Theor Appl Genet (1994) 88:407–411

N. Bahrman · M. Zivy · C. Damerval · Ph. Baradat

Organisation of the variability of abundant proteins in seven geographical origins of maritime pine (*Pinus pinaster* Ait.)

Received: 20 May 1993 / Accepted: 9 September 1993

Abstract The comparison of 42 two-dimensional protein patterns from megagametophytes of maritime pine from seven geographical origins enabled the analysis of the genetic variability of abundant proteins. More than 84% of the polypeptides were variable. The intra- and inter-origin variability levels were of a similar magnitude. Correspondence analysis and a dendrogram computed using a dissimilarity index between individuals showed three main groups. The first group included the individuals from Landes (France), Portugal, eastern Spain, and Corsica, without individualising the provenances. The second group was composed of accessions from Italy and Sardinia, and the individuals of each location were separated. The third group included all of the individuals of Moroccan origin. This clustering was in agreement with the Atlantic, Mediterranean and North African structuration of maritime pine established from terpene data.

Key words Maritime pine · Two-dimensional electrophoresis · Proteins · Genetic variability

Introduction

In forest trees, isozyme markers have largely been used to assess genetic diversity (Hamrick and Godt 1990; Müller-Starck et al. 1992), to examine relationships between populations or species (Yeh et al. 1986; Hamrick et al. 1992), and to follow gene flow (Ellstrand 1992). Dependent on the species studied, between 5 and 30 loci were examined in these investigations. However, in order to assess the extent of genetic variability for breeding purposes or to manage genetic resources, such a sampling of the genome may be too limited.

The two-dimensional gel electrophoresis (2-D PAGE) technique enables the analysis of several hundreds of gene products in a single gel (O'Farrell 1975). These proteins are relatively more abundant than isozymes and might be under different functional constraints (McConkey et al. 1979; Walton et al. 1979). They were initially found to be less genetically variable than isozymes (Leigh Brown and Langley 1979; McLellan et al. 1983), but this observation must be reevaluated since it has been shown that technical improvements can result in increased levels of variability (Damerval et al. 1986). In maize, an assessment of the qualitative variation that corresponds to allelic polymorphism revealed a level of variability between inbred lines that was of the same order as that observed with isozymes (about 2.4 alleles per locus; J. Burstin, personal communication). In maritime pine, barley, and maize (Bahrman and Damerval 1989; Gerber et al. 1993; Zivy et al. 1992; Damerval unpublished) it has been possible to map individual genes controlling the qualitative variation (presence/absence variation and position shift) of polypeptides. In these studies 28–81 loci were identified by means of the 2-D PAGE technique, thus showing that 2-D PAGE can be used to analyse the polymorphism of a large number of loci. This technique has already been used for phylogenetic reconstruction and assessment of relationships between populations and species in the animal kingdom (Goldman et al. 1987, 1989; Spicer 1988) as well as in the plant kingdom (Bahrman et al. 1988a, 1988b; Thiellement et al. 1989).

The natural area of maritime pine (*Pinus pinaster* Ait.) is limited to the western part of the Mediterranean basin (Fig. 1), extending from south-western France to Portugal, Spain, Italy, Corsica, Sardinia in the coastal regions and northern Morocco (Mirov 1967). Several lines of arguments have lead to the hypothesis that the quaternary glaciations (Würm III) have played a major role in the structuration of this species (Baradat and

Communicated by P. M. A. Tigerstedt

N. Bahrman (✉) · Ph. Baradat
INRA, Station de Recherches Forestières de Bordeaux-Cestas, Laboratoire de Génétique de Amélioration B. P. 45, F-33611 Gazinet Cédex, France

M. Zivy · C. Damerval
Station de Génétique Végétale, CNRS-INRA-UPS, Ferme du Moulon, F-91190 Gif-sur-Yvette, France

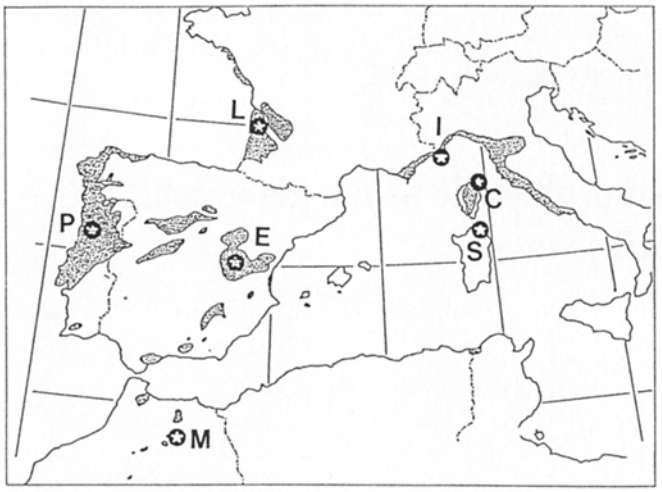

Fig. 1 Natural area and location of the seven origins of maritime pine (see Table 1 for origin code)

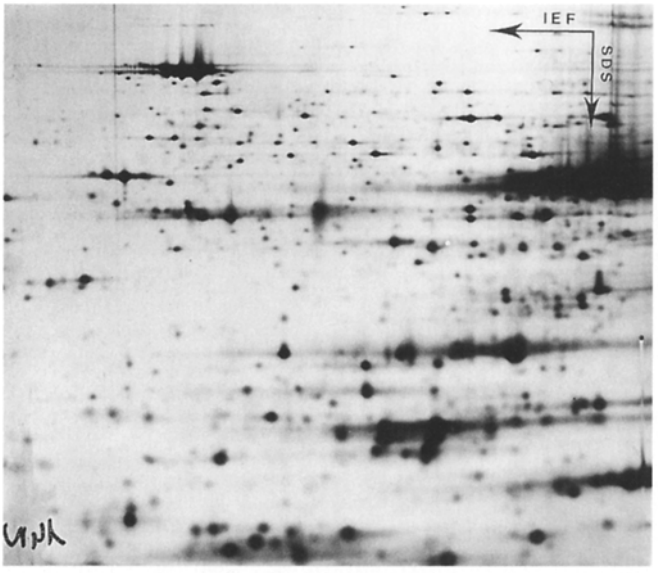

Fig. 2 Two-dimensional protein pattern obtained from the haploid megagametophyte of maritime pine

Marpeau 1988). The division of the area could have resulted in inter-population divergence through genetic drift. In order to study the structuration of genetic variability in this species, we have analysed the polymorphism of total proteins of the megagametophyte by 2-D PAGE in seven populations representative of various geographical origins.

Material and methods

Plant material, protein extraction, and electrophoresis

Forty-two megagametophytes (6 from each provenance, Table 1, Fig. 1) were analysed separately by 2-D PAGE (Fig. 2). The megagametophytes were individually extracted in 6 µl/mg 3 *M* urea, 4% FSN-100, 2% ampholytes (pharmalyte pH 3-10) and 1% dithiothreitol at room temperature according to Anderson et al. (1985) and as described in Bahrman and Damerval (1989). The isoelectrofocusing and the sodium dodecyl sulphate dimensions were performed as in Bahrman and Thiellement (1987). The gels with the Gel Bond supports (Granier and de Vienne 1986) were silver stained according to Damerval et al. (1987).

The comparisons of protein patterns were made visually by superimposition of the dried gels upon a light box. As the small amount of extract precludes repetition of the electrophoresis for each of the 42 megagametophytes only the qualitative variations (presence or absence of the spots) were taken into account. Such variations were found to be under monogenic control in a progeny of maritime pine. The genes involved may correspond either to the structural gene encoding the polypeptide or to another gene controlling the expression of the structural gene (see Bahrman and Damerval 1989 for a more extensive discussion). For convenience, the presence of a polypeptide or its absence is simultaneously referred to here as spot-characters.

Table 1 Location of the seven geographical origins of maritime pine

Origin code	Provenance	Latitude North	Longitude East/West	Elevation (m)
P	Leiria (Portugal)	39° 49′	8° 51′ W	50
M	Tamjout (Morocco)	33° 52′	4° 02′ W	1600
C	Vivario (Corsica)	42° 14′	9° 10′ E	700
L	Landes (France)	44° 20′	0° 35′ W	90
E	Cuenca (Spain)	39° 58′	1° 38′ W	1600
I	Fontanin (Italy)	43° 52′	4° 45′ E	500
S	Monte pino (Sardinia)	40° 51′	9° 08′ E	900

Statistical analyses

In order to describe the data structure, a correspondence analysis (CA) was done, with the megagametophytes as individuals and spots as variables. The presence of spot was encoded as 1 and its absence as 0. The SAS system (1990) was used. Moreover, dissimilarity indices were computed between all pairs of genotypes as:

$$D = 1 - [N(11) + N(00)]/N\text{total},$$

where $N(11)$ is the number of spots present in both individuals, $N(00)$ is the number of spots absent in both individuals, and Ntotal is the total number of spots.

A dendrogram was computed from the dissimilarity matrix showing the relationships between individuals using the SAS system (1990). In order to characterise the intra- and inter-population variability, the mean dissimilarity indices were computed:

$$D_{wI} = \{2/n_I(n_I - 1)\}\Sigma\Sigma D_{ij},$$

where i and j are different individuals of the population I, and

$$D_{bKL} = (1/n_K n_L)\Sigma\Sigma D_{kl},$$

where k is an individual from population K and l an individual from population L.

Results and discussion

Protein variability

For each of the seven geographical origins, 6 megagametophytes, each harvested from a different tree, were

analysed by 2-D PAGE. A total of 968 polypeptide spots were scored over all of the individuals. Table 2 lists the number of spots observed for each of the origins, which varied according to origin from 845 (Italian population) to 870 (Corsican population) [856 ± 11(mean ± SD)]. There were 154 invariable spots detected over all 42 megagametophytes. On average, 36.5% of the spots were fixed per origin and 63.5% were variable. The Moroccan origin had the largest number of variable spots and the Sardinian population had the smallest number of variable spots (Table 2, c). Consistent with these results, the Moroccan origin had the smallest number of invariable spots and the Sardinian origin the largest number of invariable spots (Table 2, b). We also considered the number of spots that are fixed (whether absent or present) in only one origin (Table 2, d and d'). The highest number of such cases of localized fixation (72) was observed for the Sardinian origin. This may indicate that this population diverged through genetic drift and has simultaneously lost much of its diversity as a result of a bottleneck effect.

Origin-characteristic spots (present in only one population but not in all individuals) and origin-specific spots (spots found in 6 individuals of a given origin but never in individuals of the other 6 origins) were not frequent (Table 3). The Moroccan population had more characteristic spots than the other origins (Table 3, upper part). This is also true for spots fixed in six origins and variable only in one origin (Table 3, lower part). Altogether, 43 spot-characters indicate the originality of the Moroccan population, for an average of 16 in the seven populations.

Genetic relationships

The mean dissimilarity matrix (Table 4) showed that the intra- (diagonal elements) and inter-origin average variabilities (off-diagonal elements) were very close. It seems that this is a general rule in forest tree species with large distribution ranges, high fecundities, and wind pollination (Hamrick and Godt 1990).

The first plane of the CA accounted for 50% of the variation (Fig. 3). The individuals were well grouped according to their origin for Italian, Sardinian, Moroccan, and Corsican populations. Individuals from Portugal, Spain, and Landes were mixed together. Italy and Sardinia constituted a group distant from the Landes, Spain, and Portugal group on the first axis, the Corsica origin being intermediate between the two groups. The Moroccan origin was separated from all of the other populations on the second axis.

The dendrogram computed on the dissimilarity matrix confirmed the relationships between origins (Fig. 4). Moreover it showed that individuals and populations were not very close to each other, since all clusterings occurred for the values of D ranging from 0.15 to 0.30.

Table 2 Classification of spots per origins (*a* Number of spots fixed in all origins, *b* Number of spots fixed in the origin and variable in at last one other origin, *c* number of spots variable in the origin, $a+b+c$ total number of spots present in the origin, *d* number of spots always present in the origin and never fixed in the other origins, *d'* number of spots always absent in the origin and never fixed in the other origins, $d+d'$ number of spot-characters (absent or presence) fixed only in the origin, (see Table 1 for origin code))

	P	M	C	L	E	I	S	Mean
a	154	154	154	154	154	154	154	154
b	129	**126**	147	170	131	171	**235**	158
c	566	**589**	569	522	579	520	**459**	543
a+b+c	849	869	870	846	864	845	848	856
d	17	18	14	19	14	20	46	21
d'	17	25	8	13	6	24	26	17
d+d'	34	**43**	22	32	20	44	**72**	38

Table 3 List of spot-characters characterising a single origin. Upper-part: origin-characteristic and origin-specific spots, lower part: *idem* but when the character considered is the absence of the spot (The number of individuals are indicated between ( ); total is indicated between [ ]; (see Table 1 for origin code))

P	M	C	L	E	I	S
1329(1)	1094(1)	945(1)	1345(3)	958(1)	1150(1)	2017(2)
1084(1)	397(1)	859(2)		1351(1)	2005(1)	2011(3)
442(1)	1290(2)	1339(2)			2018(2)	2007(4)
	970(3)				2014(3)	2006(4)
	401(4)				2003(4)	
	1091(6)					
3	17	5	3	2	11	13
179(1)	165(1)	1(1)	188(1)	85(1)	90(1)	
836(1)	186(1)	3(1)		192(1)	212(1)	
405(2)	234(1)	259(1)		373(1)	341(1)	
107(2)	523(1)	177(1)		520(1)	395(1)	
176(3)	385(2)	272(1)		619(1)	477(1)	
	914(2)	231(2)			799(1)	
	392(3)				878(1)	
	809(4)					
	902(5)					
	623(6)					
9	26	7	1	5	7	0
[12]	[43]	[12]	[4]	[7]	[18]	[13]

Table 4 Mean distances between and within geographical origins (see Table 1 for origin code)

	P	M	C	L	E	I	S
P	0.270						
M	0.321	0.284					
C	0.305	0.326	0.275				
L	0.282	0.315	0.291	0.252			
E	0.298	0.326	0.307	0.275	0.279		
I	0.322	0.340	0.297	0.303	0.314	0.251	
S	0.321	0.341	0.298	0.311	0.312	0.265	0.216

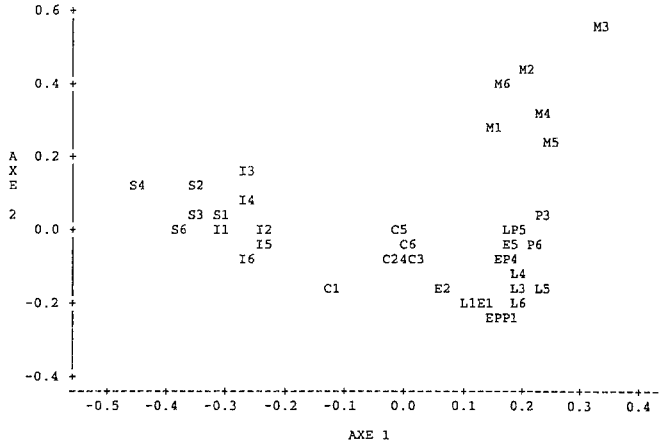

Fig. 3 Distribution of the individuals in the first plane of the correspondence analysis

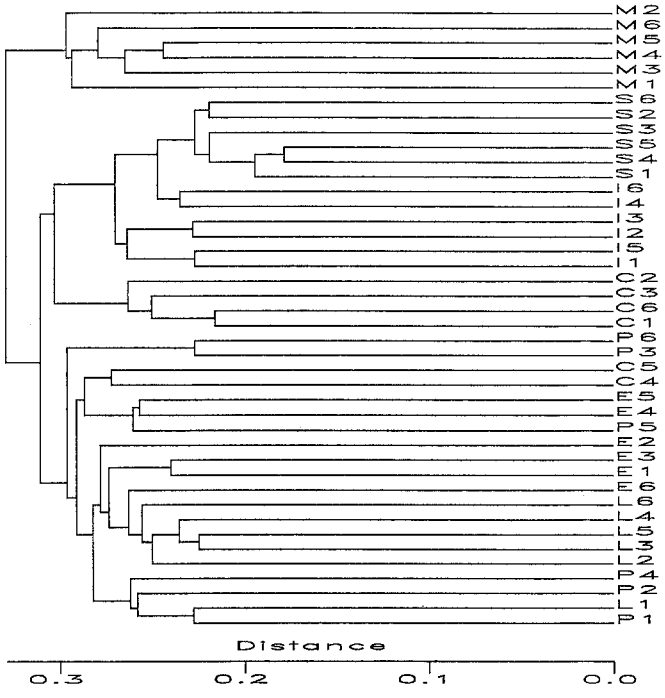

Fig. 4 Dendrogram constructed from the dissimilarity matrix

Three groups of maritime pines have been recognised from palynological, anthracological, and paleoclimatologic data (Baradat and Marpeau 1988): the Atlantic group, consisting of populations from Portugal, Spain, and France; the Mediterranean group consisting of Italian, Sardinian, and Corsican populations, and the North Africa group. These groups have also been identified using terpene polymorphism (Baradat and Marpeau 1988). The parcelling of the natural range following Würm III glaciation could have resulted in a strong divergence between groups due to genetic drift. This hypothesis is partially confirmed by our results since the three groups were well differentiated. However, while the three origins of the Mediterranean group had no overlap, the three populations of the Atlantic group could not be fully distinguished. Gene flow probably exists between the Portuguese, Spanish, and Landes populations. The Sardinian origin had the lowest level of variability and could have originated through founder effect from an Italian population and subsequent differentiation by genetic drift. The Corsican population originated from trees coming from Liguria (Italy) and eastern Spain, which conforms with its intermediate position between the Atlantic and Mediterranean groups in our analyses. The Moroccan population from Tamjout was the most distant from all other origins, which is probably a consequence of early separation of the North African provenances during the formation of the Strait of Gibraltar three million years ago.

By means of the 2-D PAGE technique, information about the relationships between populations and levels of genetic variability existing within and between origins can thus be obtained. Numerous protein markers can be scored from a single 2-D gel corresponding to one individual, thereby giving access to the polymorphism of several hundred of gene loci. This can be useful for the management of genetic resources and for breeding purposes.

Acknowledgements We thank Dr. A. Kremer and Dr. R. Petit for helpful comments on the manuscript. This research was supported in part by the contract EEC, "Forest" no. MA2B-CT 91-0040 and by the contract BRG of MRT.

References

Anderson NG, Tollaksen SL, Pascoe FH, Anderson NL (1985) Two-dimensional electrophoretic analysis of wheats seed proteins. Crop Sci 25:667–674

Bahrman N, Damerval C (1989) Linkage relationships of loci controlling protein amounts in maritime pine (*Pinus pinaster* Ait.). Heredity 63:267–274

Bahrman N, Thiellement H (1987) Parental genome expression in synthetic wheats (*Triticum turgidum* sp. × *T. tauschii* sp.) revealed by two-dimensional electrophoresis of seedling proteins. Theor Appl Genet 74:218–223

Bahrman N, Cardin M-L, Seguin M, Zivy M, Thiellement H (1988a) Variability of three cytoplasmically encoded proteins in *Triticum* genus. Heredity 60:87–90

Bahrman N, Zivy M, Thiellement H (1988b) Genetic relationships in the Sitopsis section of *Triticum* and the origin of the B genome of polyploid wheats. Heredity 61:473–480

Baradat P, Marpeau A (1988) Le pin maritime. Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce. These doctorat d'Etat ès Sciences, University of Bordeaux I, Bordeaux

Damerval C, de Vienne D, Zivy M, Thiellement H (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. Electrophorsis 7:52–54

Damerval C, Le Guilloux M, Blaisonneau J, de Vienne D (1987) Simplification of the Heukeshoven and Dernick's silver staining of proteins. Electrophoresis 8:158–159

Ellstrand NC (1992) Gene flow among seed plant populations. New For 6:241–256

Gerber S, Rodolphe F, Bahrman N, Baradat Ph (1993) 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

Goldman D, Giri PR, O'Brien SJ (1987) A molecular phylogeny of the hominoid primates as indicated by two-dimensional protein electrophoresis. Proc Natl Acad Sci USA 84:3307–3311

Goldman D, Giri PR, O'Brien SJ (1989) Molecular genetic-distance estimates among the ursidae as indicated by one- and two-dimensional protein electrophoresis. Evolution 43:282–295

Granier F, de Vienne D (1986) Silver staining of proteins: standardized procedure for two-dimensional gels bound to polyester sheets. Anal Biochem 155:45–50

Hamrick JL, Godt MJW (1990) Allozyme diversity in plant species. In: Brow HD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Assoc, Sunderland, Mass., pp 43–63

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. New For 6:95–124

Leigh Brown AJ, Langley CH (1979) Reevaluation of level of genic heterozygosity in natural population of *Drosophila melanogaster* by two-dimensional electrophoresis. Proc Natl Acad Sci USA 76:2381–2384

McConkey EH, Taylor BJ, Phan D (1979) Human heterozygosity: a new estimate. Proc Natl Acad Sci USA 76:6500–6504

McLellan T, Ames GF-L, Nikaido K (1983) Genetic variation in proteins: comparison of one-dimensional and two-dimensional gel electrophoresis. Genetics 104:381–390

Mirov NT (1967) The genus *Pinus*. Ronald Press, New York

Müller-Starck G, Baradat Ph, Bergmann F (1992) Genetic variation within European tree species. New For 6:23–47

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

SAS Institute (1990) User's guide: version 6, 4th edn. SAS Institute, Cary, N. C.

Spicer GS (1988) Molecular evolution among some *Drosophila* species as indicated by two-dimensional electrophoresis. J Mol Evol 27:250–260

Thiellement H, Seguin M, Bahrman N, Zivy M (1989) Homoeology and phylogeny of A, S, and D genomes of the Triticinae. J Mol Evol 29:89–94

Walton KE, Styer D, Gruenstein EI (1979) Genetic polymorphism in normal human fibroblasts as analyzed by two-dimensional polyacrylamide gel electrophoresis. J Biol Chem 254: 7951–7960

Yeh FC, Khalil MAK, El-Kassaby Y, Trust DC (1986) Allozyme variation in *Picea mariana* from Newfoundland: genetic diversity, population structure, and analysis of differentiation. Can J For Res 16:713–720

Zivy M, Devaux P, Blaisonneau J, Jean R, Thiellement H (1992) Segregation distortion and linkage studies in microspore-derived double haploid lines of *Hordeum vulgare* L. Theor Appl Genet 83:919–924