggested by the fossil record, no clear north–south cline in genetic variation is expected, and northern populations might contain as much genetic variation as southern ones. Furthermore, large glaciated areas seem to have existed (in the Verchojansky and Putorana mountains in the Taymir Peninsula) and acted as barriers to longitudinal migration (Fedorov *et al.*, 1998).

Materials and methods

The species investigated, the number of populations per species and their locations are given in Table 1 and Fig. 1. The identification of the species was based on characteristic morphological traits and previously published information on the species distributions (Kolesnikov, 1946; Bobrov, 1972; Dylis, 1947, 1961). For most populations, needle tissues were analysed. When only seeds were available, seedlings were grown under artificial light for 2–4 weeks and the whole seedling used for analysis.

In some populations, seeds from half-sib families were available, whereas in others the number of trees from which the seeds originated was unknown.

Allozymes

Eleven enzyme systems representing 15 protein loci were examined: glutamic-oxaloacetic transaminase (GOT, EC 2.6.1.1, two loci), glutamate dehydrogenase (GDH, EC 1.4.1.2), isocitrate dehydrogenase (IDH, EC 1.1.1.42), diaphorase (DIA, EC 1.8.1.4), phosphoglucoisomerase (PGI, EC 5.3.1.9, two loci), 6-phosphogluconate dehydrogenase (6PGD, EC 1.1.1.44), shikimate dehydrogenase (SKDH, EC 1.1.1.25), glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49), phosphoglucomutase (PGM, EC 5.4.2.2, two loci), superoxide dismutase (SOD, EC 1.15.1.1, two loci) and fluorescent esterase (EST, EC 3.1.1.1).

Protein extraction and polyacrylamide gel electrophoresis were carried out according to Shurkhal *et al.* (1992) and Semerikov & Matveev (1995). 150 mg of needles and 150 mg of insoluble PVP were ground with liquid nitrogen in a mortar, and mixed with extraction buffer. The extraction buffer was composed of 1 M sucrose, 5.7 mM L-ascorbic acid, 8.3 mM DL-cysteine, 0.02 M dithiothreitol, and 1.5 mM aminocaproic acid dissolved in electrode buffer diluted 1:1.7. 1 mL Tween-80 was added to 100 mL of this solution and after 2–14 h at 4°C the mixture of extraction buffer, ground needles and PVP was filtered through nylon filters. The supernatant was centrifuged after adding a small amount of CCl_4 in order to improve the centrifugation of the supernatant from small tissue particles. Extracts from seedlings were prepared by grinding needles in 0.15 mL extraction buffer and centrifugation with CCl_4. Electrophoresis in 7.0% polyacrylamide gel in the Tris-EDTA-borate system was conducted. The electrode buffer (pH 8.0) was the following: 116 mM Tris, 3.5 mM EDTA, 161 mM boric acid. The gel buffer (pH 8.6) was the following: 118 mM Tris, 3.5 mM EDTA, 118 mM boric acid. Histochemical staining was carried out using standard methods (Harris & Hopkinson, 1976).

Data analysis

Data were analysed using GENEPOP (version 2; Raymond & Rousset, 1995b), FSTAT (version 1.2; Goudet, 1995), PHYLIP (Felsenstein, 1993) and NTSYS-pc (Rohlf, 1988). One individual per half-sib family was used in analyses of Hardy–Weinberg and linkage disequilibrium. All individuals were used in

Table 1 Location and description of the *Larix* populations used in the present study

| Pop.	Collection site	Lat., Long.	Material analysed
L. sibirica Ledeb.			
1	Vizhaj River, near Ivdel	61°10′ N, 60°20′ E	Needles of 47 mature trees
2	Severka, near Ekaterinburg	56°50′ N, 60°15′ E	Needles of 50 differently aged trees
3	Sob River	67°00′ N, 65°40′ E	96 offspring of 35 open-pollinated trees, 3 seedlings per family on average (range 1–7)
4	Poluy River	66°10′ N, 68°30′ E	50 offspring of 19 open-pollinated trees, 2.8 seedlings per family on average (range 1–5)
5	Iremel Mt.	54°30′ N, 56°40′ E	Needles of 48 differently aged trees
6	Ajat railway station	57°05′ N, 60°15′ E	Needles of 33 mature trees
7	Ongudai settlement (Altai Mts)	51°00′ N, 89°00′ E	48 seedlings grown from seeds of bulked collection. No. of mother trees unknown but large
8	Numgi–Yaha River	66°43′ N, 73°15′ E	Needles of 48 mature trees
9	Hadutte–Yaha River	67°27′ N, 77°30′ E	Needles of 45 mature trees
10	Hadita–Yaha River	67°00′ N, 69°50′ E	44 offspring of 11 open-pollinated trees, 4 seedlings per family on average (range 1–5)
11	Yada–Yahoda–Yaha River	67°05′ N, 71°40′ E	52 offspring of 24 open-pollinated trees, 2 seedlings per family
12	Ust–Kut (Upper Lena River)	56°50′ N, 105°40′ E	Needles of 50 mature trees
30	Mergen river (tributary of Ka–Hem)	51°40′ N, 95°40′ E	60 seedlings grown from seeds of bulked collection. No. of mother trees unknown but large
31	Chadan settlement	51°10′ N, 92 °00′ E	31 seedlings grown from seeds of bulked collection. No. of mother trees unknown but large
32	Pangodi settlement	66°10′ N, 74 °00′ E	32 seedlings grown from seeds of bulked collection. No. of mother trees unknown but large
L. gmelinii Rupr.			
13	Severobaikalsk settlement	55°40′ N, 109°24′ E	Needles of 50 mature trees
14	Severomuysk settlement	56°10′ N, 113°40′ E	Needles of 50 mature trees
15	Chara settlement	56°50′ N, 118°20′ E	Needles of 48 differently aged trees
16	Nerungri settlement	56°44′ N, 124°42′ E	Needles of 48 differently aged trees
17	Skovorodino railway station	54°00′ N, 124°00′ E	Needles of 48 differently aged trees
18	Urgal settlement	51°00′ N, 132°50′ E	48 seedlings grown from seeds of bulked collection. No. of mother trees unknown, but large
L. cajanderi Mayr			
19	Bereleh River (left tributary of Indigirka River)	70°26′ N, 143°56′ E	36 offspring of 16 trees, 2.25 seedlings per family (range 1–4)
L. decidua Mill.			
20	Bluhnbachtal, Salzburg (Austria)	47°28′ N, 13°05′ E	Seedlings grown from seeds of bulked collection. No. of mother trees unknown
21	Muran (Slovakia)	49°07′ N, 20°04′ E	Seedlings grown from seeds of bulked collection. No. of mother trees unknown
L. czekanovskii Szafer			
22	Bukukun River, Sohondinskii Reserve	49°30′ N, 111°00′ E	Needles of 33 mature trees
L. olgensis A. Henry			
23	Olga Bay	43°40′ N, 135°15′ E	Needles of 50 mature trees
24	Kavalerovo settlement	44°05′ N, 135°20′ E	Needles of 51 differently aged trees
25	Shufan basaltic plateau	43°25′ N, 131°20′ E	Needles of 50 trees grown from bulked seed collection in Ussuriisk nursery. No. of mother trees unknown, but large
L. kaempferi (Lamb.) Carr.			
26	60-year-old artificially generated stand in the south of Sakhalin island.	—	Needles of 51 trees. No. of mother trees unknown
L. kamtschatica (Rupr.) Carr.			
27	Vicinity of Lebjag'e Lake	47°30′ N, 142°40′ E	Needles of 50 mature trees
L. amurensis Kolesn.			
28	Vanino Bay	49°05′ N, 140°20′ E	Needles of 45 trees in 20-year-old stand naturally regenerated after fire
L. ochotensis Kolesn.			
29	Magadan city	59°30′ N, 150°50′ E	Needles of 48 mature trees
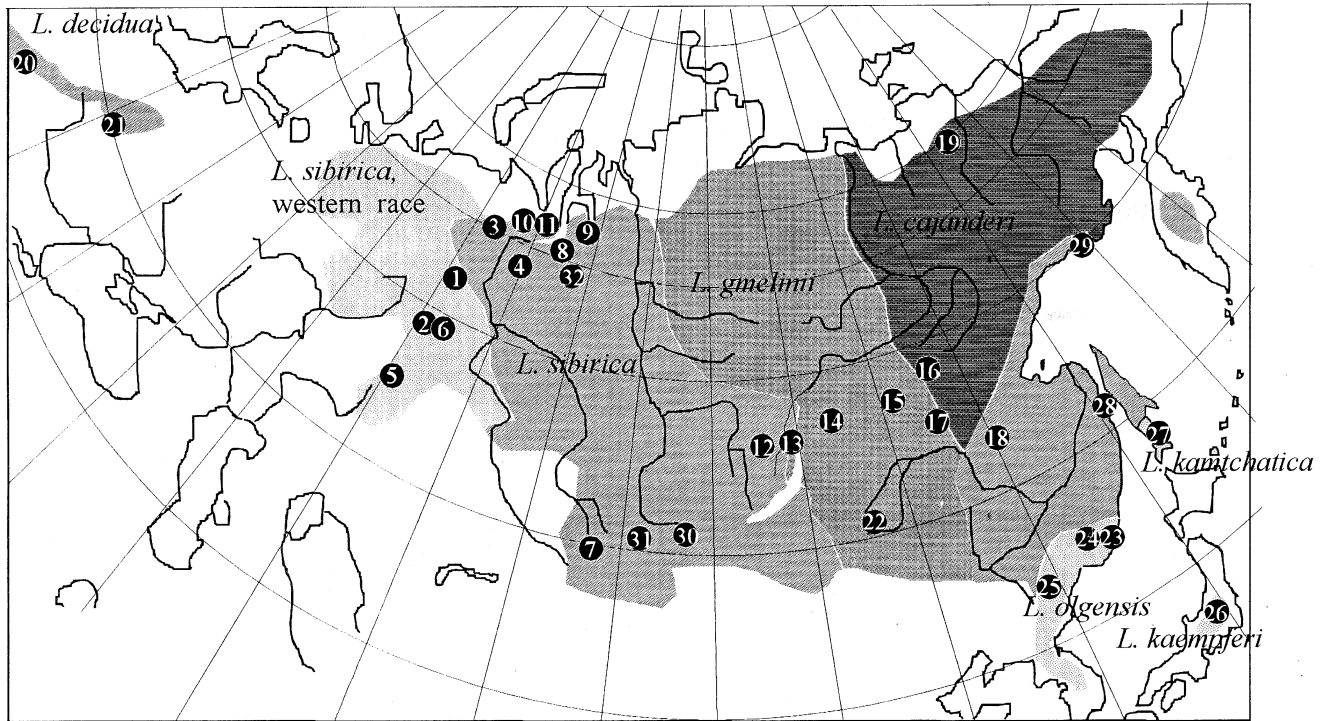
Fig. 1 Distribution map of *Larix* taxa in northern Eurasia, also indicating the locations of the 32 investigated populations. The numbers refer to the localities given in Table 1.

cluster analysis and ordination (similar results were obtained when one individual per half-sib family was used).

Hardy–Weinberg expectations The fit of genotypic distributions to Hardy–Weinberg expectations was tested by the exact test proposed by Haldane (1954). The overall significance for each locus was estimated by Fisher's combined probability test (Fisher, 1954). According to this test, if *P*-values are obtained for each locus separately under the null hypothesis, then

$$-2\sum_{i=1}^{n} \log(P_i)$$

is distributed according to a χ^2 distribution with *n* d.f., where *n* is the number of loci (Sokal & Rohlf, 1995). F_{IS} values, where F_{IS} is the correlation between two uniting gametes within a subpopulation, were estimated according to Weir & Cockerham (1984). Heterozygote deficiencies or excesses were tested using an exact test (Rousset & Raymond, 1995).

Linkage disequilibrium For each population, the nonrandom association between pairs of loci, or linkage disequilibrium, was tested using Fisher's exact test on contingency tables. Contingency tables are created for all pairs of loci in each population and an unbiased estimate of the exact probability obtained using a Markov Chain Monte Carlo method (Raymond & Rousset, 1995a). Each test is unaffected by a potential departure from Hardy–Weinberg, because each contingency table considers the genotypic composition, not the allelic composition. For each pair of loci, a global measure was obtained by averaging across populations and a global test was obtained through Fisher's combined test.

Population differentiation Genetic differentiation between populations or groups of populations was tested for each locus separately using Fisher's exact test on contingency tables. As for linkage disequilibrium, a Markov Chain Monte Carlo method permits the attainment of an unbiased estimate of the exact probability (Raymond & Rousset, 1995a). Wright's *F*-statistics, F_{IS}, F_{IT} and F_{ST}, were estimated according to Weir & Cockerham (1984) and a 95% confidence interval was estimated by bootstrapping over loci. F_{IS} and F_{IT} are the correlations between two uniting gametes relative to the subpopulation and relative to the total population, respectively, and

F_{ST} is the correlation between two gametes drawn at random from each subpopulation and measures the degree of genetic differentiation of subpopulations (Nei, 1987). Only statistically independent loci which did not depart from Hardy–Weinberg proportions were retained. Finally, genetic distances between populations were estimated with Nei's genetic distances (PHYLIP; Felsenstein, 1993). Ordination of populations was performed by Multidimensional Scaling (procedure MDSCALE in NTSYS-pc). A Mantel test (Sokal & Rohlf, 1995) was used to test for a relationship between genetic and geographical distances. The distance between populations was estimated using simple trigonometric formulae under the assumption that the Earth is a sphere with radius 6360 km.

Migration and isolation by distance The number of effective migrants (*Nm*) between populations was estimated using the relationship $Nm = (1/F_{ST} - 1)/4$ that holds for an island model (Wright, 1969) and by the private alleles method (Slatkin, 1985). Finally, isolation by distance was analysed according to Slatkin (1993).

Results

Polymorphism

Mean heterozygosity (H_{exp}) ranged from 8.8% (*L. olgensis*; Olga Bay) to 19.5% (Table 2; the allozyme frequency table is available from the first author). There were very few rare alleles, the only case being found in *L. decidua*, where *Skdh-82* and *Skdh-71* were present with average frequencies of 60% and 2%, respectively, whereas they were absent in other populations. It should be noted that in seedling tissues the products of these alleles displayed much lower activity ('seminull alleles') than other alleles. A null-allele could only be recorded in the homozygous state, as only diploid individuals were examined. Such a homozygote was observed only once in the progeny of a tree from population 11 (Yada); the seedling extract showed no activity for the analysed *Skdh* locus and was considered as a null-allele homozygote. Additional analysis of megagametophytes from this parent tree confirmed that the parental tree was a null-allele heterozygote.

Statistical independence among loci

Overall, 1199 pairs of loci were analysed. Among them only 50 pairs showed a statistically significant linkage disequilibrium at the 5% level. Thus,

globally, independence between loci cannot be rejected.

Hardy–Weinberg equilibrium

A total of 255 exact tests of Hardy–Weinberg proportions were carried out with the Markov Chain Monte Carlo method for the 13 polymorphic loci and 32 populations. Statistically significant departures ($P < 0.05$) were detected in 19 cases: the loci *Got-B*, *Pgm-B* and *Sod-A* were involved in one case each, *G6pdh*, *6Pgd* and *Skdh* in two cases each, *Pgm-A* in three cases, and *Est* in seven cases. Most (18 out of 19) of these departures were heterozygote deficits. Overall, only the *Est* locus significantly departed from Hardy–Weinberg equilibrium ($\chi^2_{62} = 126.6$, $P < 0.0001$ for Fisher's method). Among the populations only the Magadan population (*L. ochotensis*) and the population from Chadan (*L. sibirica*) had significant deviations over all loci.

Within-species differentiation

The levels of interpopulation differentiation estimated as F_{ST} are given in Table 3. The number of effective migrants *Nm*, estimated according to the equation $Nm = (1/F_{ST} - 1)/4$, was 2.91 for *L. sibirica* and 11.65 for *L. gmelinii*. The same parameter computed using Slatkin's 'private alleles' method was 5.57 for *L. sibirica* and 9.87 for *L. gmelinii*. There was no clear isolation by distance among *L. sibirica* populations ($\log(M) = -0.185$ $\log(\text{distance}) + 1.068$, where $M = (1/F_{ST} - 1)/4$).

The *L. sibirica* populations split into two groups (Fig. 2). The western group consists of the populations from the Urals and western Siberia. The eastern one includes populations located in eastern Siberia and in the north of western Siberia. The frequencies of alleles *Got-A-112*, *Dia-85*, *6Pgd-109*, *Sod-119* and *Est-108* give the strongest differentiation between these two groups. Some of the *L. sibirica* populations (populations 4, 10, 7, 31, 32), characterized by intermediate frequencies of these alleles, occur between these two groups on the scatter-plot. There is a clear dependence between Nei's genetic distance between any two populations and their longitudinal distance ($P = 0.0094$, Mantel test), but no significant relationship with latitude ($P = 0.6482$, Mantel test; Fig. 2). Moreover, the higher the latitude, the more narrow the zone of intermediate populations. For instance, in the northern part of the Siberian larch range, population 3 which is classified as 'western' and population 11 which is classified as 'eastern' are only 260 km apart.

Table 2 Average sample size per locus, average number of alleles (*A*), proportion of polymorphic loci (*P*), and observed and expected heterozygosities (H_{obs} and H_{exp}) in 32 populations of *Larix* examined at 15 allozyme loci (standard errors in parentheses)

Population	Sample size	*A*	*P*†	Mean heterozygosity	
				H_{obs}	H_{exp}‡
L. sibirica					
1	47.1	2.1	66.7	0.141	0.158
	(0.4)	(0.3)		(0.046)	(0.054)
2	46.6	1.5	46.7	0.134	0.150
	(1.3)	(0.2)		(0.054)	(0.054)
3	94.9	1.8	66.7	0.131	0.139
	(1.0)	(0.2)		(0.040)	(0.044)
4	48.5	2.0	66.7	0.180	0.195
	(1.1)	(0.3)		(0.051)	(0.055)
5	47.5	1.9	60.0	0.137	0.144
	(0.3)	(0.2)		(0.052)	(0.053)
6	32.8	1.9	60.0	0.192	0.174
	(0.2)	(0.2)		(0.061)	(0.053)
7	48.0	1.9	60.0	0.125	0.138
	(0.0)	(0.2)		(0.045)	(0.052)
8	50.3	2.2	66.7	0.173	0.176
	(0.3)	(0.3)		(0.050)	(0.049)
9	45.8	1.9	66.7	0.159	0.159
	(0.2)	(0.2)		(0.047)	(0.047)
10	42.9	1.9	53.3	0.147	0.167
	(0.8)	(0.3)		(0.048)	(0.053)
11	51.5	1.9	66.7	0.115	0.134
	(0.5)	(0.2)		(0.035)	(0.041)
12	48.2	2.1	66.7	0.172	0.170
	(0.7)	(0.2)		(0.050)	(0.048)
30	60.1	2.1	73.3	0.161	0.160
	(2.2)	(0.2)		(0.047)	(0.045)
31	30.6	2.0	60.0	0.148	0.173
	(0.2)	(0.3)		(0.045)	(0.053)
32	31.8	1.7	53.3	0.130	0.145
	(0.2)	(0.2)		(0.045)	(0.052)
L. gmelinii					
13	47.9	2.3	73.3	0.175	0.186
	(0.6)	(0.3)		(0.051)	(0.053)
14	48.2	2.1	66.7	0.161	0.146
	(0.8)	(0.3)		(0.059)	(0.052)
15	46.3	2.3	66.7	0.139	0.151
	(0.5)	(0.3)		(0.043)	(0.049)
16	46.7	2.0	60.0	0.130	0.140
	(0.9)	(0.3)		(0.046)	(0.052)
17	44.3	2.1	60.0	0.160	0.163
	(0.4)	(0.3)		(0.055)	(0.054)
18	47.5	2.5	66.7	0.161	0.163
	(0.2)	(0.4)		(0.052)	(0.052)
L. cajanderi					
19	34.8	2.0	53.3	0.126	0.120
	(0.9)	(0.3)		(0.046)	(0.041)

Table 2 *Continued*

Population	Sample size	A	P†	H_{obs}	H_{exp}‡
				Mean heterozygosity	
L. decidua					
20	67.5 (1.3)	2.1 (0.3)	53.3	0.133 (0.053)	0.134 (0.055)
21	50.1 (0.6)	2.3 (0.4)	53.3	0.182 (0.066)	0.194 (0.071)
L. czekanovskii					
22	32.6 (0.3)	2.1 (0.2)	73.3	0.147 (0.042)	0.167 (0.048)
L. olgensis					
23	49.9 (0.1)	1.5 (0.2)	40.0	0.085 (0.038)	0.088 (0.039)
24	50.0 (1.0)	2.1 (0.3)	66.7	0.103 (0.035)	0.099 (0.032)
25	47.9 (0.1)	1.9 (0.3)	60.0	0.109 (0.039)	0.108 (0.038)
L. kaempferi					
26	50.8 (0.2)	1.8 (0.2)	53.3	0.117 (0.043)	0.133 (0.050)
L. kamtschatica					
27	49.8 (0.1)	1.8 (0.2)	73.3	0.160 (0.047)	0.160 (0.047)
L. amurensis					
28	44.4 (0.3)	2.3 (0.3)	66.7	0.160 (0.040)	0.178 (0.043)
L. ochotensis					
29	47.6 (0.4)	2.1 (0.3)	66.7	0.155 (0.048)	0.181 (0.054)

†Locus is considered polymorphic if more than one allele was detected.
‡Hardy–Weinberg expectation. Unbiased estimate (see Nei, 1987).

In contrast, in southern Siberia, the allele frequencies of populations ranging from the Altai Mountains up to Tuva (populations 7, 31, 30) are intermediate between those characterizing western and eastern populations.

Between-species differentiation

Larix amurensis, *L. cajanderi*, *L. olgensis* and *L. gmelinii* are very similar genetically (Table 4; Figs 2 and 3). The frequency of allele *Pgi-B-79* in population 29 (Magadan) of *L. ochotensis* was about 50%, whereas it was less than 13% in all other populations. As illustrated by the UPGMA dendrogram (Fig. 3) and ordination of the populations using multidimensional scaling (Fig. 2), Nei's genetic distances and classical taxonomy give congruent groupings. The genetic difference between *L. ochotensis* (Magadan

population) on the one hand, and *L. olgensis*, *L. gmelinii* and *L. cajanderi* on the other hand, is mainly caused by frequency differences of allele *Pgi-B-79*. Finally, genetic differences were more pronounced between *L. kamtschatica* and *L. kaempferi* (Table 4), than between *L. kamtschatica* and *L. cajanderi*, *L. ochotensis* and *L. amurensis*.

Discussion

Three main conclusions can be drawn from our study. First, the classification of species according to Nei's genetic distances is congruent with classical taxonomy. For instance, allozymes reveal the difference between eastern and western (*L. sukaczevii* Dyl.) populations of *L. sibirica*. Secondly, whereas there is a marked dependence of the genetic distance between any two populations of *L. sibirica*

Table 3 Wright's *F*-statistics at all loci in *Larix sibirica* (populations 1–12, 30, 31, 32) and *L. gmelinii* (populations 13–18)

Locus	F_{IS}	F_{IT}	F_{ST}	P
L. sibirica				
Got-A	0.048 (0.046)†	0.148 (0.060)	0.104 (0.031)	0.0000
Got-B	0.050 (0.043)	0.103 (0.051)	0.056 (0.022)	0.0000
Dia	0.035 (0.073)	0.118 (0.058)	0.087 (0.031)	0.0000
Pgi-B	−0.004 (0.037)	0.039 (0.041)	0.043 (0.018)	0.0000
G6pdh	0.037 (0.061)	0.119 (0.057)	0.085 (0.018)	0.0000
6Pgd	0.002 (0.051)	0.059 (0.051)	0.057 (0.021)	0.0000
Skdh	0.042 (0.093)	0.096 (0.078)	0.057 (0.018)	0.0000
Pgm-A	−0.046 (0.032)	−0.011 (0.029)	0.033 (0.010)	0.0000
Pgm-B	—	—	—	0.5652
Sod-A	0.100 (0.050)	0.148 (0.049)	0.054 (0.014)	0.0000
Est	0.131 (0.045)	0.216 (0.041)	0.098 (0.019)	0.0000
Idh	—	—	—	1.0000
Mean	0.059 (0.020–0.093)‡	0.134 (0.081–0.171)	0.079 (0.058–0.091)	0.0000
L. gmelinii				
Got-A	−0.028 (0.009)	−0.025 (0.008)	0.003 (0.004)	0.3312
Got-B	0.344 (0.195)	0.391 (0.222)	0.057 (0.035)	0.0015
Dia	0.003 (0.058)	0.045 (0.053)	0.043 (0.023)	0.0002
Pgi-B	0.062 (0.081)	0.063 (0.081)	0.002 (0.007)	0.0739
G6pdh	0.012 (0.063)	0.009 (0.061)	−0.003 (0.003)	0.7909
6Pgd	0.247 (0.177)	0.259 (0.188)	0.012 (0.013)	0.0659
Skdh	−0.001 (0.073)	0.014 (0.067)	0.014 (0.015)	0.0136
Pgm-A	0.046 (0.055)	0.055 (0.056)	0.009 (0.009)	0.0138
Pgm-B	0.147 (0.169)	0.150 (0.162)	0.005 (0.008)	0.0149
Sod-A	0.036 (0.014)	0.054 (0.044)	0.019 (0.008)	0.0029
Est	−0.026 (0.038)	0.007 (0.042)	0.033 (0.028)	0.0002
Mean	0.024 (0.001–0.052)	0.044 (0.026–0.064)	0.021 (0.009–0.031)	0.0000

Supposedly intermediate populations were not included. *P* gives the corresponding probability of Fisher's exact test on contingency tables. The overall *P*-value refers to Fisher's combined probability test.
†Standard deviation in parentheses (computed by jackknifing over populations).
‡95% confidence interval in parentheses (computed by bootstrapping over loci).
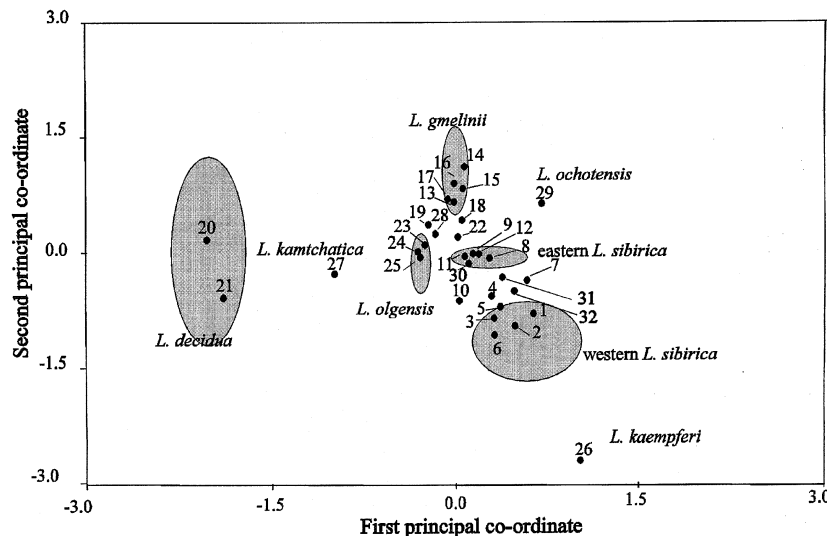

Fig. 2 Multidimensional Scaling Analysis, using Nei's genetic distance (Nei, 1987) matrix, based on 15 allozyme loci in 32 *Larix* populations. The population numbers correspond to those given in Table 1. The groupings (ellipses) correspond to different taxa.

Table 4 Mean genetic distances (*D*; Nei, 1987) within and among the *Larix* taxa

Taxon	Number of populations	Taxon										
		1	2	3	4	5	6	7	8	9	10	11
1. Western *L. sibirica*	7	0.008										
2. Eastern *L. sibirica*	5	0.025	0.008									
3. *L. gmelinii*	6	0.069	0.034	0.004								
4. *L. cajanderi*	1	0.050	0.023	0.013	—							
5. *L. decidua*	2	0.102	0.095	0.099	0.073	0.023						
6. *L. czekanovskii*	1	0.036	0.010	0.014	0.010	0.076	—					
7. *L. olgensis*	3	0.041	0.020	0.028	0.009	0.073	0.013	0.006				
8. *L. kaempferi*	1	0.088	0.121	0.159	0.115	0.171	0.136	0.112	—			
9. *L. kamtschatica*	1	0.072	0.039	0.064	0.037	0.080	0.043	0.037	0.141	—		
10. *L. amurensis*	1	0.043	0.015	0.014	0.006	0.066	0.007	0.007	0.121	0.023	—	
11. *L. ochotensis*	1	0.063	0.044	0.034	0.034	0.102	0.031	0.042	0.166	0.070	0.030	—
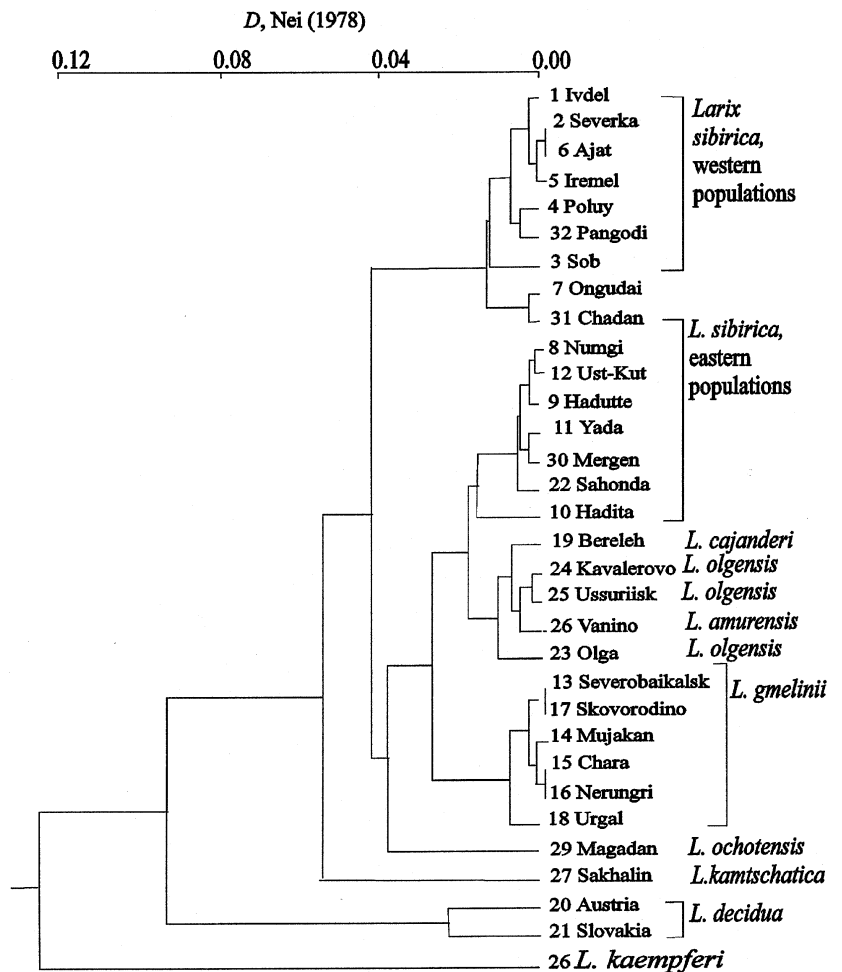
Fig. 3 UPGMA dendrogram of 32 *Larix* populations based on 15 allozyme loci and Nei's genetic distance, *D* (Nei, 1987).

and the longitudinal distance separating them, there is no significant relationship with latitude. Thirdly, there is a decrease in genetic variation in marginal populations.

Genetic variation at the margin

As in Douglas-fir (Li & Adams, 1989) and Jeffrey pine (Furnier & Adams, 1986), marginal populations of larch were genetically depauperate. For example, the populations at Olga Bay and at the Polar tree line (populations 11 and 19) had low heterozygosity. These populations are generally very small and isolated and the narrowness of the ecological niche occupied by larch at Olga Bay probably results from competition from the Manchurian broad-leaved flora and historical events (climatic fluctuation, anthropogenic factors).

Population differentiation

The F_{ST} values obtained in our study were of the same order of magnitude as those observed in other conifer species (*Picea abies* Karst., 5.2% (Lager-crantz & Ryman, 1990); *Picea maritima* (Mill.) B.S.P., 5.9% (Yeh *et al.*, 1986); *Pinus sibirica* du Tour, 1.6% (Krutovsky *et al.*, 1989); *Pinus sylvestris* L., 2.5% (Semerikov *et al.*, 1993)). In *L. occidentalis*, estimates vary from 8.6% (Fins & Seeb, 1986) to 5% (Cheliak *et al.*, 1988; overall range). The majority of *Larix* species of section Pauciseriales exhibit little genetic differentiation, as do other conifers which occupy continuous and recently recolonized areas. Substantial gene flow and history (colonization from a single refuge) are possible causes of the observed lack of genetic differentiation. Based on morphological variability, Dylis (1947) concluded that *L. sibirica* includes two species, a western one, which he named *L. sukaczewii*, and an eastern one, *L. sibirica*, with a border lying approximately along the Ob and Irtish Rivers (see Fig. 1). Allozymes reveal that there are two genetically differentiated groups in the western and eastern parts of the *L. sibirica* range, but allozyme data do not correspond exactly to Dylis's hypothesis because in the southern part of the range, allele frequencies change gradually and eastern and western populations cannot clearly be distinguished. This population differentiation pattern may be the result of the recolonization of northern Eurasia by *L. sibirica* during the late Pleistocene and Holocene. There are strong suggestions that forests remained only in restricted refuges during the last glacial maxima. Some of the refugia were preserved in southern Siberia, south of the

Urals and in northern Kazakhstan. So, at the end of glaciation, forest species extended into the north of the Urals and northern Siberia, preserving the allozyme allele frequencies of the refugia. The genetic similarity of larches from the Yamal and Taz Peninsulas (populations 8, 9 and 11) with those from the Upper Yenisey (populations 30 and 31) could, for instance, be connected with the spread of the forests along the Yenisey and Taz rivers. Fossil records also suggest that, unlike other tree species, larch glacial refugia might never have completely disappeared from higher latitudes during the last glaciation (Kremenetski, 1994), and the present genetic pattern might reflect postglacial expansion from these surviving populations. However, the results reported here are more consistent with the latter because a more pronounced longitudinal, rather than latitudinal gradient was observed. Furthermore, had the southern populations been glacial refugia one might expect to find a markedly higher number of alleles in these populations, which was not the case. On the other hand, differentiation between southern and northern populations is not pronounced and neither of the two alternatives can be unambiguously rejected. This issue may be clarified by the use of DNA markers, for instance cpDNA or mtDNA sequences, that give information on allele genealogies.

Hybridization and species differentiation

In the Russian far east only two larch species, *L. gmelinii* and *L. olgensis*, can be easily identified. Both of them have characteristic morphological traits and occupy well-defined areas. Outside these areas, morphologically variable populations occur, in which features of *L. gmelinii*, *L. olgensis* and even *L. kaempferi* can be found. Characteristic features of mature cones (flat, marginally jagged and nonpubescent cone scales) of some far-eastern larch species (*L. cajanderi*, *L. ochotensis*, *L. amurensis* and *L. kamtschatica*) are close to those of *L. gmelinii*, but other traits (large cones, numerous and sometimes convex or marginally recurved cone scales) resemble more closely those found in either *L. olgensis* or *L. kaempferi*. Therefore a hypothesis of introgressive hybridization of larches in the far east is quite plausible (Dylis, 1961; Bobrov, 1972). The introgression zone between *L. gmelinii* and *L. sibirica* probably dates from the Holocene or late Pleistocene, at least to the north of Lake Baikal. This is supported by the narrowness of this zone and the co-occurrence in this region of the marginal areas of many plant species, e.g. *Chosenia arbutifolia* (Pall.) A. Skvorts.,

Pinus pumila R.G.L., *Betula middendorfii* Trautv. et Mey, *B. exilis* Sukacz, *Populus suaveolens* Fish.; Sokolov *et al.* (1977). In contrast, the introgressive zone between *L. gmelinii* and *L. olgensis* is wide, and possibly originated in the early Pleistocene when *L. gmelinii* first occurred (Bobrov, 1972).

Larix olgensis appears to be genetically more similar to eastern populations of *L. sibirica* ($D = 0.020$) than to *L. gmelinii* ($D = 0.028$), suggesting that the *L. olgensis* – *L. sibirica* divergence took place more recently than that of *L. olgensis* and *L. gmelinii*, although the difference is not very large. However, taking into consideration the similarity of *L. olgensis* and *L. sibirica* for the main taxonomic morphological traits (Bobrov, 1972), this inference may be regarded as quite plausible. The oldest larch fossil records are from the early Pleistocene in north-eastern Siberia (Dylis, 1961). It is generally supposed that after the emergence of *L. gmelinii*, it expanded to the south and south-west and forced out both *L. sibirica* and *L. olgensis*, as they were less adapted to the new, more continental, climatic conditions of the Pleistocene (Dylis, 1961). As a result, the common *L. sibirica* – *L. olgensis* range was disrupted. The genetic closeness of these two species might therefore be a consequence of the relative short period of time that elapsed since this event, although additional investigations are needed to clarify this.

The larger genetic differences between *L. kamt-schatica* and *L. kaempferi* (Table 4) compared to those between *L. kamtschatica* and *L. cajanderi*, *L. ochotensis* and *L. amurensis* suggest that *L. kamt-schatica* originated from continental populations rather than from *L. kaempferi*. A clearer picture emerges from this allozyme study than from the cpDNA study carried out by Qian *et al.* (1995). These authors suggest that it might be a consequence of the importance of hybridization, cpDNA, which is paternally inherited in larch, crossing barrier species more readily than nuclear DNA. Whether this is actually the case remains to be shown.

Acknowledgements

References

BELOVA, V. A. 1985. *Vegetation and Climate of South of East Siberia in Later Cenozoic*. Nauka, Novosibirsk (in Russian).

BENNETT, K. D. 1997. *Evolution and Ecology. The Pace of Life*. Cambridge University Press, Cambridge.

BOBROV, E. G. 1972. *History and Systematics of* Larix. Nauka, Leningrad (in Russian).

BOJARSKAJA, T. D. AND MALAJEVA, E. M. 1967. *Development of Vegetation in Siberia and Far East in Quaternary*. Nauka, Moscow (in Russian).

CHELIAK, W. M., WANG, J. AND PITEL, J. A. 1988. Population structure and genetic diversity in tamarack, *L. laricina* (Du Roi) K. Koch. *Can. J. Forest Res.*, **18**, 1318–1324.

DUMOLIN-LAPÉGUE, S. 1998. *Variabilité chloroplastique et mitochondriale chez les chênes blancs européens à deux échelles spatiales.* Ph.D. Thesis, Université de Paris-Sud.

DYLIS, N. W. 1947. *Siberian Larch*. Moscow Society of Nature Investigators, Moscow (in Russian).

DYLIS, N. W. 1961. *Larch of Eastern Siberia and Far East*. Academy of Sciences of USSR, Moscow (in Russian).

FEDOROV, V., FREDGA, L. AND JARRELL, G. H. 1998. Mitochondrial DNA variation and the evolutionary history of chromosome races of collared lemmings (*Dicrostonyx*) in the Eurasian Arctic. *J. Evol. Biol.*, in press.

FELSENSTEIN, J. 1993. PHYLIP (*Phylogeny Inference Package*). Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FINS, L. AND SEEB, L. W. 1986. Genetic variation in allozymes of western larch (*Larix occidentalis*). *Can. J. Forest Res.*, **16**, 1013–1018.

FISHER, R. A. 1954. *Statistical Methods for Research Workers*, 12th edn. Oliver & Boyd, Edinburgh.

FURNIER, G. R. AND ADAMS, W. T. 1986. Geographic patterns of allozyme variation in Jeffrey pine. *Am. J. Bot.*, **73**, 1009–1015.

GOUDET, J. 1995. FSTAT, version 1.2. A computer program to calculate *F*-statistics. *J. Hered.*, **86**, 485–486.

HALDANE, J. B. S. 1954. An exact test for randomness of mating. *J. Genet.*, **52**, 631–635.

HARRIS, H. AND HOPKINSON, D. A. 1976. *Handbook of Enzyme Electrophoresis*. North-Holland, Amsterdam.

KISANUKI, H., KURAHASHI, A., KATO, H., TERAUCHI, R., KAWANO, S., IDE, Y. AND WATANABE, S. 1995. Interspecific relationship of genus *Larix* inferred from the RFLPs of chloroplast DNA. *J. Jap. For. Soc.*, **77**, 83–85.

KOLESNIKOV, B. P. 1946. Contribution to the taxonomy and the historical development of larches (section *Pauciseriales* Patschke). In: *Materials on the History of the Flora*

and Vegetation of the USSR, vol. 2, pp. 321–364. Nauka, Moscow (in Russian).

KREMENETSKI, C. V. 1994. Holocene history of ranges of main coniferous trees in Siberia. In: *Short Periodical and Sharp Landscape and Climatic Fluctuations During Last 15 000 Years*, pp. 160–210. Institute of Geography of Russian Academy of Science, Moscow (in Russian).

KRUKLIS, M. V. AND MILYUTIN, L. I. 1977. *Larix czekanovskii*. Nauka, Moscow (in Russian).

KRUTOVSKY, K. V., POLITOV, D. V., ALTUKHOV, YR. P., MILUTIN, L. I., KUZNETSOVA, G. V., IROSHNIKOV, A. I. ET AL. 1989. Genetic variability in Siberian Cedar pine. IV. Genetic diversity and differentiation between populations. *Genetika*, **25**, 2009–2032 (in Russian).

LAGERCRANTZ, U. AND RYMAN, N. 1990. Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. *Evolution*, **44**, 38–53.

LEPAGE, B. A. AND BASINGER, J. F. 1995. The evolutionary history of the genus *Larix* (Pinaceae). In: *Ecology and Management of* Larix *Forests: A Look Ahead*, pp. 19–29. Proceedings of an International Symposium. October, 5–9, 1992, Whitefish, MT, U.S.A.

LI, P. AND ADAMS, W. T. 1989. Ring-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). *Can. J. Forest Res.*, **19**, 149–161.

MAIER, J. 1992. Genetic variability in European larch (*Larix decidua* Mill.). *Ann. Sci. For.*, **49**, 39–47.

NEI, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

NEISHTADT, M. I. 1957. *History of Forests and Paleogeography of USSR in Holocene*. Academy of Sciences of USSR, Moscow (in Russian).

QIAN, T., ENNOS, R. A. AND HELGASON, T. 1995. Genetic relationships among larch species based on analysis of restriction fragment variation for chloroplast DNA. *Can. J. Forest Res.*, **25**, 1197–1202.

RAYMOND, M. AND ROUSSET, F. 1995a. An exact test for population differentiation. *Evolution*, **49**, 1283–1286.

RAYMOND, M. AND ROUSSET, F. 1995b. GENEPOP (version 1.2). A population genetics software for exact tests and ecumenism. *J. Hered.*, **86**, 248–249.

ROHLF, F. J. 1988. NTSYS-PC. *Numerical taxonomy and multivariate analysis system.* Exeter Publishing, New York.

ROUSSET, F. AND RAYMOND, M. 1995. Testing heterozygote excess and deficiency. *Genetics*, **140**, 1413–1419.

SEMERIKOV, V. L. AND MATVEEV, A. V. 1995. Investigation of genetic variation of allozyme loci in Siberian larch. *Russian J. Genet.*, **31**, 944–949.

SEMERIKOV, V. L., PODOGAS, A. V. AND SHURKHAL, A. V. 1993. Variability of allozyme loci in populations of common pine. *Russian J. Ecol.*, **24**, 14–20 (in Russian).

SHIGAPOV, Z. KH., PUTENIKHIN, V. P., SHIGAPOVA, A. I. AND URAZBAKHTINA, K. A. 1998. Genetic structure of the Ural populations of *Larix sukaczewii* Dyl. *Genetika*, **34**, 65–74 (in Russian).

SHURKHAL, A. V., PODOGAS, A. V. AND ZHIVOTOVSKY, L. A. 1992. Allozyme differentiation in the genus *Pinus*. *Silvae Genet.*, **41**, 105–109.

SLATKIN, M. 1985. Rare alleles as indicators of gene flow. *Evolution*, **39**, 53–65.

SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

SOKAL, R. R. AND ROHLF, F. J. 1995. *Biometry*, 3rd edn. W.H. Freeman and Co., New York.

SOKOLOV, S. J., SVAZEVA, O. A. AND KUBLI, V. A. 1977. *Areas of Trees and Shrubs in USSR*, vol. 1. Nauka, Leningrad (in Russian).

SUKACZEW, V. N. 1924. *On the Evolutionary History of* Larix. People's Commissariat of Agriculture, Leningrad (in Russian).

WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

WRIGHT, S. 1969. *Evolution and the Genetics of Populations*, vol. 2, *The Theory of Gene Frequencies*. University of Chicago Press, Chicago.

YEH, F. C., KHALIL, M. A. K., EL-KASSABY, Y. A. AND TRUST, D. C. 1986. Allozyme variation in *Picea maritima* from Newfoundland: genetic diversity, population structure and analysis of differentiation. *Can. J. Forest Res.*, **16**, 713–720.