Heredity **75** (1995) 35–44 *Received 6 October 1994*

Genetic variability of *Fagus sylvatica* L. in Italy: the role of postglacial recolonization

STEFANO LEONARDI* & PAOLO MENOZZI
Istituto di Ecologia, Università di Parma, Viale delle Scienze, 43100 Parma, Italy

Genetic variability of 21 Italian populations of beech (*Fagus sylvatica* L.) was studied using starch gel electrophoresis and nine polymorphic enzyme gene loci. Expected mean heterozygosity varied from 13.6 per cent to 20.3 per cent. Observed heterozygosity was less than expected in all but two populations. No association between allele frequencies and soil type or altitude was found. As in other forest tree species, the among-populations component of variability was low (average $F_{ST}=0.046$). Despite low genetic differentiation, principal components analysis of allelic frequencies revealed a geographical pattern. The first principal component, significantly correlated with latitude and longitude, showed a clear separation of southern and northern populations. The statistical significance of the geographical pattern was tested by a resampling technique (bootstrap). The origin of Italian beech populations from eastern and southern refugia during the last glaciation is discussed. First principal component values and the higher allele variability found in southern populations seem to concord with the palynological evidence for a southern origin of beech in the peninsular part of Italy.

Keywords: *Fagus sylvatica*, genetic variability, geographical variation, postglacial colonization.

Introduction

In Italy, European beech (*Fagus sylvatica* L.) is a dominant forest species between 800 and 1500 m a.s.l. in parts of the Alps and between 1000 m and the timberline in the Apennines.

Relatively few studies exist on the population genetics of beech, and the knowledge of Italian populations is particularly poor. Comps *et al.* (1990) include 21 samples from Italy in their extensive analysis of the variability of six allozymes in central European beech populations, but they do not report the geographical distribution of variability within Italy.

To our knowledge the first population study on the genetic variability of peroxidases and glutamate-oxaloacetate-transaminases in a group of French populations was published in 1982 (Thiébaut *et al.*, 1982). The relationship between peroxidase allozymes and temperature and moisture found by this study was confirmed in an investigation of a larger number of populations (Felber & Thiébaut, 1984).

Studying damaged and apparently healthy beech trees in Germany, a larger heterozygosity was found in stress-tolerant plants, although the genotypic response varied at different gene loci (Müller-Starck, 1985, 1989).

Gregorius *et al.* (1986) studied within-population differentiation in space and time in four demes over 2 years for five enzyme loci. Genetic diversity showed no variation pattern in space or time although differences among the considered partitions were found for single loci.

A heterogeneity of the paternal contribution to the within-population variability (with an intra- and an interdemic component) was found by Merzeau *et al.* (1989) using data from four enzyme loci (*GOT, SOD, MDH1, IDH*).

Comps *et al.* (1990) reported data for 140 beech stands from central and Mediterranean Europe. Based on a relatively low number of enzyme systems (PER-1, PER-2, GOT-1, IDH-1, MDH-1, GPI-1) the investigation found most of the variability (90–95 per cent) in the 'within-population' component, a fact already known for other long-living wind-pollinated forest tree species (Hamrick *et al.*, 1979; Loveless & Hamrick, 1984; Hamrick & Godt, 1990). The study, besides confirming the association already found at a smaller geographical scale between peroxidase allele frequencies and climate, revealed a higher inter- and intra-population diversity for the southern locations. An association between climate and allele frequencies of *PER-1, PER-2* and *GOT* was also found in a subset of populations from Croatia (Comps *et al.*, 1991).

Gömöry *et al.* (1992) stress the need for caution in interpreting clinal variations as a consequence of selec-

*Correspondence.

tion, and see the genetic variability found in their data from French beech stands as a consequence of post-glacial recolonization events.

Geographical patterns of genetic variation have often been seen as a result of past mass migratory events in forest tree populations. Postglacial reinvasion routes have been related to observed genetic variation in *Picea abies* populations from central and northern Europe (Lagercrantz & Ryman, 1990). The genetic structure of lodgepole pine (*Pinus contorta* ssp. *latifolia*) found in North America has been attributed mainly to drift during the colonization following the retreat of glaciers (Cwynar & MacDonald, 1987).

Comps *et al.* (1990) suggested a relationship between the lower variability of northern populations and the short time available for their differentiation in areas covered by ice during the last glaciation. The authors stressed the isolation of Italian beech forests, separated from other populations by the sea in the peninsular part and by the Alps in the north. The Eastern Alps, however, were probably a weaker barrier at the end of the last glaciation: with lower sea water levels, the Dinaric Alps populations were not as separated as they seem now (Comps *et al.*, 1990).

Palynological data from Italy, although scarce and not thoroughly analysed (Huntley & Birks, 1983), seem in agreement with the hypothesis that beech populations might have recolonized Italy after the last glaciation from two possible refugia, one in the south (Calabria and Sicily) and one in the east (Croatia).

Here we report an analysis of beech genetic variability in Italy and discuss the possible origin of modern Italian beech populations from southern and eastern refugia.

Materials and methods

Sampling and electrophoretic analysis

Dormant buds were collected from 21 localities to cover most of the geographical range of beech in Italy (Table 1, Fig. 3). All 21 localities are well within the species's biogeographical range in Italy. Sampled individuals, randomly selected at least 2 m apart, were drawn from large populations. In 13 localities (TAN, CAN, GIU, CER, ENT, PNA, PIS, ABE, VUL, POL, SER, REC, SIC), where distances between trees were actually measured, an average distance of 8.6 ± 5.4 m (SD) was found. In four localities (PRA, PNA, PIS, ABE) individuals were collected from transects at different altitudes indicated by multiple values in Table 1.

Table 1 Altitude, mean sample size, mean number of alleles per locus (standard errors are reported in parentheses) and percentage of polymorphic loci (99 per cent criterion) for each population of *Fagus sylvatica*

Population (province)	Altitude (m a.s.l.)	Mean sample size per locus	Mean no. of alleles per locus	Percentage of polymorphic loci
TAN Passo Tanamea (UD)	800	91.8 (4.1)	2.0 (0.2)	88.9
CAN For. Cansiglio (PN)	1230	112.2 (3.6)	2.3 (0.2)	100.0
GIU Valli Giudicarie (TN)	1250	93.8 (3.1)	2.1 (0.2)	88.9
CER Val Cervo (VC)	1040	94.9 (2.6)	2.3 (0.3)	88.9
ENT Entracque (CN)	1280	79.8 (3.0)	2.2 (0.3)	77.8
PRA Pradaccio (PR)	1200-1400-1600	193.3 (16.7)	2.3 (0.2)	88.9
PNA Passo Pradarena (LU)	1200-1400-1630	134.8 (2.6)	2.2 (0.2)	88.9
PIS Monte Pisanino (LU)	1200-1300-1400	80.0 (3.3)	2.4 (0.2)	100.0
ABE Abetone (PT)	1200-1500-1750	143.4 (9.1)	2.7 (0.2)	100.0
BAG Bagni Romagna (FO)	1030	41.2 (1.3)	2.0 (0.2)	88.9
AMI Monte Amiata (GR)	1500	82.3 (2.8)	2.2 (0.2)	88.9
SUL Sulmona (AQ)	1200	38.0 (1.4)	2.0 (0.2)	77.8
GUA Guardiaregia (CE)	1300	44.9 (1.2)	2.2 (0.3)	77.8
GIO S. Giorgio Matese (CE)	1300	49.0 (1.2)	2.3 (0.3)	77.8
VUL Monte Vulture (PZ)	900	132.6 (12.2)	2.6 (0.3)	100.0
UMB Foresta Umbra (FG)	750	119.6 (3.0)	2.4 (0.2)	100.0
POL Monte Pollino (PZ)	1500	138.9 (2.6)	2.8 (0.2)	100.0
SIL Sila (CS)	1400	76.2 (1.6)	2.2 (0.1)	100.0
SER Serra S. Bruno (CZ)	1100	157.9 (4.1)	2.6 (0.2)	100.0
REC Aspromonte (RC)	1400	137.1 (0.6)	2.9 (0.3)	100.0
SIC Sicilia (ME)	1500	132.8 (2.3)	2.6 (0.2)	100.0

Material was stored at 4°C in the dark. For electrophoresis, bud tissue was homogenized in buffer solution pH 7.3 (Müller-Starck, 1985), using a hand grinder. Buffer and staining solutions were described by Müller-Starck (1985) and Shaw & Prasad (1970).

Genetic types are designated by letters for gene loci ('*A*' indicates the fastest migrating zone) and by numbers for alleles. The generally most common allele is arbitrarily designated *100* and other alleles at the same locus are assigned numerical values representing the electrophoretic migration distance of their homomeric products relative to that of the *100* allele. Ten loci were assayed by means of horizontal starch gel electrophoresis and used to determine the multilocus genotype of each tree: *6PGD-A* (EC 1.1.1.44), *DIA-A* (1.6.4.3), *GOT-A* (2.6.1.1), *GPI-B* (5.3.1.9), *IDH-A* (1.1.1.42), *MDH-A* and *MDH-B* (1.1.1.37), *ME-A* (1.1.1.40), *PER-B* (1.11.1.7), *SKDH-A* (1.1.1.25).

The observed banding patterns are in accordance with the allelism reported in the literature (Thiébaut *et al.*, 1982; Merzeau *et al.*, 1989; Müller-Starck & Starcke, 1993).

Data analysis

Population genetics parameters were estimated using the BIOSYS-1 computer program (Swofford & Selander, 1981). Deviations from Hardy–Weinberg equilibrium were tested by χ^2, pooling genotypes whenever expected frequencies were less than five. Four multilocus measures of genetic diversity were calculated: average number of alleles per locus, percentage of polymorphic loci, mean observed and expected heterozygosities. Wright's *F*-statistics as extended by Nei (1977) were used to estimate genetic differentiation among populations. In four stations for which altitudinal subdivisions existed, *F*-statistics were calculated on four hierarchical levels: individual tree, transects within population, population and total of the four populations. Nei's unbiased genetic distances were computed (Nei, 1987). The SAS statistical package was used for statistical computations (SAS Institute, 1985). Principal components were computed using transformed (arcsin, square root) allele frequencies.

Resampling with replacement of the transformed gene frequencies ('bootstrap' technique (Efron, 1982)) was used to validate the observed geographical patterns with respect to the sampling of genes while conserving the geographical relationship among populations. Each resampling produces a new data matrix by randomly choosing the alleles from the original matrix: some will not be present, some will be present once, some two or more times. The number of alleles of the new matrix is unchanged as, of course, is the number of populations. For each new matrix principal components are computed. We repeated the process 200 times. From the distribution of the first principal component values of each population, the normal deviate (mean/standard deviation) was computed and the statistical probability of zero belonging to the distribution was calculated. The geographical location of the observed significant values supplies information on the robustness of the observed geographical pattern (Piazza *et al.*, 1995).

Principal component coefficients have arbitrary signs (Tatsuoka, 1971, p. 122). In some bootstrap cycles, as expected, a clearly reversed geographical pattern was found in comparison with the computation carried out on the observed data. To avoid a corresponding bias in the estimate of the standard deviation, signs of scores were changed in the cycles with negative correlation between bootstrap scores and scores from the original data.

Regression of latitude with the first principal component scores was computed after each sampling and the distribution of the slope estimates was prepared.

Results

Electrophoretic analysis and genetic variation within populations

Frequencies of alleles are reported in Table 2. Of the 10 enzyme systems we used in the population survey only ME-A was always monomorphic and was omitted from further computations. A good fit to Hardy–Weinberg expectations is considered a confirmation of the allelic nature of the observed electromorphs. With the exception of *GOT-A*, which showed significant deviations from Hardy–Weinberg expectations for most populations, significant differences between observed and expected genotype frequencies were found in few populations: in five for *PER-B* (in populations: CER, ABE, PRA, AMI, SER) and *SKDH-A* (PRA, ABE, BAG, REC, PIS), in three for *IDH-A* (ABE, SUL, REC), in two for *DIA-A* (TAN, VUL), in one for *6PGD-A* (PNA), *GPI-B* (BAG), *MDH-B* (ABE) and MDH-A (SIC). *GOT-A* allele frequencies are probably affected by the difficulty in distinguishing heterozygotes from homozygotes for allele *102*. Allele *102* has relatively low frequencies, and the existence of a third allele unmistakably scorable only in the Sicilian populations advised against the exclusion of this locus from the analysis.

Not considering *GOT-A*, a total of 19 out of 168 (11.3 per cent) significant deviations from Hardy–Weinberg equilibrium was found, a number not

Table 2 Frequencies of alleles for each locus and population of *Fagus sylvatica*

Locus / Allele	Population																				
	TAN	CAN	GIU	CER	ENT	PRA	PNA	PIS	ABE	BAG	AMI	SUL	GUA	GIO	VUL	UMB	POL	SIL	SER	REC	SIC
6PGD-A																					
84	0.042	0.086	0.041	0.029	0.034	0.027	0.045	0.035	0.037	0.034	0.017	0.000	0.014	0.000	0.091	0.005	0.048	0.013	0.012	0.062	0.011
100	0.958	0.914	0.959	0.971	0.966	0.973	0.955	0.965	0.963	0.966	0.983	1.000	0.986	1.000	0.909	0.995	0.948	0.987	0.988	0.938	0.989
110	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000
DIA-A																					
80	0.125	0.122	0.033	0.005	0.034	0.025	0.095	0.085	0.018	0.011	0.007	0.038	0.000	0.000	0.007	0.004	0.000	0.006	0.000	0.004	0.004
90	0.000	0.021	0.000	0.005	0.000	0.014	0.000	0.011	0.024	0.023	0.000	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.003	0.000	0.000
100	0.875	0.857	0.961	0.969	0.939	0.960	0.901	0.903	0.957	0.966	0.993	0.962	0.958	0.980	0.978	0.992	0.986	0.987	0.988	0.989	0.962
105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.004	0.000
110	0.000	0.000	0.007	0.021	0.027	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.042	0.020	0.004	0.004	0.011	0.006	0.009	0.004	0.034
GOT-A																					
97	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
100	0.840	0.950	0.915	0.835	0.827	0.842	0.852	0.701	0.764	0.894	0.597	0.889	0.900	0.856	0.636	0.855	0.774	0.805	0.798	0.850	0.808
102	0.160	0.050	0.085	0.165	0.173	0.158	0.148	0.299	0.236	0.106	0.403	0.111	0.100	0.144	0.364	0.145	0.226	0.195	0.202	0.150	0.188
GPI-B																					
95	0.026	0.013	0.010	0.030	0.000	0.048	0.034	0.040	0.016	0.250	0.113	0.013	0.000	0.000	0.013	0.046	0.039	0.101	0.039	0.047	0.033
100	0.974	0.987	0.990	0.970	1.000	0.952	0.966	0.960	0.984	0.750	0.887	0.988	1.000	1.000	0.987	0.954	0.957	0.899	0.952	0.953	0.967
105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.009	0.000	0.000
IDH-A																					
90	0.000	0.000	0.000	0.015	0.006	0.022	0.011	0.047	0.018	0.000	0.137	0.026	0.022	0.020	0.000	0.008	0.000	0.000	0.000	0.004	0.000
100	0.724	0.729	0.715	0.855	0.629	0.676	0.754	0.753	0.668	0.575	0.786	0.474	0.793	0.724	0.727	0.720	0.831	0.791	0.748	0.817	0.850
110	0.276	0.271	0.285	0.130	0.365	0.302	0.236	0.200	0.314	0.425	0.077	0.500	0.185	0.255	0.273	0.272	0.169	0.209	0.252	0.180	0.150
MDH-A																					
90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.004
97	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.004	0.000	0.000	0.000	0.000	0.000	0.010	0.004	0.021	0.000	0.000	0.000	0.000
100	1.000	0.992	1.000	1.000	1.000	1.000	1.000	0.986	0.989	1.000	1.000	1.000	1.000	0.990	0.977	0.988	0.975	0.994	0.991	0.993	0.982
102	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.010	0.013	0.008	0.000	0.006	0.009	0.007	0.015
MDH-B																					
85	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.000	0.000	0.000	0.000	0.000	0.007	0.000
90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.969	0.954	0.945	0.960	0.959	0.944	0.950	0.964	0.935	0.962	0.937	0.872	0.707	0.816	0.890	0.915	0.934	0.904	0.921	0.938	0.982
115	0.031	0.046	0.055	0.040	0.041	0.056	0.050	0.036	0.050	0.038	0.063	0.128	0.293	0.163	0.110	0.085	0.066	0.096	0.079	0.055	0.018
PER-B																					
90	0.005	0.118	0.065	0.115	0.088	0.322	0.077	0.043	0.023	0.000	0.017	0.000	0.021	0.059	0.007	0.008	0.011	0.020	0.009	0.007	0.000
95	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.011	0.000	0.028	0.004	0.055
100	0.897	0.777	0.745	0.830	0.765	0.417	0.813	0.830	0.813	0.909	0.822	0.949	0.830	0.745	0.881	0.775	0.792	0.855	0.790	0.814	0.776
110	0.098	0.105	0.190	0.055	0.147	0.261	0.109	0.128	0.163	0.091	0.161	0.051	0.149	0.196	0.109	0.216	0.187	0.125	0.173	0.175	0.169
SKDH-A																					
85	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.031	0.000	0.000	0.000	0.000	0.000	0.004	0.000
90	0.017	0.080	0.053	0.080	0.125	0.037	0.008	0.007	0.004	0.064	0.011	0.038	0.070	0.042	0.132	0.075	0.072	0.125	0.065	0.064	0.138
100	0.983	0.900	0.947	0.920	0.837	0.933	0.992	0.972	0.930	0.936	0.978	0.936	0.884	0.917	0.868	0.925	0.924	0.875	0.929	0.929	0.849
110	0.000	0.020	0.000	0.000	0.038	0.030	0.000	0.021	0.067	0.000	0.011	0.026	0.035	0.010	0.000	0.000	0.004	0.000	0.007	0.004	0.013

too far from the expected proportion (5 per cent) arising from type I statistical error. All deviations, except three cases (*GPI-B* in population BAG, *IDH-A* in populations REC and SUL), were the result of heterozygote deficiencies.

Table 1 and Fig. 1 report summary statistics of the observed genetic variability. The average number of alleles per locus is never less than 2. Southern locations tend to have higher values contributing to the overall negative correlation of this parameter with latitude ($r = -0.59$; $P = 0.005$).

No significant correlations with latitude and longitude were found for observed ($r = -0.17$; $P = 0.46$) ($r = 0.19$; $P = 0.41$) and expected heterozygosity ($r = -0.05$; $P = 0.84$) ($r = -0.07$; $P = 0.77$), nor with percentage of polymorphic loci ($r = -0.32$; $P = 0.16$) ($r = 0.38$; $P = 0.08$).

Genetic differentiation

Genetic differentiation among populations is small although in accordance with values reported for forest tree populations. Only about 5 per cent of the total genetic diversity results from differences among populations ($F_{ST} = 0.046$, Table 3). Nei's genetic distances among all possible pairwise population combinations are always less than 0.034.

Significant allele frequency heterogeneities were found among populations at all polymorphic loci (all χ^2 probabilities are less than 0.009), although differences are modest in absolute terms with no alternate fixations and all populations sharing the most frequent allele at all loci (Table 2).

Four localities allowed an estimate to be made of variability among transects taken at different altitudes. *F*-statistics indicate that 90 per cent of genetic variability is attributable to the *within-transect* component and 5.1 per cent to the *among-transect* within-population component. This suggests that the differences among altitudinal transects are almost as small as the differences among the four sites.

Environmental and geographical causes of genetic variation

No association of allele frequencies with altitude was found. Only five out of 40 χ^2 values were significant (12.5 per cent) and no clear trend of allele frequencies with altitude was found in any population.

The possible effect of soils of different geological origins was explored in two Northern Apennines populations (Pisanino and Passo Pradarena) located on opposite sides of a valley with bedrock of different geochemistry (one side made up of highly buffered calcium rocks, the other lying on more acidic quarzitic sandstone). No association was found in our data between allele frequencies and soil type: out of 10 comparisons the homogeneity χ^2 was significant only for *GOT* ($\chi^2_1 = 12.7$; $P < 0.001$).

Correlations between single allele frequencies and latitude and longitude were significant for seven alleles (*DIA-A 100, DIA-A 80, MDH-A 100, MDH-A 102, PER-B 90, PER-B 95, SKDH-A 90*) out of a total of 18 computed coefficients.

Geographical patterns of variation were revealed by the use of synthetic variables generated by multivariate statistics. Different statistics gave essentially the same picture. We report the scatter of the first two principal components scores that account for 22 per cent and 16 per cent of the total variation of allele frequencies,

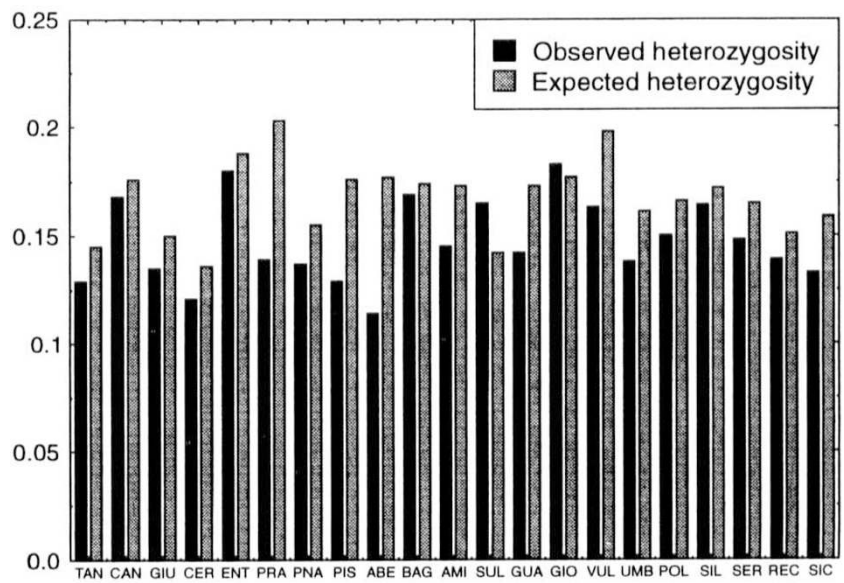

Fig. 1 Observed and expected heterozygosities for each sampling location of *Fagus sylvatica*. See Table 1 for symbol explanation.

Table 3 *F*-statistics according to Nei (1977) for all 21 populations of *Fagus sylvatica*

Locus	F_{IS}	F_{IT}	F_{ST}
6PGD-A	0.079	0.096	0.019
DIA-A	0.033	0.070	0.038
GOT-A	0.395	0.425	0.049
GPI-B	−0.097	−0.020	0.070
IDH-A	−0.012	0.033	0.045
MDH-A	0.037	0.045	0.009
MDH-B	−0.001	0.049	0.050
PER-B	0.145	0.191	0.054
SKDH-A	0.152	0.173	0.025
Mean	0.117	0.157	0.046

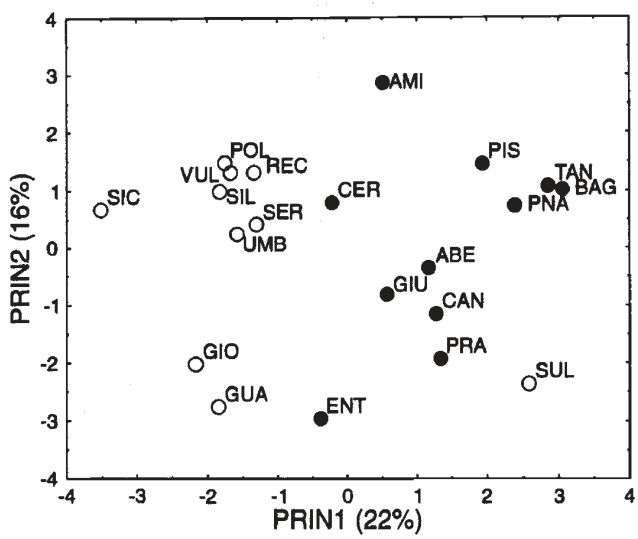

Fig. 2 Plot of first (PRIN1) and second (PRIN2) principal components calculated from transformed allele frequencies. Percentages indicate the proportion of the total variance explained by each principal component. Filled circles refer to northern populations of Italian *Fagus sylvatica*; open circles to southern populations.

respectively (Fig. 2). Northern and southern populations are clearly separated. Only Sulmona (SUL), the northernmost among the southern group with a small sample size, is misclassified.

Regressions of the first principal component scores with both latitude and longitude are significant ($b = 0.52$, $P < 0.001$; $b = -0.36$, $P = 0.01$). In a multiple regression model latitude and longitude explain 54 per cent of the variance although the contribution of longitude is not significant. This results from population scores in the peninsular part of Italy. In the same regression model for northern locations only, the clear east–west trends north and south of the Po plain (Fig. 3) are reflected in a significant contribution of longitude (62 per cent) as well as latitude (20 per cent) to the regression fit.

The geographical pattern was tested by resampling with replacement ('bootstrap' procedure (Efron, 1982)). Three out of 21 populations (14 per cent) were found to be significantly different from zero. More important, their position in the range of beech is what is expected for a north–south cline: one population is located at the extreme north-eastern end of Italy and the other two at the southern end of the Italian Peninsula, in Calabria and Sicily (Fig. 3).

Regression coefficients of latitude with the first component were significant at the 5 per cent level for 144 of 200 cycles computed on bootstrap samples. No negative significant coefficient was found.

Discussion

Genetic variability of beech in Italy seems structured in a fashion comparable with what is known for forest tree populations. For example, using only one of the variability parameters that have been computed, we found that average expected heterozygosity estimates range from 13.6 per cent to 20.3 per cent with a mean of 16.8, a value not too far from average heterozygosities reported for other members of the same family: 18.7 per cent for *Quercus macrocarpa* and 20.4 per cent for *Quercus gabbelii* (Schnabel & Hamrick, 1990), 24 per cent for *Castanea sativa* (Villani *et al.*, 1991) and approximately 21 per cent for *Quercus* species cited in Müller-Starck (1991). Comparison with data from the literature for the same species is made difficult by differences in the number of loci used and in the geographical scale considered. Comps *et al.* (1990) report values between 25.7 per cent and 31.7 per cent (average 30.8 per cent) but on larger geographical subdivisions (whole countries) using only six loci. Considering only loci in common between the two studies for Italy, we obtain much closer estimates (24.5 per cent for our study compared with 23.1 per cent). This confirms the known dependence of heterozygosity upon the loci used for the estimate (Nei, 1987).

Found in almost all populations (Fig. 1), the excess of homozygotes estimated (average $F_{IS} = 0.117$ (Table 3)) is similar to an equivalent estimate for Italian beech populations (average $F_{IS} = 0.125$) by Comps *et al.* (1990). A lack of heterozygotes is often reported among embryos of the more intensely investigated coniferous species (Yazdani, 1985; Muona *et al.*, 1987) but not in later life stages. This difference is usually considered as an effect of selection against selfed individuals (Sorensen, 1969). Cuguen *et al.* (1988)

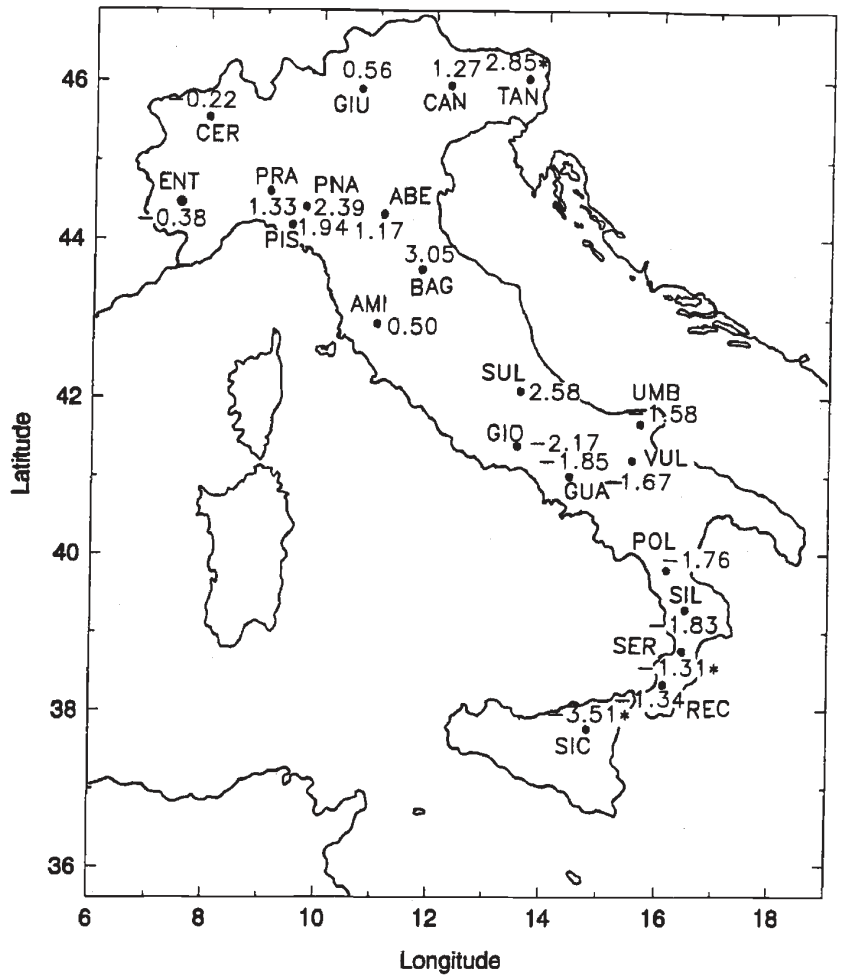

Fig. 3 Geographical locations of the sampling stations for *Fagus sylvatica* (see Table 1 for symbol explanation). First principal component values calculated on transformed allele frequencies are also reported. A star indicates populations with significant mean/SD ratios evaluated by 200 resamplings with replacement of the original data (bootstrap).

investigated the lack of heterozygotes found in 250 European stands studied for three loci. Because selfing was low, they suggest as a complementary explanation the spatial structure of genetic variability, probably generated by reproduction occurring between neighbours. The matter deserves further investigation but cannot be settled with the information of the present study.

Genetic differentiation among Italian beech populations, in agreement with the literature on forest trees, is low with almost 95 per cent of the variability found within populations. Comps *et al.* (1990) found an F_{ST} value of 0.054 among European populations and a slightly higher one (0.058) for their Italian samples using a smaller number of loci. The high proportion of variability found in subdivisions below the population level, already reported in the literature (Knowles, 1984; Gregorius *et al.*, 1986; Diebel & Feret, 1991), reinforces this observation.

The existence of a clear, statistically tested, geographical pattern in the spatial distribution of synthetic statistics is probably the most interesting result of our study. It is somewhat surprising to find a clear geographical structure in the genetic variability of beech in Italy if we consider the low level of interpopulation differentiation (F_{ST} approximately 5 per cent) and the relatively low number of significant associations between single allele frequencies and latitude.

There was little evidence for selection, with no significant number of associations of allele frequencies with soil origin and altitude. The only locus associated with both altitude and latitude was peroxidase, reported as responding to environmental variables (Thiébaut *et al.*, 1982; Felber & Thiébaut, 1984; Comps *et al.*, 1991).

The most plausible explanation for such geographical differentiation is probably to be found in the natural history of these populations. The conceptual models that predict the genetic consequences of range expansion have been worked out (a review in Hewitt (1993)). Clines in gene frequencies can be generated by colonization from two refugia that have developed differences through separation. They can also be

formed as a consequence of an expansion from one refugium by 'leap-frog' long-range dispersal from the populations at the colonizing front, given that diversification results from drift and that successive migration establishes clines.

Postglacial recolonization events have left a trace in the pollen records from Italy. The palynological data, although not as good as for other regions of Europe, has allowed different authors to advance hypotheses of beech postglacial recolonization pathways in Italy. In the south the earliest significant presence of *Fagus* has been found in Calabria in the late-glacial (12 500 to 9000 BP) (Huntley & Birks, 1983) and in the Monticchio Lake sediments (18 000 BP) (Watts, 1985), close to the Vulture population (VUL) (Fig. 3). Moving north the earliest postglacial presence of beech has been determined by ^{14}C dating for more locations (Fig. 4). A temporal trend seems quite clear up to locations in the southern part of western Italy, close to the border with France (3000 BP). From here a reverse time trend seems to exist for pollen cores from locations in the Southern Alps up to the high pollen values found in the area of the Dinaric Alps at much earlier dates (8500 BP) (Huntley & Birks, 1983).

Kral (1979), using the evidence available at the time, proposed a bidirectional postglacial recolonization route for northern Italy with an east–west direction for the Southern Alps and a north-west-bound front from the Northern Apennines. For the peninsular part of Italy a northward recolonization is generally suggested (Ferrarini, 1962; Bertoldi, 1980; Huntley & Birks, 1983). The relative importance of the two recolonization movements is not clearly established in the literature with some authors also suggesting a possible penetration from French southern refugia (Kral, 1979; Beaulieu *et al.*, 1984).

We can only underline the following observations from our genetic data. The higher level of genetic variability found in our data from the south of Italy (negative correlation of mean number of alleles per locus with latitude) is in agreement with the greater diversity expected in former centres of diffusion (Cwynar & MacDonald, 1987). Being a former centre of diffusion seems to over-ride the fact that southern Italy is

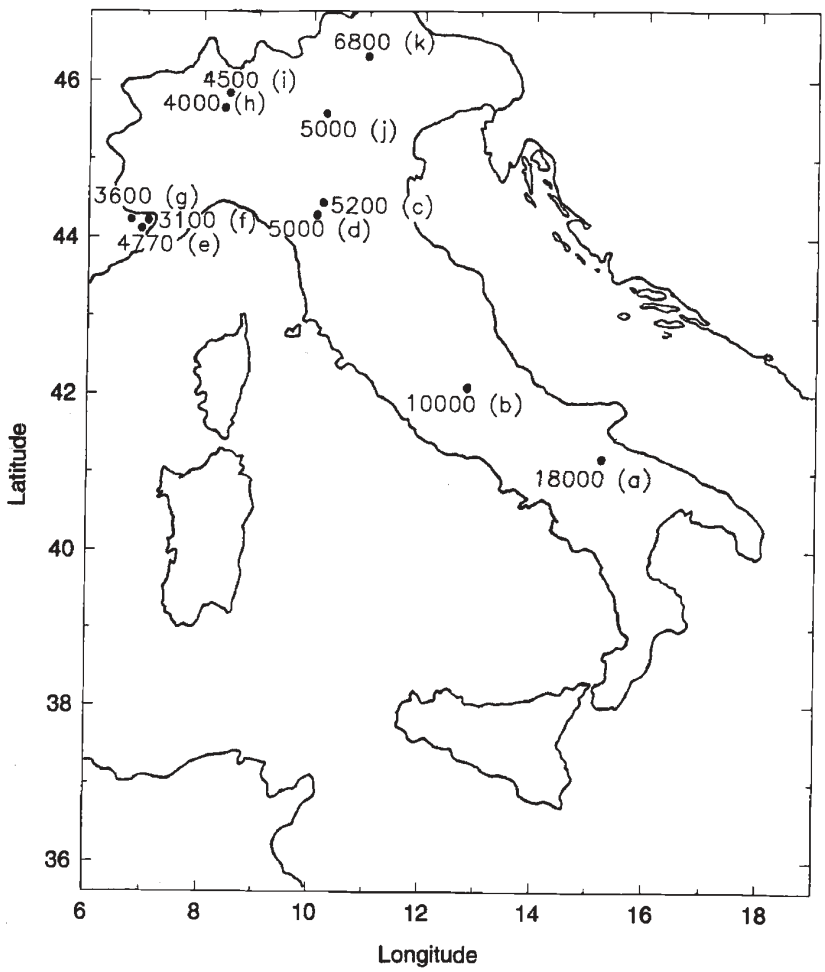

Fig. 4 Map of approximated spread date (years before present) of *Fagus* in Italy. (a) Watts (1985); (b) Follieri *et al.* (1988); (c) Lowe (1993); (d) Lowe (1992); (e)–(g) Beaulieu *et al.* (1984); (h), (i) Schneider (1990); (j) Bertoldi & Consolini (1989); (k) Wahlmüller (1990).

marginal in the biogeographical range of beech. The classical theory developed from observations in *Drosophila*, that predicts lower levels of genetic variability for marginal populations (Dobzhansky *et al.*, 1963), has been found to be true for other coniferous species (Tigerstedt, 1973; Farris & Schaal, 1983; Silander, 1984). These results and the statistically tested south–north gradient found in the principal component scores seem in agreement with the palynological evidence for northward recolonization of peninsular Italy.

Although the genetic data statistically support a north–south differentiation only, it is interesting to interpret the first principal component east–west trends in northern populations as a trace of the east–west postglacial migration suggested by palynological data.

Acknowledgments

The study was supported by the Italian National Research Council (CNR): Special project RAISA (Subproject 2, paper no. 2074). We are grateful to R. Giannini for sharing with us his knowledge of Italian beech. R. Bertoldi guided our efforts in organizing and summarizing palynological data. The authors are grateful to the members of the Italian Forest Service (Corpo Forestale dello Stato) and to S. Baroni, G. Bucci, M. Borghetti and M. Naldi for assistance during sampling. We also thank V. Rossi and two anonymous reviewers for helpful comments on the manuscript.

References

BEAULIEU, J.-L. DE, COÛTEAUX, M., PONS, A., REILLE, M. AND TRIAT LAVAL, H. 1984. Première approche d'une histoire post-würmienne de quelques taxons arborèens dans le sud-est de la France. *Revue de Paléobiologie*, Volume Spécial, 11–24.

BERTOLDI, R. 1980. Le vicende vegetazionali e climatiche nella sequenza paleobotanica würmiana e post-würmiana di Lagdei (Appennino Settentrionale). *L'Ateneo Parmense – Acta Naturalia*, **16**, 147–175.

BERTOLDI, R. AND CONSOLINI, M. 1989. Deglaciazione e sviluppo vegetazionale tardi-postglaciale nella regione del Lago d'Iseo attraverso analisi polliniche e datazione ^{14}C. *Mem. Soc. Geol. It.*, **42**, 139–145.

COMPS, B., THIÉBAUT, B., PAULE, L., MERZEAU, D. AND LETOUZEY, J. 1990. Allozymic variability in beechwoods (*Fagus sylvatica* L.) over central Europe: spatial differentiation among and within populations. *Heredity*, **65**, 407–417.

COMPS, B., THIÉBAUT, B., SUGAR, I., TRINAJSTIC, I. AND PLAZIBAT, M. 1991. Genetic variation of the Croatian beech stands (*Fagus sylvatica* L.): spatial differentiation in connection with the environment. *Ann. Sci. For.*, **48**, 15–28.

CUGUEN, J., MERZEAU, D. AND THIÉBAUT, B. 1988. Genetic structure of the European beech stands (*Fagus sylvatica* L.): F-statistics and importance of mating system characteristics in their evolution. *Heredity*, **60**, 91–100.

CWYNAR, L. C., AND MACDONALD, G. M. 1987. Geographical variation of lodgepole pine in relation to population history. *Am. Nat.*, **129**, 463–469.

DIEBEL, K. E., AND FERET, P. P. 1991. Isozyme variation within the Fraser Fir (*Abies fraseri* (Pursh) Poir.) population on Mount Rogers, Virginia: lack of microgeographic differentiation. *Silvae Genet.*, **40**, 79–84.

DOBZHANSKY, Th., HUNTER, A. S., PAVLOVSKY, O., SPASSKY, B. AND WALLACE, B. 1963. Genetics of natural populations XXXI. Genetics of an isolated marginal population of *Drosophila pseudoobscura*. *Genetics*, **48**, 91–104.

EFRON, B. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics, Philadelphia.

FARRIS, M. A. AND SCHAAL, B. A. 1983. Morphological and genetic variation in ecologically central and marginal populations of *Rumex acetosella* L. (Polygoniaceae). *Am. J. Bot.*, **70**, 246–255.

FELBER, F. AND THIÉBAUT, B. 1984. Étude préliminaire sur le polymorphisme énzymatique du hêtre *Fagus sylvatica* L.: variabilité génétique de deux systèmes de peroxydases en relation avec les conditions écologiques. *Acta Œcol. Œcol. Plant.*, **5**, 133–150.

FERRARINI, E. 1962. Analisi polliniche di depositi lacustri dell'Appennino lunigianese con osservazioni sulla vegetazione naturale. *Ann. Acc. It. Sci. For.*, **11**, 61–95.

FOLLIERI, M., MAGRI, D. AND SADORI, L. 1988. 250,000-year pollen record from Valle di Castiglione (Roma). *Pollen et Spores*, **30**, 329–356.

GÖMÖRY, D., VYŠNÝ, J., COMPS, B. AND THIÉBAUT, B. 1992. Geographical patterns of genetic differentiation and diversity in European beech (*Fagus sylvatica* L.) populations in France. *Biológia (Bratislava)*, **47**, 571–579.

GREGORIUS, H.-R., KRAUHAUSEN, J. AND MÜLLER-STARCK, G. 1986. Spatial and temporal genetic differentiation among the seeds in a stand of *Fagus sylvatica* L. *Heredity*, **58**, 255–262.

HAMRICK, J. L. AND GODT, M. J. W. 1990. Allozyme diversity in plant species. In: Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*, pp. 43–63. Sinauer, Sunderland, MA.

HAMRICK, J. L., LINHART, Y. B. AND MITTON, J. B. 1979. Relationship between life history characteristics and electrophoretically detectable genetic variation in plants. *Ann. Rev. Ecol. Syst.*, **10**, 173–200.

HEWITT, G. M. 1993. Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In: Leeds, D. R. and Edwards, D. (eds) *Evolutionary Patterns and Processes*, pp. 97–123. The Linnean Society of London/Academic Press, London.

HUNTLEY, B. AND BIRKS, H. J. B. 1983. *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 years ago*. Cambridge University Press, Cambridge.

KNOWLES, P. 1984. Genetic variability among and within closely spaced populations of lodgepole pine. *Can. J. Genet. Cytol.*, **26**, 177–184.

KRAL, F. 1979. *Spät- und postglaziale Waldgeschichte der Alpen auf Grund der bisherigen Pollenanalysen*. Veröffentlichung des Instituten für Waldbau an der Universität für Bodenkultur in Wien, Wien.

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

LOVELESS, M. D. AND HAMRICK, J. L. 1984. Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.*, **15**, 65–95.

LOWE, J. J. 1992. Lateglacial and early Holocene lake sediments from the northern Apennines, Italy – pollen stratigraphy and radiocarbon dating. *Boreas*, **21**, 193–208.

LOWE, J. J. 1993. Lateglacial and early Holocene pollen stratigraphy of the northern Apennines, Italy. *Q. Sci. Rev.*, **12**, 727–738.

MERZEAU, D., DI GIUSTO, F., COMPS, B., THIÉBAUT, B., LETOUZEY, J. AND CUGUEN, J. 1989. Genetic control of isozyme systems and heterogeneity of pollen contribution in beech (*Fagus sylvatica* L.). *Silvae. Genet.*, **38**, 195–201.

MÜLLER-STARCK, G. 1985. Genetic differences between "tolerant" and "sensitive" beeches (*Fagus sylvatica* L.) in an environmentally stressed adult forest stand. *Silvae Genet.*, **34**, 241–247.

MÜLLER-STARCK, G. 1989. Genetic implications of environmental stress in adult forest stands of *Fagus sylvatica* L. In: Scholz, F., Gregorius, H.-R. and Rudin, D. (eds) *Genetic Effects of Air Pollutants in Forest Tree Populations*, pp. 127–142. Springer, Berlin.

MÜLLER-STARCK, G. 1991. Survey of genetic variation as inferred from enzyme gene markers. In: Müller-Starck, G. and Ziehe, M. (eds) *Genetic Variation in European Populations of Forest Trees*, pp. 20–37. J. D. Sauerländer, Frankfurt am Main.

MÜLLER-STARCK, G. AND STARCKE, R. 1993. Inheritance of isoenzymes in European Beech (*Fagus sylvatica* L.). *J. Hered.*, **84**, 291–296.

MUONA, O., YAZDANI, R. AND RUDIN, D. 1987. Genetic change between life stages in *Pinus sylvestris*: allozyme variation in seeds and planted seedling. *Silvae Genet.*, **36**, 39–42.

NEI, M. 1977. *F*-statistics and analysis of gene diversity in subdivided populations. *Ann. hum. Genet.*, **41**, 225–233.

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

PIAZZA, A., RENDINE, S., MINCH, E., MENOZZI, P., MOUNTAIN, J. AND CAVALLISFORZA, L. L. 1995. Genetics and the origin of European languages. *Proc. Natl. Acad. Sci. U.S.A.* (in press).

SAS INSTITUTE 1985. *SAS User's Guide: Statistical Version*. 5th edn. SAS Institute, Cary, NC.

SCHNABEL, A. AND HAMRICK, J. L. 1990. Comparative analysis of population genetic structure in *Quercus macrocarpa* and *Quercus gambelii* (Fagaceae). *Syst. Bot.*, **15**, 244–251.

SCHNEIDER, R. E. 1990. L'influsso umano sulla vegetazione neolitica nel territorio di Varese dedotto dai diagrammi pollinici. *Monografie di 'Natura Bresciana'*, **13**, 91–97.

SHAW, C. R. AND PRASAD, R. 1970. Starch gel electrophoresis of enzymes: a compilation of recipies. *Biochem. Genet.*, **4**, 297–320.

SILANDER, J. A., Jr. 1984. The genetic basis of ecological amplitude of *Spartina patens*. III. Allozyme variation. *Bot. Gaz.*, **145**, 569–577.

SORENSEN, F. 1969. Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. *menziesii*. *Am. Nat.*, **103**, 389–398.

SWOFFORD, D. L. AND SELANDER, R. B. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *J. Hered.*, **72**, 281–283.

TATSUOKA, M. M. 1971. *Multivariate Analysis: Techniques for Educational and Psychological Research*. John Wiley, New York.

THIÉBAUT, B., LUMARET, R. AND VERNET, Ph. 1982. The bud enzymes of beech (*Fagus sylvatica* L.). Genetic distinction and analysis of polymorphism in several French populations. *Silvae Genet.*, **31**, 51–60.

TIGERSTEDT, P. M. A. 1973. Studies on isozyme variation in marginal and central populations of *Picea abies*. *Hereditas*, **75**, 47–60.

VILLANI, F., PIGLIUCCI, M. AND BENEDETTELLI, S. 1991. Genetic variation of Italian chestnut: a tool to study environmental impact. In: Giannini, R. (ed.) *Effects of Pollution on the Genetic Structure of Forest Tree Populations*, pp. 57–65. CNR, Firenze.

WAHLMÜLLER, N. 1990. Spät- und postglaziale Vegetationgeschichte des Tschögglberges (Südtirol). *Ber. nat.-met. Verein Innsbruck*, **77**, 7–16.

WATTS, W. A. 1985. A long pollen record from Laghi di Monticchio, southern Italy: a preliminary account. *J. geol. Soc. London*, **142**, 491–499.

YAZDANI, R. 1985. Genetic structure of *Pinus sylvestris* L. seed-tree stand and naturally regenerated understory. *Forest Sci.*, **31**, 430–436.