Hereditas 147: 283–292 (2010)

Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers

H. GOUTA[1], E. KSIA[2], T. BUHNER[3], M. Á. MORENO[3], M. ZARROUK[4], A. MLIKI[4] and Y. GOGORCENA[3]

[1]*Unité des Ressources génétiques et de l'Amélioration de l'Olivier, de l'Amandier et du Pistachier, Institut de l'Olivier, Sfax, Tunisia*
[2]*Laboratoire de Biotechnologie et Biologie végétale, Faculté des Sciences de Tunis, Campus Universitaire, El Manar, Tunis, Tunisia*
[3]*Departamento de Pomología, Estación Experimental de Aula Dei (CSIC), Zaragoza, Spain*
[4]*Centre de Biotechnologie Borj-Cedria, Hammam-Lif, Tunisia*

Gouta, H., Ksia, E., Buhner, T., Moreno, M. À., Zarrouk, M., Mliki, A. and Gogorcena, Y. 2010. Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers. – *Hereditas 147*: 283–292. Lund, Sweden. eISSN 1601-5223. Received 8 September 2009. Accepted 18 January 2010.

Genetic diversity of 50 Tunisian almond (*Prunus dulcis* Mill.) genotypes and their relationships to European and American cultivars were studied. In total 82 genotypes were analyzed using ten genomic SSRs. A total of 159 alleles were scored and their sizes ranged from 116 to 227 bp. The number of alleles per locus varied from 12 to 23 with an average of 15.9 alleles per locus. Mean expected and observed heterozygosities were 0.86 and 0.68, respectively. The total value for the probability of identity was 4×10^{-13}. All SSRs were polymorphic and they were able all together to distinguish unambiguously the 82 genotypes. The Dice similarity coefficient was calculated for all pair wise and was used to construct an UPGMA dendrogram. The results demonstrated that the genetic diversity within local almond cultivars was important, with clear geographic divergence between the northern and the southern Tunisian cultivars. The usefulness of SSR markers for almond fingerprinting, detection of synonyms and homonyms and evaluation of the genetic diversity in the Tunisian almond germplasm was also discussed. The results confirm the potential value of genetic diversity preservation for future breeding programs.

Hassouna Gouta, Unité des Ressources génétiques et de l'Amélioration de l'Olivier, de l'Amandier et du Pistachier, Institut de l'Olivier, PO Box.014, 4061 Sousse, Tunisia. E-mail: zallaouz@yahoo.fr

The almond tree, dating back to ancient times, has been grown extensively in Tunisia since the Carthaginian era, 8th century B.C. (Jaouani 1976). Being the granary of Rome during the Roman Empire, Tunisia was considered as one of the main trade routes along which almond was spread throughout the shores of the Mediterranean Sea (Felipe 2000). Occupying the second position after olive (*Olea europea* L.) with approximately 22 millions of trees covering more than 302 000 ha, Tunisian almond plantations are located throughout all the country in different climatic conditions. About 90% of the land devoted to this fruit crop is located in the central and southern agricultural area of the country under arid and semi-arid conditions (Fig. 1). Sfax (34°44'N, 10°46'E) and Sidi Bouzid (35°04'N, 9°49'E) are the main producing regions with 45% of the national production. In the north, Bizerte (37°16'N, 9°52'E) (humid to sub-humid climate) presents a very specific ecosystem with cultivars particularly tolerant to strong winds, high humidity and many fungal diseases such as *Monillinia* and *Gloeosporium* responsible for moniliose and anthracnose diseases, respectively.

In the framework of the international cooperation with GREMPA (Groupe de Recherches et d'Etudes Mediterranéennes pour le Pistachier et l'Amandier), an almond germplasm collection was established in the Experimental Station of Ettaous (Sfax, Tunisia) at the beginning of the 1970s, and it was the starting point for the establishment of many other almond collections. As the management of the genetic resources in ex situ germplasm banks is rather expensive, precise identification of the accessions for avoiding duplications and mislabelling is needed. Furthermore, the correct evaluation of relatedness is essential for efficient genetic resources management and for maintaining enough variability for breeding programs. In the last five years, more than three million almond trees were lost because of the long period of drought (2000–2002). This has increased the need to preserve as much as possible the Tunisian almond genetic diversity, in order to prevent genetic erosion.

Due to the lack of information on the existing germplasm originating from either chance seedlings or human selections, identification and collection of this material were carried out during the last few years through the northern and central part of Tunisia. Consequently, an important genetic diversity has been identified. Until now,

DOI: 10.1111/j.1601-5223.2009.02147.x

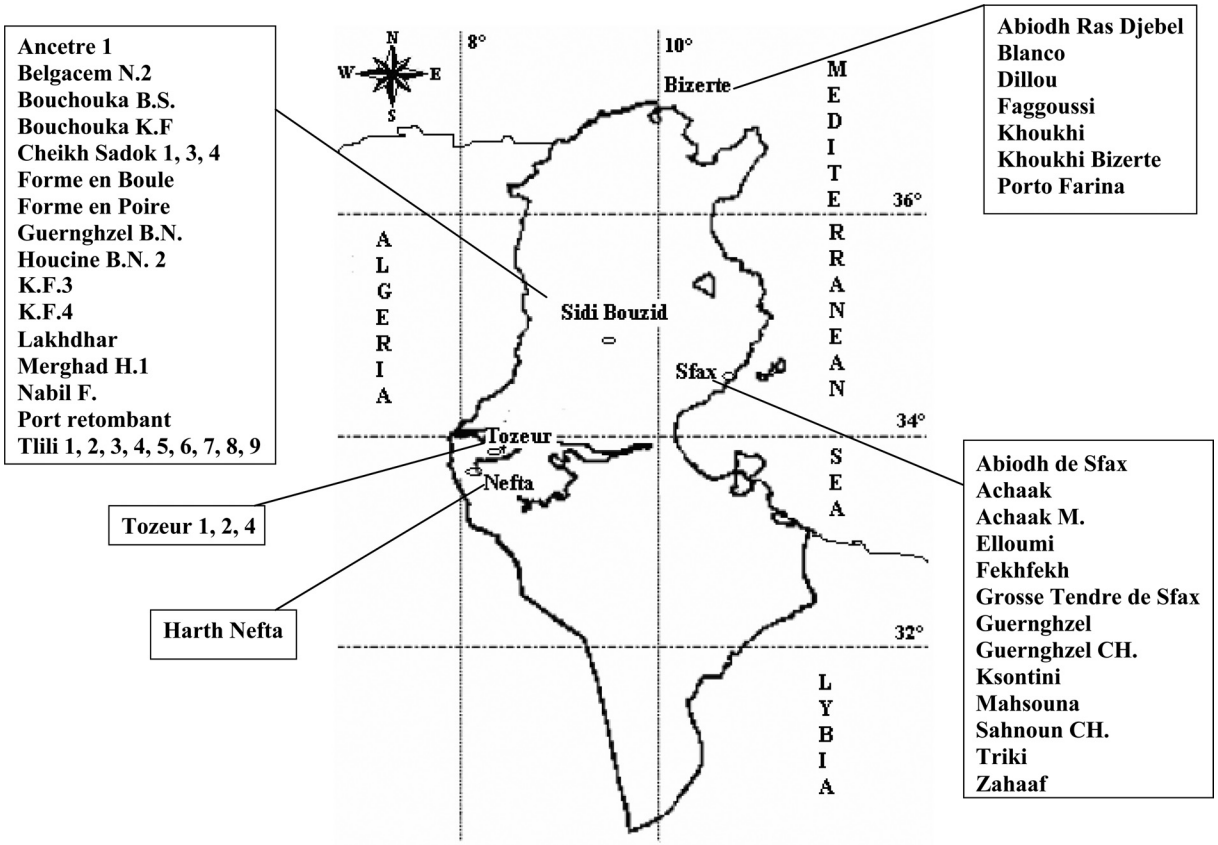

Fig. 1. Geographic position of the different accessions originated from different areas all over Tunisia.

the characterization of Tunisian almond cultivars has been done mainly through some quantitative and biochemical parameters such as: productivity, nut features (weight, shelling percentage and size), mineral composition (P, K, Ca, Mg) and lipid characteristics (oil and fat acid contents) (Ghrab et al. 2002; Ayadi et al. 2006), and more recently by using RAPD markers (Gouta et al. 2008).

The use of morphological descriptors for cultivar identification is very controversial due to environmental influences. On the other hand, the fact that RAPD is a dominant marker and non transferable between laboratories has prompted discussions about its efficiency. For these reasons, several other types of PCR-based molecular markers, such as microsatellites or simple sequence repeat markers (SSRs) are becoming the preferred marker for a wide range of applications in genetics and plant breeding. Recently, SSR primers generated from different *Prunus* species have been reported in almond (Joobeur et al. 2000; Testolin et al. 2004; Mnejja et al. 2005; Dangl et al. 2009); peach *P. persica* (L.) Batsch (Cipriani et al. 1999; Dirlewanger et al. 2002); apricot *P. armeniaca* L. (Lopes et al. 2002; Messina et al. 2004); Japanese plum *P. salicina* Lindl. (Mnejja et al. 2004) and cherry *P. avium* L. (Downey and Iezzoni 2000). These markers have been used for the molecular characterization and estimation of genetic diversity among peach, almond and other *Prunus* species (Aranzana et al. 2002; Dirlewanger et al. 2002; Martínez-Gómez et al. 2003a; Bouhadida et al. 2007), sweet cherry (Wünsch and Hormaza 2002) and apricot cultivars (Hormaza 2002; Maghuly et al. 2005). Moreover, SSRs are currently being employed for molecular characterization, estimation of genetic diversity and genetic relationships among almond cultivars and related *Prunus* species (Martínez-Gómez et al. 2003b; Xu et al. 2004; Sánchez-Pérez et al. 2006; Shiran et al. 2007; Zeinalabedini et al. 2008).

As only little information is available about the genetic diversity and relatedness within Tunisian almond cultivars and their relationship with almond cultivars originated from other countries, the aims of this work are to : 1) identify by SSR analysis the accessions preserved in the Tunisian National Collection and those collected directly from different sites of the country (Sidi Bouzid and Bizerte); 2) determine their relatedness to European and American cultivars; and 3) estimate the level of genetic diversity.

MATERIAL AND METHODS

Plant material

Eighty-two almond accessions from different origins (Table 1) were analyzed in this study. Most of them originated from Tunisia (50), the others included in the National Collection were from France (9), Italy (7), Morocco (1), Spain (8), USA (3), or were of unknown origin (4).

The 50 Tunisian local genotypes were either from the region of Bizerte, Nefta, Sfax and Tozeur that are conserved in the National Germplasm Collection of Ettaous, or originated from a recent identification and collection efforts undertaken in the regions of Sidi Bouzid and Bizerte.

Genomic DNA extraction

From all accessions, young leaves were collected for DNA extraction. Total genomic DNA was isolated using the procedure described by Doyle and Doyle (1987). DNA quality was examined by electrophoresis in 0.8% agarose and DNA concentration was quantified spectrophotometrically (Gene Quant, Amersham Pharmacia Biotech, UK). Extracted DNA was diluted to 5 ng μl^{-1} with Tris-EDTA (TE) buffer (1 mM Tris-HCl : 0.1 mM EDTA, pH 8.0) and stored at –20°C for PCR amplifications.

DNA amplification

DNA was amplified by PCR using ten primer pairs of microsatellite (Table 2), nine pairs derived from a library enriched for AG/TC motifs, constructed with the almond cultivar 'Texas' (Mnejja et al. 2005) and one pair previously cited by Joobeur et al. (2000).

Amplification reactions were carried out in a final volume of 15 µl containing 10 ng of template DNA, 1× reaction buffer (20 mM $(NH_4)_2SO_4$, 75 mM Tris-HCl, pH 8.8), 2 mM $MgCl_2$, 50 µM each of dATP, dGTP, dTTP, dCTP (Amersham Pharmacia Biotech, Spain), 0.15 mM of forward and reverse primers each, and 0.5 U of Tth DNA Polymerase (Biotools Band M Labs, S.A., Spain). PCR amplifications were carried out in a Gene Amp 2700 thermocycler (Applied Biosystems, CA, USA) using the following temperature cycles: 1 cycle of 3 min at 95°C; 35 cycles of 1 min at 94°C, 45 s at the corresponding annealing temperature (Table 2) and 1 min at 72°C. The last cycle was followed by a final incubation for 7 min at 72°C and the PCR products were stored at 4°C until analysis. Two independent SSR reactions were performed for each DNA sample. The DNA amplification products were loaded on 5% polyacrylamide sequencing gels. Gels were run for 2 h at 65 W and then silver-stained according to the protocol described by Bassam et al. (1983). Fragment sizes were estimated using 30–330 bp AFLP ladder (Invitrogen, Carlsbad, CA, USA) DNA sizing markers, and analyzed by the Quantity One program (Bio Rad, Hercules, CA, USA).

Diversity parameters

Allelic composition of each accession and total number of alleles were scored for each SSR locus from gel profile analysis. Putative alleles were indicated by the estimated size, in bp. Diversity analysis was performed for accessions with one or two bands per microsatellite using the following parameters: number of alleles per locus (A), number of genotypes per locus (Gn), observed heterozygosity (Ho), expected heterozygosity (He), effective number of alleles (Ne (1/1 – He)), discrimination power (PD $(1 - \sum g_i^2)$, where g_i is the frequency of ith genotype) (Kloosterman et al. 1993), polymorphism information content (PIC), which is the probability that an individual is informative with respect to the segregation of its inherited alleles (Botstein et al. 1980), Wright's fixation index (F (1/1 – Ho/He) (Wright 1951)), frequency of null alleles (Fna), and probability of identity (PI) $(1 - \sum p_i^4 + \sum\sum (2p_ip_j)^2$, where p_i and p_j are the frequency of the ith and jth alleles, respectively), which measures the probability that two randomly drawn diploid genotypes would be identical assuming observed allele frequencies and random assortment (Paetkau et al. 1995). Total probability of identity, defined as the probability of two cultivars sharing the same genetic profile by chance, was also calculated from the individual PI values.

IDENTITY 1.0 (Centre for Applied Genetics, Univ. of Agricultural Sciences, Vienna, Austria) was used to calculate A, Ho, He, PI, Fna and allele frequencies. Those parameters served to evaluate the information given by the microsatellite markers.

Data were analyzed as discrete variables (1) for the presence and (0) for the absence of a similar band. Genetic relationships between the genotypes were calculated using UPGMA cluster analysis of the similarity matrix obtained from the proportion of shared fragments (Nei and Li 1979). Cluster analysis was done using the sequential agglomerative hierarchical nested cluster analysis (SAHN) procedure of NTSYS, which uses the unweighted pair group method with arithmetic averages (UPGMA) to cluster the genotypes. Obtained results were used to construct a final dendrogram showing all accessions and its robustness was evaluated by the cophenetic coefficient computed after the construction of a cophenetic matrix. All analyses were computed with the program NTSYS software ver. 2.1 (Rohlf 2000). Bootstrap support values were obtained from 2000 replicates using TREECON 1.3b (Van de Peer and De Wachter 1994).

Table 1. *List of the cultivars, origin, location and main characteristics of the 82 almond genotypes studied.*

Cultivar	Origin and location	Lineage	Shell hardness	Self compatibility	Flowering date
Abiodh Ras Djebel	Bizerte (Tunisia) – E. C.	old local cultivar	soft	SI	early
Faggoussi	Bizerte (Tunisia) – E. C.	old local cultivar	semi-soft	SI	early
Khoukhi	Bizerte (Tunisia) – E. C.	old local cultivar	semi-hard	SI	early
Harth Nefta	Nefta (Tunisia) – E. C.	seedling selection	semi-hard	SI	very early
Achaak M.	Sfax (Tunisia) – E. C.	seedling of Achaak	semi-hard	SI	very early
Abiodh de Sfax	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Achaak	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Elloumi	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Fekhfekh	Sfax (Tunisia) – E. C;	old local cultivar	hard	SI	very early
Grosse Tendre de Sfax	Sfax (Tunisia) – E. C.	old local cultivar	hard	SI	very early
Guernghzel	Sfax (Tunisia) – E. C.	old local cultivar	hard	SI	very early
Guernghzel CH.	Sfax (Tunisia) – E. C.	old local cultivar	hard	SI	early
Ksontini B	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Mahsouna	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Sahnoun CH.	Sfax (Tunisia) – E. C.	old local cultivar	semi-hard	SI	very early
Triki	Sfax (Tunisia) – E. C.	old local cultivar	hard	SI	very early
Zahaaf	Sfax (Tunisia) – E. C.	old local cultivar	hard	SI	very early
Tozeur 1	Tozeur (Tunisia) – E. C.	seedling selection	hard	SI	very early
Tozeur 2	Tozeur (Tunisia) – E. C.	seedling selection	soft	SI	very early
Tozeur 4	Tozeur (Tunisia) – E. C.	seedling selection	hard	SI	very early
B200	Unknown – E. C.	unknown origin	semi-hard	SI	intermediate
B202	Unknown – E. C.	unknown origin	semi-hard	SI	intermediate
B203	Unknown – E. C.	unknown origin	semi-hard	SI	early
B204	Unknown – E. C.	unknown origin	semi-soft	SI	early
Forme en Boule	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	very early
Forme en Poire	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	very early
Houcine B.N. 2	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-soft	SI	very early
Lakhdhar	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	very early
Port retombant	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	early
Tlili 1	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	very early
Tlili 2	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Tlili 3	Ben Aoun (Sidi Bouzid – Tunisia)	shance seedling	semi-hard	SI	early
Tlili 4	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	very early
Tlili 5	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	early
Tlili 6	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Tlili 7	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Tlili 8	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Tlili 9	Ben Aoun (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	very early
Belgacem N.2	Regueb (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Guernghzel B.N.	Regueb (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	early
Cheikh Sadok 1	Regueb (Sidi Bouzid – Tunisia)	unknown origin	semi-soft	SI	very early
Cheikh Sadok 3	Regueb (Sidi Bouzid – Tunisia)	unknown origin	hard	SI	very early
Cheikh Sadok 4	Regueb (Sidi Bouzid – Tunisia)	unknown origin	hard	SI	very early
Ancetre 1	Ouled Haffouz (Sidi Bouzid – Tunisia)	unknown origin	semi-soft	SI	very early
Bouchouka B.S.	Ouled Haffouz (Sidi Bouzid – Tunisia)	unknown origin	semi-hard	SI	early
Bouchouka K.F.	Ouled Haffouz (Sidi Bouzid – Tunisia)	unknown origin	soft	SI	early
K.F.3	Ouled Haffouz (Sidi Bouzid – Tunisia)	chance seedling	semi-soft	SI	very early
K.F.4	Ouled Haffouz (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	very early
Merghad H.1	Ouled Haffouz (Sidi Bouzid – Tunisia)	chance seedling	hard	SI	very early
Nabil F.	Ouled Haffouz (Sidi Bouzid – Tunisia)	chance seedling	semi-hard	SI	very early
Porto Farina*	Ouled Haffouz (Sidi Bouzid – Tunisia)	unknown origin	soft	SI	very early
Blanco	Bizerte (Tunisia)	old local cultivar	semi-hard	SI	very early
Dillou	Bizerte (Tunisia)	unknown origin	soft	SI	very early

(Continued)

Table 1. *(Continued).*

Cultivar	Origin and location	Lineage	Shell hardness	Self compatibility	Flowering date
Khoukhi Bizerte	Bizerte (Tunisia)	old local cultivar	semi-soft	SI	very early
Bruantine	France – E. C.	old local cultivar	soft	SI	intermediate
Doree	France – E. C.	old local cultivar	hard	SI	intermediate
Ferraduel	France – E. C.	Cristomorto × Aï ♣	hard	SI	late
Ferragnes	France – E. C.	Cristomorto × Aï ♣	semi-hard	SI	late
Fournat de Breznaud	France – E. C.	Marie (1901) ♣	semi-soft	SI	intermediate
Languedoc	France – E. C.	old local cultivar	semi-soft	SI	intermediate
Lauranne	France – E. C.	Ferragness × Tuono ♣	hard	SC	late
Pointue d'Aureille	France – E. C.	old local cultivar	semi-soft	SI	intermediate
Soucaret	France – E. C.	old local cultivar	semi-hard	SI	late
Avola	Italy – E. C.	old local cultivar	hard	SI	intermediate
Cristomorto	Italy – E. C.	unknown origin	hard	SI	late
Fasciuneddu	Italy – E. C.	unknown origin	hard	SI	early
Genco	Italy – E. C.	Genco G. (1910)♣	hard	SC	intermediate
Mazetto syn. Tuono	Italy – E. C.	old local cultivar	hard	SC	late
Pizzuta	Italy – E. C.	old local cultivar	hard	SI	intermediate
Super Nova	Italy – E. C.	mutation from Fascionello♣	hard	SC	late
Ramlet	Morroco – E. C.	unknown origin	hard	SI	intermediate
Desmayo Largueta	Spain – E. C.	old local cultivar	hard	SI	very early
Desmayo Rojo	Spain – E. C.	unknown origin	hard	SI	intermediate
Guara	Spain – E. C.	old local cultivar	hard	SC	late
Malagueña	Spain – E. C.	old local cultivar	semi-soft	SI	late
Marcona	Spain – E. C.	old local cultivar	hard	SI	intermediate
Mas Bovera	Spain – E. C.	Primorskiy × Cristomorto ♣	hard	SI	late
Moncayo	Spain – E. C.	Tardive de la verdiere × Tuono	hard	SI	very late
Tarragona	Spain – E. C	unknown origin	hard	SI	late
Ne Plus Ultra	USA – E. C.	Hatch A.T. (1884)♣	semi-soft	SI	intermediate
Non Pareil	USA – E. C.	Hatch A.T. (1884)♣	soft	SI	intermediate
Peerless	USA – E. C.	unknown origin	hard	SI	late

Note: ♣A.J. Felipe (2000).*This cultivar was identified in Sidi Bouzid but Porto Farina is the native name of a city (actually Ghar El Melh) in Bizerte. E.C: accessions established at the Ettaous National germplasm Collection. SC: Self compatible, SI Self incompatible (for the local genotypes, notification is according to farmers).

RESULTS

Microsatellite polymorphism and genetic diversity

Eighty-two almond genotypes from Tunisia, Europe and USA were analyzed using ten SSRs. All microsatellites produced alleles that could be scored. The parameters of variability analyzed are shown in Table 3. A total of 159 alleles were scored with sizes ranging from 116 bp to 227 bp. The number of alleles per locus (A) varied from 12 in CPDCT022 and CPDCT033 to 23 in CPDCT042, with an average of 15.9 alleles per locus while the number of genotypes (Gn) ranged from 27 in CPDCT033 to 49 in CPDCT042. The effective number of alleles (Ne) ranged from 5 in CPDCT044 to 12 in CPDCT042, with an average of 7.5. Allelic frequencies ranged from 0.006 to 0.367 (data not shown).

Expected heterozygosity (He) ranged from 0.81 in CPDCT044 to 0.92 in CPDCT042 with a mean value of 0.86. Observed heterozygosity (Ho) ranged from 0.49 in CPDCT044 to 0.87 in CPDCT027 with a mean value of 0.68. Observed heterozygosity was slightly lower than the corresponding expected heterozygosity for all loci, except for the CPDCT027 locus, in which the situation was the opposite. Consequently, F-values ranged from –0.05 in CPDCT027 to 0.40 in CPDCT044 with an average of 0.13 showing heterozygote deficiency for the majority of the SSRs. The frequency of null alleles (Fna) ranged from –0.02 in CPDCT027 to 0.17 in CPDCT044 with a mean of 0.09.

Regarding the probability of identity (PI), the maximum (0.09) was observed for CPDCT022, CPDCT038 and CPDCT044 with a respective number of alleles of 12, 14 and 21. The minimum (0.02) was for CPDCT042 with 23 alleles. The average was 0.06, and the total probability of identity was 4×10^{-13}. This low value confirms the

Table 2. *SSR loci used to study the 82 almond genotypes.*

Locus/GenBank accession no.	Primer sequence (5'–3')	Annealing temp (°C)	Motif	Size (bp)	Reference
CPDCT022/AY862459	F: GATCGGCGTCTCCTTTATC R: AAAGCAAGCAGGCAAATGAA	62	(CT)17	133–161	Mnejja et al. 2005
CPDCT025/AY862462	F: GACCTCATCAGCATCACCAA R: TTCCCTAACGTCCCTGACAC	62	(CT)10	172–194	Mnejja et al. 2005
CPDCT027/AY862464	F: TGAGGAGAGCACTGGAGGAG R: CAACCGATCCCTCTAGACCA	62	(CT)19	156–176	Mnejja et al. 2005
CPDCT033/AY862470	F: CAAAACACAAAAACCCACCA R: ATTCGGGGAGTCAATCAGG	62	(CT)18	126–150	Mnejja et al. 2005
CPDCT038/AY862475	F: ATCACAGGTGAAGGCTGTGG R: CAGATTCATTGGCCCATCTT	62	(GA)25	149–181	Mnejja et al. 2005
CPDCT040/AY862477	F: TGATGAGGCCTAGAAATTGGA R: CACAGCAATCAGCAAAAAGC	62	(GA)24	138–170	Mnejja et al. 2005
CPDCT042/AY862479	F: ACGCGTTACAAGTGAGATGC R: TGAAAAATCTTGATGGACGTG	62	(GA)27	164–186	Mnejja et al. 2005
CPDCT044/AY862481	F: ACATGCCGGGTAATTAGCAA R: AAAATGCACGTTTCGTCTCC	62	(GA)21	163–185	Mnejja et al. 2005
CPDCT047/AY862437	F: TCAAAAACACCCATTATTGAA R: AAACATTTAGGGCTTGTTTGG	58	(CT)10	182–204	Mnejja et al. 2005
PS9f8	F: GGTTCTTGGTTATTATGA R: ACATTTCTATGCAGAGTA	60	–	156	Joobeur et al. 2000

efficiency of the microsatellites used in this study for almond genotypes fingerprinting.

Discrimination power (PD) of all loci was very high; it ranged from 0.91 to 0.97 with an average of 0.94. The highest value was found in both of CPDCT025 and CPDCT042 loci. The most informative locus was CPDCT042 with a PIC value of 91% and the highest effective number of alleles (Ne = 12), and number of genotypes (Gn = 49). The CPDCT044 locus was the least informative marker, with a PIC value of 0.79 and Ne of 5. All the 10 SSRs were polymorphic and they were able to distinguish unambiguously the 82 genotypes.

Cluster analysis and genetic relationships among accessions

The diversity among the 82 almond genotypes studied was evaluated according to their genetic similarity based on UPGMA analysis using the similarity matrix generated by the Nei and Li (1979) coefficient. Several dendrograms were possible and the one with the highest cophenetic correlation coefficient (0.73) was chosen. Genetic similarities ranged from 0 to 0.9 with an average of 0.22 (data not shown).

Figure 2 represents the dendrogram with four main clusters (A, B, C and D). In the first cluster (A), 40 of the

Table 3. *Locus name, size range of the amplified fragments, number of alleles (A), effective number of alleles (Ne), expected (He) and observed (Ho) heterozygosities, frequency of null alleles (Fna), probability of identity (PI), Wright's fixation index (F), power of discrimination (PD), polymorphism information content (PIC) and number of genotypes (Gn) were calculated for 10 SSRs markers in 82 almond cultivars.*

Locus	Range size (bp)	A	Ne	He	Ho	Fna	PI	F	PD	PIC	Gn
CPDCT022	133–175	12	6	0.83	0.59	0.13	0.09	0.29	0.93	0.81	30
CPDCT025	162–200	15	10	0.90	0.73	0.09	0.03	0.19	0.97	0.89	40
CPDCT027	156–202	13	6	0.83	0.87	–0.02	0.07	–0.05	0.94	0.81	31
CPDCT033	116–150	12	7	0.85	0.72	0.07	0.08	0.15	0.94	0.83	27
CPDCT038	147–197	14	6	0.82	0.55	0.15	0.09	0.33	0.93	0.80	33
CPDCT040	138–174	14	6	0.84	0.61	0.12	0.08	0.27	0.93	0.82	31
CPDCT042	160–212	23	12	0.92	0.70	0.11	0.02	0.24	0.97	0.91	49
CPDCT044	161–227	21	5	0.81	0.49	0.17	0.09	0.40	0.91	0.79	34
CPDCT047	170–218	20	9	0.89	0.73	0.08	0.04	0.18	0.96	0.87	45
PS9f8	126–178	15	8	0.88	0.81	0.04	0.05	0.08	0.96	0.86	41
Total		159					4×10^{-13}				360
Mean		15.9	7.5	0.86	0.68	0.09	0.06	0.13	0.94	0.84	36

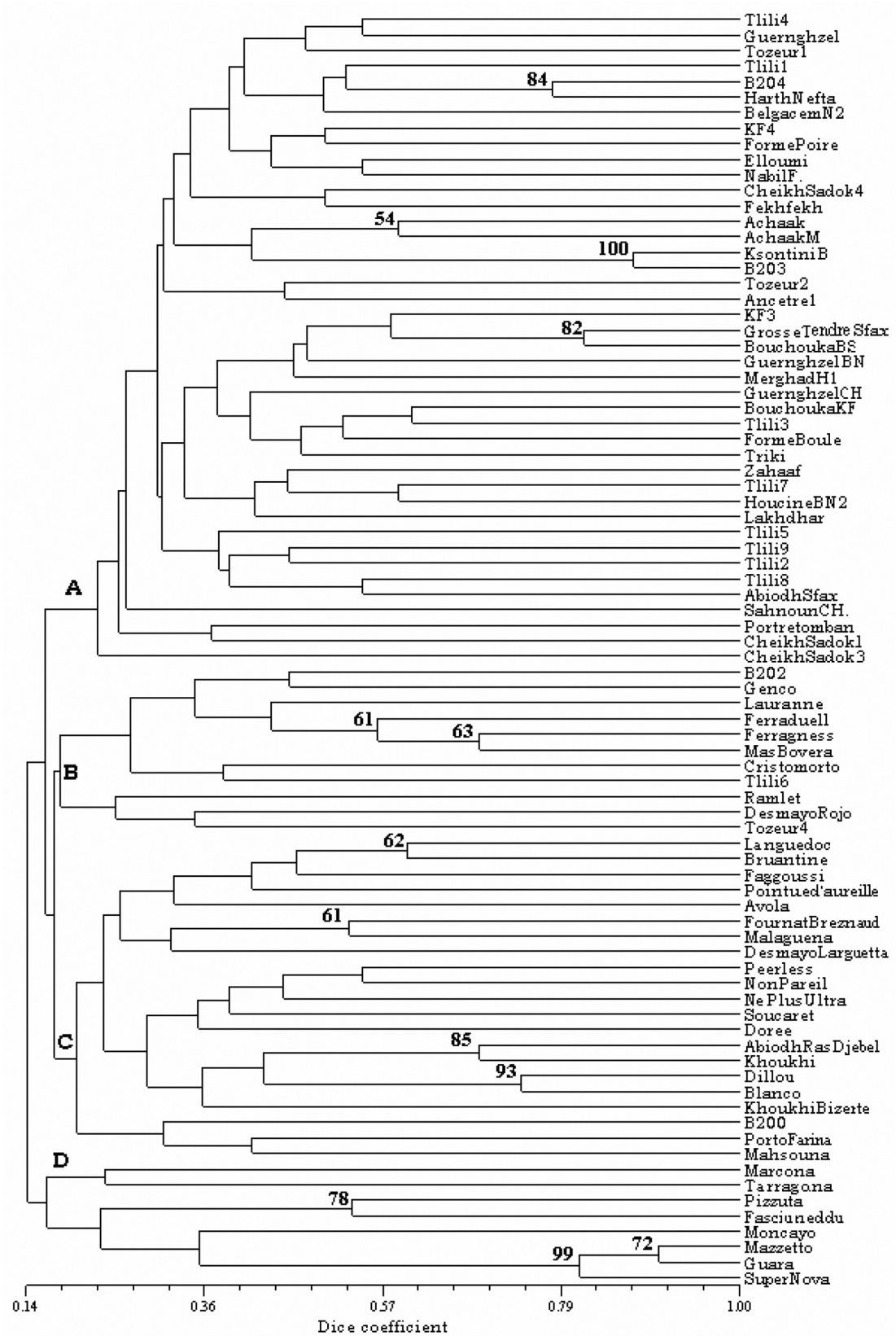

Fig. 2. Dendrogram of 82 almond genotypes based on UPGMA analysis using the similarity matrix generated by the Nei and Li (1979) coefficient with 10 pairs of SSR primers. Bootstrap values out of 2000 replicates are shown if 50% or higher.

50 Tunisian almond genotypes were grouped, in addition to the two of unknown origin ('B203' and 'B204'). The most cultivated genotypes such as 'Achaak', 'Fekhfekh', 'Ksontini B' and 'Zahaaf' (originating from Sfax), in addition to the ecotypes from Sidi Bouzid (except 'Tlili 6' that is included in cluster B, and 'Porto Farina' in cluster C) were present in this cluster. Many genotypes such as 'Guernghzel', 'Guernghzel CH.' and 'Guernghzel B.N.'; 'Bouchouka K.F.' and 'Bouchouka B.S' or 'Achaak' and 'Achaak M' having similar names but different origins, were clustered separately in different sub-clusters and seemed to be homonymous.

In cluster B, 'Lauranne' was in the same sub-cluster with 'Ferraduel', 'Ferragness', 'Mas Bovera', 'Cristomorto' and 'Tlili 6' in addition to 'B202' and 'Genco'.

In cluster C, five Tunisian cultivars: 'Abiodh Ras Djebel', 'Khoukhi', 'Dillou', 'Blanco' and 'Khoukhi Bizerte' were clustered at a genetic similarity of 0.28 with the North American cultivars: 'Peerless', 'Non Pareil' and 'Ne Plus Ultra', in addition to the French cultivars: 'Doree' and 'Soucaret'. In this same sub-cluster, four other French cultivars ('Languedoc', 'Bruantine', 'Pointue d'Aureille' and 'Fournat de Breznaud'), as well as two Spanish ('Malagueña' and 'Desmayo Largeta'), one Italian ('Avola') and one Tunisian ('Faggoussi') cultivars were included. 'B200' and the Tunisian cultivars 'Porto Farina' and 'Mahsouna' formed a second sub-cluster of this group.

Cluster D mostly included some traditional Italian ('Pizzuta', 'Fasciuneddu', 'Mazzetto' and 'Super Nova') and Spanish cultivars ('Marcona', 'Tarragona', 'Moncayo' and 'Guara') with a very high genetic similarity between 'Guara' and 'Mazzetto' (GS = 0.9) and those clustered with 'Super Nova' at a genetic similarity of 0.84.

DISCUSSION

In this work, ten SSRs were used to study the main Tunisian almond accessions preserved in the National Germplasm Collection, as well as some local genotypes collected from the field in Sidi Bouzid, being one of the most important area of almond diversity in the country, and Bizerte. Some European and North American cultivars were also included as references in this study.

An average of 15.9 (alleles locus^{-1}) was observed for the ten SSR studied (Table 3). This value is relatively high compared to the values of 8.4 and 6.6 alleles locus^{-1} obtained in almond by Testolin et al. (2004) and Mnejja et al. (2005), respectively. This may be explained by the higher number of almond genotypes included (82), while these authors used only 16 and 8 genotypes, respectively. For all the SSRs, the size range of the amplified bands was also larger than previously reported by them. Our work probably reflects the presence of new alleles in almond that were not described before.

The average values of the expected (0.86) and observed (0.68) heterozygosity as well as the PIC values (0.84), were slightly higher compared to those previously reported by Mnejja et al. (2005), using nine of the ten primer pairs used in this study. As PIC value provides an estimate of the discrimination power of a marker by taking into account not only the number of alleles at a locus but also the relative frequencies of these alleles, this fact may also be attributed to the higher number of genotypes analyzed in our study.

The power of discrimination (PD) mean value (0.94) was similar to that obtained by Sanchez-Pérez et al. (2006; 0.92), when screening 21 almond cultivars with six primer pairs derived from peach. This relatively higher value compared to that obtained by Mnejja et al. (2005; 0.84) can be explained by the fact that we have selected the nine microsatellite markers presenting the highest PD values from the 31 single locus markers evaluated by these authors.

The majority of the Tunisian genotypes were clustered together but they showed several minor groups, which revealed their high heterogeneity (cluster A of the dendrogram, Fig. 2). This is probably due to the traditional method of propagation of this species all over the country which was mainly done by seeds (open-pollinated), until the more extensively use of grafting in the Mediterranean area at the beginning of the 20th century (Grasselly and Crossa Raynaud 1980). In addition, the need of out-crossing of this species as self incompatible is assumed to be one of the main causes of the existing genetic diversity.

The clear distinction between the majority of local cultivars from the central and southern part and all of the other groups which was not previously demonstrated using RAPD (Gouta et al. 2008) is a proof of the higher discrimination power of SSRs compared to RAPD. The bootstrap values of 84% and 100% supporting the relatedness among'B204' and 'Harth Nefta' from one side and 'B203' and 'Ksontini B' from the other, confirms the hypothesis of a local origin of these two unknown cultivars.

In contrast to what has been observed in group A (Fig. 2), the local cultivars from Bizerte (north of Tunisia), which belongs to the humid and sub humid bioclimatic zone, with rainfalls greater than 700 mm year^{-1}, were clustered in the group C with some European and all the North American cultivars. In fact, the position of this area in the extreme north of Tunisia probably favoured the exchange of genotypes between both shores of the Mediterranean Sea. The presence in group C of the two cultivars: 'Porto Farina' as was the old name of Ghar El Melh (a city in Bizerte) and 'Faggoussi' could be another fact in favour of this hypothesis. The fact that the three American almond cultivars: 'Peerless', 'Non Pareil' and 'Ne Plus Ultra', clustered together with the French cultivars pool: 'Doree', 'Languedoc', 'Bruantine', 'Pointue d'Aureille' and 'Fournat de Breznaud', is an evidence for their origin as independent selections from the same initial French germplasm pool (Hauagge et al.

1987; Kester 1994; Bartholozi et al. 1998). Moreover, 'Ne Plus Ultra' and 'Non Pareil' were mentioned to originate from seedlings selected by A. T. Hatch in California in 1979, originated from material of the Languedoc region of France (Kester 1994). The high bootstrap values observed in the sub cluster grouping cultivars from Bizerte (85% for 'Abiodh Ras Djebel' and 'Khoukhi' and 93% for 'Dillou' and 'Blanco') support the specificity of this site.

Since few previous cluster analysis and genetic relationship for almond have been done using genomic microsatellites derived from almond, the present work reveals new genetic similarity values among many local and foreign cultivars.

The presence of 'Ferragness', 'Ferraduel', 'Lauranne' and 'Mas Bovera' in the same sub-cluster with 'Cristomorto', previously obtained by Martins et al. (2003) is supported by statements concerning their genetic origin (Table 1). In fact, 'Ferraduel' and 'Ferragness' were selected from seedlings of the cross-pollination between 'Cristomorto' × 'Aï', 'Lauranne' is a selection of 'Ferragness' × 'Tuono', while 'Mas Bovera' derives from a cross of 'Primorskiy' × 'Cristomorto' (Felipe 2000). Genetic similarity between 'Ferragness' and 'Ferraduel' was 0.57, which is different from the value (1.0) found by Xu et al. (2004). The same authors found a higher value of GS between 'Non Pareil' and 'Ne Plus Ultra' (0.73) compared to our value (0.45). These differences may be due to the fact that these authors have used EST derived SSRs that generally display lower polymorphism than genomic ones (Cho et al. 2000; Scott et al. 2000; Eujayl et al. 2002).

The close relationship found between 'Mazzetto syn. Tuono' and 'Guara' (cluster D) with a genetic similarity of 0.90 was already observed by Martins et al. (2003; GS = 0.87), who concluded that 'Guara' is probably a seedling of 'Tuono'. Relatedness between 'Moncayo' and 'Tuono' was already proposed by Shiran et al. (2007) and it has also been confirmed in this work. The presence of 'Super Nova' in the same sub-cluster with these last two cultivars could be due to the fact that they have some common agronomic background as happens with many cultivars showing the self-compatibility trait. This subgroup classification is also supported by the high bootstrap values (72 and 99%). Nevertheless, this could be furthermore confirmed by the use of some microsatellites from the linkage group 6 where the self-compatibility trait is located.

The cultivar 'Khoukhi Bizerte' collected directly from the field in Bizerte was added to the analysis to test the authenticity of the specimen 'Khoukhi' existing in the National Collection of Ettaous. The low genetic similarity found (0.42) between them, suggest that they are probably homonymous and as a consequence respective corrections should be made in the collection.

Finally, the presence of the genotypes 'B202', 'Tlili 6', 'Tozeur 4', 'B200' and 'Mahsouna' out of cluster A raises the question of their possible relatedness to the European genotypes and opens the way for further investigations.

The fixation index average (F = 0.13) shows a deficit of heterozygosity for nine of the ten loci. This could be explained by the population structure and/or inbreeding like effect. In fact, the need of cross-pollination for the majority of almond cultivars as self-incompatible and the historical origin of almond along the shores of the Mediterranean Sea are strong statements in favour of these hypotheses.

This study reveals the high diversity and the distinct origin of the Tunisian almond germplasm and can be considered as a first step in understanding the parental relationships and the origin of local and traditional cultivars grown in Tunisia.

In summary, SSRs analysis has been successfully used to examine the crop origin, geographic divergence and distribution as well as for revealing synonymous and mislabelling in the Tunisian germplasm. All the Tunisian genotypes except the northern cultivars from Bizerte, were genetically distant from the European and American cultivars studied. The great diversity found in the Tunisian almond germplasm supports the idea that Tunisia has a valuable source of almond genes to be exploited in further international breeding programs. We advice that the local genotypes collected from Sidi Bouzid and Bizerte, described for the first time in this study, should be included and preserved in the National Germplasm Collections of Tunisia.

Acknowledgements – This research was supported in part by the Tunisian Ministry of Higher Education, Scientific Research and Technology, the Spanish Agency for International Cooperation (A/5339/06 and A/8334/07), the Spanish MICINN (Ministry of Science and Innovation) grants AGL2008-00283/AGR, and the Regional Government of Aragon funds (A44). We are grateful to Dr Monji Msallem, the Head of the Unity of Genetic Resources and Breeding of Olive Tree, Almond and Pistachio, and to Dr Meriam Bouhadida for their helpful support and precious advices.

REFERENCES

Aranzana, M. J., Garcia-Mas, J., Carbó J. et al. 2002. Development and variability analysis of microsatellite markers in peach. – Plant Breed. 121: 87–92.

Ayadi, M., Ghrab, M., Gargouri, K. et al. 2006. Kernel characteristics of almond cultivars under rainfed conditions. – Acta Hort. 726: 377–382.

Bartolozzi, F., Warburton, M. L., Arulsekar, S. et al. 1998. Genetic characterization and relatedness among California almond cultivars and breeding lines detected by randomly amplified polymorphic DNA (RAPD) analysis. – J. Am. Soc. Hort. Sci. 123: 381–387.

Bassam, B. J., Caetano-Anoelles, G. and Gresshoff, P. M. 1983. Fast and sensitive silver staining of DNA in polyacrylamide gels. – Anal. Biochem. 196: 80–83.

Botstein, D., White, R. L., Skolnick, M. et al. 1980. Construction of genetic linkage map in man using restriction fragment length polymorphisms. – Am. J. Hum. Genet. 32: 314–331.

Bouhadida, M., Casas, M. A., Moreno, M. A. et al. 2007. Molecular characterization of Miraflores peach variety and relatives using SSRs. – Sci. Hort. 111: 140–145.

Cho, Y. G., Ishil, T., Temnykh, S. et al. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). – Theor. Appl. Genet. 100: 713–722.

Cipriani, G., Lot, G., Huang, W. G. et al. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: isolation, characterization and cross-species amplification in Prunus. – Theor. Appl. Genet. 99: 65–72.

Dangl, G.S., Yang, J., Golino, D.A. et al. 2009. A practical method for almond cultivar identification and parental analysis using simple sequence repeat markers. – Euphytica 168: 41–48.

Dirlewanger, E., Cosson, P., Tavaud M. et al. 2002. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). – Theor. Appl. Genet. 105: 127–138.

Downey, S. L. and Iezzoni, A. F. 2000. Polymorphic DNA markers in black berry cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach and sour cherry. – J. Am. Soc. Hort. Sci. 125: 76–80.

Doyle, J. J. and Doyle, J. L. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. – Phytochem. Bull. 19: 11–15.

Eujayl, I., Sorrells, M. E., Baum, M. et al. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. – Theor. Appl. Genet. 104: 399–407.

Felipe, A. J. 2000. Variedades de almendro. – In: Integrum (eds), El Almendro. Zaragoza, Spain, p. 204–279.

Ghrab, M., Ben Mimoun, M., Triki, H. et al. 2002. Yield of 24 almond cultivars in a dry area climate in Tunisia: five years of study. – Acta Hort. 591: 473–478.

Gouta, H., Ksia, E., Zoghlami, N. et al. 2008. Genetic diversity and phylogenetic relationships among Tunisian almond cultivars revealed by RAPD markers. – J. Hort. Sci. Biotechnol. 83: 707–712.

Grasselly, C. and Crossa-Raynaud P. 1980. – In: Maisonneuve and Larose (eds), L'Amandier. Paris XII, France, p. 71–142.

Hauagge, R., Kester, D. E., Arulsekar, S. et al. 1987. Isozyme variation among California almond cultivars. II. Cultivar characterization and origins. – J. Am. Soc. Hort. Sci. 112: 693–698.

Hormaza, J. I. 2002. Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. – Theor. Appl. Genet. 104: 321–328.

Jaouani, A. 1976. La culture de l'amandier en Tunisie. – Cahiers Options Méditerranéennes, 32: 67–72.

Joobeur, T., Periam, N., De Vicente, M. C. et al. 2000. Development of a second generation linkage map for almond using RAPD and SSR markers. – Genome 43: 649–655.

Kester, D. E. 1994. Almond cultivar and breeding programmes in California. – Acta Hort. 373: 13–28.

Klosterman, A. D., Budowle, B. and Daselaar, P. 1993. PCR-amplification and detection of the human DIS80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. – Int. J. Leg. Med. 105: 257–264.

Lopes, M. S., Sefc, K. M., Laimer, M. et al. 2002. Identification of microsatellite loci in apricot. – Mol. Ecol. Notes 2: 24–26.

Maghuly, F., Fernandez, E. B., Ruthner, S. et al. 2005. Microsatellite variability in apricots (*Prunus armeniaca* L.) reflects their geographic origin and breeding history. – Tree Genet. Genomes 1: 151–165.

Martínez-Gómez, P., Arulsekar, S., Potter, D. et al. 2003a. An extended interspecific gene pool available to peach and almond breeding as characterized using simple sequence repeat (SSR) markers. – Euphytica 131: 313–322.

Martínez-Gómez, P., Arulsekar, S., Potter, D. et al. 2003b. Relationships among peach and almond and related species as detected by SSR markers. – J. Am. Soc. Hort. Sci. 128: 667–671.

Martins, M., Tenreiro, R. and Oliveira, M. 2003. Genetic relatedness of Portuguese almond collection assessed by RAPD and ISSR markers. – Plant Cell Rep. 22: 71–78.

Messina, R., Lain, O., Marrazzo, T. et al. 2004. New set of microsatellite loci isolated in apricot. – Mol. Ecol. Notes 4: 432–434.

Mnejja M., Garcia-Mas, J., Howad, W. et al. 2004. Simple-sequence repeat (SSR) markers of Japanese plum (*Prunus salicina* Lindl.) are highly polymorphic and transferable to peach and almond. – Mol. Ecol. Notes 4: 163–165.

Mnejja, M., Garcia-Mas, J., Howad, W. et al. 2005. Development and transportability across *Prunus* species of 42 polymorphic almond microsatellites. – Mol. Ecol. Notes 5: 531–535.

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.

Paetkau, D., Calvert, W., Stirling, I. et al. 1995. Microsatellite analysis of population structure in Canadian polar bears. – Mol. Ecol. 4: 347–354.

Rohlf, F. J. 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system, ver. 2.1. – Exeter Publishing, Setauket, NY.

Sánchez-Pérez, R., Ballester, J., Dicenta, F. et al. 2006. Comparison of SSR polymorphisms using automated capillary sequencers, and polyacrylamide and agarose gel electrophoresis: implications for the assessment of genetic diversity and relatedness in almond. – Sci. Hort. 108: 310–316.

Scott, K. D., Eggler, P., Seaton, G. M. et al. 2000. Analysis of SSRs derived from grape ESTs. – Theor. Appl. Genet. 100: 723–726.

Shiran, B., Amirbakhtiar, N., Kiani, S. et al. 2007. Molecular characterization and genetic relationship among almond cultivars assessed by RAPD and SSR markers. – Sci. Hort. 111: 280–292.

Testolin, R., Messina, R., Lain, O. et al. 2004. Microsatellites isolated in almond from an AC-repeat enriched library. – Mol. Ecol. Notes 4: 459–461.

Van de Peer, Y. and R. de Wachter. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. – Comput. Appl. Biosci. 10: 569–570.

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

Wünsch, A. and Hormaza, J. I. 2002. Molecular characterization of sweet cherry (*Prunus avium* L.) genotypes using peach (*Prunus persica* L.) SSR sequences. – Heredity 89: 56–63.

Xu, Y., Ma, R. C., Xie, H. et al. 2004. Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. – Genome 47: 1091–1104.

Zeinalabedini, M., Majourhat, K., Khayam-Nekoui, M. et al. 2008. Comparison of the use of morphological, protein and DNA markers in the genetic characterization of Iranian wild *Prunus* species. – Sci. Hort. 116: 80–88.