Does the phenotypic selection affect the genetic structure and diversity? A study case on Walnut in eastern central Italy (the region of Marche)[§]

Fulvio Ducci[1*], Maria Emilia Malvolti[2], Roberta Proietti[1], Lorella Brandoni[3]

Accepted 12 settembre 2008

Abstract – The Persian walnut (*Juglans regia* L.) is widely planted in western Europe, either for fruit either for high quality timber production. This tree is generally considered non authoctonous, probably introduced from East some 7000 years ago and spread by several ancient civilisations. The possible artificial origin seems confirmed by the low intra-specific variation and the higher individual variability recorded by several Authors as well as by the lack of natural populations. Indeed, only wider fruit cultivation areas or small groups, lines or isolated walnut trees can be recorded in Italy. The occurrence of walnuts in forest, escaped from cultivation areas, is very rare. Due to the increased interest of planters, walnut plantations have been extended several ten thousands hectares throughout all western Europe. As a consequence of that it was evident the necessity of selected suitable basic populations in order to supply high quality reproductive materials. The conventional method based on the organisation of a wide and exhaustive seed procurement from the native range to establish provenance tests is at the present impossible. Thus it is necessary to study methods of selection which consider basic materials growing within the western European range. This study is aimed to test the efficiency of the multi-trait *Selection Index* method, in preserving levels of genetic diversity and structures compatible with the standards observed within a reference system of extended Italian populations. As a consequence of the relatively recent introduction, the genetic structure of the species shows individual variation higher than inter-population diversity. Those genetic structure characteristics were revealed also during a survey of walnut resources in the region of Marche, central Italy. The survey was the starting point for selecting and preserving basic materials for high quality woody production, possibly interesting for forest nurseries in the region. The genetic variation of Marche's population, compared to a reference system of 7 other Italian provenances, was used as a base to establish a possible improvement strategy together with basic guidelines to manage those genetic resources. Indeed, the very important individual component of the genetic variation suggested to select directly superior phenotypes in view of establishing a comparative multisite progeny test network. No substantial differences were detected concerning the genetic structure of the Italian population, neither within the Marche population, neither in the phenotypically selected material. Homozygosity was always high, probably due to genetic erosion, isolation and adaptation to extreme conditions. Given this general situation in the Marche area, a special care should be payed in the management of walnut reproductive materials, in order to maintain sufficient levels of variation in plantations.

Key words: Juglans regia, *selection index, genetic variability, isozymes.*

Riassunto – La selezione fenotipica influenza la struttura genetica e la diversità? Un caso di studio sul noce in Italia centrale (la regione Marche). Il noce comune (*Juglans regia* L.) è stato probabilmente introdotto in Europa occidentale da oriente circa 7000 anni fa. E' una specie con diffusione sporadica, ampiamente coltivata e soggetta a sostanziali manipolazioni da parte dell'uomo, utilizzata soprattutto per la produzione di frutti. Dal punto di vista genetico presenta bassi livelli di variabilità intra-specifica e alti tra individui. Secondo alcuni autori questa caratteristica confermerebbe anche la sua origine non autoctona. L'interesse per il noce è aumentato negli ultimi 20 anni, grazie alla sua utilizzazione in piantagioni per la produzione di legno pregiato. Tuttavia solo di recente sono stati avviati programmi di selezione di materiale di propagazione idoneo nell'ambito del germoplasma ancora disponibile. Questo lavoro è nato da questa esigenza e si pone l'obiettivo di individuare un indice di selezione multi-carattere, capace di mantenere alti livelli di variabilità genetica nel materiale selezionato. L'indagine è stata effettuata nel territorio della regione Marche, dove, utilizzando questo *Indice di Selezione*, sono stati identificati 83 fenotipi adatti per la produzione di legno, poi suddivisi in tre gruppi: *Acquasanta* (AP), *Feltria* (PU) e *Sibillini* (MC). L'analisi genetica, di significato esplorativo, è stata eseguita con marcatori di tipo biochimico (isoenzimi). Le provenienze delle Marche sono state poi confrontate con 7 popolazioni italiane, utilizzate come sistema di riferimento. I risultati ottenuti evidenziano che la struttura genetica delle popolazioni selezionate nella regione Marche è simile a quella delle altre popolazioni italiane sia dal punto di vista della variabilità inter-popolazione che intra-popolazione. Notevolmente superiore è invece risultato il livello medio di omozigosi probabilmente legato alle ridotte dimensioni delle sub-popolazioni artificiali ed al loro isolamento geografico, trovandosi esse nella maggior parte dei casi in aree "relativamente chiuse" fino a pochi decenni fa. Inoltre, non è possibile individuare una strutturazione geografica netta tra popolazioni, probabilmente a causa della ormai millenaria pressione selettiva di origine antropica a carico della specie. Vista la ridotta variabilità del noce, un aspetto importante da considerare è la programmazione di interventi di conservazione sia *in situ* che *ex situ* delle risorse ancora disponibili.

Parole chiave: Juglans regia, *indice di selezione, variabilità genetica, isoenzimi.*

F.D.C.: 176.1 Juglans regia : 165.6 : 165.3 : (450.57)

[§] The work reported in this paper was funded by the EU – Brains Fair III CT96-1887 Programme and by the Region of Marche (Italy) in the framework of "Selection and genetic improvement of forest species for the regional forest nursery".

[*] Correspondent author E-mail: *fulvio.ducci@entecra.it*

[1] CRA - Centro di Ricerca per la Selvicoltura, Viale S. Margherita, 80 – 52100 Arezzo (Italy) T. +39 0575 353021, Fax +39 0575 353490

[2] CNR – Istituto di Biologia Agroambientale e Forestale, Villa Paolina – Via G. Marconi, 2 - 05010 Porano (Tr - Italy) T. +39 0763 374911, Fax +39 0763 374980, E-mail *mimi@ibaf.cnr.it*

[3] Regione Marche.

Introduction

Most of the *Juglans regia* L. genetic resources in western Europe are not native (Malvolti *et al.* 1993, 1995b, 1997). Indeed, no natural populations are presently growing in this area and nowadays, walnuts are an attractive component of the rural landscape as single trees, small groups or lines. In the past, their resource value was found in fruit, oil and tannins, and wood was only seen as the final product at the end of the rotation. Anyway, studies carried out on pollen sediments in several parts of Europe suggested that probably small and rare populations managed to find suitable conditions for survival during the last Ice Age. The Balkans and southern Italy seem to have been two of the main refuge areas (Huntley and Birks 1983) but, concerning Italy, the residual population probably disappeared or was quantitatively un-significant for the following expansion. Leslie and Mc Granahan (1988) dated walnut in Italy at least before 5000 years B.C., but thought the species originated in central Asia, between the Caucase Mountains and the Himalayas.

The origin from a restricted genetic base, either natural either artificial, seems confirmed by the western Europe genetic resources which are characterised by low variation levels. The present variation is mainly affected by serious genetic erosion and inbreeding problems. Indeed, most of the present gene pool were originated from fruit varieties introduced by several ancient civilisations and/or probably from the few mother trees remaining after the Ice Age (Malvolti *et al.* 1996).

Well known and grown by the Greeks (Esopus 1978), Romans, Arabs etc. for nut production, most of walnuts were presumably spread by these civilisations throughout Europe.

During the last twenty years, the walnut has assumed prominence as a valuable source of high quality wood for industry and has became the most important tree used for intensive forest tree farming in central and southern Europe. It was estimated that 80 % of 105.000 ha planted in Italy between 1992 and 2000 were established with walnuts as the main species (Colletti 2001).

In recent years, a number of problems have been noted with regard to seed supplies for nurseries. The seeds and seedlings for walnut propagation which are found in the nursery trade do not always come from selected stock: their provenance cannot be certified nor, in many cases, are their genetic and phenotypic characteristics known (Ducci 1989, Ducci *et al.* 1997). Selecting new resources with "forest" characteristics, which make them more suitable for timber products, is thus necessary in order to establish a wider and better adapted genetic base.

The best strategy for improving tree stocks would be a preliminary survey and seed supply of provenances within the Asiatic zones of origin (Hemery 2000), in order to set up traditional field experiments. Due to organisational problems and short-term difficulties in raising a comprehensive sampling of Persian walnut from its present natural range, the selection of local resources in western Europe has became the most important challenge. In addition, the dramatic migration of rural populations to industrial areas, followed by a progressive reduction of the existing walnut stock, has made the preservation of genetic resources a matter of urgency. There is also a pressing requirement from forest nurseries for seed supply determined by forest tree farming programs. This should give focus to the study of existing varieties and to the selection of phenotypically suitable basic materials for wood production.

In general, most of the variation found in western Europe is related to individual effects (Malvolti *et al. op. cit.*). This aspect, together with the absence of true populations, convinced us to by-pass the traditionally-made provenance tests and to start individual phenotypic selection immediately (Zobel and Talbert 1984). Indeed, it is impossible defining walnut populations as communities of individuals of this species evolved within their own environment[4]. The whole species in western Europe could be considered as a wide homogeneous artificial population. For this reason "provenances" or "groups" are reported instead of "populations".

This paper shows the methods used in order to: *i)* establish a genetic base for this endangered and greatly manipulated resource; *ii)* supply seeds for forest nurseries; *iii)* test a rapid and cheap method

[4] A population can be defined on several ways according to the subject, in 1992 Lynne Corne supplyed the following: "a group of organisms coexisting at the same time and place and capable for the most part of interbreeding" and as a group of conspecific organisms that occupy a more or less well defined geographic region and exhibit reproductive continuity from generation to generation; it is generally presumed that ecological and reproductive interactions are more frequent among these individuals than between them and the members of other populations of the same species (Futuyuma 1979, Colinvaux, 1986).

to estimate how great could be the impact of selecting basic materials on the genetic structure of walnut and *iv)* examine relationships with a reference set of other populations. Another aim is to see how problems relating to preserving and selecting these genetic resources can be addressed and how the forest nursery trade in the central Italian region of Marche can be improved.

Material and Methods

Walnut materials

An initial survey of the walnut resources that might be useful for forestry purposes was carried out as a preliminary. Materials were carefully selected tree by tree within the existing stock of walnut trees in the region of Marche. The main genetic parameters of the total selected Marche artificial "population" was compared to a reference system of 8 Italian walnut provenances sampled in northern, central and southern Italy (figure 1).

Only trees whose phenotypes were characterised by typical forest shapes were inventoried on a description sheet and their location was noted by GIS methods (Ducci and Veracini, 1992). Historical information was also collected near the owners in order to exclude clones or fruit varieties or material introduced too recently from other areas.

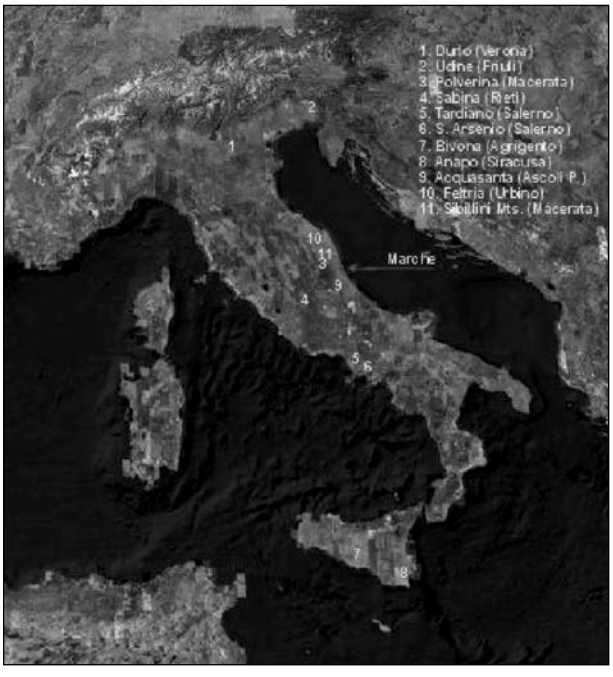

Figure 1 - *Distribution of analyzed populations (Google Earth, 2007).* Distribuzione delle popolazioni analizzate (Google Earth, 2007)

In total, 83 walnut trees showing the requisite traits were surveyed, defined as "forest shape population". These trees were already at least 35 years old, before younger trees were introduced by recent intensive forest tree farming or re-afforestation programmes. They were distributed into three provenances defined *a priori* by geographical criteria: *Acquasanta* (Ascoli Piceno), *Feltria* (Pesaro-Urbino) and *Sibillini Mountains* (Macerata). *Acquasanta* included 2 small sub-groups, *Arquata* and *Faete*, isolated and dense, as well as *Feltria* included 2 sub-groups *Camerino* and *Pesaro-Urbino*, relatively scattered on a wide area.

Trees within each of these provenances were distributed by isolated individuals, by small groups or by lines.

Selection methods

On a second phase, a set of 39 superior phenotypes was selected within the total "forest shape population". The main phenotypic traits scored were: stem form, branch architecture, dominance, fruit production, health and vigour (expressed by height and Dbh).

The adoption of single trait analysis for selecting phenotypically superior trees makes the work very hard. It has been used, therefore, a multitrait selection method, namely the *Selection Index* (SI) described by Zobel and Talbert *op. cit.*. This method was also used successfully to evaluate phenotypes in wild cherry (Ducci *et al.* 2005, 2006). Traits synthesised within the *Selection Index* (SI) were scored and weighted decreasingly according to their economic (stem form), cultivation value (branch architecture for pruning), adaptation/health state and growth parameters.

The distribution of trees among Dbh and SI classes was examined. The selection was carried out by the same team who had selected walnuts around Italy for CRA-ISSEL during previous surveys (Ducci and Veracini *op. cit.*, Ducci *et al. op. cit*). The selected groups will form the genetic base core used in order to proceed with improvement programmes in the region.

Concerning the genetic analysis, the pooled population of Marche was sub-divided into a "selected part", namely *Population 1* and a "non-selected part", *Population 2*. The genetic traits of both those sub-groups were compared with each other population included the reference system, in order to estimate how the original genetic structure of the basic material had been altered.

Allozyme analysis

Genetic variability was studied by readable and reproducible enzyme *loci* via starch gel electrophoresis. Loci were resolved from 8 enzymatic systems: Diaphorase (DIA, E.C. 1.6.4.3), Shikimic Dehydrogenase (SKDH, E.C. 1.1.1.25), 6-Phospho-gluco-dehydrogenase (6-PGD, E.C. 1.1.1.43), Phosphoglucomutase (PGM, E.C. 2.7.5.1), Glutamate oxaloacetate transaminase (GOT, E.C. 2.6.1.1), Phosphoglucose isomerase (PGI, E.C. 5.3.1.9), Malate dehydrogenase (MDH, E.C. 1.1.1.37), Isocitric dehydrogenase (IDH, E.C. 1.1.1.42). Enzymes were extracted from winter buds tissues according to Kim (1979, 1980) and Malvolti *et al.* (1993, 1995a, 1995b)

Within the 3 Marche provenance materials it was possible to examine 16 *loci*, 11 of these being polymorphic (DIA-1, DIA-3, SKDH-1, SKDH-2, 6PGD-2, GOT-2, GOT-3, PGI-2, IDH-1, MDH-2 and PGM-1). In total 27 alleles were revealed. Starch gel electrophoresis was carried out according to Ashton and Braden (1961), Aletà *et al.* (1993) and Malvolti *et al.* (1994). Zymograms were stained after Conkle *et al.* (1982) and Vallejos (1983).

Only 8 common *loci* were used for making comparisons to the reference provenances: DIA-1, DIA-3, SKDH-1, SKDH-2, 6PGD-2, PGI-1, GOT-1 and PGM-1, by revealing 17 alleles.

The genetic analysis was carried out according to Arulsekar *et al.* (1985, 1986), Aletà *et al.* (*op. cit.*) and Malvolti *et al.* (*op. cit.*). *Loci* were labelled sequentially, with those migrating closest to the anodal end designated as "1". Within a single zone of activity the allozyme variants were labelled according to their mobility and the faster allele was labeled "a".

Allozymes are considered in general as neutral markers with respect to evolution, then it is not possible to establish a correlation with adaptive traits. They rely on a relatively low number of loci and their distribution through the genome is relatively irregular. Anyway, they are codominant and the allelic variation can be easily detected in order to collect preliminary information about the genetic structure of selected materials. These biochemical markers can be therefore adopted to supply an easy and cheap method of evaluation of the genetic variation to improvers and breeders during the first steps of the phenotypic selection. On this way useful information for the management and conservation of improved genetic resources will be available.

Data analysis

Most of the following genetic variation indices were computed using Biosys 1.7 statistical software (Swofford and Selander 1989). Genetic diversity and population-level homozygosity were estimated by: mean number of alleles per *locus* (n), % of polymorphic *loci* ($P_{5\%}$; a *locus* was considered polymorphic when the percentage of the most common allele was less of 95%), expected heterozygosity He, observed heterozygosity Ho, genetic distance among populations (Nei 1978).

The Wright's inbreeding coefficient (Wright 1978) was estimated in order to measure deviations from the Hardy-Weinberg equilibrium[5]. F was calculated at each polymorphic locus and significant deviation from Hardy-Weinberg equilibrium was tested by the x^2 test (Nei *op. cit)*. The average fixation indices were calculated for each population and tested for significant difference from zero. The x^2 test of heterogeneity was carried out to test differences among population (Goudet 2001).

The examined provenances were clustered by UPGMA method (Unweighted Pair Group Method with Arithmetic averaging – Sneath and Sokal 1973), based on Nei (*op. cit.*) genetic distance, in order to explore possible relationships with geographic distance (Farris 1972).

Differentiation indices (Nei 1987, Hartl and Clark 1989) H_T[6] (total gene diversity), H_S (diversity within populations component), D_{ST} (diversity among populations component) and G_{ST} (genetic differentiation) were also estimated. The genetic structure of the studied population was also analyzed in terms of F statistics (Weir and Cockerham 1984): F_{IS} (genetic variability on total population), F_{ST} (genetic variation among populations component) and F_{IS} (genetic variation within population component). Both Nei's and Weir and Cockerham's indices are fixation indices and estimate the possible excess or defect of homozygosity respect to a reference value.

[5] F = 0, shows equilibrium, while negative or positive F values indicate respectively excess of hetorozygotes or homozygotes.

[6] $H_T = H_S + D_{ST}$; $G_{ST} = D_{ST}/_{HT}$. G_{ST} ranges from 0 (all the genetic variation maintained within populations) to 1 (all genetic variation maintained among populations).

The exact test procedure was performed according to the bootstrapping analysis (Raymond and Rousset 1995) in order to assess the effect of missing data and sampling errors on the estimation of genetic distances and fixation indices (Weir and Cockerham *op. cit.*, Cavalli – Sforza *et al* 1994,), using F-STAT software, 2.9.3 version (Goudet *op. cit.*).

The principal component analysis (PCA) on the dispersion matrix, obtained via the Correspondence Analysis (El-Kassaby 1991), was carried out on the total population of Marche with software NTSYS-pc (Numerical Taxonomy and Multivariate Analysis System, Rholf 1994). This procedure was used in order to determine *loci* the most related to the genetic diversity and to discriminate, *a posteriori*, possible genetically-homogeneous groups.

Results

Phenotypic selection

The phenotypic selection carried out per single traits is impossible when both timber and wood quality are required together with growth. It has been therefore necessary to adopt the method of SI, which allows the evaluation of the whole phenotype based on quantitative (i.e. growth) and qualitative traits of economical and technological meaning.

After the phenotypic selection was concluded the selected population was compared with the total one. Figures 2a and 2b show the initial structure of the population when only diameters were considered. The distribution seems to be relatively normal, but it should be considered that the age of trees is unknown and probably it is not strictly correlated to diameters. Most of the selected population is concentrated between 25 and 50 cm.

When trees among *Selection Index* classes were examined, a multi-modal distribution appeared (Figure 3a). This was probably due to the characteristic of plantations where walnut is cultivated and their own age, to the soil fertility and also to cultivation methods. The selected trees, were distributed on a wide range of diameter classes. Height and diameter were included in the SI with relatively lower weights. Figure 3b confirms the very low relationships between diameters and SI values. SI values of selected trees (Figure 3c) varied between 60 and 250 with a range of Dbh between 30 and 72 cm.

Genetic markers

The Marche region populations and the other Italian reference populations

Of the 8 considered loci (SKDH-1, SKDH-2, DIA-1, DIA-3, PGM-1, 6PGD-2, GOT-1, PGI-1), 2 (GOT-1 and PGI-1) were monomorphic in all populations, while 6 were polymorphic. The allele frequencies and the main genetic parameters of the populations examined are reported respectively in tables 1 and 2.

The allele distribution was similar in the different populations and the mean number of allele *per locus* was 2 everywhere, except for the Sicilian population of *Bivona*. Indeed, the value was 1.9, probably because

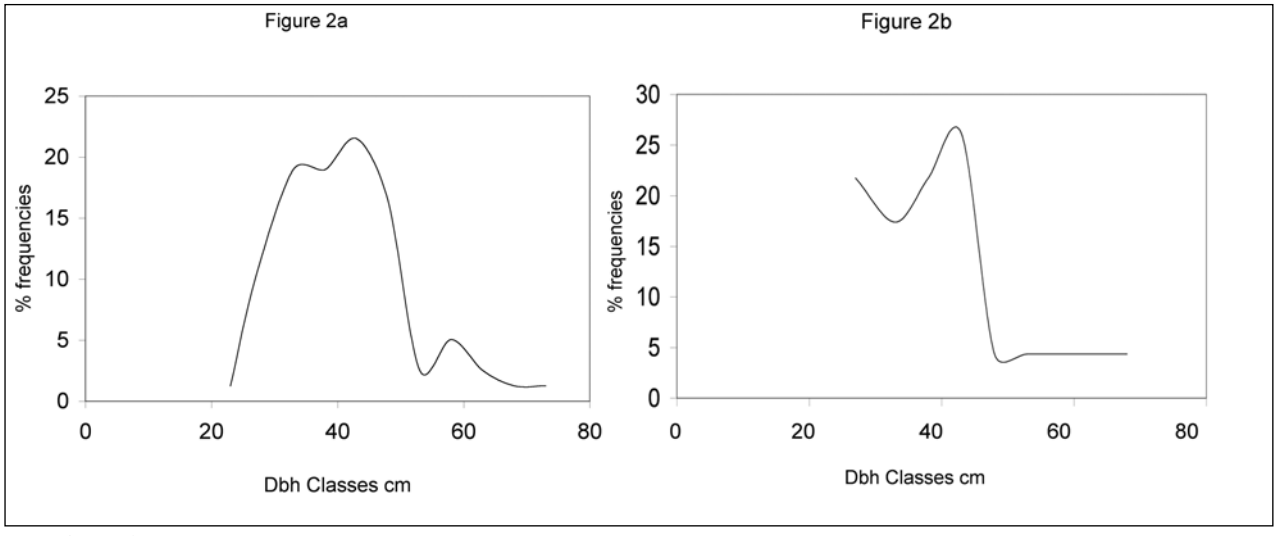

Figure 2- *% frequencies by Dbh classes in the whole walnut (2a) and in the selected walnut (2b).*
Frequenze per classi diametriche nella popolazione totale (2a) e in quella selezionata (2b)

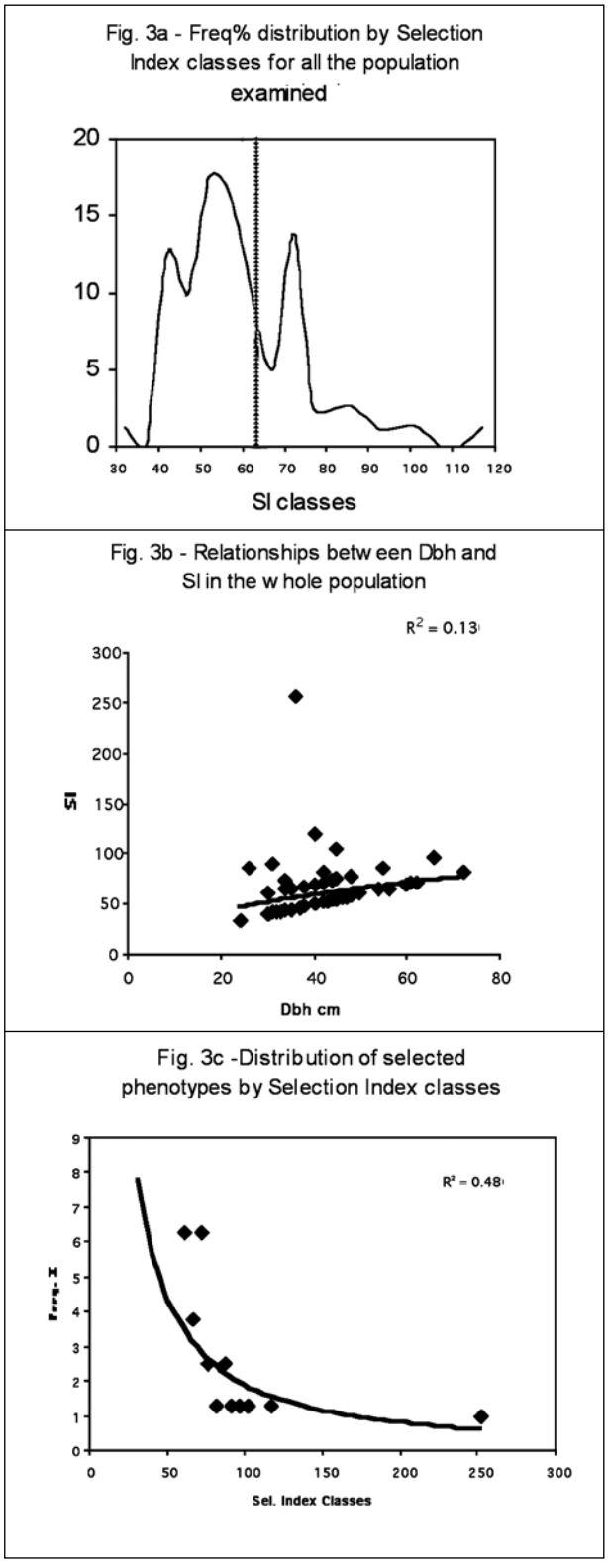

Figure 3 - *Selection Index.*
Indice di selezione. Figura 3a: frequenza per classi di Indice di Selezione (IS) per l'intera popolazione esaminata. Figura 3b: Relazioni tra Dbh e IS nell'intera popolazione. Figura 3c: Distribuzione dei fenotipi selezionati per classi di IS.

of the absence of allele *a* in *locus* PGM-1. This allele was also absent in *Tardiano*, characterised by the rare allele d in PGM-1.

Loci were polymorphic ($P_{5\%}$) at 62.5% in *Durlo*, *Polverina* and *Bivona*, less than in the other populations (75%).

The expected heterozygosity values (H_e) were higher for the Marche populations and varied between .375 (*Feltria*) and .359 (*Acquasanta*). The other populations showed values between .305 (*Polverina*) and .356 (*S. Arsenio*). On the other hand, the observed heterozygosity values (H_o) were smaller in Marche compared to the reference system. The mean fixation index values (F) were therefore relatively higher (table 2) in the Marche range.

The F values (Wrigth *op. cit.*) were negative only for *Durlo* and *Tardiano* (excess of the heterozygotes), whilst all the other populations were characterised by low excess of homozygotes. Both those deviations from the Hardy-Weinberg equilibrium were no significant. Deviations were noted as significant for *Feltria* and *Sibillini* at SKDH-1 *locus*, for *Durlo*, *Friuli*, *Acquasanta* and *Sibillini* at SKDH-2 *locus*, for *S. Arsenio*, *Bivona* and *Anapo* at DIA-3 *locus*, for *Polverina* and *Acquasanta* at PGM-1 *locus* and at 6PGD-2 *locus* for *Sabina* and *Acquasanta* (table 3).

The mean values of the three F indices (Weir and Cockerham *op. cit.*) were positive (table 4a); an excess of homozygotes was thus evident in both inter-populations (F_{ST}: .054) and intra-populations (F_{IS}: .092). SKDH-2 (F_{IT}: .281) and PGM-1 (F_{IT}: .205) gave the highest contribution to this result. Low level of differentiation (Table 5) were noted among populations (G_{ST} = .038). The most important amount of the total gene diversity (H_T = .353) was observed within populations (H_S = .340).

The genetic distance analysis showed the minimum distance (.001) between the two northern populations (*Durlo* and *Friuli*) and between the two southern populations (*Tardiano* and *S. Arsenio*), while the maximum distance was of *Acquasanta* with *Bivona* (.063) and *Tardiano* (.059). However, a geographical gradient variation of genetic distance does not exist. In fact *S. Arsenio*, a southern population, was closer to *Friuli* than *Sabina*, a population in central Italy. The same happened for the two Sicilian populations which the distance was greater compared with respect to other central-northern populations. Concerning the populations of Marche, *Sibillini* presented the

Table 1 - *Estimated allele frequencies to examined* loci *in Italian populations used as reference system, Marche populations and Marche in all.*
Frequenze alleliche stimate per i *loci* esaminati nelle popolazioni italiane impiegate come sistema di riferimento, in quelle delle Marche e nella popolazione Marche nel suo complesso.

	Population											
***Locus*/ alleles**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
$Skdh^{-1}$												
(N)	30	46	26	31	27	30	35	32	19	14	39	72
A	.367	.315	.288	.419	.370	.383	.357	.406	.474	.321	.436	.424
B	.633	.685	.712	.581	.630	.617	.643	.594	.526	.679	.564	.576
$Skdh^{-2}$												
(N)	30	20	27	31	26	30	36	32	19	14	39	72
A	.033	.075	.037	.177	.173	.200	.042	.063	.263	.357	.231	.264
B	.967	.925	.963	.823	.827	.800	.958	.938	.737	.643	.769	.736
Dia^{-1}												
(N)	30	46	27	31	27	30	36	32	20	14	39	73
A	.400	.522	.611	.710	.500	.533	.472	.828	.475	.429	.500	.479
B	.600	.478	.389	.290	.500	.467	.528	.172	.525	.571	.500	.521
Dia^{-3}												
(N)	30	46	27	31	27	30	35	31	20	14	39	73
A	.750	.739	.611	.823	.833	.683	.543	.597	.650	.536	.590	.596
B	.250	.261	.389	.177	.167	.317	.457	.403	.350	.464	.410	.404
Pgm^{-1}												
(N)	30	46	27	31	27	30	36	32	19	14	39	72
A	.033	.152	.037	.113	.000	.033	.000	.063	.079	.107	.141	.118
B	.650	.554	.815	.419	.519	.567	.486	.688	.789	.607	.744	.729
C	.317	.293	.148	.468	.407	.400	.514	.250	.132	.286	.115	.153
D	.000	.000	.000	.000	.074	.000	.000	.000	.000	.000	.000	.000
$6Pgd^{-2}$												
(N)	30	46	27	31	27	30	34	32	5	13	38	56
A	.333	.348	.333	.419	.370	.333	.309	.297	.200	.154	.316	.268
B	.417	.457	.426	.339	.593	.567	.588	.500	.200	.692	.513	.527
C	.250	.196	.241	.242	.037	.100	.103	.203	.600	.154	.171	.205
Got^{-1}												
(N)	30	46	27	31	27	30	36	32	20	14	39	73
A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Pgi^{-1}												
(N)	30	46	27	31	27	30	36	32	20	14	39	73
A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

1 - Durlo (VR)
2 - Friuli2 (UD)
3 - Polverina (MC)
4 - Sabina (RI)
5 - Tardiano (SA)
6 - S. Arsenio (SA)
7 - Bivona (AG)
8 - Anapo (SR)
9 - Acquasanta (AP)
10 - Feltria (PU)
11 - Sibillini (MC)
12 - Marche

minimum genetic distance (.015) with *S. Arsenio* and *Polverina*, and the maximum with *Sabina. Feltria* too presented the minimum distance value with *S. Arsenio*, while higher values were observed with *Sabina*, *Anapo* and *Acquasanta. Acquasanta* showed the minimum genetic distance with *Sibillini*, *Durlo* and *Polverina*, while the maximum values were with *Bivona* and *Tardiano*.

The cluster analysis showed that a well-defined geographical subdivision does not exist between groups. The cluster reported three main branches. The first was subdivided into three parts: one included the two northern provenances *Durlo* and *Friuli*, another two southern materials from *Tardiano* and *S. Arsenio* and the the Sicilian one of *Bivona*. In the second branch there were two Marche groups, *Feltria* and *Sibillini*, while in the third there were *Polverina* (Marche) and *Anapo* (Sicily). Each of the two populations of central Italy, *Sabina* and *Acquasanta*, formed a single cluster (figure 4a).

The pooled population (83 trees) of the Marche Region

All trees selected as a basic population were pooled to form one single population named *Marche* and were then compared with the 8 Italian provenances.

In the *cluster* of figure 4b, *Marche* occupied a position similar to *Feltria* and *Sibillini* in the *cluster* of figure 4a. *Marche* presented the maximum value

Table 2 - *Genetic variation to analysed common* loci *in Italian populations (used as reference system), Marche and Marche in all. Standards errors in brackets.*
Principali parametri genetici stimati per i *loci* comuni alle popolazioni italiane (sistema di riferimento) alle popolazioni delle Marche e alla popolazione Marche nel suo complesso. Tra parentesi l'errore standard.

Population	N/l	N	$P_{5\%}$	H_o	H_e	F
1. Durlo (VR)	30.0 (.0)	2.0 (.3)	62.5	.321 (.099)	.319 (.092)	-.006
2. Friuli 2 (UD)	42.8 (3.3)	2.0 (.3)	75.0	.313 (.087)	.338 (.091)	.074
3. Polverina (MC)	26.9 (.1)	2.0 (.3)	62.5	.280 (.082)	.305 (.089)	.082
4. Sabina (RI)	31.0 (.0)	2.0 (.3)	75.0	.306 (.074)	.346 (.088)	.116
5. Tardiano (SA)	26.9 (.1)	2.0 (.3)	75.0	.339 (.084)	.331 (.081)	-.024
6. S. Arsenio (SA)	30.0 (.0)	2.0 (.3)	75.0	.338 (.085)	.356 (.082)	.051
7. Bivona (AG)	35.5 (.3)	1.9 (.2)	62.5	.315 (.092)	.327 (.089)	.037
8. Anapo (SR)	31.9 (.1)	2.0 (.3)	75.0	.290 (.081)	.311 (.087)	.068
Mean value		*1.99 (.01)*	*70.3*	*.313 (.007)*	*.329 (.006)*	*.050*
9. Acquasanta (AP)	17.8 (1.8)	2.0 (.3)	75.0	.248 (.064)	.359 (.083)	.309
10. Feltria (PU)	13.9 (.1)	2.0 (.3)	75.0	.307 (.079)	.375 (.083)	.181
11. Sibillini (MC)	38.9 (.1)	2.0 (.3)	75.0	.261 (.071)	.361 (.083)	.277
12. Marche pooled	*70.5 (2.1)*	*2.0 (.3)*	*75.0*	*.267 (.069)*	*.365 (.083)*	*.268*

N/L: Mean sample size *per locus*
n: Mean number of alleles *per locus*
$P_{5\%}$: Percentage of polymorphic *loci* (a *locus* is considered polymorphic if the frequency of most common allele is lower to 95%)
H_o: Mean observed eterozygosities
H_e: Mean expected eterozygosities (based on Nei, 1978)
F: Mean fixation index *per* population F=1-(H_o/H_e)

of genetic distance with Sabina and the minimum with *S. Arsenio* and *Polverina*. The latter was not included in the regional data set, but was considered as a standard reference population together with the other Italian sets.

As expected, *Marche* showed a high statistical value F (.268) (table 2) and, therefore, it strayed from the Hardy-Weinberg equilibrium owing to an excess of homozygosity, significant for *loci* Skdh-1, Skdh-2, Dia-1, Pgm-1 and 6Pgd-2 (table 3). The three mean F statistics also remained positive for *Marche*, but they were relatively smaller (table 4b). The F_{IT} value (.102) suggested that the lack of heterozygous individuals was common to all the Italian populations analysed, while the F_{IS} value (.062) indicated that nearly 94% of the diversity observed depended on the genetic diversity within populations. The F_{ST} value (.042) confirmed this data because it showed that 95.8% of the total variability observed was common to all populations (.008 ‹F_{ST} ‹ .064).

The selected population and the released part

To conclude the study, a genetic diversity analysis was made within the *Marche* sample, by selecting the best phenotypes (*Population 1*) as an improvement population for the next phases and by dividing them from the others (*Population 2*). For these two groups were analysed a number of *loci* (16) higher then in the previously described analysis. 5 of these *loci* were monomorphic (Got-1, Dia-2, Pgi-1, 6Pgdh-1 and Mdh-1), while 11 were polymorphic (Got-2, Got-3, Idh, Skdh-1, Skdh-2, Pgi-2, Dia-1, Dia-3, Pgm-1, 6Pgd-2 and Mdh-2).

Table 3 - *Mean fixation index (F) and probability of deviation from Hardy-Weinberg equilibrium estimated using x^2 test between observed and expected frequencies at polymorphic* loci *(*P≤ 0.05; **P≤ 0.01).*
Indice medio di fissazione (F) e significatività della deviazione dall'equilibrio di Hardy-Weinberg stimata con il test del x^2 fra le frequenze alleliche osservate ed attese ai *loci* polimorfici (* P≤ 0.05; **P ≤ 0.01).

	Population											
Locus	**Durlo**	**Friuli 2**	**Polverina**	**Sabina**	**Tardiano**	**San Arsenio**	**Bivona**	**Anapo**	**Acquasanta**	**Feltria**	**Sibillini**	**Marche**
SKDH-1	.139	-.158	-.031	.073	.047	-.058	-.182	-.036	.367	.509 *	.479 **	.460 **
SKDH-2	1.000 **	.640 **	-.038	.005	.059	-.042	-.043	-.067	.457 *	.067	.567 **	.428 **
DIA-1	-.250	.042	-.013	.217	-.259	-.205	-.337	-.208	.298	.125	.231	.232 *
DIA-3	-.156	-.240	-.013	-.216	-.200	.461 **	.424 **	.397 *	-.099	-.292	-.166	-.166
PGM-1	.020	.180	.298 **	.238	.074	-.094	.055	-.017	.404 **	.336	.257	.323 **
6PGD-2	-.021	.175	.145	.108 *	-.016	.163	.141	.043	.286 *	.187	.307	.289 **

Table 4 - *Mean F-statistic at all* loci. *4a) mean F-statistic (for common* loci*) for Italian populations and for Marche populations subdivided in* Acquasanta, Sibillini *and* Feltria; *4b) mean F-statistic for Italian populations and* Marche *population (for common* loci*); 4c) mean F-statistic for sample selected in Marche Region subdivided in* Popolazione 1 *(chosen phenotypes) and* Popolazione 2 *(excluded phenotypes).*
F-statistici medi calcolati per tutti i *loci*. 4a)F-statistici medi (stimati per i *loci* comuni) nelle popolazioni italiane e nelle popolazioni delle Marche suddivise in Acquasanta, Sibillini e Feltria;4b) F- statistici medi (stimati per i *loci* comuni) nelle popolazioni italiane e nella popolazione *Marche*;4c) F-statistici medi per i campioni selezionati nella Regione Marche suddivisi in *Popolazione 1* (fenotipi scelti) e *Popolazione 2* (fenotipi esclusi).

4 a

Locus	$F_{(IS)}$	$F_{(IT)}$	$F_{(ST)}$
Skdh-1	.108	.119	.012
Skdh-2	.215	.281	.084
Dia-1	-.029	.032	.059
Dia-3	.012	.057	.046
Pgm-1	.148	.205	.067
6Pgdh-2	.137	.193	.066
Mean	**.092**	**.141**	**.054**

4 b

Locus	$F_{(IS)}$	$F_{(IT)}$	$F_{(ST)}$
Skdh-1	.033	.041	.008
Skdh-2	.177	.229	.063
Dia-1	-.087	-.017	.064
Dia-3	.066	.109	.046
Pgm-1	.112	.167	.062
6Pgdh-2	.119	.138	.022
Mean	**.062**	**.102**	**.042**

4c

Locus	$F_{(IS)}$	$F_{(IT)}$	$F_{(ST)}$
Got-2	.753	.764	.042
Got-3	.444	.449	.008
Idh	.320	.339	.028
Skdh-1	.461	.466	.009
Pgi-2	.501	.520	.040
Dia-1	.230	.232	.002
Dia-3	-.178	-.168	.008
Pgm-1	.330	.342	.018
6Pgdh2	.305	.359	.078
Mdh-2	.307	.321	.021
Skdh-2	.382	.388	.010
Mean	**.339**	**.357**	**.027**

Table 5 - *Genetic differentiation (assessed for common* loci*) in Italian and Marche populations.*
Differenziazione genetica (stimata per i *loci* comuni) nelle popolazioni italiane ed in quelle delle Marche.

Locus	H_S	H_T	D_{ST}	G_{ST}
SKDH1	0.474	0.470	-0.004	-0.008
SKDH2	0.240	0.257	0.018	0.069
DIA-1	0.475	0.499	0.024	0.047
DIA-3	0.431	0.445	0.014	0.032
PGM1	0.493	0.521	0.028	0.053
6PGD-2	0.605	0.633	0.028	0.043
GOT-1	0.000	0.000	0.000	0.000
PGI-1	0.000	0.000	0.000	0.000
Mean	**0.340**	**0.353**	**0.013**	**0.038**

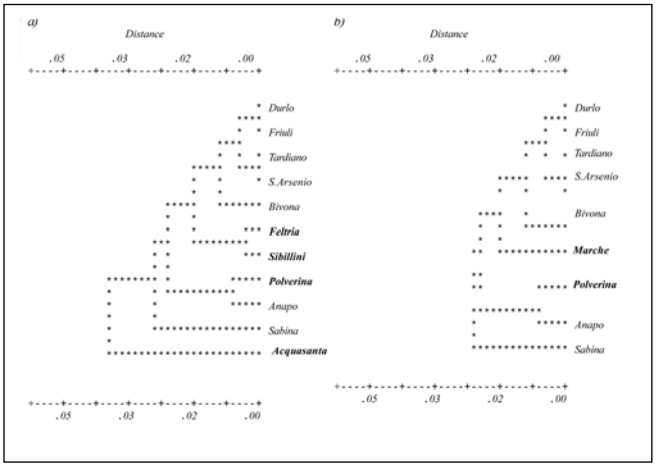

Figure 4 - *a) UPGMA dendrogram based on similarity matrix (Nei 1978). The* Marche *material was subdivided in 3 populations (*Sibillini, Feltria *and* Acquasanta*). b) UPGMA dendrogram based on similarity matrix (Nei 1978) for Italian populations, used as comparison, and* Marche *pooled population.*
a) Dendrogramma UPGMA stimato sulla matrice di similarità (Nei 1978). I campioni delle Marche sono stati suddivisi in 3 popolazioni:Sibillini, Feltria e Acquasanta. b) Dendrogramma UPGMA stimato sulla matrice di similarità (Nei 1978) per le popolazioni italiane (sistema di riferimento) e per la popolazione *Marche* nel suo complesso.

Allele frequencies were reported in table 6.

For both groups, the mean number of allele *per locus* remained 2, while the polymorphic *loci* percentage dropped to 68.8%. The Wright's index F reached .377 for *Population 1* and .327 for *Population 2* (table 7). As shown in table 8, the excess of homozygous trees in *Population 1* was not significant only for Dia-1, Dia-3, Pgm-1 *loci*, while in *Population 2* it was not significant for Idh, Dia-1, Dia-3, 6Pgd-2, Mdh-2.

The F_{IS} (.339) and F_{ST} (.027) values showed that, also in this case, genetic variability depended mainly on differences between single individuals and not between the two groups (table 4c). Most of the *loci* tended to be fixed because the F_{IT} values (table 4c) varied between .764 (Got-2) and .232 (Dia-1). On the contrary *locus* Dia-3 showed negative value (-.168).

The Common Principal Component Analysis showed the existence of a low degree of variability between the individuals selected in Marche. In fact 6 components were necessary to explain about 50% of the total variance (table 9) and the first 9 components were needed to raise it to 66%.

The alleles with higher discriminant value, in order of importance, were: 6Pgd-2c, Got-2a, Got-2b, Idh-1b, Got-2c, Idh-1a, 6Pgd-2b and Skdh-2c. The components

Table 6 - *Estimated allele frequencies to examined* loci *in Marche in all, Marche populations, selected (*Population 1*) and not selected (*Population 2*) phenotypes.*
Frequenze alleliche stimate per tutti i *loci* nella popolazione *Marche* nel suo complesso, nelle tre popolazioni marchigiane, nella *Popolazione 1* (fenotipi selezionati) e nella *Popolazione 2* (fenotipi non selezionati).

	Population					
Locus	**Marche**	**Acquasanta**	**Feltria**	**Sibillini**	**Population 1**	**Population 2**
GOT-1						
(N)	73	20	14	38	37	36
Monomorphic	1.000	1.000	1.000	1.000	1.000	1.000
GOT-2						
(N)	73	20	14	38	37	36
A	.130	.000	.214	.171	.081	.181
B	.774	.850	.571	.803	.784	.764
C	.096	.150	.214	.026	.135	.056
GOT-3						
(N)	73	20	14	38	37	36
A	.137	.200	.143	.105	.162	.111
B	.863	.800	.857	.895	.838	.889
IDH						
(N)	73	20	14	38	37	36
A	.123	.075	.286	.092	.135	.111
B	.767	.775	.607	.829	.716	.819
C	.110	.150	.107	.079	.149	.069
SKDH-1						
(N)	72	19	14	38	36	36
A	.424	.474	.321	.434	.444	.403
B	.576	.526	.679	.566	.556	.597
SKDH-2						
(N)	72	19	14	38	36	36
A	.264	.263	.357	.237	.306	.222
B	.736	.737	.643	.763	.694	.778
PGI-1						
(N)	73	20	14	38	37	36
Monomorphic	1.000	1.000	1.000	1.000	1.000	1.000
PGI-2						
(N)	73	20	14	38	37	36
A	.219	.300	.429	.105	.284	.153
B	.644	.575	.429	.750	.622	.667
C	.137	.125	.143	.145	.095	.181
DIA-1						
(N)	73	20	14	38	37	36
A	.479	.475	.429	.500	.486	.472
B	.521	.525	.571	.500	.514	.528
DIA-2						
(N)	73	20	14	38	37	36
Monomorphic	1.000	1.000	1.000	1.000	1.000	1.000
DIA-3						
(N)	73	20	14	38	37	36
A	.596	.650	.536	.592	.649	.542
B	.404	.350	.464	.408	.351	.458
PGM-1						
(N)	72	19	14	38	36	36
A	.118	.079	.107	.145	.083	.153
B	.729	.789	.607	.750	.722	.736
C	.153	.132	.286	.105	.194	.111
6PGD1						
(N)	56	5	13	37	22	34
A	.268	.200	.154	.324	.273	.265
B	.527	.200	.692	.500	.477	.559
C	.205	.600	.154	.176	.250	.176
6PGD2						
(N)	56	5	13	37	22	34
Monomorphic	1.000	1.000	1.000	1.000	1.000	1.000
MDH-1						
(N)	73	20	14	38	37	36
Monomorphic	1.000	1.000	1.000	1.000	1.000	1.000
MDH-2						
(N)	73	20	14	38	37	36
A	.192	.275	.143	.158	.257	.125
B	.808	.725	.857	.842	.743	.875

Table 7 - *Genetic variation to analysed* loci *in Marche in all, Marche populations and selected (*Population 1*) and not selected (*Population 2*) phenotypes. Standards errors in brackets.*
Variazione genetica per i *loci* analizzati nella popolazione *Marche* nel suo complesso, nelle popolazioni marchigiane, nella *Popolazione 1* (fenotipi selezionati) e nella *Popolazione 2* (fenotipi non selezionati). In parentesi l'errore standard.

Population	N/l	N	$P_{5\%}$	H_o	H_e	F
1. Acquasanta	17.9	1.9	68.8	.184	.302	.391
	(1.3)	(.2)		(.043)	(.057)	
2. Feltria	13.9	2.0	68.8	.220	.332	.337
	(.1)	(.2)		(.050)	(.063)	
3. Sibillini	37.9	2.0	68.8	.183	.275	.335
	(.1)	(.2)		(.042)	(.054)	
Mean value		1.97	68.8	.196	.310	.369
		(.03)		(.012)	(.011)	
4. Marche pooled	70.7	2.0	68.8	.193	.297	.350
	(1.4)	(.2)		(.043)	(.056)	
5. Population 1 ("Selected trees")	34.9	2.0	68.8	.195	.313	.377
	(1.3)	(.2)		(.042)	(.058)	
6. Population 2 ("Non selected trees")	35.8	2.0	68.8	.189	.281	.327
	(.2)	(.2)		(.045)	(.055)	

N/L: Mean sample size per *locus*
n: Mean number of alleles per *locus*
$P_{5\%}$: Percentage of polymorphic *loci* (a *locus* is considered polymorphic if the frequency of most common allele is lower to 95%)
H_o: Mean observed eterozygosities
H_e: Mean expected eterozygosities (based on Nei, 1978)
F: Mean fixation index per population $F=1-(H_o/H_e)$

Table 8 - *Mean fixation index (F) and probability of deviation from Hardy-Weinberg equilibrium estimated using x^2 test between observed and expected allele frequencies at polymorphic* loci *for* Popolazione 1 *(chosen phenotypes) and* Popolazione 2 *(excluded phenotypes. *P≤ 0.05; **P≤ 0.01).*
Indice medio di fissazione e significatività della deviazione dall'equilibrio di Hardy-Weinberg, stimato tramite il test del x^2 sulle frequenze alleliche osservate e attese ai *loci* polimorfici, nella *Popolazione 1* (fenotipi scelti) e nella *Popolazione 2* (fenotipi non scelti.*P ≤ 0.05; **P ≤ 0.01)

Locus	Population 1	Population 2
SKDH-1	.437 **	.480 **
SKDH-2	.345 **	.518 **
DIA-1	.243	.220
DIA-3	-.186	-.175
PGM-1	.167	.474 **
6PGDH-2	.499 **	.198
GOT-2	.700 **	.781 **
GOT-3	.403 **	.437 **
IDH	.395 **	.197
PGI-2	.536 **	.500 **
MDH-2	.504 **	-.143

had correlation coefficients between .69 and .50. Got-2b, 6Pgd-2c and Got-2c had a negative correlation, while all the others were positive.

The cluster analysis carried out on all the Marche genotype set (*Tree procedure* in NTSYS, Rholf *op. cit.*), presented 4 main clusters. Most of the phenotypes selected for the progeny test were grouped in the first two. The small sub-population of *Acquasanta*, *Faete*, was distributed throughout all 4 clusters, which showed that this population had some variability even though the number of trees was limited. Those of *Camerino* and *Feltria* were concentrated in the third and fourth clusters, while those of *Sibillini* and *Acquasanta* were found above all in the first and second.

This distribution is confirmed in figure 5. It shows the dispersion of genotypes based on the matrix of genotype similarity, obtained using Common Principal Component Analysis. For the first two components, most of the selected trees tended to be grouped in the first and second quadrant. Those belonging to the popolation *Sibillini* were more homogeneous and therefore less dispersed and more concentrated in the fourth quadrant. The small sub-population of *Faete*, very nice from the phenotypic point of view, had a certain level of variability. The few trees were dispersed in at least three quadrants.

Discussion and conclusions

As already reported in literature (Malvolti *et al. op. cit.*, Fornari *et al.* 1999), the genetic variability of walnut in Italy is relatively reduced and its structure is quite complex to explain because this species is

Table 9 - *Common component analysis (estimated on matrix of 27 allele of 11 polymorphic* loci*) for the sample selected in Marche.*
Analisi delle componenti principali (stimata sulle matrici di 27 alleli di 11 *loci* polimorfici) per i fenotipi selezionati nelle Marche.

Component	Autovalue	Simple %	Cumulate	Allele	Correlation Coefficient
1	2.963	10.975	10.975	6pgd-2c	0.66
2	2.613	9.677	20.652	got-2a	-0.62
3	2.435	9.017	29.670	got-2b	-0.69
4	2.135	7.907	37.576	idh-1b	0.54
5	1.853	6.863	44.439	got-2c	0.53
6	1.638	6.047	50.486	idh-1a	0.53
7	1.619	5.996	56.482	6pgd-2b	0.53
8	1.346	4.984	61.466	got-2c	-0.51
9	1.237	4.581	66.047	skdh-2c	0.5

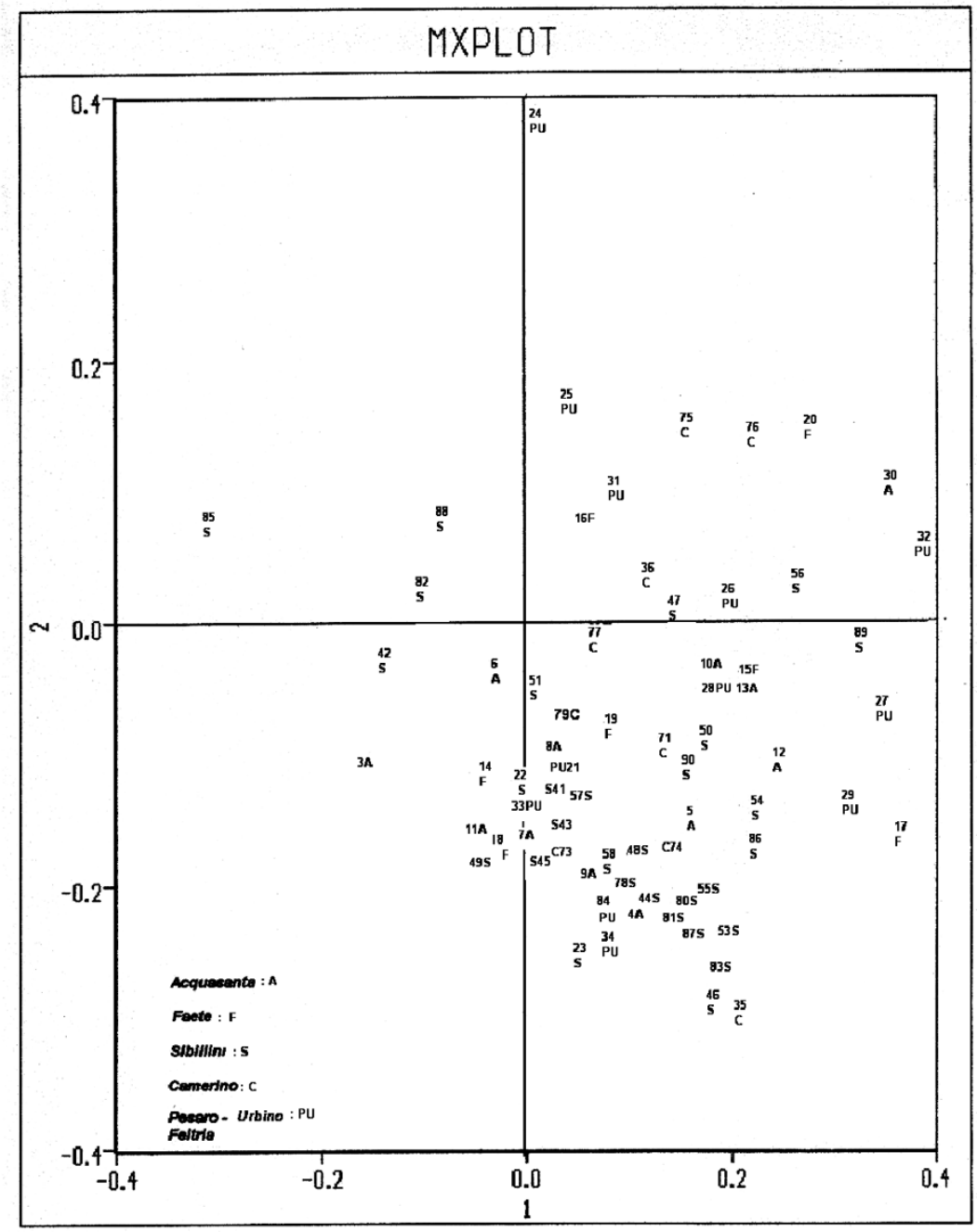

Figure 5 - *Analysis of Correspondences: distribution of enzyme genotypes of walnuts from Marche according to the first two functions (Rholf 1994).*
Analisi delle corrispondenze: distribuzione dei genotipi enzimatici dei noci selezionati nelle Marche in base alle prime due funzioni (Rholf 1994).

hardly synanthropic.

The material studied in Marche showed a genetic structure relatively similar to the other populations, with lower genetic differentiation. This situation is probably accounted by the species' high degree of dispersal in this region, by the human influence on its distribution and reproduction, by its geographical isolation in some areas and by its origin from a few mother trees, single or from fruit varieties.

In general, some parameters of the genetic variability, such as the mean number of alleles *per locus* and the percentage of polymorphic *loci*, were similar to the data obtained by Malvolti *et al.* (*op. cit.*, Fornari *et al. op. cit.*), either for the selected trees, either within a single populations, or provenances were considered.

The mean number of alleles *per locus*, was lower than the mean number for other temperate-zone hardwood trees, but was similar to the European (2.4) mean value, supplied after Fornari *et al.* (*op. cit.*) for this species. The greater uniformity within the selected group of trees was shown, by the ratio between the observed heterozygosity (H_o) and the expected heterozygosity (H_e). For all the groups examined, the first figure was smaller than the second. In general, the statistic index F for *Acquasanta*, *Sibillini*, *Feltria* and the pooled population of *Marche* was higher than the other reference populations. As a result, the deviation from the Hardy-Weinberg equilibrium showed excess of homozygotes.

A number of factors could be responsible for this result. *Acquasanta*, *Sibillini* and *Feltria* had a smaller number of fertile individuals compared to the other populations, and the population could be affected by inbreeding. It can be supposed that only some genotypes might have been reproduced, the most adapted to face Adriatic climate which is strongly influenced by the continental climate of Balkans.

Moreover, in order to collect the greatest number of specimen in the reference Italian populations, the sampling was carried regardless to their forest habit: "fruit crown form" trees and "forest form" trees were sampled. This could have improved the genetic variation.

What can be safely said is that, over the centuries, as a result of the hard pressure exerted by the human activity on this species, the original gene pool has been greatly eroded, thus leading to the fixation of particular characteristics and therefore also to the fixation at random of some of the *loci* studied.

Walnut is typically dispersed in the Marche territory by small, geographically-isolated populations or individual trees growing under climatic conditions not always suitable for cross pollination of this allogamic species. This situation reduces presumably pollen exchanges. Because those factor, the homozygosity percentage slowly rises while the degree of polymorphism becomes reduced. The 3 provenances from Marche meet these characteristics. Indeed, *Acquasanta* includes phenotypes of two "micro-populations", *Faete* (AP) and *Acquasanta* (AP), two isolated valley-floor areas. *Sibillini*, on the contrary, is greater, but the trees are mostly isolated and dispersed within a relatively wide area near the Sibillini Range. *Feltria*, with the smallest mean number of individuals (13.9), showed a lower F index. This was probably due to its trees spread over an area without significant geographic barriers and where the milder environmental conditions might have favoured cross pollination and exchanges of reproductive materials between the rural communities.

A similar situation has also been observed for chestnut. Like walnut, this species did not originate in Europe but gradually became naturalised (Villani *et al.* 1991) and, as a consequence, has low genetic variability.

The genetic information of a few walnut genotypes survived probably the last Ice Age in some sheltered areas of the Italian peninsula (Huntley and Birks *op. cit.*). However, the low values of the genetic distance between the Italian populations indicate that even initially only a relatively small gene pool existed. The central position of the Marche populations in the *cluster* could be explained in part by the exchanges of walnut propagation material between northern and southern human populations.

The higher genetic distance between *Sabina* and the provenances geographically nearer of Marche, could derive from the isolation determined by narrow valleys. Moreover these two areas are separated by the Apennine watershed with altitudes above 1500 m. Both those factors could have been prevented natural migration as well as propagation material flows. In addition, walnut trees were comparatively scarce and the only substantial possibility of migration lay in seed exchange between rural communities, which in these areas was very limited until the 20th century. These factors may explain the remarkable differences between the material found on the two sides of Apennines.

Although the genetic variability of walnut in Italy is low, examination of the allozyme variation using Correspondence Analysis and Common Principal Component Analysis showed a high individual component within the material from Marche. In order to reduce the erosion of the genetic resources still existent in this area, a programme of *in situ* and *ex situ* conservation is necessary .

Unpublished data about adaptive and phenotypic characteristics of walnuts showed the same structure of the genetic traits. Genetic improvement can be undertaken following individual phenotypic selection without considering preliminary provenance tests.

For this reason, a multi-site comparative progeny

trial was set up in Marche region, Tolentino (MC) and Amandola (AP), and Vitiano near Arezzo, Tuscany, in order to test individual basic material to be possibly used to establish seed orchards.

The basic material sampled was divided into two wide seed harvesting areas: one more northern, within the Pesaro-Urbino province, and one southern, including the Sibillini Range. These two areas are not very distant genetically, but are situated, respectively, in medium-sized hills and in higher hills or mountains. Moreover, the northernmost area is characterised by higher heterozygosity levels, making this area maybe the most suitable for seed collections and for its possibly greater plasticity.

These seed harvesting areas should be kept separated for future genetic resource management and for seed supplies. In additions, there are other two smaller areas: *Polverina* and *Arquata-Faete.* Both these populations are clearly distinct from the nearby and larger population of Sibillini. Of them, *Polverina* is the most suitable as basic material for breeding, while *Arquata-Faete* is more suitable for local use. *Arquata-Faete* includes two subpopulations with common origin and were left isolated because of the rugged orography. Their isolation made them well-adapted to the local environment, on the one hand, as shown by the high phenotypic quality of trees, but from the genetic point of view it also produced strong genetic erosion by the fixation of some alleles and the loss of others. For this reason we suggest that seeds should be collected from the largest possible number of trees in order to ensure the greatest possible genetic diversity.

Compared to other populations, *Faete* is composed of a small number of plants, but it has a very high level of individual diversity which reduces the risk of genetic poverty. This area, therefore, could be used as a small seed harvesting population.

In any case, no matter what the provenance or where the population, seed collection for regional nurseries should be picked from at least 10 trees in order to guarantee seed mixture. The trees should be different every year. Only in this way it will be possible to supply seeds from selected basic material but preserving a minimum variation. 10 trees is really lower than would be needed with usual methods, but the particular characteristics of walnut distribution have to be considered.

The basic material needed for nursery supplies could be given the following trade classifications:

- *"Identified at source"*: subdivided into *seed source* (when there are at least 10 plants) or populations-*collection areas* (when there are at least 30 plants);
- *"Selected"*: referring to *population-collection areas* where a population is selected with higher-than-average phenotypic characteristics;
- *"Qualified"*: referring to *parent* trees, where an individual phenotypic selection is being carried out, as in the case of *Population 1* and preliminary information about performances can be obtained from tests.

Preliminary genetic analysis, using cheap and rapid methods to assess variation, reduces the risk of starting experimental tests without information or estimates about the genetic structure of *Juglans regia* in the examined region. The higher individual component of variation detected can be used by foresters to avoid preliminary provenance tests and to pass directly to phenotypic selection and to establish progeny tests.

On the one hand, this reduces the period for making improvement. On the other hand, however, this approach can be dangerous when only a few trees/genotypes are used for seed collection: it can contribute dramatically to the genetic erosion of the species in that area. For this reason seeds should be supplied from a minimum number of selected trees or from whole populations even if they are small. This is the case, in example, of the small *Faete* population, where phenotypes are very good in general and the within variation is also higher than other.

The higher level of homozygosity detected in general can be a real sign of erosion and of adaptation to the local micro-conditions. The use of this kind of material should be limited to districts showing similar conditions to the site of origin, since the environment pushed the material to be specialised.

These aspects might be considered too prudent or too complicated for planning field works but, to ensure that all *a priori* choices are well made, the high revenue which walnuts can earn also needs to be taken into account.

Results issued after the phenology selection showed the low impact of improvement on genetic pattners of the basic population. If this seems to be true for neutral markers, independent from environment traits, it should be considered the effect on the genetic component of adaptive traits, as phenology

or stress tolerance/resistance factors of phenotypes. That will be the deal of future researches (Vendramin and Morgante 2005).

Aknowledgements

A special thank is due to the State Forest Service (CFS) operating in the region Marche wich made possible the accurate exploration and the inventory of walnut resources and assited us in the field surveyes.

Another special thank is due to the two anonymous referees who reviewed this paper for their suggestions and useful comments.

References

Aletà N., Rovira M., Ninot A., Arùs P., 1993 - *Inheritance of four isozymes in walnut.* Acta Horticulturae 311: 62–67.

Ashton G. C., Braden W. H., 1961 - *Serum globulin polymorphism in mice.* J. Biol. Sci. 14: 248–253.

Arulsekar S., Parfitt D. E., McGranahan G., 1985 - *Isozyme gene markers in* Juglans *species. Inheritance of GPI and AAT in* Juglans regia *L. and* Juglans hindsii. J. of Heredity 76: 103–106.

Arulsekar S., Parfitt D. E., McGranahan G., 1986 - *Inheritance of phosphoglucomutase and esterase isozymes in Persian walnut.* J. of Heredity 77: 220–221.

Cavalli – Sforza L. L., Menozzi P., Piazza A., 1994 - *Storia e geografia dei geni umani.* Adelphi edizioni: 791 p.

Colinvaux P., 1986 - *Ecology.* New York. John Wiley & Sons.: 134 p.

Colletti L., 2001. *Risultati dell'applicazione del regolamento CEE 2080/92 in Italia.* Sherwood, Foreste ed Alberi Oggi, 70: 23–31.

Conkle M. T., Hodgkiss P. D., Nunnally L. B., Hunter S. C., 1982 - *Starch gel electrophoresis of Conifer seeds: a laboratory manual.* Pacif. SW Forest and Range Experiment Station, Berkley USDA – FS Gen. Techn. Rep. PSW 64: 18 p.

Ducci F., 1989 - *Noce e ciliegio da legno: esame della produzione vivaistica nell'Appennino centro settentrionale.* Ann. Ist. Sperim. Selvic., Arezzo XVIII (1987): 177–213.

Ducci, F., De Rogatis A., Proietti R., 1997 - *Protezione delle risorse genetiche di* Juglans regia *L.*. Ann. Ist. Sperim. Selvic., Arezzo XXV and XXVI (1994-1995): 35–55.

Ducci F., Proietti R., Calvo E., Correale F., 2005. *2.5 La valutazione di cloni e provenienze.* In: Monografia sul ciliegio selvatico (*Prunus avium* L.), a cura di F. Ducci, CRA - Ist. Sper. Selv., Arezzo (ISBN 88-901923-0-5): 45 – 52.

Ducci F., Veracini A., 1992 - *Criteri di scelta e sistema di valutazione di fenotipi superiori nel miglioramento genetico di latifoglie a legname pregiato.* Ann. Ist. Sperim. Selv., Arezzo XXI (1990): 57–80.

Ducci F., Germani A., Janin G., Proietti R., Signorini G.– 2006 – *Clone selection for wild cherry (*Prunus avium *L.) with special reference to some traits used.* Bozzano M., Rusanen M., Rotach P., Koskela J. (compilers). Euforgen – Noble Hardwoods Network, Report of VI (9-11 June 2002, Alter do Chao, Portugal) and VII Meeting (22-24 April 2004, Arezzo, Italy). International Plant Genetic Resources Institute, Rome, Italy: 53-60.

El Kassaby Y. A.,1991 - *Genetic variation within and among conifer population: review and evaluation of methods.* In: Biochemical markers in the population genetics of forest trees, Fineschi S., Malvolti M. E., Cannata F., Hattemer H. H. Ed., SPB Academic Publishing, The Hague, The Netherlands: 61–65.

Esopus, 1978 - *Favole.* Ed., Rizzoli, Bibl. Universale Rizzoli, Milano: 179 p.

Farris J. S., 1972 - *Estimating phylogenetic trees from distance matrices.* Am. Nat., 106: 645-688.

Fornari B., Cannata F., Spada M., Malvolti M. E., 1999 - *Allozyme analysis of genetic diversity and differentiation in European and Asiatic walnut (*Juglans regia *L.) populations.* Forest Genetics 6 (2): 115-127.

Futuyma D. J., 1979. *Evolutonary Biology.* Sunderland, Mass. Sinauer Associates, Inc.: 156 p.

Goudet J., 2001 - *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)*

Hartl D.L., Clark A.G., 1989 - *Principles of population genetics.* Sinauer, Sunderland, MA.

Hemery G. E., 2000 - Juglans regia *L.: genetic variation and provenance performance.* Dep. of Plant Science, Univ. of Oxford: 176 p. (Unpublished D Phil. Thesis).

Huntley B., Birks H. J. B., 1983 - *An Atlas of past and present pollen maps for europe: 0-3000 years ago.* Cambridge University Press. N. Y.: 238 – 242.

Kim Z.S., 1979 - *Inheritance of leucine aminopeptidase and acid phosphatase in beech (*Fagus sylvatica *L.)*. Silvae genetica 28: 68–71.

Kim Z.S., 1980 - *Veränderung der genetischen struktur von buchenpopulationen durch viabilitätsselektion in keimlingstadium.* Göttingen Res. Notes in For. Genetics, 3 (Phd thesis).

Leslie C. A., McGranahan G., 1988 - *Native populations of* Juglans regia *L.. A draft.* In: Proceedings of the International Conference on Walnut, Yalova, Turkey, September, 19 – 23, 111–124.

Lynne Corn M.. 1992. The Listing of a of Species: Legal Definition and Biological Realities, 92-944 ENR, CRS Report for the Congress. National Library for the Environment, USA, (http://www.cnie.org/NLE/CRSreports/biodiversity/biodv-10.cfm#1).

Malvolti M. E., Paciucci M., Fineschi S., Cannata F., 1993 - *Genetic variation in Italian populations of* Juglans regia *L.*. Acta Horticulturae 311: 86–94.

Malvolti M. E., Fineschi S., Pigliucci M., 1994 - *Morphological Integration and Genetic Variability in* Juglans regia *L.*. Journal of Heredity 85: 389–394.

Malvolti M. E., Fineschi S., Morgante M., Vendramin G., 1995a - *Mating system of a naturalized* Juglans regia *L. population in Italy.* Population genetics and genetic conservation of forest trees. Baradat Ph., Adams W. T. & Müller–Starch G. Ed., SPB Academic Publishing, Amsterdam, The Netherlands, 305–308.

Malvolti M. E., Spada M., Beritognolo I., Cannata F., 1995b - *Differentiation of* Juglans regia *L. european stands and italian cultivars.* In: Proc. Workshop of E. U. Air/Walnut Project: European development of walnut trees for wood and fruit production as an alternative and extensive system to agricultural crops. Groups C and D. Thessaloniki, Greece, March 27-30 , 1995: 34–51.

Malvolti M. E., Beritognolo I., Spada M., 1996 - *Diversità genetica in* Juglans regia *L.: valutazione delle risorse genetiche in Europa per uno sviluppo agricolo e forestale sostenibile.* Sherwood – Foreste e Alberi Oggi 15: 11–17.

Malvolti M. E., Beritognolo I., Spada M., Cannata F, 1997 - *Ricerche sulle risorse genetiche e sulla biologia riproduttiva di* Juglans regia *L. in Italia mediante marcatori molecolari.* Ann. Ist. Sperim. Selv. Arezzo XXV – XXVI (1994-1995): 9–34.

Nei M., 1978 - *Estimation of average heterozygosity and genetic distance from a small number of individuals.* Genetics 89: 583–590.

Nei M., 1987 - *Molecular evolutionary genetics.* Columbia University Press, NY, USA

Raymond M., Rousset F., 1995 - *An exact test for population differentiation.* Evolution 49: 1280–1283.

Rholf F. J., 1994 - NTSYS-pc. *Numerical taxonomy and Multivariate Analys System, vers. 1.80.* Exeter Software–Applied Biostatistic Inc., N. Y..

Sneath P.H.A., Sokal R. R., 1973 - *Numerical Taxonomy.* Freeman, S. Francisco (USA): 573 p.

Swofford D.L., Selander R.B., 1989 - *Biosys-1.* Illinois Natural History Survey: 43 p.

Vallejos C. E., 1983 - *Enzyme activity staining.* S. D. Tanksley and T. J. Orton Ed., Isozymes in Plant Genetics and Breeding , part A. Elsevier Science Publishers B. V. Amsterdam: 469–516.

Vendramin G. G., Morgante M., 2005 - *Genetic diversity in forest tree populations and conservation: analysis of neutral and adaptive variation.* Proc. Workshop on: "The role of biotechnology", Villa Gualino, Turin, Italy. 5-7 March, 2005: 129-130.

Villani F., Benedettelli S., Paciucci M., Cherubini M., Pigliucci M., 1991 - *Genetic variation and differentiation between natural populations of chestnut.* Biochemical markers in the population genetics of forest trees. Edited by Fineschi S., Malvolti M.E., Cannata F., Hattemer H. H., SPB Acad. Publishing, bv, the Hague, the Netherlands: 91–103.

Weir B. S., Cockerham C. C., 1984 - *Estimating F-statistic for the analys of population structure.* Evolution 38: 1358–1370.

Wright, S. 1978 *Evolution and genetics of populations.* Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago.

Zobel B., Talbert, J., 1984 - *Applied Forest Tree Improvement.* J. Wiley and Sons Ed., N.Y.: xv + 505 p.