Resources. Sinauer Associates, Sunderland, M. A., 43–63. — HAMRICK, J. L., GODT, M. J. and SERMAN-BROYLES, S. L.: Factors influencing levels of genetic diversity in woody plant species. New Forest **6**, 95–124 (1992). — HARRY, D. E.: Identification of a locus modifying the electrophoretic mobility of malate dehydrogenase isozymes in incense-cedar (*Calocedrus decurrens*), and its implications for population studies. Biochem. Genet. **21**, 417–434 (1983). — HARRY, D. E.: Inheritance and linkage of isoenzyme variants in incense-cedar. J. Hered. **77**, 261–266 (1986). — HAVRYLENKO, M. ROSSO, P. A. and FONTENLA, S. B.: *Austrocedrus chilensis*: contribución al estudio de su mortalidad en Argentina. Bosque **10**, 29–36 (1989). — HUECK, K.: Los bosques de Sudamérica. Soc. Alemana de Cooperación Técnica GTZ, Eschborn. pp. 474 (1978). — JOHN, R. and JONES, R.: The nature of the multiple forms of cytoplasmatic aspartate aminotransferasa from pig and sheep heart. Biochemical J. **141**, 401–406 (1974). — KRUTOVSKII, K. V. and BERGMANN, F.: Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) KARST., and Siberian, *P. obovata* LEDEB., spruce species studied by isozyme loci. Heredity **74**, 464–480 (1995). — LEWANDOWSKI, A. and MEJNARTOWICZ, A. L.: Levels and patterns of allozyme variation in some European larch (*Larix decidua*) populations. Hereditas **115**, 221–226 (1991). — LOVELESS, M. D. and HAMRICK, J. L.: Ecological determinants of genetic structure in plant populations. Ann. Rev. Ecol. Syst. **15**, 65–95 (1984). — MILLAR, C. I. and MARSHALL, K. A.: Allozyme variation of Port-Orford-cedar (*Chamaecyparis lawsoniana*): implications for genetic conservation. For. Sci. **37**, 1060–1077 (1991). — MÜLLER-STARCK, G., BARADAT, PH. and BERGMANN, F.: Genetic variation within European tree species. New Forests **6**, 23–47 (1992). — SCANDALIOS, J.: Genetic control of multiple molecular forms of enzymes in plants: A review. Biochem. Genet. **3**, 37–79 (1969). — SELANDER, R. K., SMITH, M. H., YANG, S. Y., JOHNSON, W. E. and GENTRY, G. B.: Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*). Univ. Tex. Publ. 7103, 49–90 (1971). — SHAW, C. R. and PRASAD, R.: Starch gel electrophoresis of enzymes – A compilation of recipes. Biochem. Genet. **4**, 297–320 (1979). — SPIESS, E. B.: Genes in Populations. John Wiley and Sons, New York (1990). — YEH, F. C.: Isozyme variation of *Thuja plicata* (CUPRESSACEAE) in British Columbia. Biochem. Syst. and Ecol. **16**, 373–377 (1988).

Characterization and Variability of Seed Storage Protein Subunits in *Sophora japonica* (Fabaceae)

By C. GONZÁLEZ, C. GARMA and N. HENRIQUES-GIL

Unidad de Genética, Universidad San Pablo CEU, Montepríncipe, Boadilla del Monte, 28660 Madrid, Spain

(Received 14th June 1995)

Summary

The seed storage globulins of *Sophora japonica* correspond to 7 main groups of polypeptide monomeres – SP98, SP9*, SP74, SP7*, SP37, SP23 and SP22 – with molecular weights ranging from 98 kDa to 22 kDa that can be ascribed to 6 different loci (SP9* and SP7* polypeptydes may show different allelic forms of more than 90 kDa and 70 kDa respectively). SP9* form both disulphur-linked dimers and tetrameres, SP37 forms an heterodimer with SP23, and SP22 an homodimer. The polypeptides from 79 kDa to 71 kDa seem to be subject of some post-traductional partial modification.

Variability, in what molecular weight of monomere polypeptides was concerned, was found to be affected with loci *sp9*, sp7*, sp37*. Although this species was imported to Spain in the present century, 2 stands (populations) analysed show a high degree of variability (mean heterozygosity as 0.213 and polymorphic index as 0.288).

Key words: Sophora japonica, seed storage proteins, SDS-PAGE, polymorphism.

FDC: 165.53; 161.4; 161.6; 176.1 *Sophora japonica.*

Introduction

The pagoda tree, *Sophora japonica,* belongs to the subfamily Cæsalpinioideae and was originated from east Asia. During the last century *S. japonica* has been imported to Europe because it is well adapted in urban forestry and, at present, it is even frequent in natural or semi-natural woods. However, the facts that neither the populations of origin nor the existing levels of variability were known, the actual culture of this species may face problems such as inbreeding or to establish the best stands for use as seed source.

Legume seeds have been subject of intensive work due to their high protein content. These proteins are usually classified into 2 main groups, legumins and vicilins (DERBYSHIRE *et al.,* 1976). Most results about seed protein genetics refer to crops employed in human and animal feeding like the genera *Vicia, Glycine, Pisum* or *Phaseolus* (for review, see CASEY *et al.,* 1986) contrasting with the scarce data concerning tree species. In some cases, seed protein variation has been employed in taxonomic or evolutionary studies of certain tree groups (GIFFORD, 1988; JENSEN and BERTHOLD, 1989; JENSEN and LIXUE, 1991). Long generation periods make difficult traditional genetic studies, so that, for instance, a method designed for gene mapping in human families have been employed to establish linkage maps in *Pinus pinaster* (GERBER *et al.,* 1993). In *Cercis siliquastrum,* a cæsalpinoid tree, the genetic control of seed storage proteins and the linkage relationships of the correspondent loci have been recently described (GONZÁLEZ and HENRIQUES-GIL, 1994). It is important to keep in mind that many Fabaceae trees produce large amounts of edible legumes or seeds. Although this specific use as food is restricted to a few cases like the bean tree, *Ceratonia siliqua*, or some *Cassia* species (SÁNCHEZ-MONGE, 1991), legume seeds may have a stronger ecological importance as a protein source for a number of animals in natural conditions.

The levels of variability found in this type of proteins are usually high (CASEY *et al.,* 1986; TUCCI *et al.,* 1991) since their

function as storage products may suffer little change despite large differences in sequence and size. In this work we define the main features of the different storage protein subunits in the seeds of *S. japonica*. The different alleles found for the correspondent loci allow an approach to the genetic variability of a Spanish population of this species.

Materials and Methods

One hundred and one seeds of *Sophora japonica* were collected during April 1992, in 2 different stands (populations) with ages of about 40 years, in Madrid: Parque del Retiro and Parque del Oeste (51 and 50 trees respectively). The allelic frequencies among the seed populations were estimated from analysis of 1 single seed per adult tree.

For protein extractions, a piece of cotyledon was crushed inside an Eppendorf tube and suspended in 50 µl of a buffer with 2% sodium dodecyl sulfate (SDS), 6.25% Tris HCl pH 6.8, 10% glycerol, 0.01% pyronine Y and 5% 2-mercaptoethanol (2ME) per mg of sample; 2ME was not employed in some extractions for the detection of disulphide-linked polymeres. After 1 hour with occasional shaking, the tubes were heated to 100 °C for 1.5 minutes.

Polypeptide separation was made in 10% polyacrylamide gels plus SDS (SDS-PAGE) as described by PAYNE *et al.,* (1981) with an electrode buffer of 14% glycine, 3% Tris and 1% SDS in distilled water. The gels were run at a constant voltage of 60 V for half an hour and then at 50 V for 14 hours. Two-dimensional electrophoresis was performed in order to reveal more precisely which polypeptides form disulphur links: the samples without 2ME were inserted in 10% polyacrylamide gels in 2 mm diameter tubes; once completed the first running (8 hours), the gels were incubated in 10.3% glycerol, 0.07 M Tris pH 6.8, SDS 2.4% and 1% 2ME buffer at 37 °C for 1.5 hour; the second dimension running was performed in slab gels as described above.

Staining was carried out in a solution of 0.02% Coomassie Brilliant Blue R, 5% ethanol and 6% trichloroacetic acid (TCA) for about 20 hours.

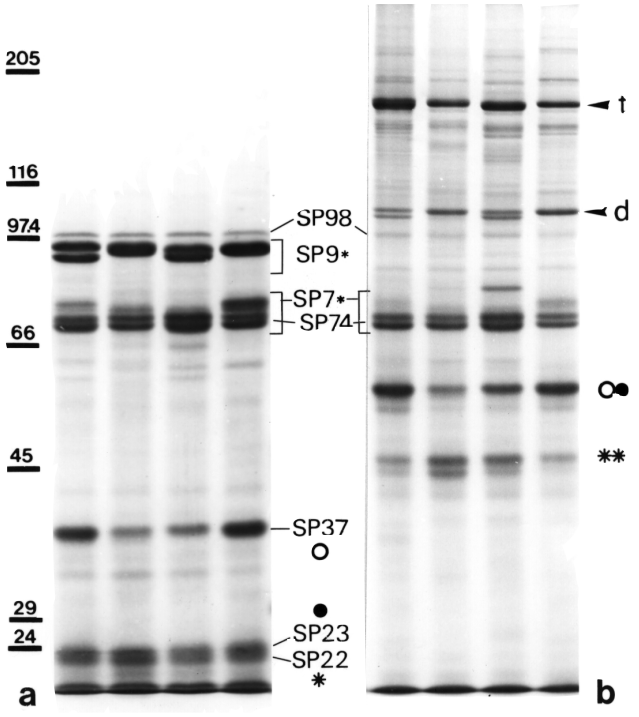

Figure 1. – Proteinograms obtained by SDS-PAGE from seed globulin extracts in presence (a) and absence (b) of 2-ME, for the same 4 seeds. The molecular weigths are given in kilodalton. The polypeptides that show different mobilities in both conditions must have some disulphur linkage: (t) tetramere of SP9*; (d) dimer of SP9*; (○) SP37; (●) SP23; (*) SP22.

Results and Discussion

The proteins of *S. japonica* seeds are salt soluble globulins as it commonly occurs in legume species (CASEY *et al.,* 1986); extractions based on acid or apolar media produce very poor protein extraction (unpublished results).

1) Loci codifying the seed protein subunits

Figure 1 shows a typical SDS-PAGE obtained from seed extracts of *S. japonica*. There were 7 main groups of polypeptides in the material studied. We named the monomeres of *Sophora* seed proteins as SP plus a number that indicates their mean molecular weight (MW). Because of their biochemical differences and variability displayed, each of those groups can be ascribed to 6 different loci. Namely, the loci *sp98, sp74, sp37* and *sp22* encode the polypeptides SP98, SP74, SP37 and SP22 respectively; the locus *sp9** has 3 different alleles with the products SP95, SP94 and SP91; *sp7** also shows 3 variants encoding SP79, SP73 and SP71.

SP23 could come from a 7th locus; however, because it forms an heterodimer with SP37 (see below) a single locus (*sp37*) could be responsible for both polypeptides.

The correspondence is less clear for the polypeptides ranging from 79 kDa to 73 kDa, because the products of 2 different loci overlap in the electrophoresis. Indeed, 2 different loci are necessary to explain the biochemical variability: one, *sp74*, is monomorphic since SP74 appears in all seeds analysed: if it was an allelic form of the others, the 101 seeds would necessarily be heterozygotes; moreover many seeds had SP74 plus 2 of SP79, SP73 and SP71 which is impossible to obtain with a single locus in a diploid species like *S. japonica.* The other locus, *sp7** displayed a typical triallelic variability with all possible genotypes (i.e., homozygotes *sp79,* heterozygotes *sp79-sp73,* and so on).

2) Disulphur links

The polypeptides SP98, SP74 and the group SP79-73-71 showed identical mobilities both in reduced and unreduced conditions (*Figures 1* and *2*) which indicates that no disulphur links exist in these protein subunits. In contrast, the product of the gene *sp9** forms dimers and, more abundantly, tetrameres (*Figures 1* and *2*). This coexistence of multimeric arrangements was also found for the units of higher molecular weight of *Cercis siliquastrum* (GONZÁLEZ and HENRIQUES-GIL, 1994), in *Pisum sativum* (MATTA et al., 1981) and in *Vicia faba* by TUCCI et al. (1991) who concluded that part of the heterogeneity had a genetic basis; a not well defined rule for assembly of native proteins proposed by CASEY et al. (1986) could probably apply for a large number of leguminous (crop and tree species) seed proteins.

Our results also suggest that the monomeres have a non-globular shape that becomes perfectly assembled in the dimer: this would explain an apparent MW of dimers and tetrameres (104 kDa to 107 kDa and 174 kDa respectively) much lower than 2 or 4 times that of the monomere (91 kDa to 95 kDa).

The SP37 and SP23 monomeres form disulphur-linked associations. The heterodimer SP37-SP23 can be clearly demonstrated by 2D-electrophoresis (*Fig. 2*) as the only dimeric band obtained from unreduced extracts is dissociated in their 2 different components. In contrast, SP22 forms an homodimer.

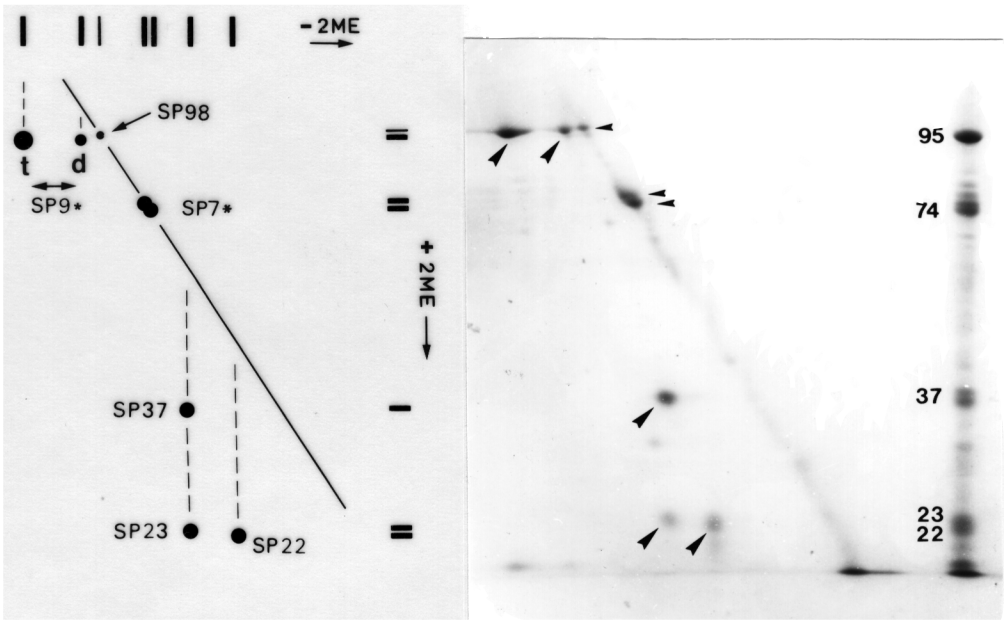

Figure 2. – 2D electrophoresis: the first dimension (left to right) was made from unreduced extracts and the second one (top to bottom) after a treatment with 2ME. At right, a single running of a reduced extract of the same seed. Small arrowheads indicate the polypeptides having the same mobilities in both conditions and bigger arrowheads those that form disulphur links.

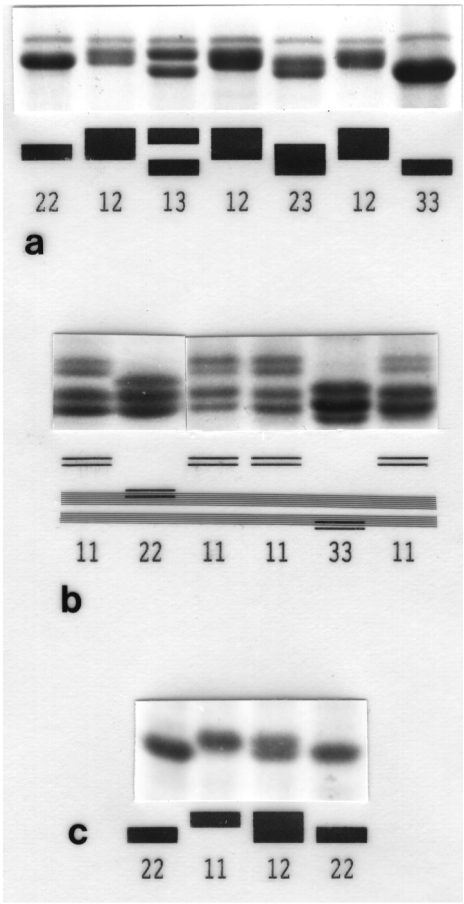

Figure 3. – Allelic variation of proteins SP9* (3a), SP7* (3b) and SP37 (3c) with the correspondent schemes and genotypes; alleles *1*, *2*, and *3*, correspond to polypeptides from higher to lower MW (see text for the exact weight in each case). In the scheme of *fig. 3b*, the monomorphic SP 74 double band (overlapping with SP7*) is represented as a shadow to make easier the interpretations.

3) Post-traductional modification

SP74 and the group SP79, SP73 and SP71 appear each invariably as a double band (*Fig. 3*). In the first case, 2 different loci could be claimed to explain both subunits with slightly different molecular weight; alternatively and keeping in mind their similarity, the double band may be simply a consequence of post-traductional modification. Signal sequence elimination and a variable degree of glycosidation usually occur in nascent storage polypeptides. Some of them are also post-translational modified by proteolytic cleavage (HIGGINS, 1984). This is clearer in the 2nd case (SP79, SP73 and SP71), because despite the interseed variability, 2 bands are still observed for each variant: it would be extremely unprobable that the products of different loci showed an exactly identical pattern of variation.

If SP37 and SP23 are encoded by a single locus, then a single primary product becomes intrachain-linked and after the polypeptidic chain is broken originating the heterodimer observed.

4) Allelic variants and population analysis

As shown in *figure 3*, the polypeptide SP37 may be found in 2 allelic forms of 37.5 kDa and 36.5 kDa; for *sp9** and *sp7** 3 different alleles exist as it was already stated. The polypeptide SP23 and perhaps SP22 also seemed to be polymorphic.

Table 1. – Allelic frequencies among the sample of 101 seeds. Alleles 1 to 3 correspond to slower (higher MW) to faster (lower MW) polypeptide migration.

	Locus				
	sp98	sp9*	sp7*	sp74	sp37
Allele					
1	1.000	0.347	0.505	1.000	0.129
2	-	0.470	0.121	-	0.871
3	-	0.183	0.374	-	-

Nevertheless their fast migration in 10% SDS-PAGE often produces some distortion that makes impossible a systematic analysis; keeping in mind that SP23 may also depend the locus *sp37*, the allelic variation could be a consequence of that described above.

The 2 stands from which seeds were collected showed very similar allelic frequencies with no significant differences in any locus ($\chi^2=6.258$, 4 d.f., N.S.; $\chi^2=4.804$, 2 d.f., N.S.; $\chi^2=0.188$, 1 d.f., N.S. for *sp9**, *sp7** and *sp37* respectively). *Table 1* gives the allelic frequencies among the whole sample of 101 seeds for all loci except *sp22* (for the reason mentioned above). As far as the SDS-PAGE may distinguish, no inter-individual differences were recognized for SP98 and SP74. The other loci showed a strong polymorphism. The mean heterozygosity (H) was H=0.213 and the polymorphic index (PI) as defined by HAMRICK (1979) was PI=0.288. Keeping in mind that we are analysing only a specific type of proteins (seed storage) and that only the variation in MW is detected, those values can be considered high (see the review by HAMRICK, 1979). *S. japonica* is an entomophil pollination species and hence most probably allogamous. The genotypic frequencies for loci *sp9** and *sp37* showed no significant differences with respect to HARDY-WEINBERG expectations, but for locus *sp7** there was a significant increase of homozygotes ($\chi^2=32.617$; d.f.=2; $p<0.001$). A certain degree of selfing could be responsible for this fact, although it is not clear why the heterogeneity of genotypes may affect some loci but not others, such as described in *Robinia pseudoacacia* (SURLES *et al.,* 1990) and in the Douglas-fir (RITLAND and EL-KASSABY, 1985).

European populations of *S. japonica* are quite recent and, in cases of the Spanish populations, data concerning the precise origin and number of seed trees or seed lots imported lack completely (many reforestations are carried out by private companies after an official assignment). Therefore, the diversity in seed protein constitution detected here is important since it implies that these populations of *Sophora* contain a remarkable genetic variability making them an appropriate material for any further selection and breeding programs.

Acknowledgements

The authors are indebted to Ms. VERÓNICA GONZÁLEZ for her excellent and kind assistance in the recollection of the material.

References

CASEY, R., DOMONEY, C. and ELLIS, N.: Legume storage proteins and their genes. Oxford Surv. Plant Mol. Cell Biol. **3**, 1–95 (1986). — DERBYSHIRE, E., WRIGHT, D. J. and BOULTER, D.: Legumin and vicilin, storage proteins of legume seeds. Phytochemistry **15**, 3–24 (1976). — GERBER, S., RODOLPHE, F., BAHRMAN, N. and BARADAT, PH.: Seed-protein variation in maritime pine (*Pinus pinaster* AIT.) revealed by two-dimensional electrophoresis: genetic determinism and construction of a linkage map. Theor. Appl. Genet. **85**, 521–528 (1993). — GIFFORD, D. J.: An electrophoretic analysis of the seed proteins from *Pinus monticola* and eight other species of pine. Can. J. Bot. **66**, 1808–1812 (1988). — GONZÁLEZ, C. and HENRIQUES-GIL, N.: Genetics of seed storage proteins in the love tree *Cercis siliquastrum* L. Theor. Appl. Genet. **89**, 895–899 (1994). — HAMRICK, J. L.: Genetic variation and longevity. In: Topics in plant population biology. O. T. SOLBRIG, S. JANG, G. B. JOHNSON and P. H. RAVEN. Columbia University Press (1979). — HIGGINS, T. J. V.: Synthesis and regulation of major proteins in seeds. Ann. Rev. Plant Physiol. **35**, 191–221 (1984). — JENSEN, U. and BERTHOLD, H.: Legumin-like proteins in gymnosperms. Phytochemistry **28**, 1289–1394 (1989). — JENSEN, U. and LIXUE, C.: *Abies* seed protein profile divergent from other Pinaceae. Taxon **40**, 435–440 (1991). — MATTA, N. K., GATEHOUSE, J. A. and BOULTER, D.: Molecular and subunit heterogeneity of legumin of *Pisum sativum* L. (Garden pea) – a multidimensional gel electrophoretic study. J. Exp. Bot. **32**, 1295–1307 (1981). — PAYNE, P. I., CORFIELD, K. G., HOLT, L. M. and BLACKMAN, J. A.: High-molecular-weight subunits of glutenin and bread making quality in progeny of six crosses of bread wheat. J. Sci. Food Agric. **32**, 51–60 (1981). — RITLAND, K. and EL-KASSABY, Y. A.: The nature of inbreeding in a seed orchard of Douglas-fir as shown by an efficient multilocus model. Theor. Appl. Genet. **71**, 375–384 (1985). — SÁNCHEZ-MONGE, E.: Flora Agrícola. I. M. Agricultura, Pesca y Alimentación. Madrid (1991). — SURLES, S. E., HAMRICK, J. L. and BONGARTEN, B. C.: Mating systems in open-pollinated families of black locust (*Robinia pseudoacacia*). Silvae Genetica **39**, 35–40 (1990). — TUCCI, M., CAPPARELLI, R., COSTA, A. and RAO, R.: Molecular heterogeneity and genetics of *Vicia faba* seed storage proteins. Theor. Appl. Genet. **81**, 50–58 (1991).

Allozyme Variation in Natural Populations of Eurasian Pines

IV. Population Structure and Genetic Variation in Geographically Related and Isolated Populations of *Pinus nigra* ARNOLD on the Crimean Peninsula

By A. E. SILIN and G. G. GONCHARENKO

Department of Molecular Genetics, Institute of Forestry of the Academy of Sciences of Belarus,
71 Proletarskaya Str., 246654 Gomel, Belarus

(Received 4th July 1995)

Summary

Eight natural populations of *Pinus nigra* ARN. occurring on the Crimean peninsula were investigated by starch-gel electrophoresis. A total of 70 alleles were observed at 24 loci. Interpopulation genetic diversity was about 2% of the total genetic diversity. NEI' s genetic distance coefficient ranged from 0.005 to 0.022 among populations and averaged 0.012. The level of gene flow was 18.98 migrants per generation. The results of the analysis of gene diversity, genetic differentiation and gene flow demonstrated that the *P. nigra* populations studied had similar genetic structures. Estimated parameters of genetic variation showed that in *P. nigra* more than 66% of loci were polymorphic and the mean expected and observed heterozygosity values were 0.248 and 0.241, respectively, which allowed to regard the *P. nigra* as one of the most highly variable members of the genus *Pinus*.