Genetic diversity in Tunisian Crataegus azarolus L. var. aronia L. populations assessed using RAPD markers

Chayma Rajeb, Chokri Messaoud, Hnia Chograni, Afef Bejaoui, Abdennacer Boulila, Mohamed Nejib Rejeb, Mohamed Boussaid

Ann. For. Sci. 67 (2010) 512
© INRA, EDP Sciences, 2010
DOI: 10.1051/forest/2010014

Available online at:
www.afs-journal.org

Original article

Genetic diversity in Tunisian *Crataegus azarolus* L. var. *aronia* L. populations assessed using RAPD markers

Chayma RAJEB[1], Chokri MESSAOUD[1], Hnia CHOGRANI[1], Afef BEJAOUI[1], Abdennacer BOULILA[1], Mohamed Nejib REJEB[2], Mohamed BOUSSAID[1]*

[1] National Institute of Applied Sciences and Technology (INSAT) Department of Biology Laboratory of Plant Biotechnology Centre Urbain Nord, BP 676, 1080 Tunis Cedex, Tunisia
[2] National Institute of Research in Agricultural Engineering, Water and Forests, BP 10, Ariana-2080, Tunisia

(Received 19 March 2009; accepted 25 November 2009)

Keywords:
RAPD
***Crataegus azarolus* /**
genetic diversity /
natural populations /
conservation

Abstract

• The genetic diversity of nine wild Tunisian *Crataegus azarolus* var. *aronia* L. populations from different bioclimates was assessed using RAPD markers.

• Eight selected primers generated a total of 105 bands, 81 of which were polymorphic. Shannon's index (H') ranged from 0.222 to 0.278 according to a population with an average of 0.245. The genetic variation within the species (H_{SP} = 0.423) was relatively low. A high differentiation (G_{ST} = 0.421) among populations coupled with a low level of gene flow (N_m = 0.472) were observed. The analysis of molecular variance (AMOVA) revealed also significant differentiation among populations (Φ_{ST} = 0.371), even at a low scale space. The majority of variation occurred within populations (63.31%). The Mantel test performed on genetic (Φ_{ST}) and geographic distance matrices among population pairs did not reveal an isolation by distance.

• Interpretation of Neighbour-joining tree based on Nei's and Li's genetic distance among individuals showed distinct population groupings. The UPGMA dendrogram based on Φ_{ST} values revealed two population sub-clusters, each including populations from different bioclimates and/or geographic regions.

• The low level of genetic diversity and the high genetic structure of populations resulted from genetic drift caused both by habitat fragmentation and the low size of populations.

• The high differentiation among populations and the similar low level of diversity within populations suggest that in situ conservation should interest all populations. The ex situ conservation should be based on the collection of seeds rather within than among populations because of the maximum of variation was revealed within populations.

Mots-clés :
RAPD /
***Crataegus azarolus* /**
diversité génétique /
populations naturelles /
conservation

Résumé – Analyse de la diversité génétique de populations tunisiennes de *Crataegus azarolus* L. var. *aronia* L. par des marqueurs RAPD.

• La diversité génétique de neuf populations naturelles de *Crataegus azarolus* var. *aronia* L. en Tunisie, appartenant à différents étages bioclimatiques, a été analysée par des marqueurs RAPD.

• Huit amorces retenues ont généré 105 bandes dont 81 sont polymorphes. L'indice de Shannon (H') varie de 0,222 à 0,278 selon les populations avec une moyenne de 0,245. La variabilité génétique au sein de l'espèce est relativement faible (H_{SP} = 0,423). Une différenciation importante entre les populations (G_{ST} = 0,421) et un faible flux de gènes entre elles (N_m = 0,472) ont été observés. L'analyse de la variance moléculaire (AMOVA) a révélé, elle aussi, une différenciation significative entre les populations considérées ensemble (Φ_{ST} = 0,371) ou regroupées selon leur localisation bioclimatique. La majeure proportion de la variabilité réside à l'intérieur des populations (63,31 %). Le test de Mantel, effectué sur les matrices des distances génétiques (Φ_{ST}) et géographiques entre les paires des populations, n'a pas révélé une isolation par distance.

• Le dendrogramme, établi à partir des distances génétiques de Nei et Li, a montré des regroupements des individus dans leurs populations respectives. Le dendrogramme UPGMA, construit à partir des valeurs de Φ_{ST} entre les paires des populations, révèle deux groupes de populations. L'agrégation des populations au sein des groupes ne s'opère pas toujours selon leur proximité géographique et/ou leur localisation bioclimatique.

* Corresponding author: mohamed.boussaid@insat.rnu.tn

Article published by EDP Sciences

• La faible diversité génétique intrapopulation et le niveau élevé de différenciation des populations résulteraient de la dérive génétique due à la fragmentation des habitats et à la taille réduite des populations.
• Toutes les populations devraient bénéficier d'une priorité pour leur conservation in situ, vu le niveau élevé de leur différenciation et leur faible diversité génétique. La conservation ex situ devrait s'appuyer sur le prélèvement d'un maximum de semences plutôt au sein qu'entre les populations étant donné que la majeure proportion de la variabilité réside à l'intérieur des populations.

1. INTRODUCTION

The genus *Crataegus* L. (Rosaceae) is widely distributed in temperate regions of the Northern hemisphere including North America, Europe, Northern Asia and the Mediterranean region. It is represented by over 265 species (Christensen, 1992). Species naturally occur as small trees and shrubs, mainly diploid ($x = 17$) (Evans and Campbell, 2002). The systematic of the genus based on morphological (Phipps, 2005), cytological (Dickinson and Talent, 2007) and reproductive traits (Christensen, 1992; Evans and Campbell, 2002) is still unclear owing to interspecific hybridization, polyploidy and apomixis. Studies of phylogenic relationships among species based on molecular markers are currently under progress (Albarouki and Peterson, 2007; Evans and Campbell, 2002; Fineschi et al., 2005; Lo et al., 2007; 2009).

In Tunisia, the genus *Crataegus* L. includes two species: *Crataegus monogyna* Jacq. var. *monogyna* Jacq. (=*Crataegus oxyacanthus* L. subsp. *monogyna* (Jacq.) Rouy and Camus) and *Crataegus azarolus* L. which can be sympatric over a wide area of the Tunisian Dorsal Mountain (Pottier-Alapetite, 1979). Hybridization between the two species may occur (Albarouki and Peterson, 2007; Christensen, 1992). *Crataegus azarolus* L. (*orientalis* series, Crataegus section) is cultivated, with mixed trees (i.e. Almond, Olive) in small orchards (1-6 individuals/field) for fruit use. It comprises two botanical varieties (Pottier-Alapetite, 1979): *C. azarolus* L. var. *aronia* L. [=var. *aronia* (Willd.) Batt.] and *C. azarolus* var. *azarolus* L. The latter is extremely rare and occurs in few individuals mixed with *Crataegus monogyna* and/or *C. azarolus* var. *aronia*. The two varieties are similar in vegetative and floral traits but differ by fruit colour. When completely ripe, fruits of *C. azarolus* var. *azarolus* and *C. azarolus* var. *aronia* are typically red and yellow, respectively.

C. azarolus var. *aronia* is predominantly diploid ($2n = 2x = 34$) and outcrosser (Christensen, 1992; Talent and Dickinson, 2005). It reproduces via seeds and shows a vegetative propagation through sprouts. The species grows in humid and semi-arid bioclimates, on marneous soil usually below 800 m, in edges of forests and in garrigues dominated by *Pinus halepensis* Mill., *Juniperus phoenicea* L., *Quercus ilex* L. and *Tetraclinis articulata* Benth. The flowering period extends from mid-March to mid-May. Flowers are hermaphrodite, pentamerous with numerous stamens, and grouped in racemes. They are pollinated by bumble bees and smaller hymenoptera (Dickinson and Campbell, 1991). The fruit ripening period extends from August to December. Fruits (pomes), that differ in size (10–25 mm in diameter), contain 2–3 pyrenes each including one seed. Birds, small mammals and ungulates are the major seed dispersal vectors (Dickinson and Campbell, 1991).

Tunisian natural *C. azarolus* var. *aronia* populations are fragmented and represented by scattered individuals. All populations are highly affected by human practices (i.e. clearing, charcoal production, over harvesting) and characterized by a low size (often below 50 cm). Many populations are lost or endangered. Population isolation and the decreasing size increase genetic erosion, reducing therefore adaptability to environmental changes. Thus, knowledge of genetic diversity and genetic structure of populations is required for the development of appropriate in situ and ex situ conservation strategies.

The purpose of this paper is to assess, by RAPDs (Random Amplified Polymorphic DNA) (Williams et al., 1990), the genetic diversity within and among Tunisian populations of *Crataegus azarolus* var. *aronia* from different sites, with a focus on developing conservation strategies for the species. RAPD markers are a valuable tool for genetic studies in natural populations for woody plants and can yield a large number of loci by providing a wide representative sample of the genome (Ferrazzini et al., 2008; Fournier et al., 2006).

2. MATERIALS AND METHODS

2.1. Analysed populations

Nine Natural populations collected from different geographical regions were assessed (Fig. 1 and Tab. I). They belong to upper semi-arid (populations 1, 2, 3, 4, 5 and 6) and sub-humid (populations 7, 8 and 9) bioclimates according to Emberger's (1966), pluviothermic coefficient Q_2 ($Q_2 = 2000P/(M^2 - m^2)$); where P is the average of annual rainfall (mm), M is the mean of maximal temperature (Kelvin) for the warmest month (July) and m is the average of minimal temperature (Kelvin) for the coldest month (February). Altitudes of sites ranged from 150 to 865 m. The average annual rainfall varied from 400 to 600 mm. All populations were collected in *Pinus halepensis* Mill. margin forests.

Six to twelve trees in each population (at fruit ripening period), with 50 to 300 m between individuals, were sampled. Sample numbers were dictated by the scarcity of individuals and the small size of the populations encountered. From each individual, branches with leaves were taken for RAPD analyses.

2.2. DNA extraction

Three hundred milligrams of leaves from each plant were grounded to fine powder in liquid nitrogen and mixed with 1.5 mL of preheated extraction buffer (100 mM Tris-HCl, pH 8; 20 mM EDTA;

Table I. Properties of the 9 Tunisian *Crataegus azarolus* var. *aronia* populations analysed.

Population	Code	Bioclimatic zone[a]	Q_2[b]	Altitude (m)	Rainfall (mm/year)	Sample size
Kesra	1	Upper semi-arid	45.50	400	400–500	10
Bargou Jb. Mt[c]	2		45.70	865	400–500	9
Le Kef	3		49.50	350	400–500	9
Zaghouan Jb. Mt	4		55.44	300	400–500	8
Touiref	5		49.77	300	400–500	12
El Mnihla	6		57.30	250	400–500	6
Ezzit Jb. Mt	7		63.57	350	500–600	8
Oued Abid	8	Sub-humid	88.20	200	500–600	7
Ain Tbornek	9		67.70	150	500–600	8

[a] Bioclimatic zones were defined according to Emberger's (1966).
[b] Q_2 was calculated for each site using P, M and m average values for the period 1953–2003 from data provided by the Tunisian National Institute of Meteorology (see text).
[c] Jb. Mt = Jbel Mountain.

Figure 1. Map of Tunisia: Geographical distribution of the 9 Tunisian *Crataegus azarolus* var. aronia populations analysed. ●: Sub-humid; ▲: Upper semi-arid. 1, 2, 3, ..., 9: Population code.

2% CTAB; 1.4 M NaCl; 1% β-mercaptoéthanol) and 100 mg PVP 40000. Samples were then incubated at 65 °C for 1 h with slow shaking every 5 min. Subsequently the mixture was treated twice with 600 µL chloroform-isoamyl alcohol (24:1). Cold ethanol 95% and NaCl (5M) were used to precipitate the DNA following the method described by Lodhi et al. (1994). The DNA pellet was re-suspended in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA). The RNA was eliminated by adding 2 µL of Rnase-Dnase free solution (10 mg/mL). The quality of the DNA was estimated on an agarose gel (0.8%) stained with ethidium bromide. DNA quantity was evaluated spectrophotometrically by measuring absorbance at 260 nm.

2.3. Random amplification and sampling primers

Reactions were standardized and all PCR reactions were run on the same thermal cycler (Mark Maximum-Gene). For every 25 µL of volume reaction, (50 ng of DNA), 2.5 µL of 10X Taq polymerase buffer, 40 pmoles of primer, 2.5 mM of $MgCl_2$, 200 µM of dNTP and 1.5 U of Taq polymerase were included. Each reaction was overlaid with an equal volume of mineral oil. The PCR program was set as follow: an initial denaturation step of 94 °C for 2 min, followed by 45 cycles of 30 s at 94 °C, 1 min at 36 °C (annealing step) and 2 min at 72 °C (elongation step). An additional 10 min period for elongation at 72 °C followed this cycle.

Amplification products were separated on 1.5% agarose gel in TAE buffer (1X; pH 8), stained with ethidium bromide, and photographed under UV light using a DOC PRINT Photo Documentation System. Molecular weights were estimated using a 200 bp DNA Promega ladder. To ensure the reproducibility within and between runs, DNA from the same two additional individuals was included and amplified twice in every PCR run.

Sixteen primers (Operon Technologies) were tested for RAPD amplification. Eight primers (OPB08, OPB13, OPB14, OPB16, OPC17, OPJ07, OPJ16 and OPJ17), giving reliable banding patterns with high reproducibility and clear band resolution, have been selected.

2.4. Data analysis

RAPD bands were recorded as presence (1) or absence (0) of a band for each marker in each individual sample, and data were entered as a binary matrix. As RAPDs limitations (i.e. reproducibility of bands, dominant expression) may have a biased effect on the estimation of the genetic diversity and population structure, we have therefore (i) removed from the data bands with a frequency lower

than $3/N$ (N = 77 plants analysed; frequency > 0.249), and (ii) used, jointly to classical genetic parameters, the molecular variance (AMOVA) considered to be the less unbiased differentiation coefficient for RAPD analyses (Excoffier et al., 1992; Holsinger and Wallace, 2004).

The percentage of polymorphism ($P\%$ = (number of polymorphic bands/number of total bands) ×100), at the population and the ecological group (populations from the same bioclimate) levels, was calculated. The genetic diversity was determined by calculating Shannon's index (H') for each population and each group ($H' = -\Sigma p_i \log 2p_i$; where p_i is the frequency of the presence or absence of a RAPD band in a population or in a group). A variance analysis (ANOVA procedure) (SAS, 1990) was used to estimate the significance of variation of H' among populations and among ecological groups. The correlation among Shannon's indices (H') and altitude or Emberger's Q_2 matrices was evaluated using Kandell's test. The average diversity over all populations (H_{pop}) was calculated as: $H_{pop} = -1/n\Sigma H'$; where n is the number of populations. The total diversity among all individuals within the species was estimated as: $H_{sp} = -\Sigma p_s \log 2p_s$; where p_s is the frequency of presence or absence of the RAPD band in the whole sample.

The differentiation among populations was assessed by G_{ST} [$G_{ST} = (H_{SP} - H_{POP})/H_{SP}$]. The POPGENE computer package (Yeh et al., 1999) was used to calculate the different indices. The genetic structure among populations was also determined by analyses of the molecular variance (AMOVA) using the program WINAMOVA 1.55 (Excoffier et al., 1992). Φstatistics were calculated to estimate Φ_{ST} (differentiation among populations), Φ_{CT} (differentiation among ecological groups) and Φ_{SC} (differentiation among populations within groups). The significance of variance components was determined using 1000 independent permutations runs. Pairwise genetic distance (Φ_{ST}) between the nine populations were used to estimate the gene flow as the number of individuals migrating among populations per generation, using Wright's (1951) migrate number N_m[$N_m = 1/4(1/\Phi_{ST} - 1)$]. The correlation between the matrices of genetic and geographic distances among pairs of populations was estimated by a Mantel test (Mantel, 1967) using ZT program (Bonnet and van de Peer, 2002).

A Neighbour-joining cluster was produced to illustrate the relationship between individuals using Nei's and Li's (1979) genetic distance D_{xy} generated from the similarity coefficient S_{xy} [$S_{xy} = 2m_{xy}/(m_x + m_y)$; where m_{xy} is the number of bands shared by samples x and y, and m_x and m_y are the number of bands in samples x and y, respectively]. The genetic distance D_{xy} between individuals was estimated using the complementary value S_{xy}; [$D_{xy} = 1 - S_{xy}$]. The program MEGA version 2.0 (Kumar et al., 2001) was used to construct the dendrogram. UPGMA tree based on pairwise Φ_{ST} was also generated to compare similarities among populations using the program MVSP version 3 (Kovach, 1999).

3. RESULTS

A total of 105 RAPD bands were scored by the eight selected primers, among which 81 (Pr = 77.1%) are polymorphic (Tab. II). The revealed number of bands per primer varied from 11 (OPC17, OPJ07, OPJ17) to 16 (OPB16). The highest number of polymorphic bands was evidenced by OPB13 (12 bands) and OPB16 (13 bands). The polymorphism within a population ($P\%$) ranged from 34.3% (population 8 from the sub-humid bioclimate) to 47.6% (populations 2 and 4 from the upper semi-arid zone), with an average of 42.1% (Tab. III). The polymorphism was slightly higher in the semi-arid (76.2%) than in the sub-humid populations (64.8%).

Table II. Selected RAPD primers, number of polymorphic bands and percentages of polymorphic loci ($Pr\%$) per primer.

Primer	Sequence (5′ → 3′)	Total bands	Polymorphic bands	$Pr\%$
OPJ 16	TTC CCC CCA G	14	11	78.50
OPC 17	GTC CAC ACG G	11	9	81.80
OPB 08	TTC CCC CGC T	15	10	66.60
OPB 13	CTG CTT AGG G	15	12	80.00
OPJ 07	CCT CTC GAC A	11	10	90.90
OPJ 17	ACG CCA G TT C	11	6	54.54
OPB 14	TCC GCT CTG G	12	10	83.30
OPB 16	TTT GCC CGG A	16	13	81.25
	Total	105	81	77.14

Shannon's diversity index (H') within a population ranged from 0.218 (population 8, sub-humid zone) to 0.278 (population 2, upper semi-arid zone). However, the level of variation among populations was not significant (ANOVA test, $P > 0.05$). Besides, there was no correlation between Shannon's diversity index (H') and altitude matrices (Kendall's test; $r = 0.43$; $P > 0.05$) or between H' and Emberger's Q_2 matrices (Kendall's test; $r = -0.42$; $P > 0.05$). Averages H' did not differ between the sub-humid and the upper semi-arid groups (0.371 and 0.414, respectively). The average diversity within populations (H_{pop}) was 0.245, and that within ecological groups 0.393. Diversity within the species was relatively low (H_{sp} = 0.423) (Tab. III).

A high level of differentiation was observed among populations (G_{ST} = 0.421). This value was higher than that observed among the ecological groups (G_{STg} = 0.122). AMOVA indicated also a substantial level of differentiation among populations (Φ_{ST} = 0.371), and 63.31% of the genetic variation is apportioned within populations (Tab. IV). The level of gene flow is low (N_m = 0.423). N_m values between pairs of populations ranged from 0.278 (between populations 5 and 8) to 1.07 (between populations 1 and 9). Among the 36 Nm values, 35 were below 1 (Tab. V).

The differentiation among ecological groups was low (Φ_{CT} = 0.014; $P > 0.05$), that among populations within groups was important (0.363).The maximum variance (62.86%) was found within populations (Tab. IV).

Pairwise comparison of Φ_{ST} values were all highly significant ($P < 0.001$ after 1 000 permutations). The highest Φ_{ST} value (0.473) has been observed among populations 5 and 8 belonging to the upper semi-arid and the sub-humid bioclimates, respectively. These two populations were separated by 213 km each from another. The lowest value (0.189) was noted between the sub-humid population 9 and the upper semi-arid population 1, geographically 135 km apart each from another. At a low scale space (less than 20–25 km), Φ_{ST} values ranged from 0.312 to 0.406. Genetic distance (Φ_{ST}) and geographical distance matrices (km) did not show significant correlation ($r = 0.297$; $P > 0.05$). For populations grouped into

Table III. Polymorphism (*P%*), and genetic diversity parameters within and among populations and ecological groups.

Population			1	2	3	4	5	6	7	8	9
	P%		46.67	47.62	40.95	47.62	37.14	44.76	40.95	34.29	39.05
	H′		0.269[a]	0.278[a]	0.222[a]	0.271[a]	0.222[a]	0.260[a]	0.235[a]	0.218[a]	0.229[a]
	H_{POP}	0.245									
	H_{SP}	0.423									
	G_{ST}	0.421									
Ecological group				Sub-humid			Upper semi-arid				
	P%			64.76			76.19				
	H'_{GE}			0.371[a]			0.414[a]				
	H_{GE}	0.393									
	H_{SPg}	0.447									
	G_{STg}	0.122									

1, 2, 3, . . . , 9: Population code. Values with the same letter in line are not significantly different (ANOVA, $P > 0.05$).

Table IV. Nested analysis of molecular variance (AMOVA) at different hierarchical levels.

	Source of variation	df	SSD	MSD	Variance	% of total variance	Φ statistics
Population	Among populations	8	455.04	56.88	5.55	36.69	
	Within population	68	651.75	9.58	9.58	63.31	$\Phi_{ST} = 0.371^{*}$
	Total	76	1106.80				
Ecological group	Among groups	1	60.95	60.96	0.20	1.35	
	Among populations Within group	7	394.08	56.29	5.45	35.78**	$\Phi_{CT} = 0.014^{ns}$
	Within population	68	651.75	9.58	9.58	62.86[*]	$\Phi_{SC} = 0.363^{*}$
	Total	76	1106.80				

df, Degree of freedom; SSD, sum of squared deviation; MSD, mean squared deviation.

[*] Highly significant ($P < 0.001$ after 1 000 permutations); [ns] not significant ($P > 0.05$ after 1 000 permutations).

Table V. Geographical distance (km) (above diagonal) and genetic (Φ_{ST}) distance (below diagonal) among population pairs of *Crataegus azarolus* var. *aronia*. Nm values are given in parentheses.

Population	1	2	3	4	5	6	7	8	9
1		38	69	73	93	144	120	168	133
2	0.201[*] (0.993)		83	61	102	101	83	131	94
3	0.328[*] (0.513)	0.399[*] (0.377)		129	25	148	150	192	154
4	0.259[*] (0.716)	0.292[*]	0.377[*] (0.413)		152	57	20	71	40
5	0.328[*] (0.512)	0.439[*]	0.406[*]	0.369[*]		168	172	213	175
6	0.355[*] (0.455)	0.386[*] (0.397)	0.447[*] (0.309)	0.387[*] (0.396)	0.430[*] (0.331)		48	53	29
7	0.269[*] (0.678)	0.333[*] (0.501)	0.387[*] (0.396)	0.371[*] (0.424)	0.445[*] (0.312)	0.393[*] (0.386)		51	22
8	0.329[*] (0.509)	0.433[*] (0.328)	0.464[*] (0.289)	0.390[*] (0.391)	0.473[*] (0.278)	0.344[*] (0.478)	0.443[*] (0.314)		40
9	0.189[*] (1.071)	0.331[*] (0.505)	0.422[*] (0.342)	0.369[*] (0.428)	0.350[*] (0.464)	0.427[*] (0.336)	0.312[*] (0.551)	0.357[*] (0.450)	

[*] Highly significant at $P < 0.001$ (after 1000 permutations). 1, 2, 3, . . . , 9: Population code.

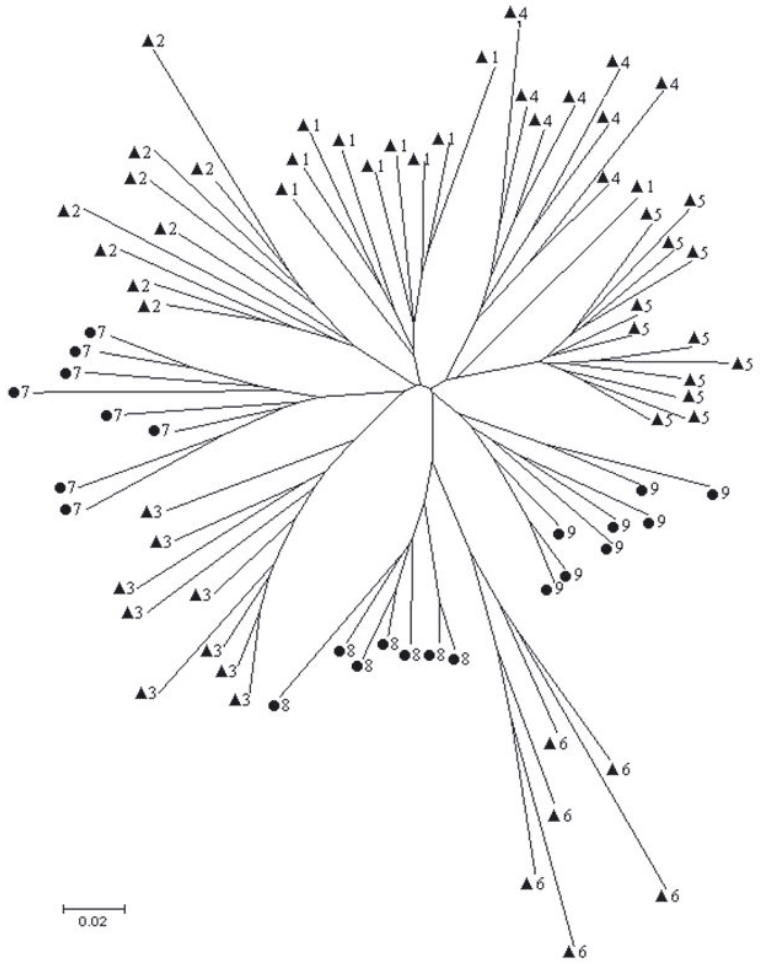

Figure 2. Neighbour-joining dendrogram generated for all individuals of *C. azarolus* var. aronia analysed using genetic distances of Nei and Li (1979). 1, 2, 3, . . . , 9: Population code.

two ecological groups, the isolation by distance was not revealed neither for the upper semi-arid populations (r = 0.497; $P > 0.05$), nor for sub-humid ones (r = 0.950; $P > 0.05$).

The dendrogram based on Nei's and Li's genetic distance showed that all individuals from the same population clustered together except for one individual from the population 1 grouped with samples from the population 5 (Fig. 2). The populations 6 and 8 from the upper semi-arid and the sub-humid bioclimates, respectively (53 km distant) were more isolated from the other populations. Results of UPGMA cluster analysis based on Φ_{ST} distance matrix revealed two main population groups (Fig. 3). The first contains populations 6 (upper semi-arid zone) and 8 (sub-humid zone). The second cluster is constituted by seven populations from different geographical and bioclimatic areas.

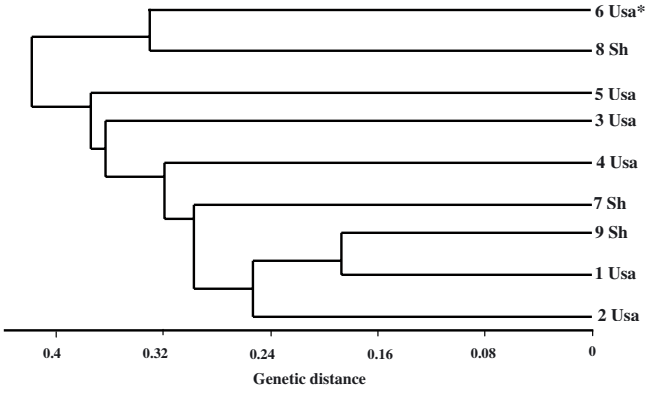

Figure 3. Dendrogram of the 9 populations based on Φ_{ST} pairwise values. 1, 2, 3, . . . , 9: Population code. * Usa: Upper semi-arid; Sh: Sub-humid.

4. DISCUSSION AND CONCLUSIONS

The analysis of the genetic diversity within and among populations of *Crataegus azarolus* var. *aronia* in Tunisia is prerequisite for conservation and improvement strategies. Our data based on RAPDs revealed a low level of variation within populations and high differentiation among them. RAPD markers are dominantly inherited (Williams et al., 1993). The

analysis of population genetic variation with RAPDs could be hampered by a loss of a part of genetic information (Kremer et al., 2005; Lynch and Milligan, 1994; Nybom, 2004). However, this disadvantage could be buffered by the detection of a high number of loci (Aagaard et al., 1998) and the use of appropriate genetic population parameters such as Φ_{ST} which reduces the bias in estimating the genetic variation. In our study, the eight selected primers revealed 105 loci (Frequency > 0.249), of which 81 were polymorphic. This large number of loci could be considered sufficient to compensate loss of genetic information content at loci.

Crataegus azarolus var. *aronia* populations maintain a relatively low level of genetic diversity as observed in other *Crataegus* (Ferrazini et al., 2008; Fineschi et al., 2005) or Rosaceae species (Petit et al., 2003). The low value of genetic diversity recorded in our study may result from loss of genetic variation through inbreeding and/or genetic drift due to the restriction of the species to small and degraded populations. However, amounts of the vegetative reproduction and selfing reported in *Crataegus* species could contribute to the low level of diversity detected. Thus, analyses including allo-autopollination experiments were required for collecting more information about the reproductive biology of the species and understanding the dynamic of its populations.

The observed averages of gene diversity at the population (H_{pop} = 0.245) and species (H_{sp} = 0.423) levels are within those reported from RAPDs for outcrossing trees (Hogbin and Peakall, 1999), but relatively lower compared to values observed in several gymnosperms with outcrossing mating system and wind-dispersed pollen (Begona et al., 2005).

The analysis of the molecular variation according to geographical and ecological factors (altitude and Q_2) has not revealed clinal structuring. Populations or ecological groups have similar values of average genetic diversity. This could be explained by the neutrality of RAPD markers, the large number of revealed loci and probably by the genomic characteristics of the species (high repetitive and non encoding DNA). However, the genetic similarity among populations, from RAPDs, may mask special morphological and/or physiological adaptations to bioclimates that will not be revealed through RAPDs.

The G_{ST} estimates using Shannon's index revealed a high differentiation among populations (G_{ST} = 0.421), and the majority of the genetic variation was attributable to variation within populations. The observed G_{ST} value was higher than that based on Shannon's index for outbreeding species (14.5 < G_{ST} < 38%) (Bussel, 1999) and Rosaceae species generally showing low level of genetic differentiation (Fineschi et al., 2005; Petit et al., 2003). With AMOVA, the amount of variation (Φ_{ST} = 0.371) was higher than that for outcrossers and wind-dispersal species (Φ_{ST} = 0.22 – 0.28, Nybom, 2004; Nybom and Bratish, 2000).

In most cases, the high differentiation among populations was caused by factors such as breeding system, isolation of populations, seed and pollen dispersal distance. This study showed that the population genetic structure, at all scale spaces, was not correlated to geographical distance. Thus isolation by distance could not alone explain the genetic structure observed. The high differentiation might result from habitat fragmentation, which led to the isolation of populations, the decreasing of their size and the limitation of gene flow among them. Similar results were obtained in Italian *C. monogyna* (Ferrazini et al., 2008).

The UPGMA dendrogram based on Φ_{ST} genetic distances among populations did not clearly separate populations according to their bioclimate and/or geographical location. Two groups of populations were revealed. The populations 6 and 8 constitute one group and seem to show low genetic divergence among them despite their different geographical and bioclimatic localization. The second group was composed of populations from different origins.

The isolation of populations 6 and 8 from the other populations could be explained by two hypothesis: firstly, it is possible that they have undergone a very high degree of genetic drift due to their low size (6 and 7 individuals) and their isolation, and are now very different from the others populations. Secondly, the two populations may correspond to hybrids among *Crataegus azarolus* and *Crataegus monogyna*. In fact, they grow with several individuals of *Crataegus monogyna*, and introgression between the two taxa may occur (Albarouki and Peterson, 2007; Christensen, 1992). In order to elucidate the hybrid character of these populations and their relationship with the two taxa, analyses using species specific molecular (i.e. ISSR, chloroplast microsatellite markers) and morphological (i.e. style and ovule numbers per flowers, leaf and fruit shape, pyrene number per fruit) markers are necessary to gain further insight into the hybrid character of these populations.

The aggregation of populations, in the second group, without any relationship to geographic distance and bioclimate, provides further evidence of genetic drift which can not be easily confirmed by RAPDs only. These markers show a dominant expression and are assumed to possess two alleles per locus, which may bias the estimation of some genetic parameters (i.e. frequency of null homozygotes, F index) mainly if selfing is occurring or if the populations are not randomly mating. On the other hand, RAPDs are dispersed throughout the genome and their association with ecological traits (i.e. climate, soil) is influenced by selection only in the region under selection pressure. The other loci are subject to random genetic drift. Thus, the genetic relationship between populations revealed by RAPDs did not necessary reflect natural selection process leading to similar local environment adaptation of populations.

Most *Crataegus azarolus* var. *aronia* populations in Tunisia were represented by few individuals. They suffer from a loss of genetic variation, and were highly differentiated. The long term viability of populations could be affected by an increasing habitat destruction. Taking into account these points, as well as the small of number of populations, efforts should be made to protect all populations and limit human impact. Given the high genetic differentiation among populations and the low level of genetic variation recorded, any in situ conservation strategy should aim to include populations from both sub-humid and upper semi-arid bioclimates. Such strategy could include restoration within a population through cuttings from the same population. Ex situ approaches may also be appropriate as apart of an overall conservation strategy. Seeds

may be collected rather within than between populations, because the maximum amount of genetic diversity within population was high. Populations from the upper semi-arid exhibiting a slightly high level of diversity must be protected in the first place.

Acknowledgements: The authors thank the Tunisian Ministry of Scientific Research and Technology and the National Institute of Applied Science and Technology for their financial support (Research grant 99/UR/09-10).

REFERENCES

Aagaard J.E., Krutovskii K.V., and Strauss S.H., 1998. RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. Heredity 81: 69–78.

Albarouki E. and Peterson A., 2007. Molecular and morphological characterization of *Crataegus* L. species (Rosaceae) in southern Syria. Bot. J. Linn. Soc. 153: 255–263.

Begona R., Sergio G.N., Ester S., Joel A., Peter C., and Juan S., 2005. Genetic diversity and structure of natural and managed populations of *Cedrus atlantica* (Pinaceae) assessed using random amplified polymorphic DNA. Am. J. Bot. 92: 875–884.

Bonnet E. and van de Peer Y., 2002. zt: a software tool for simple and partial Mantel tests. J. Stat. Software 7: 1–12.

Bussell J.D., 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). Mol. Ecol. 8: 775–789.

Christensen K.I., 1992. Revision of *Crataegus* Sect. *Crataegus* and Nothosect. *Crataeguineae* (Rosaceae-Maloideae) in the Old World. Syst. Bot. Monogr. 1–199.

Dickinson T.A. and Campbell C.S., 1991. Population structure and reproductive biology in the Maloideae (Rosaceae). Syst. Bot. 16: 350–362.

Dickinson T.A. and Talent N., 2007. Polyploidy, reproductive biology and Rosaceae understanding evolution and making classification. Plant. Syst. Evol. 266: 59–78.

Emberger L., 1966. Une classification biogéographique des climats. Recherches et Travaux des Laboratoires de Géologie, Botanique et Zoologie, Faculté des Sciences Montpellier (France) 7: 1–43.

Evans R.C. and Campbell C.S., 2002. The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated GBSSI genes. Am. J. Bot. 89: 1478–1484.

Excoffier L., Smouse P.E., and Quattro J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Ferrazzini D., Monteleone I., and Belletti P., 2008. Small-scale genetic diversity in one seed hawthorn (*Crataegus monogyna* Jaq.). Eur. J. For. Res. 127: 407–414.

Fineschi S., Salvini D., Turchini D., Pastorelli R., and Vendramin G.G., 2005. *Crataegus monogyna* Jacq. & *C. laevigata* (Poir.) DC (Rosaceae, Maloidese) display low levels of genetic diversity assessed by chloroplast markers. Plant. Syst. Evol. 250: 187–196.

Fournier N., Rigling A., Dobbertin M., and Gugerli F., 2006. Faible différenciation génétique, à partir d'amplification aléatoire d'ADN polymorphe (RAPD), entre les types de pin sylvestre (*Pinus sylvestris* L.) d'altitude et de plaine dans les Alpes à climat continental. Ann. For. Sci. 63: 431–439.

Hogbin P.M. and Peakall R., 1999. Evaluation of the contribution of genetic research to the management of the endangered plant, *Zieria prostrata*. Conserv. Biol. 13: 514–522.

Holsinger K.E. and Wallace L.E., 2004. Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). Mol. Ecol. 13: 887–894.

Kremer A., Caron H., Cavers S., Colpaert N., Gheysen G., Gribel R., Lemes M., Lowe A.J., Margis R., Navarro C., and Salgueiro F., 2005. Monitoring genetic diversity in tropical trees with multilocus dominant markers. Heredity 95: 274–280

Kovach W.L., 1999. A Multivariate Statistical Package for Windows, ver 3.1., Kovach Computing Services, Pentraeth, UK.

Kumar S., Tamura K., Jakobsen I.B., and Nei M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17: 1244–1245.

Lo E.Y.Y., Stefanovic S., and Dickinson T.A., 2007. Molecular reappraisal of relationships between *Crataegus* and *Mespilus* (Rosaceae, Pyreae) – Two genera or one? Syst. Bot. 32: 596–616.

Lo E.Y.Y., Stefanovic S., Christensen K.I., and Dickinson T.A., 2009. Evidence for genetic association between East Asian and western North American *Crataegus* L. (Rosaceae) and rapid divergence of the eastern North American lineages based on multiple DNA sequences. Mol. Phylogen. Evol. 51: 157–168.

Lodhi M.A., Ye G.N., and Weeden N.F., 1994. A simple and efficient method for DNA extraction from grapevine cultivars and *Vitis* species. Plant. Mol. Biol. Rep. 12: 6–13.

Lynch M. and Milligan B.G., 1994. Analysis of population genetic structure with RAPD markers. Mol. Ecol. 3: 9–99.

Mantel N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Nei M. and Li W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA. 76: 5269–5273.

Nybom H. and Bartish I.V., 2000. Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. Perspectives in plant ecology. Evol. Syst. 3: 93–114.

Nybom H., 2004. Comparaison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol. Ecol. 13: 1143–1155.

Petit R., Aguinagalde I., de Beaulieu J.-L., Bittkau C., Brewer S., Cheddadi R., Ennos R., Fineschi S., Grivet D., Lascoux M., Mohanty A., Muller-Starck G., Demesure-Musch B., Palme Â., Martîn J.P., Rendell S., and Vendramin G.G., 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. Science 300: 1563–1565.

Phipps J.B., 2005. A review of hybridization in North American hawthorns – Another look at "the *Crataegus* problem". Ann. Missouri Bot. Gard. 92: 113–126.

Pottier-Alapetite G., 1979. Flore de la Tunisie: angiospermes dicotylédones, Dialypétales. Publications Scientifiques Tunisiennes, Tunis, 654 p.

SAS (Statistical Analysis System), 1990. SAS user's guide: SAS STAT, SAS BASIC. Version 6 fourth edition. SAS incl, Box 8000. Cary, NC 27512-8000, Cary: NC: SAS institut Inc.

Talent N. and Dickinson T.A., 2005. Polyploidy in *Crataegus* and *Mespilus* (Rosaceae, Maloideae): evolutionary inferences from flow cytometry of nuclear DNA amounts. Can. J. Bot. 83: 1268–1304.

Williams J.G.K., Hanafey M.K., Rafalski J.A., and Tinjey S.V., 1993. Genetic analysis using random amplified polymorphic DNA markers. Methods Enzymol. 218:704–740.

Williams J.G.K., Kubelik A.R., Rafalski K.J., and Tingey S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531–6535.

Wright S., 1951. The genetical structure of populations. Ann. Eugenics 15: 323–354.

Yeh F., Yang R., and Boyle T., 1999. Popgene, version 1.31. Microsoft window-based freeware for population genetic analysis. University of Alberta, Edmonton.