© NIAB 2013
ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization; 1–11
doi:10.1017/S1479262113000543

Bottleneck and gene flow effects impact the genetic structure of seed-propagated apricot populations in Moroccan oasis agroecosystems

Ali Mamouni[1,2], Ahmed El Bakkali[1,3], Patrick Lambert[4], Lamia Krichen[5], Ahmed Oukabli[1], Jean Marc Audergon[4], Philippe Chatelet[3], Cherkaoui El Modafar[2] and Bouchaib Khadari[3,6]*

[1]*INRA, UR Amélioration des Plantes et Conservation des Ressources Phytogénétiques, Meknés, Morocco,* [2]*Laboratoire AgroBiotech L02B005, Faculté des Sciences et Techniques Guéliz, Université Cadi Ayyad, 40000 Marrakech, Morocco,* [3]*INRA, UMR 1334 Amélioration Génétique et Adaptation des Plantes (AGAP), F-34398 Montpellier, France,* [4]*INRA Centre PACA – UR1052 Génétique et Amélioration des Fruits et Légumes GAFL, Domaine St Maurice, CS60094, 84143 Montfavet Cedex, France,* [5]*Laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie, Faculté des Sciences de Tunis, Université Tunis El Manar, Campus Universitaire El Manar, 2092 Tunis, Tunisia and* [6]*CBNMed, UMR 1334 AGAP, F-34398 Montpellier, France*

Received 19 September 2013; Accepted 11 November 2013

Abstract

In order to highlight the genetic status and origin of Moroccan apricot populations, trees were collected from ten oasis agroecosystems and analysed with AFLP markers. A total of 87 accessions and 12 cultivars grown in Moroccan orchards, including 'Canino' and 'Del Patriarca' cultivars, were surveyed and compared with *in situ* Tunisian and *ex situ* Montfavet (France) collections. Our results highlighted a narrow genetic diversity in the Maghreb region (Tunisia and Morocco) associated with a strong differentiation from the other groups, which supports a bottleneck effect. A similar model was illustrated at a finer geographical scale, i.e. the Draa Valley in Morocco. Genetic structure appeared as two major clusters subdivided into six sub-clusters in which Moroccan germplasm constituted specific groups in comparison with other Mediterranean apricots. Moroccan germplasm was classified into three sub-clusters, two of which were formed by genotypes related to 'Del Patriarca' and 'Canino', respectively. The present study highlights the wide Moroccan apricot's diversity in traditional agroecosystems, and also suggests a substantial gene flow occurring from recently introduced cultivars ('Canino' and 'Del Patriarca') to local apricot populations, thus leading to local germplasm diversification through seedling propagation. If we consider its geographical position, the historical diffusion of the species and farming practices, Morocco could be viewed as an additional centre of secondary diversification for apricot. Understanding the origin and specificity of local apricot populations is crucial for managing local collections in regard to adaptive traits for arid and Saharan conditions as well as for introducing local genetic resources into current breeding programmes.

Keywords: amplified fragment length polymorphism; genetic diversity; genetic drift; introgression; *Prunus armeniaca* L

* Corresponding author. E-mail: khadari@supagro.inra.fr

Introduction

Apricot (*Prunus armeniaca* L.) originates from a wide area between the north-east of China and Uzbekistan (Couranjou, 1975). Taking into account agronomic and geographic criteria, a genetic study of apricot using amplified fragment length polymorphism (AFLP) markers has allowed for the resolution of four groups termed 'diversification (D)', 'geographical adaptation (C)', 'continental Europe (B)' and 'Mediterranean (A)' (Hagen *et al.*, 2002). A significant loss of genetic diversity has been revealed, declining from the 'Irano-Caucasian' gene pool, considered as a secondary centre of diversification, to the northern and south-western Mediterranean Basin (Hagen *et al.*, 2002; Bourguiba *et al.*, 2012a). Such a finding highlights a marked domestication bottleneck in Mediterranean apricots.

According to apricot dissemination routes from its centre of origin, Faust *et al.* (1998) identified three major groups: (1) an Asian group, close to the centre of origin; (2) the European group; (3) the North African group. Bourguiba *et al.* (2012a) highlighted the ancestor of these latter groups, which probably issued from the 'Irano-Caucasian' gene pool through two diffusion routes: a Northern route and a South Mediterranean one. A probable exchange between the northern and southern Mediterranean countries during the Andalusian period would have diversified apricot populations in the Maghreb region (namely Morocco, Algeria, Tunisia and Libya), as attested by the presence of accessions admixed between the local gene pool in North Africa and introduced cultivars (Bourguiba *et al.*, 2013).

In the traditional oasis agroecosystems in Morocco, apricot is seed-propagated and grown under contrasting ecological conditions; through human selection pressure, we expect to find populations adapted to arid and Saharan conditions. These populations probably display a narrow genetic basis since they result from processes of diffusion into an extreme western geographic zone, combined then to bottleneck effects as has been shown previously in Tunisia (Khadari *et al.*, 2006; Bourguiba *et al.*, 2010) and in the Maghreb region (Bourguiba *et al.*, 2012b). Moreover, significant historical differences (no Ottoman presence in Morocco and a relatively short period of French presence) would have probably given rise to a distinct genetic status of apricot populations in Morocco compared with Algeria and Tunisia. Selected cultivars such as 'Canino' and 'Del Patriarca' have been introduced in the middle of the 20th century and cultivated in modern orchards in the region of Marrakech, Morocco (Barbeau and Bouami, 1980). These cultivars were sometimes found growing in the traditional oasis agroecosystems where, traditionally, apricot is exclusively seed-propagated. Thus, the coexistence of these two genetic entities and the use of seed propagation would have probably resulted in genetic admixture.

Focusing on the genetic structure of apricot under the agroecological conditions of the Moroccan oasis agroecosystems, we are confronted with three main questions: (1) what is the genetic status of seed-propagated populations in Morocco? (2) Do they constitute a specific group compared with the apricot reference germplasm maintained in INRA Montfavet France? (3) Do they reveal a specific centre of diversification? In this study, we aimed to determine the origin of local seed-propagated populations called 'Mech-mech' and understand the process of apricot diversification in this extreme zone of diffusion.

Materials and methods

Plant material

A total of 99 Moroccan apricot accessions were analysed using AFLP markers. Among these, 87 local seed-derived trees were collected from ten oasis agroecosystems located in the four main geographical regions in the south-east of Morocco, each with distinct ecological conditions (Fig. 1; Table S1, available online; Barbeau and Bouami, 1980): Moulouya Valley [Guercif (2 apricot trees), Outat El-Haj (6) and Missour (11)]; Ziz Valley [Er-Rich (12), Aoufous (6) and Goulmima (6)]; Dades Valley [Boumalen (10) and El-Kalaa M'Gouna (9)]; Draa Valley [Skoura (12) and Agdez (13)]. In addition, twelve cultivars grown at the Ain Taoujdate Experimental Station (INRA-Meknes, Morocco) were studied, including eight accessions identified by Barbeau and Bouami (1980) ('Ait Gmat n.5', 'Amez n.1', 'Khorbat n.6', 'Mansouri n.15' and 'Marouch n.1, 3, 4 and 16') and four introduced cultivars ('Canino', 'Del Patriarca', 'Gélitano' and 'Maoui', a seedling of 'Del Patriarca'). Besides these accessions, 84 additional cultivars were used, i.e. 31 Tunisian accessions described in Khadari *et al.* (2006) and 53 accessions from the INRA Montfavet (France) collection originating from different countries: Europe; North Africa; Turkey; Iran; China. These cultivars have already been studied by Hagen *et al.* (2002) and Khadari *et al.* (2006) and were analysed in this study with the same markers for comparison with local accessions.

DNA preparation and AFLP genotyping

DNA of Moroccan apricot accessions was extracted from lyophilized young leaves according to protocols described by Bernatzky and Tanksley (1986) and Lefevre *et al.* (1993). AFLP analysis was conducted with

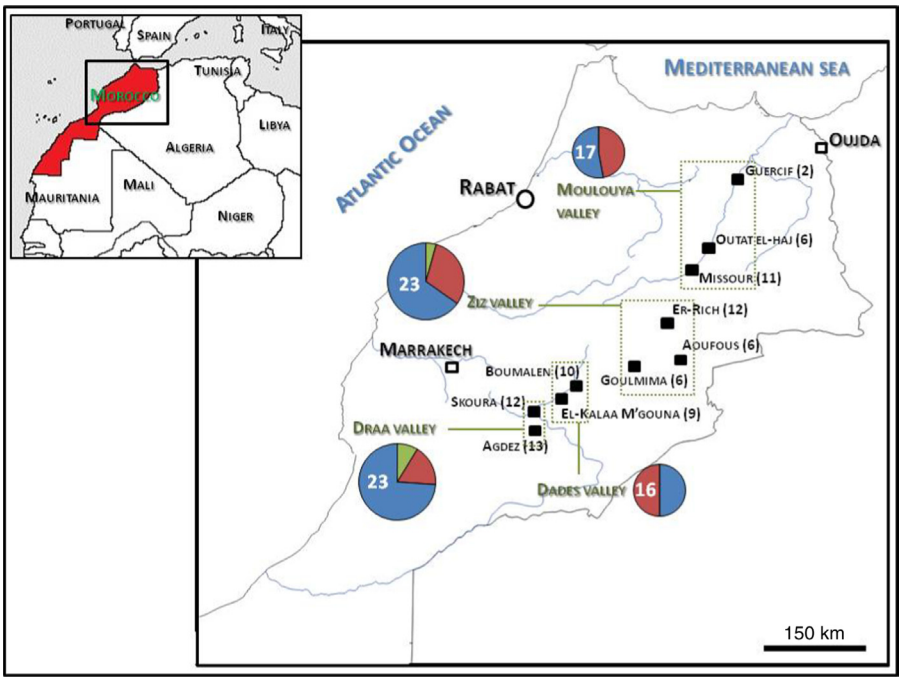

Fig. 1. (colour online) Sampling locations of Moroccan apricot trees in ten geographical sites and four zones, respectively. The numbers in brackets indicate the number of trees sampled within each site. The pie diagrams indicate the proportion of genotypes assigned to 'Del Patriarca'-related genotypes (green), 'Canino'-related genotypes (brown) and seed-propagated genotypes (blue) in each sampling zone. The number within the pie diagrams indicates the number of genotypes assigned at a membership probability ≥ 0.8.

154 markers resulting from the same *Eco* RI–*Mse* I primer combination, as has been previously used in Hagen *et al.* (2002) and Khadari *et al.* (2006) on apricot germplasm from INRA Montfavet France and on Tunisian apricots, respectively. AFLP markers were read on the autoradiograms by two individuals independently and scored for presence/absence. Only data validated by the two individuals were kept for further analyses.

Data analysis

An AFLP binary matrix without missing data was established by scoring the presence of an AFLP band as 1 and its absence as 0. Allelic frequencies of AFLP markers were computed by a non-uniform prior distribution method using the AFLP-SURV v.1.0 package (Vekemans, 2002) from the observed frequencies of fragments using the Bayesian approach proposed by Zhivotovsky (1999) for diploid species, assuming that populations reached Hardy–Weinberg equilibrium. Estimated allelic frequencies were used to compute the parameters of genetic diversity in different groups defined according to genetic data or to the area of origin: the number and proportion of polymorphic markers at the 5% level, genetic diversity H_j (the average of expected heterozygosity under Hardy–Weinberg equilibrium) and its standard deviation. Furthermore, the parameters of genetic diversity within and between groups were calculated at different levels of analysis using the method proposed by Lynch and Milligan (1994): overall (H_T) and intra-group (H_S; within the group of cultivars) genetic diversity and Wright's fixation index (F_{ST}; proportion of genetic differentiation among the group of cultivars). The significance of genetic differentiation between groups was tested by comparison of the fixation index F_{ST} observed with a distribution of F_{ST} with no genetic structure using 1000 permutations. The similarity between different groups was assessed using F_{ST} genetic differentiation according to the neighbour-joining algorithm. A bootstrap analysis with 1000 replicates was performed to evaluate the robustness of genetic relationships.

To assess apricot genetic structure, three analyses were carried out: (1) construction of a Ward dendrogram (Ward, 1963) based on the simple matching coefficient (Sokal and Sneath, 1963) between genotype pairs using Clustering Calculator (http://www2.biology.ualberta.ca/jbrzusto/cluster.php) and FigTree v.1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/); (2) a principal coordinate analysis (PCoA) as implemented in the Darwin v.5.0.137 program (Perrier *et al.*, 2003) using the simple matching coefficient as the distance matrix to describe the spatial distribution of genotypes; (3) a model-based Bayesian clustering analysis as implemented in the Structure

program (Pritchard *et al.*, 2000) using two independent analyses – (1) the complete apricot sample including Moroccan, Tunisian and *ex situ* germplasm collections and (2) the Moroccan apricot sample only. Both analyses were performed without prior information about the geographical origin of the material. The Structure program was run using the admixture model with correlated allele frequencies, with the number of genetic clusters (K) varying from $K = 1$ to $K = 10$, and ten trials for each K cluster. Each run involved a burn-in/Markov chain Monte Carlo of 200,000 iterations and a post-simulation length of 1,000,000. The reliability of the number of K clusters was examined using the *ad hoc* measure ΔK of Evanno *et al.* (2005) with the R v2.13.0 program (R Development Core Team, 2011), and the similarity coefficient between runs (H') for the same K clusters was calculated by the Clumpp program (Jakobsson and Rosenberg, 2007).

Results

Characterization of apricot accessions

Based on the five AFLP primer combinations previously used by Hagen *et al.* (2002) and Khadari *et al.* (2006), 183 accessions including the *ex situ* germplasm and the *in situ* Tunisian and Moroccan apricots were analysed using 154 AFLP markers which identified 180 distinct AFLP profiles (genotypes). All cultivars from the INRA Montfavet collection were found to be distinct. By contrast, one and two accession pairs, respectively, were found to be redundant in the Tunisian and Moroccan germplasm: 'Bangui 1'/'Bangui 2' for the Tunisian germplasm; 'Skoura 2'/'Skoura A1' and 'Boumalen Kh2'/'Boumalen A1' for the Moroccan germplasm.

Based on the 180 genotypes identified, pairwise comparisons were conducted using a simple matching coefficient. Two accession pairs from the *ex situ* germplasm were found to be genetically close: 'A2210 Goldrich'/'A2218 Goldrich' and 'Tokaloglu'/'Veecot' were distinct by one and three AFLP markers, respectively. Within the *in situ* Tunisian germplasm, the apricot pairs 'Zbidi'/'Messelmani' and 'Bou Herra'/'Oud Ras Jbel' from Kairouan were observed to be distinct by only one AFLP marker, while the pair 'Oud Nakhla'/'Ouled El Oud' from Testour were found to be distinct by three AFLP markers. By contrast, the closest Moroccan accessions were found to be distinct by four AFLP markers: 'El-Kalaa M'Gouna 1'/'El-Kalaa M'Gouna 3' and 'Agdez G1'/'Er-Rich TA1'.

Compared with graft-propagated cultivars, 'Er-Rich TA1' from the Ziz Valley was distinct from 'Maoui' and 'Del Patriarca' by one and two AFLP markers, respectively. These two grafted cultivars were found to be genetically close (three dissimilar AFLP markers), indicating that 'Maoui' was probably derived from 'Del Patriarca'. The seed-propagated accession 'Outat El-Haj 2' from the Moulouya Valley was found to be closest genetically to 'Canino' in the Ain Taoujdate collection (only seven dissimilar AFLP markers).

Genetic structure within and among the groups

Based on the simple matching similarity and Ward algorithm, the 180 apricot genotypes were classified into two major clusters [foreign cultivars (C1) and Moroccan germplasm (C2) clusters] and six sub-clusters (Fig. 2). The first cluster C1 included all the groups defined by Hagen *et al.* (2002) and the Tunisian germplasm. The cultivars were assigned to the same groups determined by Hagen *et al.* (2002); however, 'Précoce de Tyrinthe' and 'A114 Bergeron', previously classified in the C and B groups, respectively, were moved to the Mediterranean group A in the present analysis. 'Andswee' (assigned previously to group D), 'Polonais' (A) and 'Hamidi' (A) were moved to group B, and 'Chinese' and 'Harcot', previously assigned to group D, were moved to group C, whereas '2121A4', initially assigned to group A, was moved to group D. Based on the dendrogram in Fig. 2, the diversification 'D' and geographically adaptable 'C' groups clustered into one sub-cluster (C1-1) and continental Europe group 'B' in a second one (C1-2), while the Tunisian genotypes were clearly close to the Mediterranean group 'A' and thus considered as a third sub-cluster (C1-3). In fact, among the 30 identified cultivars, 29 were assigned to the sub-cluster C1-3.

With the exception of 'Mansouri n.15', 'Er-Rich TIL 2' and 'Outat El-Haj 7', which were classified into the second sub-cluster (C1-2), all the other Moroccan genotypes (94) were classified into the second cluster (C2; Moroccan germplasm) together with the cultivars 'Rouge du Roussillon' (French) and 'Khad Halima' (Tunisian). Within this cluster, three sub-clusters were resolved with no clear structure according to the geographical origin of accessions. The first sub-cluster (C2-1) was constructed with 53 genotypes including 'Rouge du Roussillon'. The sub-cluster C2-2 termed the 'Del Patriarca' group included four seed-propagated accessions and four cultivars: Tunisian 'Khad Hlima', Spanish 'Del Patriarca' and the local selected 'Maoui' and 'Khorbat n.6'. The sub-cluster C2-3 termed the 'Canino' group included 31 seed-propagated accessions and four cultivars: 'Canino', 'Gélitano', 'Amez n.1' and 'Marouch n.16'.

Similar results were observed using PCoA (Fig. S1, available online). In fact, the first two axes explained 16.1% of the total variation and the first axis allowed for the

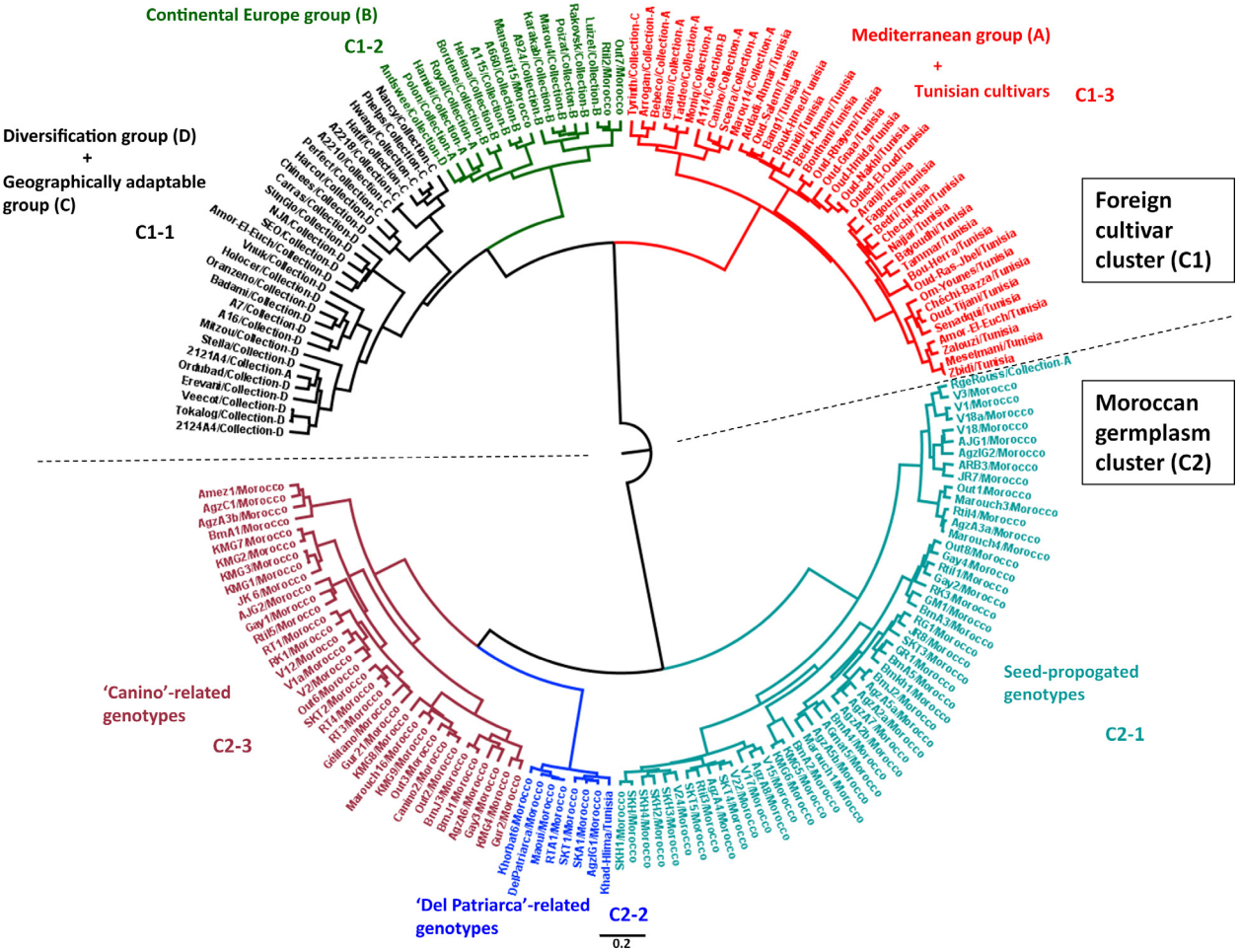

Fig. 2. (colour online) Ward's dendrogram of apricot material (180 genotypes) as constructed based on the simple matching similarity of 154 polymorphic AFLP markers. The scale bar represents the simple matching similarity distance. A total of six sub-clusters were identified: C1-1 to C2-3.

separation of apricots into two main clusters, similar to that shown by the dendrogram (Fig. 2). The sub-cluster C1-1 spanned a wide surface in the scatter plot, indicating its high diversity, while both sub-clusters C1-2 and C1-3, characterized by a low diversity, were plotted in the restricted zones of the plot. Unlike the sub-cluster C2-2 ('Del patriarca' group), the sub-clusters C2-1 and C2-3 ('Canino' group) spanned a wide diversity range.

The results obtained using the Structure Bayesian analysis program for all the 180 genotypes indicated that the $K = 2$ and $K = 3$ clusters were the best genetic structure models based on the *ad hoc* measure ΔK of Evanno *et al.* (2005) and the similarity coefficient among the ten runs (H'; Jakobsson and Rosenberg, 2007; Fig. 3). Based on this premise, at $K = 2$ ($\Delta K = 302.99$; $H' = 0.998$), the Moroccan germplasm was observed to be distinct from the other genotypes, while at $K = 3$ ($\Delta K = 160.02$; $H' = 0.998$), the following three clusters were identified: (1) D–C groups; (2) B–A and *in situ* Tunisian apricots; (3) Moroccan apricot accessions. Four Moroccan accessions showed an admixture signature with other clusters at $K = 2$ and $K = 3$.

Focusing on the Moroccan germplasm, no clear genetic structure was observed according to the area of origin using the Structure program, whereas the *ad hoc* measure ΔK and the similarity coefficient H' indicated that $K = 2$ and $K = 3$ were the best genetic structure model (Fig. 4). Similar results were obtained when allelic frequencies were computed assuming a deviation of $F_{is} = 0.2$ from Hardy–Weinberg equilibrium (data not shown). At $K = 3$, a clear separation was observed between 'Canino' (brown)- and 'Del Patriarca' (green)-related genotypes. The first cluster comprised six genotypes including 'Del Patriarca' and the local cultivar 'Maoui', the second cluster held 31 accessions including the 'Canino' and 'Gélitano' cultivars, whereas the third cluster with the largest size ($n = 53$) comprised seed-propagated accessions especially from the Ziz and Draa valleys (65.3 and 73.9%

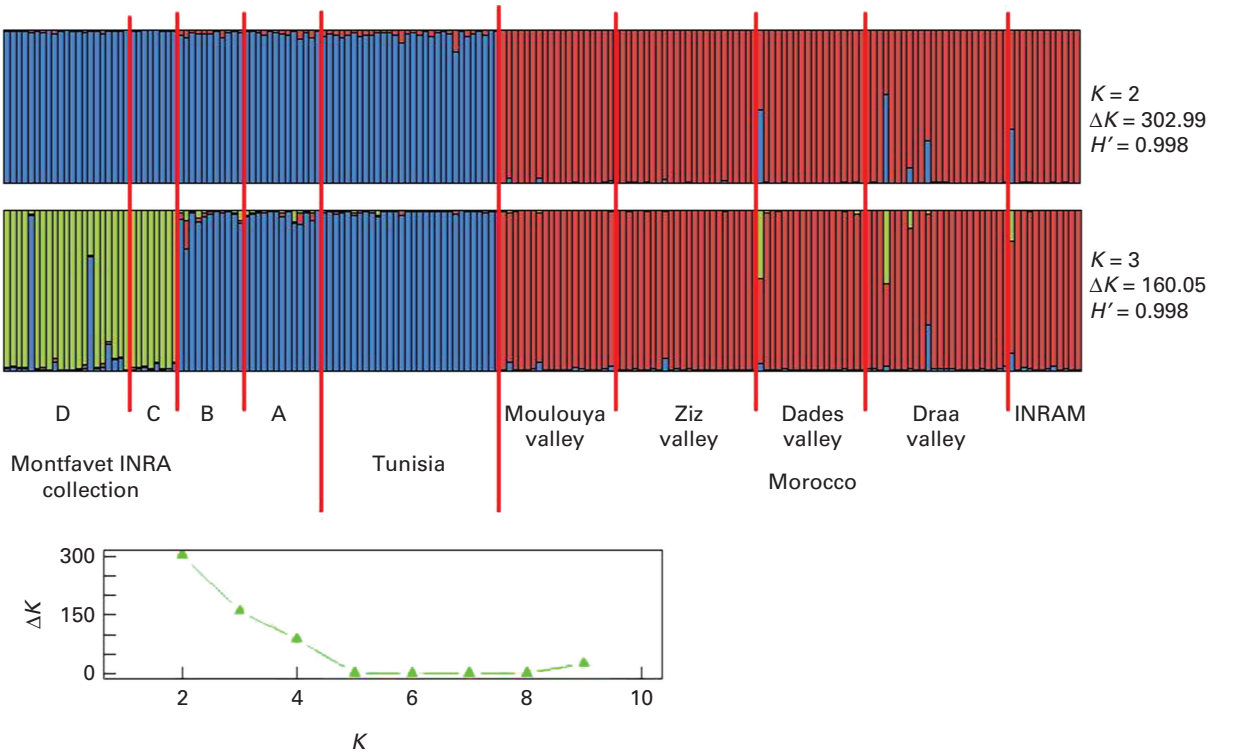

Fig. 3. (colour online) Inferred genetic structure for $K = 2$ and $K = 3$ for the total apricot genetic diversity (180 genotypes). H' represents the similarity coefficient between runs, whereas ΔK represents the *ad hoc* measure of Evanno *et al.* (2005). Based on ΔK and H', $K = 2$ clusters were identified to be the best genetic structure model.

of genotypes assigned, respectively) and four known local varieties (Tables S2 and S3, available online). Approximately 50% of the genotypes sampled from all the different valleys were assigned with a membership probability ≥ 0.80 (Fig. 1; Table S3, available online). Among the 97 genotypes, only seven (7.2%) were shared by different clusters with a membership coefficient lower than 0.80. The 'Canino'-related genotypes were found to be mainly located in the Moulouya and Dades valleys, while the 'Del Patriarca'-related group was found to be mainly located in the Ziz and Draa valleys (although in this latter case, the small size of the sample reduced statistical significance).

Genetic diversity within apricot groups

Using the six groups based on their area of origin, the highest number of polymorphic loci was observed in the diversification group as defined by Hagen *et al.* (2002) ('D' group; 114 loci) and the lowest in the Mediterranean group ('A'; 57 loci; Table 1). Compared with the other groups, groups D and C displayed 49 unique AFLP markers, whereas groups B and A displayed only two markers (data not shown). No unique AFLP markers were observed within the Tunisian and Moroccan apricot groups. Genetic diversity within groups D and C showed the highest values: $H_j = 0.227$ and 0.168, respectively. The values decreased to 0.15, 0.146, 0.141 and 0.131 within the groups A and B, *in situ* Moroccan and Tunisian apricot, respectively (Table 1). Compared with the D and C gene pools, the Moroccan germplasm displayed a reduction in genetic diversity of about 37.8 and 16.07%, respectively, whereas its reduction of diversity was only 3.4 and 6% compared with the continental Europe 'B' and Mediterranean 'A' pools, respectively. The Tunisian germplasm showed a lower overall diversity than the Moroccan germplasm.

The accessions from the Dades Valley showed the highest genetic diversity and number of polymorphic loci within the Moroccan germplasm ($H_j = 0.162$ and 66 loci, respectively), while those from the Draa Valley displayed the lowest values ($H_j = 0.133$ and 56 loci, respectively) (Table 1). Among the 88 AFLP markers observed in the Moroccan genotypes, 57 (64.7%) were shared by the four geographical zones (data not shown).

Focusing on the six sub-clusters defined previously (Fig. 2), the highest number of polymorphic loci and genetic diversity values were observed at the sub-cluster C1-1 composed by the diversification (D) and geographical adaptable (C) groups (105 loci and $H_j = 0.22$, respectively), whereas the sub-cluster C2-2 formed by

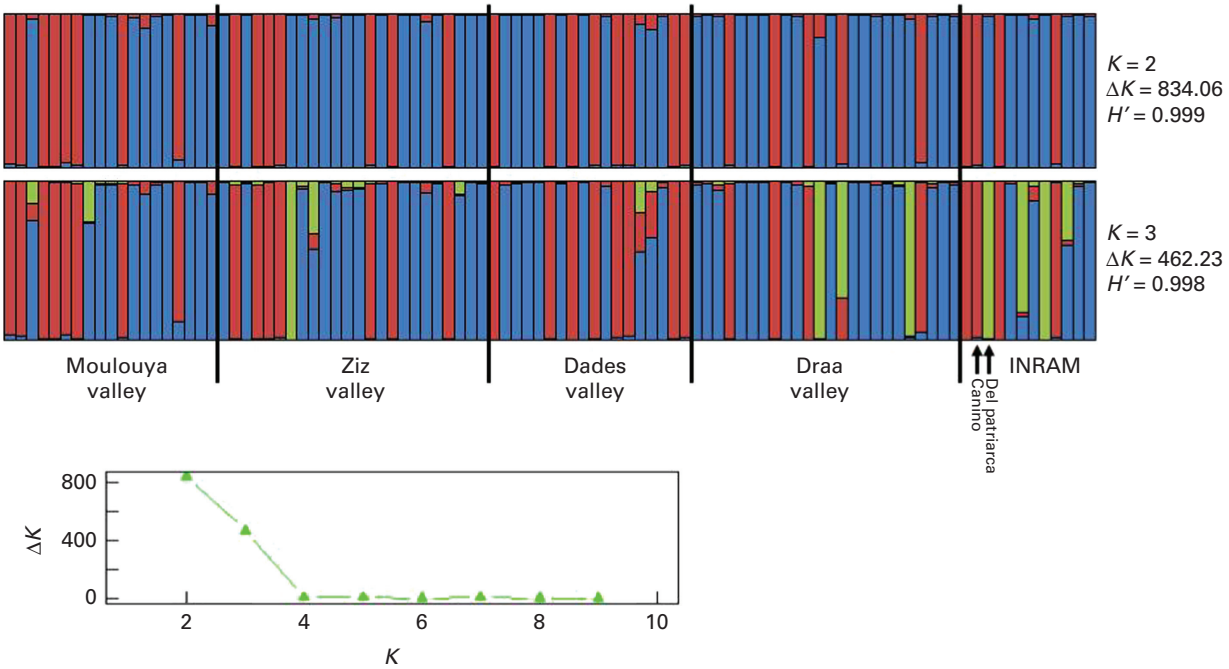

Fig. 4. (colour online) Inferred genetic structure for $K = 2$ and $K = 3$ for the Moroccan apricot material (97 genotypes). H' represents the similarity coefficient between runs, whereas ΔK represents the *ad hoc* measure of Evanno *et al.* (2005). Based on ΔK and H', $K = 2$ clusters were identified to be the best genetic structure model.

eight genotypes including the 'Del Patriarca' cultivar displayed the lowest values (Table 1). The sub-cluster C2-3 termed the 'Canino' group showed higher diversity and polymorphic loci number than the sub-cluster C1-3 which was made up of the Mediterranean group (A) and Tunisian cultivars (Table 1).

To avoid biased results, genetic variation among and within the groups was computed after excluding cultivars grown at the Ain Taoujdate collection because of the unknown origin of the accessions (12 cultivars; Table 2). Genetic differentiation among all the groups was about 16.1% and decreased to 8.2 and 2.8% within the Tunisian

Table 1. Genetic diversity within the different apricot groups

Cultivar group	Sample size	Number and percentage of polymorphic loci	Number and percentage of private bands	Genetic diversity, H_j (± SE)
Diversification group (D)[a]	21	114 (74%)	44 (28.5%)	0.227 (±0.012)
Geographically adaptable group (C)[a]	8	74 (48.1%)	5 (3.2%)	0.168 (±0.014)
Continental Europe group (B)[b]	11	65 (42.2%)	1 (0.64%)	0.146 (±0.014)
Mediterranean group (A)[c]	13	57 (37%)	1 (0.64%)	0.15 (±0.015)
In situ Tunisian apricots	30	61 (39.6%)	0	0.131 (±0.014)
In situ Moroccan apricots	97	64 (41.6%)	0	0.141 (±0.014)
Moulouya Valley group	19	61 (39.6%)	2 (1.28%)	0.138 (±0.014)
Ziz Valley group	24	61 (39.6%)	3 (1.92%)	0.141 (±0.014)
Dades Valley group	18	66 (42.9%)	1 (0.64%)	0.162 (±0.015)
Draa Valley group	24	56 (36.4%)	2 (1.28%)	0.133 (±0.014)
Sub-cluster C1-1[a]	28	105 (68.2%)	49 (31.8%)	0.22 (±0.013)
Sub-cluster C1-2[b]	17	64 (41.6%)	2 (1.2%)	0.15 (±0.014)
Sub-cluster C1-3[c]	39	61 (39.6%)	1 (0.64%)	0.137 (±0.014)
Sub-cluster C2-1[d]	53	58 (37.7%)	1 (0.64%)	0.132 (±0.014)
Sub-cluster C2-2[d]	8	48 (31.2%)	0	0.08 (±0.012)
Sub-cluster C2-3[d]	35	69 (44.8%)	0	0.141 (±0.013)

[a] C and D groups as defined by Hagen *et al.* (2002). [b] B group as defined by Hagen *et al.* (2002). [c] A group as defined by Hagen *et al.* (2002) and *in situ* Tunisian germplasm. [d] Moroccan genotypes.

and Moroccan apricots, respectively. Genetic variation in the Moroccan as well as Tunisian apricots occurred within the groups rather than between the groups, indicating that all the four Moroccan groups were closely related. Unlike group B and Tunisian apricots, lesser genetic differentiation was observed between group B and Moroccan apricots (9.6%), while a similar differentiation were found for both Moroccan and Tunisian apricots compared with the Mediterranean group (8.3 and 8.7%, respectively). Finally, a high differentiation level was observed between the Tunisian and Moroccan apricots (13.9%; Table 2).

Genetic differentiation among the 11 groups was about 16.1% and the pairwise F_{ST} ranged from 0.011 (between the Ziz and Moulouya valleys) to 0.301 (between the C and Testour groups; Table S4, available online). Genetic relationships between these groups, based on the F_{ST} values and the neighbour-joining algorithm, showed that the *ex situ* germplasm and the *in situ* Tunisian apricot were found to be clearly distinguished from the Moroccan apricots with a set of closely related genotypes for both Tunisian and Moroccan apricots. The Mediterranean group was shown to be closer to the Tunisian apricots than to the Moroccan germplasm (Fig. S2 and Table S4, available online). Focusing on the genetic differentiation between the six sub-clusters identified, the pairwise F_{ST} was found to be 19.4%. The sub-cluster C2-2 ('Del Patriarca' group) was found to be the most differentiated from all the other sub-clusters (Table S5, available online); this result could be explained by the sample size since it contained only eight genotypes. Low F_{ST} values were observed among the sub-clusters C2-1 and C2-2, on the one hand, and among the sub-clusters C1-1 and C1-3, and C1-1 and C1-2, on the other hand (F_{ST} = 0.109, 0.086 and 0.087, respectively; Fig. S3 and Table S5, available online).

Discussion

Except for three accession pairs, all genotypes were found to be distinct by at least one AFLP marker, using the 154 AFLP marker set. Therefore, our results confirm that AFLP markers are effective tools for the molecular characterization of apricot as has been attested in previous studies (Panaud *et al.*, 2002; Hagen *et al.*, 2002; Geuna *et al.*, 2003; Krichen *et al.*, 2008).

A clear distinction was observed between the group including the *ex situ* and *in situ* Tunisian germplasm and the group including only the *in situ* Moroccan germplasm. This separation was confirmed by multivariate analyses as well as by the model-based Bayesian clustering. In regard to the *in situ* Tunisian apricots, Khadari *et al.* (2006) reported a distinction between these local genetic resources and the *ex situ* germplasm, whereas in our study, the *in situ* Tunisian apricot germplasm was classified as being within the Mediterranean group. These results indicate that Moroccan apricot genetic resources are characterized by an evolution process distinct from the other germplasm as expected according to geographical position, historical diffusion and farming practices.

Genetic differentiation of Moroccan apricot populations as a result of seed propagation

A number of factors support the hypothesis that Moroccan apricots have a distinct genetic lineage from other

Table 2. Genetic structure within the different groups: diversification (D), continental Europe (B), geographically adaptable (C), Mediterranean (A), *in situ* Tunisian and Moroccan

Groups (n)[a]	Number of genotypes	Number and percentage of segregating fragments	Overall locus diversity, H_T	Average locus diversity within groups, H_S	F_{ST}
A, B, C, D, Tunisian and Moroccan (11)	168	154 (100%)	0.178	0.149	0.161***
C, D and Morocco (6)	114	151 (98.1%)	0.189	0.162	0.145***
B and Moroccan (5)	96	91 (59.1%)	0.159	0.144	0.096***
A and Moroccan (5)	98	93 (60.4%)	0.158	0.145	0.083***
A, Tunisian and Moroccan (8)	128	96 (62.3%)	0.16	0.138	0.136***
Tunisian and Moroccan (7)	116	93 (60.4%)	0.158	0.136	0.139***
Moroccan (4)	85	85 (55.2)	0.148	0.144	0.028***
C, D and Tunisian (5)	59	150 (97.4%)	0.187	0.154	0.171***
B and Tunisian (4)	41	75 (48.7%)	0.151	0.131	0.134***
A and Tunisian (4)	43	74 (48.1%)	0.145	0.132	0.087***
Tunisian (3)	30	61 (39.6%)	0.137	0.126	0.082***

***Highly significant at $P < 0.001$.

[a] Number of groups within each geographical zone were compared. Tunisia has three groups (Testour, Ras Jbel and Kairouan) and Morocco four groups (Moulouya, Ziz, Dades and Draa).

available germplasm. For example, two apricot groups are distinguished by their propagation mode. Apricots from the *ex situ* and the *in situ* Tunisian germplasm are propagated by grafting, whereas the Moroccan apricots are propagated by seed, except for the 'Canino', 'Gélitano', 'Del Patriarca' and 'Maoui' cultivars. Contrary to graft propagation, which fixes the genotype, seed propagation entails numerous successive reproductive events that can result in genetic drift within a population of limited size, with a further effect of selection by local farmers. According to this evolution model, we expect to find both low genetic diversity and high genetic differentiation in the seed-propagated apricot populations. First, our results showed that Moroccan apricots indeed displayed a narrow genetic basis when compared with the other gene pool. Similar results based on simple sequence repeat (SSR) markers were observed by Bourguiba *et al.* (2012b). Second, we observed a high genetic differentiation between the *ex situ* germplasm and the *in situ* Moroccan germplasm ($F_{ST} = 0.113$). Similar results were obtained for the local Tunisian varieties; indeed, Khadari *et al.* (2006) proposed the 'bottleneck' hypothesis in which these local varieties were derived from seedlings of a few introduced genotypes. This hypothesis was verified by Bourguiba *et al.* (2010) supporting the assumption that grafted and seed-propagated apricots shared the same origin in the Tunisian region. Strikingly, the genetic differentiation between the *in situ* Tunisian and Moroccan apricots was the highest ($F_{ST} = 0.17$), while the *ex situ* germplasm differentiated from the Tunisian and Moroccan germplasm by lower F_{ST} values (0.086 and 0.113, respectively). These results indicate a genetic drift effect for the Tunisian and Moroccan populations as has been proposed in previous studies (Khadari *et al.*, 2006; Bourguiba *et al.*, 2010). Therefore, in Tunisia, Morocco and probably Algeria, the genetic drift observed in the seed-derived apricot populations probably resulted from a genetic bottleneck. This hypothesis has recently been confirmed by analysing seed-propagated populations from the Maghreb region (Algeria, Morocco and Tunisia) using SSR markers (Bourguiba *et al.*, 2012b).

Moreover, we observed a significant reduction in genetic diversity between the diversified D and C gene pools and the *in situ* Moroccan apricot (37.8% for the D gene pool and 16.07% for the C gene pool, respectively), whereas a limited diversity loss was observed compared with the continental Europe group (B; 3.4%) and the Mediterranean group (A; 6%). Unlike Moroccan apricots, Tunisian apricots showed a significant reduction in genetic diversity compared with the A and B gene pools (12.6 and 10.2%, respectively), associated with a high genetic differentiation from the B group ($F_{ST} = 0.134$; Table 2). Taking into account the history of Morocco and its geographical position at the extreme west of the Mediterranean, close to Europe, Moroccan apricots may have diversified through two diffusion routes from the 'Irano-Caucasian' secondary diversification centre as proposed by Bourguiba *et al.* (2012a): (1) North Africa (Mediterranean group), during the Islamization period (Kostina, 1969; Bailey and Hough, 1975), and (2) continental Europe under Andalusian and European influences (Valdeyron and Crossa-Raynaud, 1950). This finding is supported by the example of 'Rouge du Roussillon', a French cultivar which was classified with Moroccan germplasm, and by the impact of 'Canino' and 'Del Patriarca', two recently introduced Spanish cultivars, whose genomes have clearly been introgressed into the local apricot germplasm.

Evolution of 'Mech-mech' populations at a fine scale under the bottleneck effect versus *gene flow*

Our results also provide support for explaining how local Moroccan apricot evolved at a finer geographical scale. Genetic differentiation among the four geographical zones, Dades, Draa, Moulouya and Ziz valleys (87 genotypes), was about 2.8%, with the most differentiated group being the Draa Valley, which also displayed the lowest genetic diversity ($H_j = 0.133$). Therefore, the Draa group displayed a genetic signature that points to a combined effect of population bottleneck and genetic drift. This genetic structure at the fine scale is similar to that at the regional level since we observed low genetic diversity and high differentiation when comparing the Moroccan apricot with the *ex situ* germplasm. Focusing on the genetic structure without any geographical consideration, three sub-clusters were revealed, two of which appeared to be highly impacted by the 'Del Patriarca' and 'Canino' genomes. No clear genetic structure was displayed according to the geographical origin of accessions. However, combining these two analyses (genetic differentiation and structure), we identified two phenotypic groups related to both genetic structure and geographical location. The first group includes local apricots and their related variety 'Del Patriarca' from the Ziz (one genotype) and Draa (two genotypes) valleys. These apricots are grown under Saharan climatic conditions with a short winter period and high temperatures starting in early spring (Barbeau and Bouami, 1980). They are characterized by having low chilling requirements and are early blooming, with flat-shaped, white flesh fruit; they, however, display a limited ability for fresh fruit conservation and processing. The second group includes local apricots with 'Canino' phenotypic traits from the Dades and Moulouya valleys

(with eight genotypes each). These apricots are grown under arid conditions with a relatively long winter period and lower temperatures than the first group (Barbeau and Bouami, 1980). These accessions display higher chilling requirements than the first group and are characterized by oval-shaped, relatively large-sized fruit with yellow flesh, similar to the fruits of 'Canino'. In these areas and especially in the Moulouya Valley, local farmers produce apricot from seed-derived trees but also from grafted 'Canino' individuals (Barbeau and Bouami, 1980), explaining why local apricots from these valleys are classified with the 'Canino' variety. Thus, our results strongly suggest a substantial gene flow from 'Canino' to local apricot populations in the Moulouya and Dades valleys. A similar gene flow from 'Del Patriarca' to local apricot populations in the Draa and Ziz valleys is also strongly suspected, as has already been observed with different markers on a larger set of apricots in North Africa by Bourguiba *et al.* (2013).

Conclusion

Our results highlight the substantial Moroccan apricot diversity and emphasize the importance of oasis agroecosystems as active incubators of genetic variability, containing both considerable crop species diversity and high within-crop genetic diversity. Two antagonistic processes probably impacted the evolution of 'Mech-mech' populations at a fine scale: a genetic drift due to a bottleneck effect *versus* the likely gene flow from the clonal varieties 'Canino' and 'Del Patriarca' into local apricot populations. The present study provides insights to improve management and conservation approaches for local germplasm. The substantial diversity thus identified could supply very valuable adaptive traits for arid and Saharan conditions in current apricot breeding programmes.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1479262113000543

Acknowledgements

The authors thank M. Lahlou for his help in apricot sampling from the oasis agroecosystems in Morocco and H. Achtak for his kind remarks on an earlier version of the manuscript. This study was supported by the French-Moroccan cooperation program PRAD 03-02.

References

Bailey CH and Hough LF (1975) Apricots. In: Janick J and Moore JN (eds) *Advances in Fruit Breeding*. Lafayette, IN: Purdue University Press, pp. 367–383.

Barbeau G and El Bouami A (1980) Apricot Surveys in the southern Morocco. *Fruits* 35: 115–122.

Bernatzky R and Tanksley SD (1986) Genetics of action-related sequences in tomato. *Theoretical and Applied Genetics* 72: 314–321.

Bourguiba H, Khadari B, Krichen L, Trifi-Farah N, Santoni S and Audergon JM (2010) Grafting *versus* seed propagated apricot populations: two main gene pools in Tunisia evidenced by SSR markers and model-based Bayesian clustering. *Genetica* 138: 1023–1032.

Bourguiba H, Audergon JM, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D'Onofrio C, Murat Asma B, Santoni S and Khadari B (2012*a*) Loss of genetic diversity as signature of apricot domestication and diffusion into the Mediterranean Basin. *BMC Plant Biology* 12: 49.

Bourguiba H, Audergon JM, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S and Khadari B (2012*b*) Genetic diversity and differentiation of grafted and seed propagated apricot (*Prunus armeniaca* L.) in the Maghreb region. *Scientia Horticulturae* 142: 7–13.

Bourguiba H, Khadari B, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S and Audergon JM (2013) Genetic relationships between local apricot germplasm in North Africa and recent introduced varieties. *Scientia Horticulturae* 152: 61–69.

Couranjou J (1975) Genetic improvement of apricot by the National Institute of Agronomic Research in France. *Bulletin Technique d'Information* 300: 397–410.

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software Structure, a simulation study. *Molecular Ecology* 14: 2611–2620.

Faust M, Suranyl D and Nyujto F (1998) Origin and dissemination of apricot. *Horticultural reviews* 22: 225–266.

Geuna F, Toschiand M and Bassi D (2003) The use of AFLP markers for cultivar identification in apricot. *Plant Breeding* 122: 526–531.

Hagen LS, Khadari B, Lambert P and Audergon JM (2002) Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. *Theoretical and Applied Genetics* 105: 298–305.

Jakobsson M and Rosenberg NA (2007) Clumpp: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

Khadari B, Krichen L, Lambert P, Marrakchi M and Audergon LM (2006) Genetic structure in Tunisian apricot, *Prunus armeniaca* L., populations propagated by grafting: a signature of bottleneck effects and ancient propagation by seedlings. *Genetic Resources and Crop Evolution* 53: 811–819.

Kostina KF (1969) The use of varietal resources of apricots for breeding. *Trudy Gosudarstvennogo Nikitskogo botanicheskogo Sada* 40: 45–63.

Krichen L, Martins JMS, Lambert P, Daaloul A, Trifi-Farah N, Marrakchi M and Audergon JM (2008) Using AFLP markers for the analysis of the genetic diversity of apricot cultivars in Tunisia. *Journal of American Society for Horticultural Science* 133: 204–212.

Lefevre V, Palloix A and Rives M (1993) Nuclear RFLP between pepper cultivars (*Capsicum annuum* L.). *Euphytica* 71: 189–199.

Lynch M and Milligan B (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3: 91–99.

Panaud O, Chaib A and Sarr A (2002) Dynamic conservation of apricot *Prunus armeniaca* in Saharian oases: use of AFLP markers to assess genetic diversity in traditional orchards. *Euphytica* 128: 301–305.

Perrier X, Flori A and Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X and Glaszmann JC (eds) *Genetic Diversity of Cultivated Tropical Plants*. Montpellier: Enfield, Science Publishers, pp. 43–76.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure from multilocus genotype data. *Genetics* 155: 945–959.

R Development Core Team R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available from http://www.R-project.org/2011

Sokal RR and Sneath PHA (1963) *Principles of Numerical Taxonomy*. San Francisco: Freeman, p. 359.

Valdeyron G and Crossa-Raynaud P (1950) The fruits of Tunisia. *Annals de l'Institut Natational de la Recherche Argonomique de Tunis* 23: 65–82.

Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Ward JH (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–244.

Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* 8: 907–913.