Theor Appl Genet (1996) 92: 832–839

A. El Mousadik · R. J. Petit

High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco

Received: 28 October 1995 / Accepted: 17 November 1995

Abstract Genetic diversity at nine isozyme loci was surveyed in an endangered tree species, the argan tree, endemic to south-western Morocco. The species is highly diverse (3.6 alleles/locus) with populations strongly differentiated from each other (F_{ST}=0.25). This example is used to illustrate a method for standardizing measures of allelic richness in samples of unequal sample sizes, which was developed for the estimation of the number of species and relies on the technique of rarefaction. In addition, it is shown that the measure of subdivision, ρ_{ST}, obtained when allelic richness is used in place of h (Nei's index of diversity), is much larger than the F_{ST} [e.g. $\rho_{ST(40)}$=0.52, where (40) indicates the specified sample used to estimate the allelic richness]. This suggests that rare alleles (which strongly influence measures of allelic richness) have a more scattered distribution than more frequent ones, a result which raises special conservation issues for the argan tree.

Key words Desertification · Genetic diversity · Fragmentation · Genetic conservation · Rarefaction method

Introduction

The argan tree [*Argania spinosa* (L.) Skeels, Sapotaceae] is a multi-purpose tree species endemic to South-Western Morocco where it plays a key role in the local economy (Morton and Voss 1987; Benchekroun and Buttoud 1989). The high-quality oil extracted from its seeds represents a significant contribution in fatty acids to the diet of the local population in the Souss valley. In addition, the leaves and fruits are used as forage for goats and camels and the wood is transformed into charcoal. Individual trees can live up to 250 years (Ehrig 1974). The architecture of the tree varies from a fully erect to a weeping habit. The flowering and fruiting phenologies are also very variable, and the nuts, globular or oval, are produced from early spring to mid summer. The species is now regenerated nearly exclusively by sprouts after clear cutting (Monnier 1965). Indeed, although the seeds can germinate readily, overgrazing generally prevents regeneration by seeds.

Its main range is limited to the south-west of the country, in the Agadir region, north of Oued Draa and south of Oued Tensift (Emberger 1925; Pelletier 1982) (Fig. 1). The area covered by this species (800 000 hectares, Ayad 1989) probably represents less than half of its distribution at the beginning of this century (Monnier 1965). The presence of a small, isolated population in the high valley of Oued Grou near Rabat, and the existence of a larger population growing on the northern slopes of the Beni-Snassen mountains north-east of Oujda near the Mediterranean coast (Maire 1939), could be evidence of this former more-widespread distribution (Tregubov 1963).

The situation of the argan tree forests is becoming very serious. In 1949 Dupin mentioned average densities ranging from 150 to 250 trees per hectare in the Souss plain and 50 trees per hectare in the Anti Atlas. Now the density has shrunk to about 30 trees per hectare and some "forests" have less than ten trees per hectare (Nouaïm et al. 1991). Damage caused by human pressure and by goats and camels has probably been accelerated by climatic hazards.

Since the disappearance of argan trees will often result in a rapid desertification process, its conservation represents a high priority. In particular, given the various climates under which it grows (arid or sub-arid, with coastal or more continental influences), and the variety of edaphic conditions encountered across its range, the description and conservation of its genetical resources seem particularly important. Except for the preliminary study of

Communicated by P. M. A. Tigerstedt

A. El Mousadik
Laboratoire d'Agroforesterie, Faculté des Sciences, Département de Biologie, Université Ibnou Zohr, BP 28S, Agadir, Morocco

R. J. Petit (✉)
Institut National de la Recherche Agronomique, Laboratoire de Génétique et Amélioration des Arbres Forestiers, BP 45, 33611 Gazinet Cedex, France

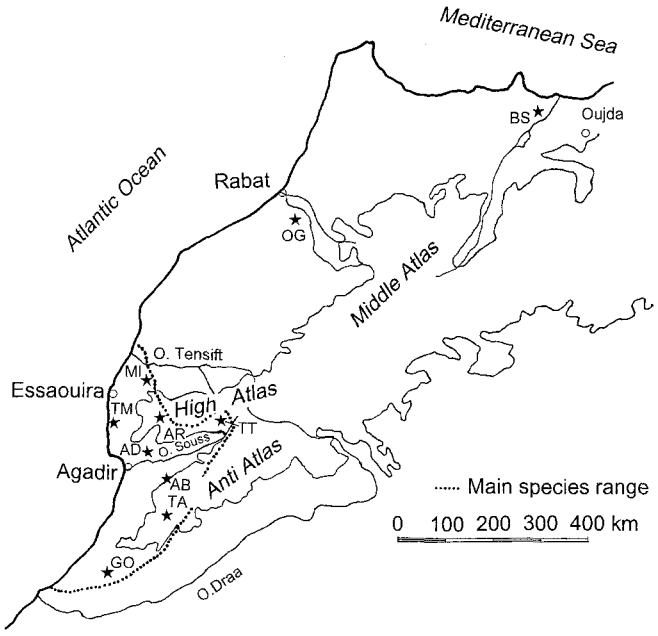

Fig. 1 Map of the location of the populations of argan trees and main range of the species

Msanda (1993), who identified two polymorphic enzyme systems, the present study provides the first data on the genetic diversity of this species across its range, and includes the two marginal populations mentioned above.

In the population genetic literature, very few measures of gene diversity are used. The most successful is the Simpson (1949) index popularized in genetics by Nei (1973). However, this measure emphasizes differences in allelic frequencies rather than in the number of alleles. Since for conservation purposes an emphasis on the number of alleles seems more adequate, measures such as the percentage of polymorphic loci or allele numbers per locus may be more appropriate. However, these measures depend on sample size and can not therefore be compared across studies. Hence, although often reported, measures of allelic richness are little discussed and exploited. Here, we apply the method of rarefaction, previously described in the ecological literature (Hurlbert 1971), to the allelic data of argan tree populations and show that simple, immediately interpretable, and comparable measures of allelic richness can be easily obtained.

Materials and methods

Ten populations were sampled from diverse ecogeographic sites throughout the species range. Among them, two represent the relic populations of Oued Grou and Beni-Snassen (Table 1 and Fig. 1). The sampled area was 1 to 2 hectares, except for the low-density Saharan population of Goulimine where about 100 hectares had to be sampled to have enough trees represented. Seeds were collected from each site as open-pollinated families and germinated. The multilocus genotype of 292 seedlings (one per mother tree) was obtained at nine different polymorphic isozyme loci.

Protein extracts were produced by grinding 200 mg/0.2 ml of young leaves in Tris-HCl (pH 7.6) buffer containing: 0.1 M Tris, 10% glycerol, 1% polyethyleneglycol 20000, 1% polyvinylpyrrolidone and 2.5% 2-mercaptoethanol. After centrifugation at 20000 g for 30 mn at 2°C, the supernatants were stored at –20°C before use.

Electrophoresis was performed on 8.5% polyacrylamide gels at 300 V during 4 h, in a cooled Biorad vertical gel-apparatus. For the *Lap* system, a continuous buffer (Tris Borate EDTA pH 8.3) was used (Peacock and Dingman 1967), while for the four other enzymatic systems, a discontinuous mode was preferred (Ornstein and Davis 1964). The staining procedures were those of Gabriel (1971) for *Est*, Brunel (1979) for *Amy* and *Lap*, and Pasteur et al. (1987) for *Got* and *Adh*. A total of nine putative polymorphic loci were interpreted from the zymograms. These loci are characterised by a classical monomeric or dimeric quaternary structure depending on the enzyme system (Table 2)

F-statistics were computed according to Pons and Chaouche (1995) with a computer program provided by O. Pons. Nei's genetic distances were computed with the GD program of Ritland (1989).

Allelic richness was computed using the rarefaction method as described in Hurlbert (1971) for ecological diversity when sample sizes are unequal. Here we adapted this method to population genetics. Let us first define the allelic richness at a locus, for a fixed sample size, as $r_{(n)}$, the number of different alleles found when n genes are examined at this locus. The corrected allelic richness at this locus is then $r'_{(n)}=r_{(n)}-1$, since the diversity must be zero when there is a single allele fixed in the population (Patil and Taillie 1982). If a total of N (>n) genes are examined at this locus, the expected number of different alleles in a sample of n genes is obtained with the rarefaction method of Hurlbert (1971):

$$\hat{r}_{(n)} = E\left[r_{(n)}\right] = \sum_i \left[1 - \binom{N-N_i}{n} \Big/ \binom{N}{n}\right] \quad (1)$$

where N_i represents the number of occurrences of the ith allele among the N sampled genes. The average within-population allelic richness is simply the mean across populations of the allelic richness for the locus, $r_{S(n)}$, whereas the allelic richness of the total population, $r_{T(n)}$, is obtained by first computing a weighted sum for each allele across all k populations and then applying formula (1) above; that is, computing the expected number of different alleles when n genes are sampled in the total population. Populations are weighted in inverse proportion to their sample sizes, in order to give the same weight to all populations. Because the rarefaction method supposes the use of whole numbers, the weighted total is rounded to the nearest integer.

Table 1 Geographic origins of the studied populations

Population	Abbreviation	Geographic origin	Altitude (m)	Climate
Ademine	AD	Souss plain	90	Arid
Aït Baha	AB	Northern side of the Anti Atlas	550	Arid
Argana	AR	Southern side of the High Atlas	600	Arid
Beni-Snassen	BS	North-East of Oujda	260	Arid
Goulimine	GO	Northern limit of the Sahara	530	Arid
Mijji	MI	Chichaoua plateau	350	Arid
Oued Grou	OG	South-East of Rabat	400	Sub-arid
Tafraout	TA	Northern side of the Anti Atlas	1150	Arid
Tamanar	TM	Atlantic coast	170	Sub-arid
Tizint'est	TT	Southern side of the High Atlas	1200	Arid

A measure of differentiation for allelic richness follows as: $\rho_{ST(n)} = 1 - r'_{S(n)}/r'_{T(n)}$. Mean measures of allelic richness and differentiation are obtained by averaging over loci, $R'_{S(n)} = \sum_l r'_{S(n)}/L$, $R'_{T(n)} = \sum_l r'_{S(n)}/L$, and $\rho_{ST(n)} = 1 - R'_{S(n)}/R'_{T(n)}$, where l indexes the loci and L is the total number of loci. Note that the Simpson index may be considered as a particular measure of corrected allelic richness since it becomes identical to the expected number of different alleles in a sample of two genes, minus one, when N is large enough. As a consequence, $\rho_{ST(2)} \approx F_{ST}$.

Table 2 Description of the enzyme systems used and number of alleles detected

Enzyme (EC number)	Locus	Quaternary structure	Alleles
Est (EC 3.1.1.6.)	*Est*-1	Monomeric	3
	Est-2	Monomeric	5
Got (EC 2.6.1.1.)	*Got*-1	Dimeric	2
	Got-2	Dimeric	5
	Got-3	Dimeric	5
Amy (EC 3.2.1.1.)	*Amy*-1	Monomeric	5
Lap (EC 3.4.11.1.)	*Lap*-1	Monomeric	3
	Lap-2	Monomeric	2
Adh (EC 1.1.1.1.)	*Adh*-3	Dimeric	2
Total	9		32

Results

Altogether, 32 alleles were identified for the nine loci (2–5 per locus, with a mean of 3.6). Allelic frequencies are given in Table 3. Note the presence of eight specific alleles: the two populations growing at altitudes higher than 1 000 m, Tizint'est (High Atlas) and Tafraout (Anti Atlas), both have three specific alleles. The population of Goulimine, at the northern limit of the Sahara, has one specific allele. Moreover, it has a second allele (*Adh*-3A) at high frequency (0.688) which is found elsewhere only in the Ademine forest, at low frequency. The relic population of Beni-Snassen also has one specific allele, *Lap*-1B, with a frequency of 0.433. In addition, it is fixed for the *Got*-1A allele, which is found only at much lower frequency in the Tafraout population. Finally, both northern relic populations (Oued Grou and Beni-Snassen), are characterised by a high frequency of the *Amy*-1D allele (higher than 0.45) which is

Table 3 Allelic frequencies in the ten Moroccan populations. Alleles of special interest are in bold characters. Population-specific alleles are underlined

Allele	Population									
	TT	AR	AD	AB	TA	GO	MI	TM	OG	BS
Est-1A	0.520	0.475	0.512	0.475	0.517	0.047	0.475	0.550	0.517	0.467
Est-1B	0.040	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.017	0.000
Est-1C	0.440	0.525	0.450	0.525	0.483	0.953	0.525	0.450	0.467	0.533
Est-2A	0.070	0.050	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.083
Est-2B	0.510	0.550	0.438	0.400	0.483	0.688	0.525	0.475	0.467	0.917
Est -2C	0.050	0.000	0.087	0.075	0.117	0.000	0.000	0.050	0.000	0.000
Est-2D	0.370	0.400	0.438	0.525	0.400	0.203	0.475	0.475	0.533	0.000
Est-2E	0.000	0.000	0.000	0.000	0.000	**0.109**	0.000	0.000	0.000	0.000
Got-1A	0.000	0.000	0.000	0.000	**0.250**	0.000	0.000	0.000	0.000	**1.000**
Got-1B	1.000	1.000	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.000
Got-2A	0.000	0.000	0.000	0.000	**0.050**	0.000	0.000	0.000	0.000	0.000
Got-2B	0.520	0.450	0.375	0.525	0.467	0.922	0.525	0.500	1.000	0.517
Got-2C	0.060	0.050	0.237	0.375	0.433	0.000	0.000	0.050	0.000	0.000
Got-2D	0.410	0.500	0.387	0.100	0.050	0.078	0.475	0.450	0.000	0.483
Got-2E	**0.010**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Got-3A	**0.040**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Got-3B	0.450	0.525	0.538	0.525	0.133	0.000	0.000	0.000	0.000	0.000
Got-3C	0.000	0.000	0.000	0.000	**0.017**	0.000	0.000	0.000	0.000	0.000
Got-3D	0.000	0.000	0.000	0.000	**0.033**	0.000	0.000	0.000	0.000	0.000
Got-3E	0.510	0.475	0.463	0.475	0.817	1.000	1.000	1.000	1.000	1.000
Amy-1A	0.530	0.900	0.538	0.850	0.517	0.922	0.575	0.550	0.550	0.533
Amy-1B	**0.450**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Amy-1C	0.000	0.100	0.063	0.050	0.000	0.000	0.000	0.000	0.000	0.000
Amy-1D	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.450**	**0.467**
Amy-1E	0.000	0.000	0.400	0.100	0.483	0.078	0.425	0.450	0.000	0.000
Lap-1A	0.000	0.000	0.000	0.000	0.033	0.078	0.100	0.000	0.000	0.000
Lap-1B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.433**
Lap-1C	1.000	1.000	1.000	1.000	0.967	0.922	0.900	1.000	1.000	0.567
Lap-2A	0.970	1.000	1.000	1.000	1.000	1.000	1.000	0.575	1.000	1.000
Lap-2B	0.030	0.000	0.000	0.000	0.000	0.000	0.000	0.425	0.000	0.000
Adh-3A	0.000	0.000	0.063	0.000	0.000	**0.688**	0.000	0.000	0.000	0.000
Adh-3B	1.000	1.000	0.938	1.000	1.000	0.313	1.000	1.000	1.000	1.000

nearly absent in the other populations. Hence, it appears from this preliminary survey that important genetic differences exist among populations of the argan tree.

Intrapopulation genetic structure

The mean observed heterozygosity, the mean expected genetic diversity, and the multilocus heterozygote deficit (F_{IS}) were computed for each of the ten populations (Table 4). In addition, the mean number of alleles per locus was computed for all these populations, as well as the expected mean number of alleles per locus, in samples of 40 genes or of ten genes, with the rarefaction method described above (Table 4). A detail of the computations involved is provided for one of the loci (*Est*-2, Table 5). In each population, the probability of sampling each of the five alleles of this locus is computed if only ten genes had been scored. The sum of these probabilities is the expected number of different alleles in a sample of ten genes observed and corresponds to formula (1).

For each population, the mean number of alleles per locus for uniform sample sizes of ten and 40 genes can then be compared with the results observed with uneven sample sizes. The effect of the correction is clear: for instance, the Tizint'est population, where 50 trees had been sampled (more than in any other population), was characterised by the highest number of alleles per locus. This was clearly an artefact since, after correction, two populations where the sample sizes were smaller turned out to be actually more diverse. Figure 2 further illustrates the relationship between the three measures of diversity, *He*, $R'_{T(10)}$, and $R'_{T(40)}$. As expected, the best correlation is obtained between the allelic richness $R'_{T(10)}$ and Nei's diversity *He*: r=0.95 ($P<0.01$) (recall that *He* is nearly identical to $R'_{T(2)}$, as mentioned in the Materials and methods section). The correlation between $R'_{T(40)}$ and *He* is smaller, though still very significant (r=0.70, $P<0.01$). Populations with

Table 4 Intrapopulation genetic measures. N: number of diploid individuals sampled per population. Nb. All./locus: means the number of alleles per locus observed (Obs.) and expected in samples of 40 genes and 10 genes. H_o and H_e: observed and expected heterozygosity. F_{IS}: heterozygote deficit

Pop.	N	Nb. All./locus			H_o	H_e	F_{IS}
		Obs.	40	10			
TT	50	2.44	2.05	1.85	0.276	0.314	0.12
AR	20	1.78	1.78	1.62	0.233	0.258	0.09
AD	40	2.22	2.19	1.91	0.306	0.333	0.08
AB	20	1.89	1.89	1.60	0.267	0.273	0.02
TA	30	2.33	2.27	1.90	0.27	0.333	0.19
GO	32	1.78	1.77	1.42	0.156	0.161	0.03
MI	20	1.56	1.56	1.52	0.200	0.248	0.19
TM	20	1.78	1.78	1.65	0.267	0.294	0.09
OG	30	1.44	1.41	1.35	0.163	0.170	0.04
BS	30	1.56	1.56	1.51	0.244	0.242	–0.01
Mean	29	1.88	1.82	1.63	0.214	0.238	0.09

Table 5 An illustration of the use of the rarefaction technique using the *Est*-2 locus

Population		*Est*-2A	*Est*-2B	*Est*-2C	*Est*-2D	*Est*-2E	Total
TT	Obs. Nb.	7	51	5	37	0	100
	Prob. (10)	0.533	1.000	0.416	0.993	0.000	2.942
AR	Obs. Nb.	2	22	0	16	0	40
	Prob. (10)	0.442	1.000	0.000	0.998	0.000	2.440
AD	Obs. Nb.	3	35	7	35	0	80
	Prob. (10)	0.334	0.998	0.623	0.998	0.000	2.953
AB	Obs. Nb.	0	16	3	21	0	40
	Prob. (10)	0.000	0.998	0.589	1.000	0.000	2.587
TA	Obs. Nb.	0	29	7	24	0	60
	Prob. (10)	0.000	0.999	0.741	0.997	0.000	2.737
GO	Obs. Nb.	0	44	0	13	7	64
	Prob. (10)	0.000	1.000	0.000	0.916	0.715	2.631
MI	Obs. Nb.	0	21	0	19	0	40
	Prob. (10)	0.000	1.000	0.000	1.000	0.000	2.000
TM	Obs. Nb.	0	19	2	19	0	40
	Prob. (10)	0.000	1.000	0.442	1.000	0.000	2.442
OG	Obs. Nb.	0	28	0	32	0	60
	Prob. (10)	0.000	1.000	0.000	1.000	0.000	2.000
BS	Obs. Nb.	5	55	0	0	0	60
	Prob. (10)	0.612	1.000	0.000	0.000	0.000	1.612
Mean	Prob. (10)	0.192	0.999	0.281	0.890	0.071	2.434
Total	Weight. tot.	14	318	22	223	6	583
	Prob. (10)	0.217	1.000	0.321	0.992	0.099	2.629
						$r'_{S(10)}$	1.434
						$r'_{T(10)}$	1.629
						$\rho_{ST(10)}$	0.119

similar values of *He* do have differences in allele numbers, as for instance in the case of Goulimine (*He*=0.161, $R'_{T(40)}$=1.77) and Oued Grou (*He*=0.170, $R'_{T(40)}$=1.41). This latter relic population is actually characterised by the smallest allelic richness. Finally, all populations except the relic one of Beni-Snassen are characterised by a deficit of heterozygotes (F_{IS}>0).

Interpopulation genetic structure

The summary of the interpopulation statistics is given in Table 6. In addition to the traditional measures of diversity (i.e. the observed heterozygosity, the mean intrapopulation and the total diversity h_O, h_S and h_T, and the different *F*-statistics), measures of allelic richness, r' and ρ_{ST}, are also provided. The details of their computation is provided in Table 5 for the *Est*-2 locus. The parameter $r_{S(10)}$=2.43 corresponds to the mean within-population allelic richness. The expected number of distinct alleles when ten genes from the total population are sampled at this locus is $r_{T(10)}$=2.63. The differentiation is then $\rho_{ST(10)}$=1–(2.43–1)/(2.63–1)=0.12.

The mean estimate of F_{IS}=0.09 indicates limited, but general, departure from panmixia. On the other hand, the level of subdivision, measured by F_{ST}, is very high for a forest-tree species, 0.25. However, the two other measures of differentiation, $\rho_{ST(10)}$=0.38 and $\rho_{ST(40)}$=0.52, indicate that the partitioning of genetic diversity as measured by the allelic richness is even stronger. In Fig. 3, the relationship between the three measures of differentiation is illustrated by plotting, the ρ_{ST} values as a function of the F_{ST} for each locus. In most cases, $\rho_{ST(40)}>\rho_{ST(10)}>F_{ST}$ [$\approx\rho_{ST(2)}$], although there are exceptions. For instance, for the *Got*-1 locus, $\rho_{ST(10)}<F_{ST}$ whereas for the *Got*-3 locus, $\rho_{ST(40)}<\rho_{ST(10)}$. Overall, the correlation at the single locus level between F_{ST} and $\rho_{ST(10)}$ is high: 0.87 (P<0.01); between F_{ST} and $\rho_{ST(40)}$ it is 0.81 (P<0.01).

The genetic relationship between the populations is illustrated by the dendrogram with standard error bars constructed with Nei's measure of genetic distances (Fig. 4). The clusters are significant if the bars are roughly less than half the branch length (Ritland 1989). All populations can

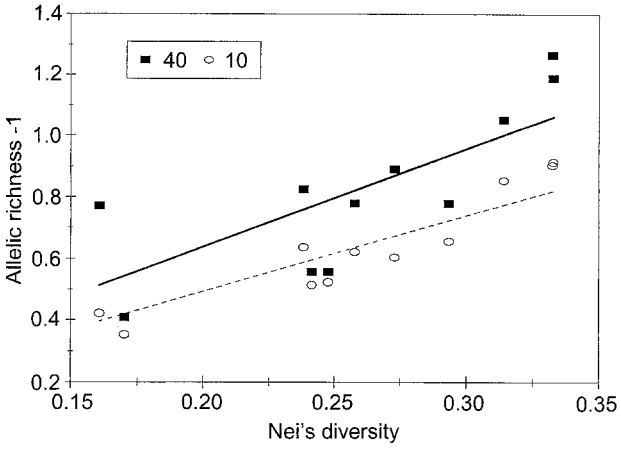

Fig. 2 Correlation between genetic diversity measured with Nei's index in each population and corrected allelic richness R' of that population (i.e. the mean number of alleles per locus in samples of ten and 40 genes, minus one). Note the better correlation when allelic richness is measured in samples of ten genes

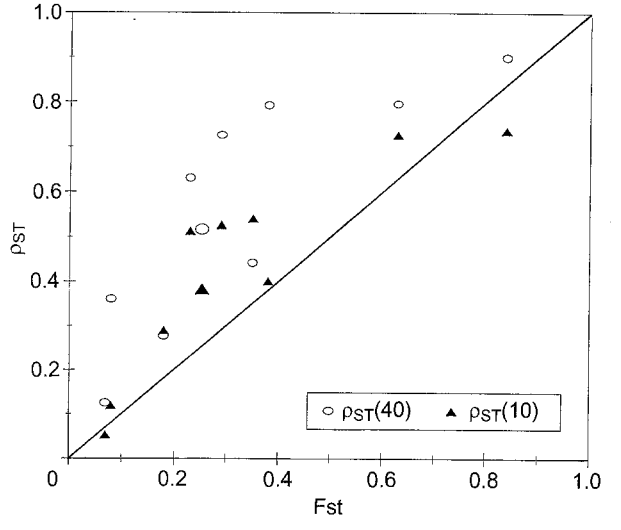

Fig. 3 Relation between single-locus measures of subdivision for allelic richness [$\rho_{ST(10)}$ and $\rho_{ST(40)}$] and F_{ST} at this locus. In most cases, differentiation for allelic richness is higher than F_{ST}

Table 6 Interpopulation genetic measures: *F*-statistics and allele-richness statistics

Locus	h_O	h_S	h_T	F_{IS}	F_{ST}	F_{IT}	$r'_{S(10)}$	$r'_{T(10)}$	$\rho_{ST(10)}$	$r'_{S(40)}$	$r'_{T(40)}$	$\rho_{ST(40)}$
Est-1	0.46	0.48	0.51	0.04	0.07	0.11	1.02	1.08	0.05	1.14	1.30	0.13
Est-2	0.46	0.52	0.56	0.10	0.08	0.17	1.43	1.63	0.12	1.78	2.78	0.36
Got-1	0.02	0.04	0.24	0.39	0.84	0.90	0.20	0.74	0.73	0.10	1.00	0.90
Got-2	0.42	0.47	0.57	0.10	0.18	0.26	1.25	1.76	0.29	1.63	2.26	0.28
Got-3	0.22	0.24	0.37	0.10	0.35	0.41	0.46	1.00	0.54	0.74	1.33	0.44
Amy-1	0.39	0.42	0.55	0.07	0.23	0.28	1.01	2.07	0.51	1.26	3.41	0.63
Lap-1	0.06	0.09	0.13	0.31	0.29	0.51	0.26	0.55	0.52	0.40	1.41	0.73
Lap-2	0.05	0.06	0.09	0.09	0.38	0.44	0.23	0.38	0.40	0.18	0.86	0.79
Adh-3	0.06	0.06	0.15	–0.01	0.63	0.62	0.15	0.55	0.73	0.20	0.96	0.79
Mean	0.24	0.26	0.35	0.09	0.25	0.32	0.67	1.08	0.38	0.82	1.70	0.52

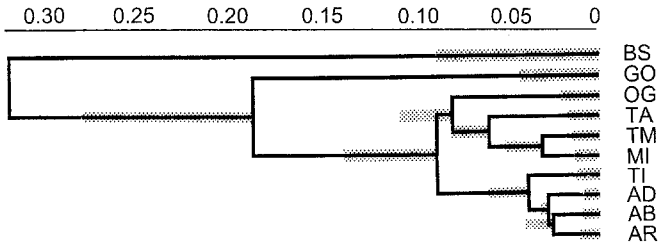

Fig. 4 Dendrogram of Nei's genetic distances between populations

be significantly distinguished. The relic Mediterranean population of Beni-Snassen is well separated from the other populations (d=0.31). The southern population of Goulimine is also strongly differentiated (d=0.18). Otherwise the closest groupings (d<0.05) identify groups of geographically close populations.

Discussion

Hamrick et al. (1992) have summarized parameters of genetic diversity in plant species in general with a particular emphasis on trees. The data obtained here for the argan tree can be compared with this survey. If the results of Table 6 (in particular h_T=0.35, h_S=0.26, and F_{ST}=0.25) are compared with those of the published survey, it appears that the argan tree has a higher level of total diversity and subdivision than many Angiosperms where on average h_T=0.29, h_S=0.25, and $G_{ST}[\approx F_{ST}]$=0.10. Some of the factors that could contribute to these genetic features can be deduced from the examination of Table 4 in the review of Hamrick et al. (1992) and from Moran (1992). In particular, the high value of differentiation among populations could originate from the combination of some of the following factors, all associated with a high G_{ST}: a reduced geographic range, the presence of disjunct populations, and a temperate-tropical distribution. Insect pollination, which is often associated with a reduced gene flow among populations, could also be involved. Moreover, the low density of argan-tree populations, particularly under the Saharan and arid interior climates, could lower the effective population sizes, in conjunction with the asynchronous flowering among and within sites. Hence, despite the fact that these trees are apparently largely outcrossing, as seen from the limited heterozygote deficit found in the populations, strong divergence between populations could quickly build up.

In conservation biology, immediate threats are generally due to demographic factors. Genetic criteria such as genetic diversity are more relevant for the long-term viability of the species. Measures of genetic diversity based on allelic richness have long been considered as important, especially in the field of conservation genetics. Indeed, for conservation purposes, it is essential to safeguard the largest possible collection of alleles (e.g., Marshall and Brown 1975; Asins and Carbonell 1987; Millar and Westfall 1992; Schoen and Brown 1993). Allelic richness is particularly vulnerable to a decrease in population size, such as a bottleneck or a founder effect (Nei et al. 1975), contrary to those measures of diversity which rely mostly on the more frequent alleles, such as Nei's measure. For instance, Nei et al. (1975, their Table 3) showed that for two representative loci, 100 individuals sampled in the source population will have 99.5% of the average heterozygosities of the parental population but only 42–53% of its allelic richness. Surprisingly, to-date, no helpful method to compare allelic richness across populations or species is available. The effective number of alleles *Ae* (Kimura and Crow 1964) is independent of sample size but measures the evenness of allelic frequencies, not allelic richness. The mean number of alleles per locus for the whole species or averaged over populations, which is often reported (e.g. Hamrick et al. 1992), is obviously dependent on the sample size. There are two possible solutions to obtain comparable measures of allelic richness. One of them would consist in extrapolating the number of alleles from a sample to the whole population. In ecology, the problem of estimating species richness through extrapolation has been given recent attention (Colwell and Coddington 1994). But much earlier, ecologists realized that the converse (estimating the species richness in smaller samples of specified sizes) was relatively easy. Sanders (1968) first introduced the method of rarefaction, which was later refined by Hurlbert (1971). Although cited in the well-known work of Marshall and Brown (1975) on the optimal sampling strategies for genetic conservation, the paper of Hurlbert does not seem to have received much attention from geneticists.

In the present paper, an application of this method of rarefaction to genetic data of special conservation interest was presented. The data consisted of a set of populations of an endangered tree species assessed for isozyme markers, in which the allelic richness is relatively high. Like the traditional measure of diversity, allelic richness combined with the technique of rarefaction may be used to derive a measure of the subdivision of gene diversity. An analogue to G_{ST} can be obtained, which we called $\rho_{ST(n)}$. Interestingly, we found that in the case of the argan tree when n increases, $\rho_{ST(n)}$ also increases. Since with larger samples rare alleles are increasingly taken into account, these results indicate that the rarer alleles have a less uniform distribution than the more frequent ones. Although expected on theoretical grounds, this is not a trivial result. In a previous paper Petit et al. (1995) concluded that the F_{ST} measure was poorly suited to test the hypothesis that rare alleles, on average of more recent origin, should have more circumscribed distributions than more frequent (i.e. older) ones. Indeed these rare alleles may not have had enough time to spread across the whole range of the species since they appeared through mutation. In addition, as mentioned above, they are much more susceptible to drift than more-frequent ones.

The classical opinion that in many cases "most of the genetic diversity resides within populations" seems to have been largely misleading, since it is valid for a particular interpretation of diversity which underestimates the impor-

tance of the less frequent alleles, and is therefore not very relevant in conservation genetics. Gregorius and Baradat (1992) already warned that "in the field of conservation of genetic resources, where knowledge of the distribution of genetic variation is of the utmost importance, decisions based on ambiguous measurements may have grave and irreversible consequences", and recalled that the use of the Shannon-Weaver measure of diversity (Shannon and Weaver 1949), if used in place of Nei's diversity index, will yield a different measure of subdivision. Here we demonstrate that if allelic richness is used instead, the measure of subdivision obtained will increase as a function of n, the number of genes sampled.

We are not aware of any previous attempt to compare allelic richness within and among populations. Schoen and Brown (1991) have noted that estimates of N_e (obtained using a population genetic model) are more variable among populations of inbreeding plants as compared to outbreeding species. Since this parameter is related to allele numbers, they concluded (Schoen and Brown 1993) that inbreeders exhibit much more variation in allelic richness among populations than do outbreeders. But this result is obtained using several important hypotheses, whereas our approach does not require any such assumptions and is also more straightforward.

Since rarer alleles turn out to be more localized than more frequent ones, the consequences of fragmentation are likely to be more drastic than what could have been expected. Fragmentation, by impeding migration, and hence the cohesiveness of the populations (Templeton et al. 1990), leads to the demographic independence of the forests which become increasingly susceptible to drift and then to extinction. Therefore, the study of gene flow among the remaining forests of argan trees would be particularly important to assess.

The so-called relic populations of the argan tree have a clearly reduced allelic richness. However, it is difficult, on the basis of such data, to evaluate the time since they became isolated from the main population. Indeed, the transfer of a few seeds by man during historical time could lead to an intense drift at the time of establishment, and this can not be distinguished, with such markers, from a long-lasting, but less severe, drift. Additional markers, for which the alleles can be ordered phylogenetically, may give an answer to this problem.

Acknowledgements The authors are indebted to Saïd Fakir who was involved in the collection of the material, especially the relic populations in the north of the country. Moreover, he greatly contributed to the equipment of the laboratory in Agadir. His encouragement throughout the study were much appreciated. We also thank Rémi Chaussod and Rachida Nouaïm who also largely contributed to the practical realization of this study. Odile Pons gave numerous statistical advice. Finally we thank Thomas Bataillon, Rémi Chaussod, Antoine Kremer and Christophe Plomion for critical reading of an earlier draft of the manuscript.

References

Asins MJ, Carbonell EA (1987) Concepts involved in measuring genetic variability and its importance in conservation of plant genetic resources. Evol Trends in Plants 1:51–62

Ayad A (1989) Présentation générale de l'arganier. In: Formation Forestière continue, Thème "l'Arganier". Forest Research Station, Rabat, Maroc, pp 9–18

Benchekroun F, Buttoud G (1989) L'arganeraie dans l'économie rurale du Soud Ouest marocain. Forêt Méditerranéenne 2:127–136

Brunel D (1979) Recherche du déterminisme génétique de quelques systèmes enzymatiques chez la luzerne (*M. sativa* L.) di et tétraploïde. Thesis, Université Paris XI, Paris

Colwell RK, Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. Phil Trans R Soc Lond B 345:101–118

Dupin (1949) L'arganier survivant de la flore tertiaire, providence du Sud marocain. Elevages et Cultures 3:28 p

Ehrig FR (1974) Die Arganie: Charakter, Ökologie und wirtschaftliche Bedeutung eines Tertiär Reliktes in Marroko. Petermanns Geogr Mitteilungen 118:117–125

Emberger L (1925) Le domaine naturel de l'arganier. Bull Soc Bot Fr 72:770–774.

Gabriel O (1971) Locating enzymes in gels. Methods Enzymol 22: 578–604

Gregorius H-R, Baradat Ph (1992) A commentary on current approaches to forest population genetics. New Forests 6:409–420

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

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. Genetics 49:725–738

Maire R (1939) Les arganiers des Beni-Snassen. Bot Not, 447–484

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, pp 53–80

Millar CI, Westfall RD (1992) Allozyme markers in forest genetic conservation. New Forests 6:347–371

Monnier Y (1965) Les problème actuels de l'arganeraie marocaine. Rev For Fr 17:750–767

Moran GF (1992) Patterns of genetic diversity in Australian tree species. New Forests 6:49–66

Morton FF, Voss GL (1987) The Argan tree (*Argania sideroxylon*, Sapotaceae), a desert source of edible oil. Econ Bot 41:221–233

Msanda F (1993) Ecologie et cartographie des groupements végétaux d'Anzi (Anti-Atlas occidental, Maroc) et contribution à l'étude de la diversité génétique de l'Arganier [*Argania spinosa* (L.) Skeels]. Thesis, Université Grenoble

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29:1–10

Nouaïm R, Chaussod R, El Aboudi A, Schnabel C, Peltier JP (1991) L'arganier. Essai de synthèse des connaissances sur cet arbre. In: Groupe d'étude de l'arbre (eds) Physiologie des arbres et arbustes en zones arides et semi arides. John Libbey Eurotext, Paris, pp 373–388

Orstein L, Davis B (1964) Disc electrophoresis. Ann NY Acad Sci 121:321–347

Pasteur N, Pasteur G, Bonhomme F, Catalan J, Britton-Davidiam J (1987) Manuel technique de génétique par électrophorèse de protéines. Lavoisier, Paris

Patel GP, Taillie C (1982) Diversity as a concept and its measurement. J Am Stat Assoc 77:548–561

Peacock AC, Dingman CW (1967) Resolution of multiple ribonucleic acid species by polyacrylamide-gel electrophoresis. Biochemistry 6:1818–1827

Pelletier JP (1982) La végétation du bassin versant de l'oued Souss (Maroc). Thesis, Université Grenoble

Petit RJ, Bahrman N, Baradat Ph (1995) Comparison of genetic differentiation in maritime pine (*Pinus pinaster* Ait.) estimated using isozyme, total protein and terpenic loci. Heredity 75:382–389

Pons O, Chaouche K (1995) Estimation, variance and optimal sampling of gene diversity. II. Diploid locus. Theor Appl Genet 91:122–130

Ritland K (1989) Genetic differentiation, diversity and inbreeding in the mountain monkeyflower (*Mimulus caespitosus*) of the Washington Cascades. Can J Bot 67:2017–2024

Sanders HL (1968) Marine benthic diversity: a comparative study. Am Nat 102:243–282

Schoen DJ, Brown AHD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. Proc Natl Acad Sci USA 88:4494–4497

Schoen DJ, Brown AHD (1993) Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proc Natl Acad Sci USA 90:10623–10627

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana

Simpson EH (1949) Measurement of diversity. Nature 163:688.

Templeton AR, Shaw K, Routman E, Davis SK (1990) The genetic consequences of habitat fragmentation. Ann Missouri Bot Gard 77:13–27

Tregubov V (1963) Etude des groupements végétaux du Maroc Oriental. Bull Mus Hist Nat Marseille 23:121–194