from Truckee. Both studies were conducted under the excellent technical assistance of PAT WELLS and MELVIN MORTON. MARY MAHALOVICH and ALVIN YANCHUK thoughtfully criticized the manuscript, and FREDRICH LAURIA provided guidance through the taxonomic maze of the *Ponderosae.*

Literature

AVISE, J. C.: Introduction: The scope of conservation genetics. In: HAMRICK, J. L. and AVISE, J. C., (eds.). Conservation genetics, case histories from nature. Chapman & Hall, New York. pp. 1–9 (1996). — BRAYSHAW, T. C.: Washoe and ponderosa pines on Promontory Hill near Merritt, B. C. Canada. Ann. Naturhist. Mus. **99**: 673–680 (1997). — CALLAHAM, R. Z. and LIDDICOET, A. R.: Altitudinal variation at 20 years from the California elevational transect study of ponderosa pine. J. Forestry **59**: 814–820 (1961). — CONKLE, M. T.: Genetic Diversity – seeing the forest through the trees. New Forests **6**: 5–22 (1992). — CONKLE, M. T. and CRITCHFIELD, W. B.: Genetic variation and hybridization of ponderosa pine. In: BAUMGARTNER, D. M. and LOTAN, J. E., (Eds.). Ponderosa pine – the species and its management. Washington State University, Pullman, Wash. pp. 27–43 (1988). — CRITCHFIELD, W. B.: Crossability and relationships of Washoe pine. Madroño **31**: 144–170 (1984). — CRITCHFIELD, W. B. and ALLENBAUGH, G. L.: Washoe pine on the Bald Mountain Range, California. Madroño **18**: 63–64 (1965). — CRITCHFIELD, W. B. and LITTLE, E. L., Jr.: Geographic distribution of the pines of the world. USDA For. Serv., Misc. Publ. 991 (1966). — GRIFFIN, J. R. and CRITCHFIELD, W. B.: The distribution of forest trees in California. USDA For. Serv. Res. Paper PSW-82. Pacific Southwest Forest and Range Experiment Station, Berkeley, Calif. (1976). — HALLER, J. R.: Taxonomy, hybridization, and evolution in *Pinus ponderosa* and *P. jeffreyi.* Ph.D. dissertation, Univ. California, Los Angeles (1957). — HALLER, J. R.: Some recent observations on ponderosa, Jeffrey and Washoe pines in northeastern California. Madroño **16**: 126–132 (1961). — HALLER, J. R.: *Pinus washoensis* in Oregon: taxonomic and evolutionary implications. Amer. J. Bot. **52**: 646 abstr. (1965). — LAURIA, F.: Taxonomy, systematics and phylogeny of *Pinus*, subsection *Ponderosae* LOUDON *(Pinaceae).* Alternative concepts. Linzer Biol. Beit. **23**: 129–202 (1991). — LAURIA, F.: The taxonomic status of *Pinus washoensis* H. MASON & STOCKW. (Pinaceae). Ann. Naturhist. Mus. **99**: 655–671 (1997). — LYNCH, M.: A quantitative genetic perspective on conservation issues. In: HAMRICK, J. L. and AVISE, J. C., (eds.). Conservation genetics, case histories from nature. Chapman & Hall, New York. pp. 471–500 (1996). — MASON, H. L. and STOCKWELL, W. P.: A new pine from Mount Rose, Nevada. Madroño **8**: 61–63 (1945). — MITTON, J. B., LATTA, R. G. and REHFELDT, G. E.: The pattern of inbreeding in Washoe pine and survival of inbred progeny under optimal environmental conditions. Silvae Genetica **46**: 215–219 (1997). — NIEBLING, C. R. and CONKLE, M. T.: Diversity of Washoe pine and comparisons with allozymes of ponderosa pine races. Can. J. For. Res. **20**: 298–308 (1990). — PATTEN, A. M.: A molecular phylogenetic analysis of *Pinus* (section *Diploxylon)* subsection *Ponderosae:* Interspecific and intraspecific genetic relationships. M. S. thesis. University of Idaho, Moscow, ID (1999). — REHFELDT, G. E.: A model of genetic variation for *Pinus ponderosa* in the Inland Northwest (U.S.A.): applications in gene resource management. Can. J. For. Res. **21**: 1491–1500 (1991). — REHFELDT, G. E.: Early selection in *Pinus ponderosa:* Compromises between growth potential and growth rhythm in developing breeding strategies. For. Sci. **38**: 661–677 (1992). — REHFELDT, G. E.: Genetic variation in the *Ponderosae* of the Southwest. Amer. J. Bot. **80**: 330–343 (1993). — REHFELDT, G. E.: Systematics and genetic structure of ponderosae taxa (Pinaceae) inhabiting the mountain islands of the Southwest. Amer. J. Bot. **86**: 741–752 (1999). — REHFELDT, G. E. and WYKOFF, W. R.: Periodicity of shoot elongation among populations of *Pinus contorta* from the northern Rocky Mountains. Ann. Bot. **48**: 371–377 (1981). — SAS: SAS user's guide, release 6.12 edition. SAS Institute, Cary, N. C. (1996). — SMITH, R. H.: Variations in the monoterpene composition of the wood resin of Jeffrey, Washoe, Coulter and lodgepole pines. For. Sci. **13**: 246–252 (1967). — SMITH, R. H.: Xylem monoterpenes in *Pinus ponderosa, Pinus washoensis,* and *Pinus jeffreyi* in the Warner Mountains of California. Madroño **21**: 26–32 (1971). — SMITH, R. H.: Monoterpenes of ponderosa pine xylem resin in western United States. USDA For. Serv., Technical Bull. 1532, Washington, D. C. (1977). — SORENSEN, F. C.: Genetic variation and seed transfer guidelines for ponderosa pine in central Oregon. U.S. Dept. of Agric., Forest Service, Pacific Northwest Research Stn., Portland, OR. Res. Pap. PNW-RP-472 (1994a). — SORENSEN, F. C.: Genetic variation and seed transfer guidelines for ponderosa pine in the Ochoco and Malheur National Forests of central Oregon. U.S. Dept. of Agric., Forest Service, Pacific Northwest Research Stn., Portland, OR. Res. Pap. PNW-RP-468 (1994b). — WELLS, O. O.: Geographic variation in ponderosa pine. I. The ecotypes and their distribution. Silvae Genetica **13**: 89–103 (1964).

Multilocus Analysis of *Pinus halepensis* MILL. from Spain: Genetic Diversity and Clinal Variation

By D. AGÚNDEZ1), B. DEGEN2)3), G. VON WUEHLISCH2) and R. ALIA1)4)

(Received 17th February 1999)

Abstract

Fifteen populations of *Pinus halepensis* from Spain were analysed employing 5 polymorphic loci by starch gel electrophoresis. A multilocus approach was used, in order to detect changes in the level of polymorphism and to reveal the pattern of variation of the species. The multilocus analyses show that large differences exist when island, inland and marginal populations are considered. These differences are quite high in terms of the effective number of genotypes, and in subpopulation differentiation. The multilocus differentiation is mainly caused by relations of allelic frequencies among populations, the most likely cause being that, by chance, alleles were associated differently in the refugia of the species. The population differentiation is large in comparison with that observed in other conifers (F_{ST} = 0.11; CF_{ST} = 0.17). The canonical variate corresponding to the largest differentiation among populations is positively correlated to the frequency of the most common allele of Mdh4 and Aco and negatively to the frequency of Hk and Pgm2. A clear pattern of variation related to latitude and longitude has been detected by the high correlation with the first canonical variate (r = –0.907 and r = –0.832 respectively). In conclusion, the clinal variation could be attributed to a

---

1) Departamento de Mejora y Biotecnología, CIFOR-INIA, Apdo 8111, E-28080 Madrid, Spain
Phone: +34 91 34 76 857; Fax: +34 91 35 72 293; email: alia@inia.es

2) Institute for Forest Genetics and Forest Tree Breeding, Sieker Landstrasse 2, D-22927 Grosshansdorf, Germany

3) Present address: INRA, Station de Recherche Forestiere, Campus Agronomique, BP 709, 97287 Kourou Cedex, Guyane Francaise.
e-mail: bdegen@gmx.net

4) corresponding author

recolonization of the species in a North-South direction, leading to a loss of variability among the island and marginal populations.

Key words: Pinus halepensis, allozymes, differentiation, multilocus diversity, clinal variation.

Introduction

Aleppo pine *(Pinus halepensis* MILL.) spreads all over the Mediterranean basin, where it plays an important role in forestry and in the conservation of the Mediterranean ecosystems. Several studies have been conducted in order to assess the genetic variability of the species and to evaluate its genetic resources. The studies at a general scale reveal a low level of polymorphism, as shown by electrophoretic analysis (SCHILLER *et al.,* 1986; CONKLE *et al.,* 1988; TEISSEIRE *et al.,* 1995; AGÚNDEZ *et al.,* 1997) and other genetic markers (MORGANTE *et al.,* 1998; GÓMEZ, 1998), especially in the western range of the species. The low level of polymorphism has been interpreted due to a bottleneck effect during glaciations and a rapid recolonization with material of Greek origin (MORGANTE *et al.,* 1998).

The adaptation of Aleppo pine to drought and forest fires (GIL *et al.,* 1996), together with human activities over time, has led to important changes in the importance and distribution of the species in the Mediterranean basin (BARBERO et al., 1998), and in the pattern of variation of Aleppo pine. In order to examine these matters more fully it is possible to focus attention on the Iberian Peninsula, an area where no introduction of material of foreign origin has been made in the past (BARADAT *et al.,* 1995; GIL *et al.,* 1996), and with an important area of the species (806,000 ha of native stands). A clinal variation seems to be detected in studies of several Spanish populations, using different types of markers (AGÚNDEZ *et al.,* 1997; GÓMEZ, 1998).

A multilocus approach with a large number of populations will allow us to increase our understanding of the geographical pattern of variation of the species, and of how genetic variability varies among populations. The main causes of multilocus differentiation (KREMER and ZANETTO, 1997) are derived from the single locus differentiation of each locus and the allelic associations among and within populations.

In forestry species, linkage disequilibrium is usually of scant importance, but on some occasions could reveal the influence of genetic drift or systematic effects in the distribution of the species (KREMER *et al.,* 1997). Of more importance are the allelic associations at population level.

The objective of this study was to analyse the pattern of variation using a multilocus approach with 5 polymorphic loci of 15 natural populations of Aleppo pine from the Western range of distribution.

Material and Methods

Seed samples were obtained from fifteen natural populations of *Pinus halepensis,* covering the natural range of the species in Spain *(Figure 1).* The geographical origins of the sampled populations are presented in *table 1.* The bulked seed samples were composed of seeds from 25 trees with a minimum distance of 30 m in between.

Fifty to sixty germinated seeds per population were analysed employing five polymorphic enzyme systems both in megagametophytes and the corresponding embryo from each seed. Thus, ordered genotypes were obtained by comparing them. The enzyme systems are the following: Phosphoglucomutase (PGM, EC 2.7.5.1), Hexokinase (HK, EC 2.7.1.1), Aconitase (ACO, EC 4.2.1.3), 6 Phosphogluconate dehydrogenase (6PGDH, EC 1.1.1.44) and Malate dehydrogenase (MDH, EC 1.1.1.37). Staining procedure and electrophoretic methods have been described elsewhere (CONKLE *et al.,* 1982; CHELIAK and PITEL, 1985; WENDEL and WEEDEN, 1989; AGÚNDEZ *et al.,* 1997).

Description and segregation analysis of HK, MDH, 6PGDH and PGM, have previously been reported (AGÙNDEZ *et al.,* 1997). ACO shows a single locus with two alleles and no deviation from the expected 1:1 MENDELian ratio.

Measures of variability at a single locus level

There are no significant deviations from the assumption of neutrality for the five loci analysed, as shown by the EWENS-WATTERSON test for neutrality (MANLY, 1985).

The HARDY-WEINBERG equilibrium has been tested using a chi-square test, for each of the five loci for every population and for the overall data. There is an excess of homozygotes for the overall data in three out of the five loci (Hk, Pgm2 and Aco), but this was not detected in most of the populations. Departures from the HW equilibrium exists in 3 populations, but in different loci.

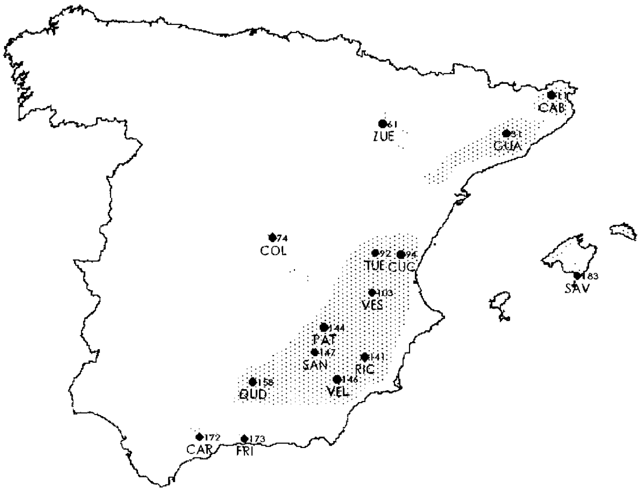

Figure 1. – Location of the 15 populations of *Pinus halepensis* analysed.

Table 1. – Origin of the fifteen sampled *Pinus halepensis* populations from Spain.

Code	Location	Name	Prov. Region	Altitude	Latitude	Longitude
011	Cabanellas	CAB	01	210	42° 14' 08" N	2° 47' 24" E
031	Guardíola	GUA	03	420	41° 40' 04" N	1° 45' 41" E
061	Zuera	ZUE	06	575	41° 55' 00" N	0° 55' 04" W
074	Colmenar	COL	07	750	40° 05' 20" N	3° 20' 10" W
092	Tuejar	TUE	09	600	39° 49' 13" N	1° 09' 31" W
094	Cucalon	CUC	09	575	39° 47' 15" N	0° 36' 44" W
103	Villa de Ves	VES	10	850	39° 10' 44" N	1° 14' 52" W
141	Ricote	RIC	14	780	38° 08' 33" N	1° 25' 50" W
144	Paterna	PAT	14	1000	38° 37' 37" N	2° 16' 37" W
146	Velez	VEL	14	750	37° 47' 21" N	2° 00' 55" W
147	Santiago	SAN	14	680	38° 13' 33" N	2° 28' 00" W
158	Quentar	QUD	15	750	37° 45' 15" N	3° 45' 15" W
172	Carratraca	CAR	17	650	36° 50' 28" N	4° 50' 04" W
173	Frigiliana	FRI	17	570	36° 49' 00" N	3° 55' 18" W
183	S'Avall	SAV	18	10	39° 17' 14" N	3° 02' 52" E

For assessing departures from random mating considering the paternal and maternal allele frequency, a chi-square test of homogeneity was performed among maternal and paternal contributions of the allozyme frequencies (GILLET, 1994). This test did not reveal significant departures from the null hypothesis.

Genotypic disequilibria were estimated for all pairwise combinations of loci, taking into account gametic frequencies in all 15 populations (WEIR, 1996). The gametic disequilibria over the 15 populations were subdivided into within- and between-population components according to the method given by OHTA (1982a and b). A series of five different components were calculated and compared.

Measures of genotypic variability at multilocus level

Four different parameters have been derived from the multilocus approach: the genotypic diversity ν (GREGORIUS, 1978), similar to the effective number of genotypes present in each population; the total population differentiation δ_T (GREGORIUS, 1987), related to the expected heterozygosity; the evenness of the distribution (e), which is considered to measure the degree to which the genotypes are equally represented (the maximum ($e=1$) holds only for uniform distribution, and as e approaches the lower bound, the unevenness increases); and the subpopulation differentiation (D_j), which is the proportion of genetic elements (genotypes in this case) by which a deme differs from the remainder of the populations. The overall subpopulation differentiation is a weighted mean of the subpopulation differentiation.

Differentiation among populations

The WRIGHT (1965) decomposition of variability has been employed, in order to estimate the coefficient of differentiation among populations. Gene flow (Nm) was estimated as $Nm = (1\text{-}G_{ST})/4G_{ST}$ (NEI, 1977).

Analysis of gametic differentiation in multiple loci has been described by KREMER *et al.* (1997), using a multivariate ANOVA. If HWE is assumed, F^*_{ST} can be derived from the model developed at the gametic level (1). However, when HW equilibrium is absent, only composite measures can be calculated at the zygotic level and these will be always larger than the F_{ST} values (KREMER and ZANETTO, 1997). The differentiation, therefore, at zygotic level is defined as CF_{ST}, as a composite measure of differentiation.

To reveal the pattern of variation, NEI's (1978) genetic distances among populations were computed considering every locus. Cluster analyses were performed using the unweighted pair group method algorithm (UPGMA). In relation to geographic data, a canonical analysis of populations has been performed to test differences among populations. For this purpose canonical variates z corresponding to the largest differentiation among populations where computed for each population.

$$z = x'u$$

where x' is the transpose of the vector of the original data and u is the two first eigenvectors with the largest differentiation (CF_{STm}) of matrix $CF_{ST.}$

An average of correlations between canonical variates and altitude, longitude and latitude have been computed.

The analyses were performed using version 1.1b of GSED programme (GILLET, 1994), POPGENE (YEH *et al.*, 1997), SAS (SAS version 6.0, 1996) and the authors own programmes.

Results

Frequencies of the most common allele of the five polymorphic loci are shown in *table 2*. There are significant differences between populations in all these systems, as shown by the chi-square analysis of populations.

Table 2. – Frequencies of the most common allele by population.

Population	Location	N	Mdh4	6Pgdh2	Hk	Pgm2	Aco
11	Cabanellas	112	0.9375	0.9107	0.8929	0.6875	0.6875
31	Guardiola	104	0.8750	0.9038	0.8846	0.7212	0.8942
61	Zuera	104	0.9038	0.8654	0.9135	0.8077	0.8558
74	Colmenar	94	0.9468	0.9468	0.4149	0.9362	0.8085
92	Tuejar	110	0.8364	0.9364	0.6818	0.7455	0.8364
94	Cucalón	106	0.8774	0.9340	0.6698	0.9340	0.7925
103	Villa de Ves	118	0.8390	0.8983	0.6271	0.8644	0.6017
141	Ricote	98	0.8571	0.9184	0.5918	0.9388	0.7143
144	Paterna	92	0.8370	0.9565	0.7174	0.9891	0.6087
146	Velez	102	0.6961	0.8529	0.6667	0.9804	0.5980
147	Santiago	116	0.9138	0.9914	0.5603	0.9483	0.6810
158	Quesada	112	0.6875	0.9554	0.6429	1.0000	0.6429
172	Carratraca	104	0.6731	0.8077	0.1731	1.0000	0.3365
173	Frigiliana	110	0.4273	0.8818	0.5545	0.9909	0.6000
183	S'Avall	104	1.0000	0.9135	0.7115	0.9231	0.8269
Overall mean	Allele A	1586	0.9062	0.8938	0.8969	0.7375	0.8094

There are no zygotic disequilibria within populations, following the method of WEIR (1979, 1980). None of the populations showed a linkage disequilibrium for the pairwise comparison of alleles. The decomposition of disequilibrium in subdivided populations *(Table 3)*, following the method given by OHTA (1982a and b), shows that the variance of the expected allelic associations among populations (D'^2_{IS}) is larger than the variances of the disequilibrium within the populations (D^2_{IS}) or within the overall populations (D'^2_{IS}); the interpretation based on analytical calculations shows that disequilibria are created mainly by random effects (D^2_{IS} is less than D^2_{ST} and D'^2_{IS} is greater than D'^2_{ST}). These results are not caused by the nonexistence of populations showing significant disequilibria for these combinations, as was shown by KREMER and ZANETTO (1997) in *Quercus petraea*.

Table 3. – Ohta's two-locus analysis of population subdivision.

LocusA / LocusB	$(DIT)^2$	$(DIS)^2$	$(D'IS)^2$	$(DST)^2$	$(D'ST)^2$
Mdh4 / 6Pgdh2	0.03765	0.00019	0.03762	0.03719	0.00003
Mdh4 / Hk	0.06933	0.00042	0.06875	0.07041	0.00058
Mdh4 / Pgm2	0.04498	0.00014	0.04488	0.04504	0.00010
Mdh4 / Aco	0.05373	0.00029	0.05314	0.05321	0.00059
6Pgdh2 / Hk	0.05299	0.00026	0.05294	0.05353	0.00005
6Pgdh2 / Pgm2	0.02098	0.00038	0.02093	0.02146	0.00005
6Pgdh2 / Aco	0.03223	0.00033	0.03205	0.03353	0.00018
Hk / Pgm2	0.06380	0.00021	0.06334	0.06274	0.00046
Hk / Aco	0.06101	0.00064	0.06007	0.06298	0.00093
Pgm2 / Aco	0.04256	0.00058	0.04190	0.04136	0.00067
Overall Average	0.04793	0.00034	0.04756	0.04815	0.00037

With regard to the genotypes, the descriptive parameters of differentiation among populations are shown in *table 4*. There are large differences between populations in terms of diversity

(Δ^2) and total diversity (δ_T) at the genotypic level. The differences are quite large in comparison to that observed when individual loci are considered. In this sense, a reduction in diversity is found in island (no. 183) or marginal (no. 61, 84) populations. (D_j) shows clearly the differences among populations. Populations in Southern Spain (no. 172 and 173) present the maximum differentiation both in terms of genetic diversity (Δ) and differentiation among populations. One of the populations (103) presents a high level of genetic diversity, but a small amount of differentiation between populations, which could be attributed to its central position in the range of the species in the studied area.

The differences between populations are highly significant, as shown by the MANOVA MANTEL test (MANLY, 1997), and the parameters of differentiation among populations are shown in *table 5*. The average inbreeding coefficient (F_{IT}) ranges from 3.1% to 31.8% depending on the loci – a large amount of variation between populations. The average inbreeding within populations (F_{IS}) shows large differences between loci, as a measure of the deviations from the HW equilibrium is important in Pgm2 and to a lesser extent in the other loci. Values of genetic differentiation between populations (measured in terms of F_{ST}) are of the same value for the different loci, with a mean value of around 11%. The multilocus approach takes into account both the HW disequilibrium in some of the loci and the non-significant linkage differences. This is the reason for the increment in the value of CF_{ST} (mean differentiation) with respect to F_{ST}, which reduces the importance of the within-population component in the total differentiation.

Multilocus differentiation may be caused by allelic associations either within populations or between different populations. The component within a population is given by the zygotic disequilibria, or by correlation at individual level (r_w). For all two-locus associations the data *(Table 6)* clearly show that associations among alleles at different loci are stronger at the between-population (r_b) level than at the within-population level. The discrepancy between r_w and r_b has a strong effect on the level of multilocus differentiation.

Table 5. – Summary of F-Statistics and Gene flow for all Loci (NEI, 1987).

LOCUS	Sample Size	F_{IS}	F_{IT}	F_{ST}	Nm*	CF_{ST}
Mdh4	1586	-0.0755	0.0683	0.1337	1.6192	0.2252
6Pgdh2	1586	0.0061	0.0309	0.0250	9.7646	0.0338
Hk	1586	0.0579	0.1946	0.1451	1.4731	0.2394
Pgm2	1586	0.2280	0.3179	0.1164	1.8974	0.1733
Aco	1586	0.0926	0.1764	0.0924	2.4568	0.1477
Average	**1586**	**0.0596**	**0.1627**	**0.1120**	**1.9812**	**0.1689**

*) Nm = Gene flow estimated from F_{ST} = 0.25 (1 - F_{ST})/F_{ST}.

Table 6. – Correlation coefficients of measure of variability between *(r_b, below diagonal)* and within populations (r_w, above diagonal) based on allelic frequencies.

	Mdh4	6Pgdh2	Hk	Pgm2	Aco
Mdh4	–	-0.005	-0.036	0.014	0.000
6Pgdh2	0.404	–	-0.035	0.064	0.058
Hk	-0.383	-0.207	–	-0.028	0.012
Pgm2	-0.473	-0.01	0.661	–	-0.078
Aco	0.587	0.432	-0.668	-0.558	–

Canonical variates were calculated when the MANOVA was computed on all 5 loci. The correlation between these two variates and geographical variates is shown in *table 7*. The canonical analysis of populations shows a quite clear separation between populations. The first two canonical variates

Table 4. – Descriptive parameters for 5 polymorphic loci, of 15 natural populations of Aleppo pine from Spain.

			Unilocus			Multilocus			
Pop	Location	No individuals	Observed	Expected	Fis	Diversity (Δ)	δ_T	absolute *e*	D_j
11	Cabanellas	56	0.2179 *(0.0851)*	0.2661 *(0.1517)*	0.1811	14.933	0.950	0.732	0.561
31	Guardiola	52	0.1962 *(0.0459)*	0.2376 *(0.0935)*	0.1742	10.242	0.920	0.678	0.547
61	Zuera	52	0.2154 *(0.0498)*	0.2245 *(0.0612)*	0.0405	9.013	0.906	0.665	0.567
74	Colmenar	47	0.1745 *(0.1057)*	0.2232 *(0.1711)*	0.2181	8.801	0.906	0.660	0.524
92	Tuejar	55	0.1891 *(0.0633)*	0.2960 *(0.1207)*	0.3611	17.690	0.961	0.714	0.469
94	Cucalón	53	0.2340 *(0.1097)*	0.2466 *(0.1383)*	0.0510	13.188	0.942	0.744	0.372
103	Villa de Ves	59	0.3322 *(0.1401)*	0.3269 *(0.1375)*	-0.0162	20.598	0.968	0.724	0.391
141	Ricote	49	0.2612 *(0.1372)*	0.2802 *(0.1605)*	0.0678	13.720	0.946	0.667	0.400
144	Paterna	46	0.2522 *(0.2077)*	0.2519 *(0.1975)*	-0.0011	10.475	0.925	0.766	0.470
146	Velez	51	0.3412 *(0.2087)*	0.3275 *(0.1842)*	-0.0418	17.938	0.963	0.710	0.504
147	Santiago	58	0.2448 *(0.2102)*	0.2400 *(0.2111)*	-0.0200	9.667	0.912	0.667	0.409
158	Quesada	56	0.3571 *(0.2899)*	0.2867 *(0.2251)*	-0.2455	10.116	0.918	0.711	0.480
172	Carratraca	52	0.2577 *(0.1792)*	0.2967 *(0.1812)*	0.1314	18.521	0.965	0.715	0.768
173	Frigiliana	55	0.3418 *(0.2161)*	0.3380 *(0.2160)*	-0.0112	21.154	0.970	0.772	0.672
183	S'Avall	52	0.1962 *(0.1668)*	0.1994 *(0.1556)*	0.0160	5.564	0.836	0.603	0.573
	Overall mean	793	0.2525 *(0.1099)*	0.3043 *(0.1335)*	0.1702	13.441	0.933	0.702	0.513

explain 62.3% and 20.5% of the total variation respectively. A large geographical trend associated with longitude and latitude is observed in the highly significant correlation with the first canonical variate, which is not related to the other variates involved in genetic variability. High values are found in northern Spain. The first canonical variate is correlated (at the between-population level) with the frequency of the most common allele of Mdh4 (r = 0.6100), Hk (r = –0.6477), Pgm2 (r = –0.4882) and Aco (r = 0.4319).

Table 7. – Correlation analysis between geographical data and allele frequency, two first canonical variates of allele frequency, genetic diversity *(Δ)*, total diversity *(δ_T)* evenness *(e)* and expected heterozygosity.

	ALTI	X-UTM	Y-UTM
Mdh4	-0.305 ns	0.653 **	0.688 **
6Pgdh2	0.172 ns	0.074 ns	0.114 ns
Hk	0.346 ns	-0.729 **	-0.723 **
Pgm2	0.478 ns	-0.696 **	-0.862 ***
Aco	-0.392 ns	0.611 *	0.715 **
CAN1	0.447 ns	-0.832 ***	-0.907 ***
CAN2	0.130 ns	-0.251 ns	-0.328 ns
Δ	0.275 ns	-0.354 ns	-0.386 ns
δT	0.473 ns	-0.436 ns	-0.286 ns
e	0.431 ns	-0.380 ns	-0.254 ns
He	0.439 ns	-0.489 ns	-0.658 **

The analysis of variation shows a clear relationship between frequencies and geographical data *(table 7)* that can be expressed in *figure 2.* The correlation between latitude (measured by Y-UTM) and longitude (X-UTM) is quite high (0.72***) and therefore, both of the variates could be used. A clear clinal pattern from North to South in which the frequency of Hk and Pgm2 increases, and the frequency of Mdh4 and Aco decreases, is observed. The variation of 6Pgdh2 does not show a clear relationship with any of the geographical variates.

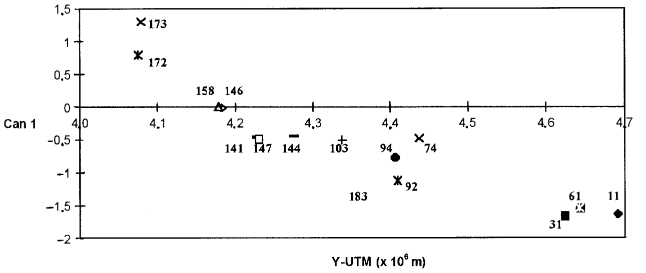

Figure 2. – Relationship between latitude (measured by Y-UTM in m) and value of the first canonical variate obtained from 5 isozyme loci in *Pinus halepensis* MILL..

Discussion and Conclusion

Levels of multilocus diversity

The study of multilocus measures of diversity reveals large differences among populations, that could be interpreted as differences in diversity at the population level. These differences in diversity (v), relative evenness and differentiation among populations (D_j) express more clearly than in a unilocus approach the large differences that exist in a species with a low level of polymorphism.

The non-existence of linkage disequilibrium among loci is a general rule in the conifers, but scant differences in HW equilibrium (unrelated to a different contribution among paternal and maternal components) could affect the differences among populations.

The amount of zygotic disequilibria in *Pinus halepensis* is of the same level as that found in some species as *Pseudotsuga menziesii* (YEH and MORGAN, 1987), *Eucalyptus obliqua* (BROWN *et al.,* 1975), and lower than that found in *Q. petraea* (KREMER and ZANETTO, 1997). In our case, the study of seedlings has not conducted to an extensive level of disequilibria at this age (as was found in *Liriodendron tulipifera,* ROBERDS and BROTSCHOL, 1985); this has been interpreted as related to the decrease in homozygosity with age which has been reported in forest trees.

OHTA's two-locus analyses of population subdivision reveal that mainly random effects are responsible for the disequilibria. This is an important result, because no random selection is involved in the disequilibria.

Geographical pattern of variation among populations at the multilocus level

In this study, more information has been provided about patterns of genetic differentiation among different types (inland, marginal, island) of populations than was previously obtained by means of multilocus genotypic differences. In contrast with previous reports of lack of clinal variation in Aleppo pine in particular, this study reveals a clear pattern of variation related to latitude and longitude. These data complete those which are currently available for this species.

In general, there were stronger allelic associations between loci at population level than at individual level. The data indicates that correlations of allelic frequencies among populations are not caused by zygotic disequilibria. The most likely cause is that, by chance, alleles were associated differently in the refugia, but without any disequilibrium; thus population correlation may just originate from common historical causes. Two different areas can be determined (both in the Northern and Southern range of distribution), and do not probably originate from postglacial recolonization.

Studies of *Pinus contorta,* (YANG and YEH, 1993) and *C. lanceolata* (YEH *et al.,* 1994) have found gametic disequilibria in most of the populations analysed. In the last study the gametic disequilibria were not related with geographical data, but correlated with longitude and latitude in the case of two of the subspecies of *P. contorta.* In *Pinus rigida* (FRYER, 1987), the results using a multivariate analysis of isozyme data provide the same information on patterns of geographic variation as that obtained from morphological measurements in early tests.

When the information from all five loci was used in a MANOVA, 16.9% of the total variability could be attributed to population subdivision. This is a composite measure of differentiation that takes into account allelic associations at the zygotic (and gametic) level, and allelic associations among populations as shown in KREMER *et al.* (1996).

The differences among populations in the area analysed represent a high level in comparison to other conifers (F_{ST} = 0.11, or CF_{ST} = 0.17). This large degree of variation – in the case of a species with a relatively low total variation, as revealed by isozymes – could be attributed to the existence of different refugia during the glaciations in Spain, and a reduction of the area leading to a local bottleneck effect of the species. In this sense, the clear pattern of variation revealed by the dendrogram based on NEI's genetic distance, and the correlation among canonical variates and geographical coordinates have to be interpreted as due to the pathways of postglacial recolonization. The reduction on genetic variability measured at the multilocus level of the inland and island populations (no 74 and 183 respectively) indicates that the main line of recolonization followed the Mediterranean coastline.

The low number of loci analysed are, however, sufficient to allow us to study geographical gradients, while in order to study the pattern of variation, a small number of canonical variates reveal the most important migration paths of the species. In other studies, most of the significant variation is revealed by only a few number of clinal patterns. In *Pinus contorta,* Yeh *et al.* (1985) found that two significant canonical discriminant functions accounted for 38% of the total variance in 20 polymorphic loci, which were associated with geography (altitude and latitude).

A clinal pattern of variation has been revealed in different species in the Iberian Peninsula, such as Scots pine (Prus-Glowacki and Stephan, 1994) and Maritime pine (Salvador *et al.,* 1997).

The low level of polymorphism detected in France (Teisseire *et al.,* 1997) reveals that in this area recolonization took place with material of extraneous origin, and not from French material. In Spain, the likely refugia were located in Northern and Southern Spain (Catalonia and Andalucia). These refugia are confirmed with data on RAPDs and Cp microsatellites (Gomez, 1998), where some special variants have been detected. The clear relationship and small genetic distance between the Mallorca island population and those from Valencia, clearly indicate a recent colonization of the island that has not allowed a large level of differentiation.

The existence of marginal populations in the inner part of the range in Spain, explains how the recolonization took place. In these areas, a reduction in variability could be the main reason to regard the coldest habitat of the species in Spain, in which it is not possible to distinguish a continuous range of the species, as an area of recolonization.

Unlike univariate measures of population differentiation such as genetic distance and F statistics which weight all loci equally, but do not account for multilocus genotypic structure, the analysis performed takes into account the relationship between all the loci, and the genotypic variability. The study strengthens the thesis that clinal patterns of variation with latitude and longitude, possibly the expression of underlying genetic processes such as past migration patterns, have had a significant role in the generation and maintenance of the genetic structure of natural populations of Aleppo pine.

Acknowledgements

The study was funded by the European Commission, Brussels, project FAIR CT95-0097 "Mediterranean Pinus and Cedrus for sustainable afforestation of marginal lands".

References

Agúndez, D., Degen, B., von Wuehlisch, G. and Alia, R.: Genetic variation of Aleppo pine *(Pinus halepensis* Mill.) in Spain. Forest Genetics **4**(4): 201–209 (1997). — Baradat, Ph., Michelozzi, M., Tognetti, R. and Khouja, M. L.: Geographical variation in the terpene composition of *Pinus halepensis* Mill. In: Baradat, Ph., Adams, W. T. and Müller-Starck, G. (eds.). Populations genetics and genetic conservation of forest trees. pp. 141–158 (1995). — Barbero, M., Lisel, R., Quezel, P., Richardson, D. M. and Romaine, F.: Pines of the mediterranean basin. In: D. M. Richardson (Ed). Ecology and biogeography of Pinus. Cambridge University Press. Chapter 8: 153–170 (1998). — Cheliak, W. M. and Pitel, J. A.: Techniques for starch gel electrophoresis of enzymes from forest tree species. Agriculture Canada. Canadian Forestry Service. Information Report Pl-X-42. Petawawa National Forestry Institute. 49 pp. (1985). — Conkle, M. T., Hodgkiss, P. D., Nunnally, L. B. and Hunter, S. C.: Starch gel electrophoresis of conifer seeds: a laboratory manual. United States Department of Agriculture. Forest Service. Pacific Southwest Forest and Range Experiment Station. General Technical Report PSW–64 (1982). — Conkle, M. T., Schiller, G. and Grunwald, C.: Electrophoretic analysis of diversity and phylogeny of *Pinus brutia* and Closely related taxa. Systematic Botany **13**(3): 411–424 (1988). — Fryer, J. H.: Agreement between patterns of morphological variability and isozyme band phenotypes in pitch pine. Silvae Genetica **36**(5/6): 199–206 (1987). — Gil, L., Diaz-Fernandez, P., Jimenez, M. P., Roldan, M., Alia, R., Agúndez, D., De Miguel, J., Martin, S. and De Tuero, M.: Regiones de procedencia de *Pinus halepensis* Mill. ICONA, Madrid (1996). — Gillet, E.M.: GSED. Genetic Structures from Electrophoretic Data. Universität Göttingen, Germany. (Unpublished) (1994). — Gomez, A.: Variabilidad genética de *Pinus pinaster* en España mediante el uso de marcadores de ADN: RAPDs y Cpmicrosatélites. PhD Thesis. UPM, Madrid (1998). — Gregorius, H. R.: The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. Mathematical Bioscience **41**: 253–271 (1978). — Gregorius, H. R.: The relationship between the concepts of genetic diversity and differentiation. Theoretical and Applied Genetics **74**: 397–401 (1987). — Kremer, A. and Zanetto, A.: Geographical structure of gene diversity in *Quercus petraea* (Matt.) Liebl. II: multilocus pattern of variation. Heredity **78**: 476–489 (1997). — Kremer, A., Zanetto, A. and Ducousso, A.: Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. Genetics **145**: 1229–1241 (1997). — Manly, B. F. J.: The Statistics of Natural Selection. Chapman and Hall, London. p. 272–282 (1985). — Morgante, M., Felice, N. and Vendramin, G. G.: Analysis of hypervariable chloroplast microsatellites in *Pinus halepensis* Mill. reveals a dramatic genetic bottleneck. In: A. Karp, P. G. Isaac and D. S. Ingram (Eds.). Molecular tools for screening biodiversity. Chapman and Hall, London. 407–412 (1998). — Nei, M.: F-statistics and analysis of gene diversity in subdivided populations. Ann. Human Genet. **41**: 225–233 (1977). — Nei, M.: Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 283–290 (1978). — Ohta, T.: Linkage disequilibrium with the island model. Genetics **101**:139–155 (1982a). — Ohta, T.: Linkage disequilibrium due to random genetic drift in finite subdivided populations. PNAS **79**: 1940–1944 (1982b). — Prus-Glowacki, W. and Stephan, B. R.: Genetic variation of *Pinus sylvestris* from Spain in relation to other European populations. Silv. Genet. **43**(1): 7–14 (1994). — Salvador, L.: Estudio de la variabilidad genética de *Pinus pinaster* en España usando marcadores proteicos e isoenzimáticos. U.P.M. ETSI Montes, Madrid. PhD Thesis (1997). — SAS: SAS-STAT. User's manual. Version 6.0. SAS Inc. (1996). — Schiller, G., Conkle, M. T. and Grunwald, C.: Local differentiation among Mediterranean populations of Aleppo pine in their isoenzymes. Silvae Genetica **35**(1): 11–19 (1986). — Schiller, G. and Grunwald, C.: Resin monoterpenes in range-wide provenance trials of *Pinus halepensis* Mill. in Israel. Silvae Genetica **36**(3–4): 109–114 (1987). — Smouse, P. E., Spielman, R. S. and Park, M. H.: Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. Amer. Natur. **119**(4): 445–463 (1982). — Teisseire, H., Fady, B. and Pichot, Ch.: Allozyme variation in five French populations of Aleppo pine (*Pinus halepensis* Miller). Forest Genetics **2**(4): 225–236 (1995). — Weir, B. S.: Genetic Data analysis II. Sinauer Associates, Inc. Publishers, Massachusetts. 445 p. (1996). — Wendel, J. F. and Weeden, N. F.: Visualization and interpretation of plant isozymes. In: Isozymes in Plant Biology. Chapter 1. Ed. Dioscorides Press, Oregon. 40 pp. (1989). — Wright, S.: The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19**: 395–420 (1965). — Yang, R. C. and Yeh, F. C.: Multilocus structure in *Pinus contorta* Dougl. Theoretical and Applied Genetics **87**(5): 568–576 (1993). — Yeh, F. C., Cheliak, W. M., Dancik, B. P., Illingworth, D. C., Trust, D. C. and Pryhitka, B. A.: Population differentiation in lodgepole pine, *Pinus contorta* spp. *latifolia:* a discriminant analysis of allozyme variation. Can.J.Gent.Cytol. **27**: 210–218 (1985). — Yeh, F. C., Shi, J., Yang, R., Hong, J. and Ye, Z.: Theoretical and Applied Genetics **88**(3/4): 465–471 (1994). — Yeh, F. C., Yang, R. C. and Boyle, T.: Popgene version 1.21. University of Alberta (1997).