

## Molecular phylogeny to specify *Zalmati* and *Chemlali Tataouine* Tunisian olive cultivars

R. BEN AYED<sup>1\*</sup>, K. ENNOURI<sup>1</sup>, H. BEN HASSEN<sup>2</sup>, A. REBAI<sup>1</sup>

<sup>1</sup>: Molecular and Cellular screening Processes Laboratory. Centre of Biotechnology of Sfax, Tunisia

<sup>2</sup>: Laboratory of Physics, Mathematics and Applications. Faculty of Sciences of Sfax, Tunisia

\* Corresponding author: raydabenayed@yahoo.fr

**Abstract** - Olive tree (*Olea europaea* L.) is mainly grown in the Mediterranean area, which is the most important oil-producing crop. In spite of its cultural, economic and nutritional importance, numerous DNA markers accomplished to compositional analyzes by bioinformatic tools have been used to illustrate and reveal homonyms, synonyms and unknown accessions. The objectives of this study were (1) to analyze 5 SNP and 8 SSR markers to characterize 12 olive genotypes originating from different regions of Tunisia and (2) to resolve the denomination ambiguity of *Zalmati* and *Chemlali Tataouine* cultivars. Analyses indicated the presence of two erroneous denominations of cultivars.

**Keywords:** Olive tree, SNP, SSR, UPGMA, Authenticity, PCA, Misnaming, bioinformatic.

### 1. Introduction

Olive is a member of the Oleaceae family, which contains the genus *Olea* (Martin and Sibbett 2005). The section *Olea* comprises the group of *Olea europaea* L., dispersed over widespread areas in the Mediterranean region, known as the most olive species cultivated for oil extraction and production. Besides, according to Rugini et al (2011), *Olea europaea* exceeds one thousand varieties, although several of these might be just distinct landraces stemmed from the original genetic stock and originating from genetic mutations and probable hybridizations. Habitually, the preliminary identifications of olive varieties were based on the sustained employ of phenotypic characters, used as descriptors. However, reliance on phenotypic characters has possibly led to uncertainty about the existing classification of olive cultivars in many sites.

It is fundamental that cultivars be identified using efficient techniques. That's why analyses with molecular markers supported the morphological identifications from preceding works (Trujillo et al. 1995; Belaj et al. 2001; 2002). However, various genotypes with identical profiles of molecular markers have shown some evident morphological differences. Moreover, unclear morphological accessions in the Germplasm Banks frequently showed small variations in their molecular profiles, still with different accessions of the same cultivars, which reveal the existence of intracultivar genetic variability for studied cultivars (Belaj et al. 2004).

A different case is the association of the less observed molecular and morphological divergences among accessions with the identical denomination (Martins-Lopes et al. 2009). Here the evidence of homonyms (The same denomination given to different cultivars) or synonyms (the same cultivar can be called with different designations) are probably the cases. Therefore, the possibility of confusion between denominations and true-to-type well-known and associated cultivars is small. Nevertheless, tests of identity should be a requirement in any publication on agronomic characters to supply right information (Taamalli et al. 2007; Rekik et al. 2008; Muzzalupo et al. 2009).

Currently, the difficulty of classification and the drawbacks of the morphological and biochemical markers, influenced by the environment and the developmental stage, was the cause for the increased use of molecular markers, such as RFLP, RAPD, AFLP, SSR and SNP, for cultivar characterisation and classification. In fact, some strategies based on a consensus list of minimum morphological characters (Barranco et al. 2000; 2005) and Simple Sequence Repeat (SSR) (Baldoni et al. 2009) have been established and are in development. Restriction Fragment Length Polymorphism (RFLP) markers have been successfully used to differentiate wild and cultivated olive varieties, corroborating the Mediterranean basin as the location of olive domestication (Besnard et al. 2001; 2002). For these reasons, genotyping by molecular markers and by morphological descriptors are techniques that are complement.

In Tunisia, olives have commercial, economic, and social importance. In addition, Tunisia is a major producer and exporter of olive oil. Several olive oil varieties were grown in Tunisian territory such as Zalmati, Chemlali Tataouine, Oueslati, Chemchali and Zarrazi (Ben Ayed et al. 2009), but the most important variety is Chemlali. (Ben Ayed et al. 2014a).

The major aim of this present work is the identification of two Tunisian local olive oil cultivars (Zalmati and Chemlali Tataouine) based on phenotypic, physico-chemical and genetic characteristics by using bioinformatic tools.

## 2. Experimental Methods

### 2.1 Plant Material

A total of twelve Tunisian olive tree cultivars (table 1) were chosen and used in this study. They were selected from different geographical regions of Tunisia. For each cultivar, 3 trees were used, and from each tree, DNA was extracted from young leaves (Ben-Ayed et al. 2012).

**Table 1:** List of olive accessions studied.

Number	Abbreviation	Variety	Purpose	Average weight of fruit (g)	Origin country of cultivation (region/town)
1	Chem_chàal	Chemlali Chàal	Oil	1	Sfax/Chàal
2	Chem_Blett	Chemlali Blettech	Oil	1	Sfax/Blettech
3	Chem_SB	Chemlali Sidi Bouzid	Oil	1	Sidi Bouzid
4	Chem_Dok	Chemlali Dokhane	Oil	1	Sfax/Dokhane
5	Chem_Nab	Chemlali Nabeul	Oil	1	Nabeul
6	Zarrazi	Zarrazi	Oil and table	3	Zarzis
7	Oueslati	Oueslati	Oil and table	1.6	Kairouan
8	Chem_Tat	Chemlali Tataouine	Oil	0.8	Tataouine
9	Chem_Sous	Chemlali Sousse	Oil	1	Sousse
10	Chem_Mon	Chemlali Monastir	Oil	1	Monastir
11	Zalmati	Zalmati	Oil	1.2	Zarzis
12	Chemch	Chemcheli	Oil and table	3	Gafsa

### 2.2 DNA Extraction

DNA was extracted separately from leaves using the CTAB methods described by Rekik et al. (2008) and an additional purification was introduced; it consists in washing and eluting once with the QIAamp DNA stool (Qiagen™) to eliminate contaminant molecules and to generate a high quality DNA for specific, reproducible and consistent amplifications (Ben-Ayed et al. 2009).

### 2.3 Genotyping analysis

The eight studied SSR markers (DCA1, DCA3, DCA4, GAPU59, GAPU71A, GAPU71B, UDO09, UDO12) were genotyped by PCR method used a fluorescent primers and capillary sequencer.

Five microliters of PCR product was mixed with 0.3 IL of marker 420 pb and 14.7 IL of deionized H<sub>2</sub>O, centrifuged at 2,000 rpm for 1 min, denatured at 94 °C for 3 min, cooled in ice and analyzed on a (3130XL Genetic analyzer of Applied Biosystems) capillary sequencer.

The SNP SOD (insertion/deletion type) was genotyped by a simple polymerase chain reaction and agarose gel migration. The other four SNPs (FAD2.1, FAD2.3, ANTHO3 and CALC) were genotyped by a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method (Ben Ayed et al. 2014b). All digestion products were separated by electrophoresis on 3% Nusieve ethidium bromide-stained agarose gels and visualized under UV light.

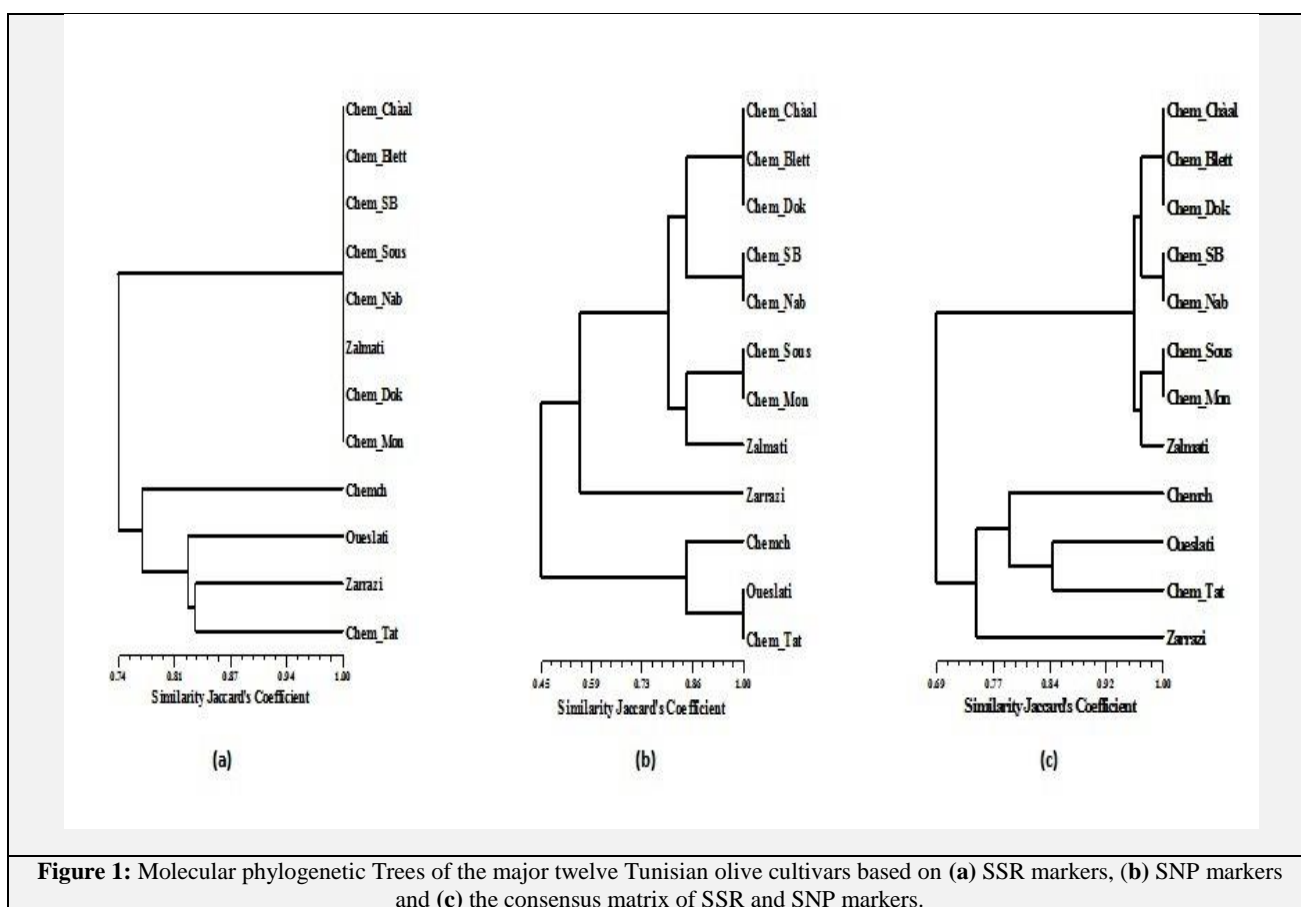
### 2.4 Statistical Analysis

The alleles detected for each SNP or/and SSR were recorded and a binary data matrix is established with (1) for presence of bands (1) (each allele representing a band) and (0) for absence of bands (each allele representing a band). Similarity matrices were generated using the SIMQUAL sub-program of NTSYS-PC software (Rohlf, 1999). Similarity coefficients were used for cluster analysis of varieties using the SAHN sub-program of NTSYS-PC software and dendrograms were inferred using an arithmetic average (UPGMA) clustering algorithm.

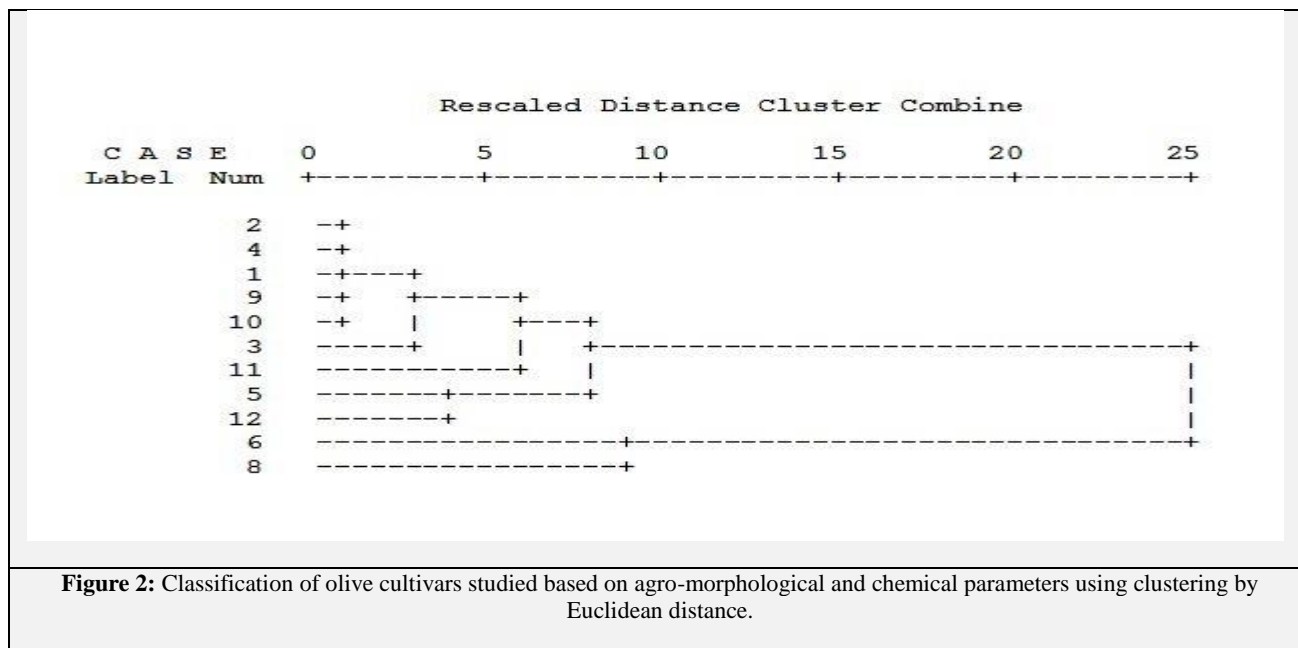
Studies on genetic diversity among cultivars were based on agro-morphological and chemical parameters (Khlif and Grati, 2001). Data were analyzed using the Euclidean distance matrix and the Principal Components Analysis (PCA). The analyses were carried out by using SPSS software (version 17.0).

### 3. Results and discussion

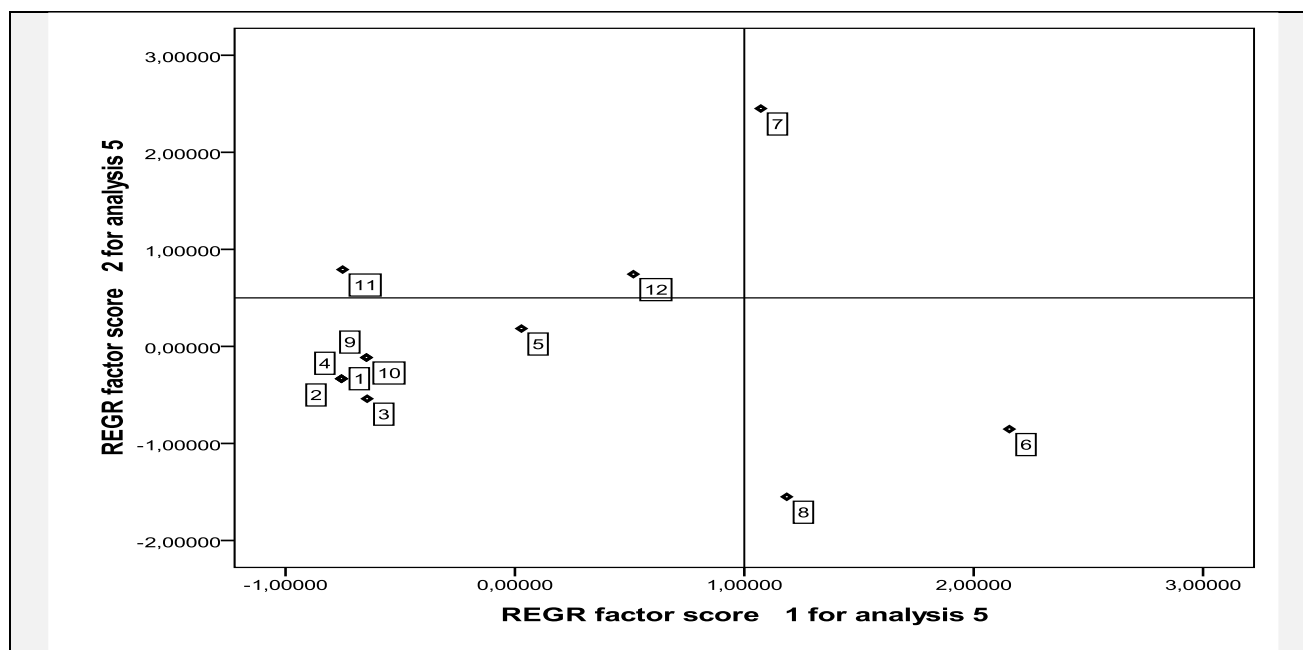
The Tunisian olive collection is a useful source of the diversity of genetic material which can be used by olive breeders, researchers, and backyard hobbyists to improve and preserve the olive crop in Tunisia. Therefore, it is vital to verify the genetic identity of olive germplasm before its distribution to users. Many previous studies showed that although ‘Chemlali’ and ‘Zalmati’ cultivars show a high similarity, the ‘Chemlali’ varieties themselves demonstrate a distinction between ‘Chemlali’ and ‘Chemlali Tataouine’ cultivars. In fact, ‘Zalmati’ and ‘Chemlali Tataouine’ were previously named by the agricultures based on morphological characteristics of the trees, the leaves and the fruits. The ‘Zalmati’ variety has almost the same morphological and chemical characteristics with ‘Chemlali’ varieties (except ‘Chemlali Tataouine’). Actually, these two varieties could not be separated even using multivariate statistical analyzes based on these characters. Building on the AFLP markers, both varieties have a low resemblance with a similarity coefficient below 0.6 ( $G_s = 0.58$ ) (Grati Kamoun 2007). However, the results found with SSR markers showed that these two species were very similar (98%) (Figure1a). This fact highlights the effectiveness of SSR markers since these two varieties are almost identical in terms of morphological and chemical properties. The present work shows that the dendrograms based on SNP markers (Figure1b) and SNP + SSR (Figure1c) prove that the ‘Zalmati’ variety belongs to the same group of ‘Chemlalis’ varieties (except ‘Chemlali Tataouine’). Indeed, used SSR markers, illustrated the total homology (100%) between these cultivars. However, the use of a limited number of SNP markers, five, is not sufficient to detect the strong similarity between these cultivars and shows only partial homology (70%). This underlines the effectiveness of SSR markers selected in the previous study. These results can therefore be explained by two assumptions; the ‘Chemlali’ varieties (except ‘Chemlali Tataouine’) and ‘Zalmati’ varieties are identical but they have two different names, or these two cultivars have a common ancestor that is very close in time.



Euclidean distances were calculated between all pairs of individual trees based on standardized values of ten quantitative traits and the matrix distance was used for cluster analysis (UPGMA) in SPSS version 17, as shown on figure 2. The graphical representation of the matrix of Euclidean distances, which lies between the studied varieties based on agro-morphological and chemical characteristics, revealed two different groups. The first group combined the Chemlali and Zalmati varieties, while the second combined three varieties (Chemchali, Zarrazi and Chemlali Tataouine).



Moreover, the diversity analysis was performed by the technique of principal component analysis (PCA) using SPSS program as illustrated in Figure 3. The first two components of the PCA explained 98.743% of the total variation among a subset of 12 Tunisian olive varieties included in the analysis. The first component (axis1) explained 95.025% of the variation, followed by 3.718% for the second components (axis 2). Axis 1 was positively correlated with polyphenol content (0.998), oxidative stability (0.912) and monounsaturated fatty acids (0.817), and negatively correlated with polyunsaturated fatty acids (-0.825), and the rate of linoleic acid (C18: 2) with a value of -0.805. On the other hand, axis 2 was positively correlated with the content of  $\beta$ -sitosterol (0.703), insaturated fatty acids (0.675) and cholesterol content (0.625), and negatively correlated with the palmitoleic acid (C16:1) with a value of -0.713, rate of unsaponifiable matter (-0.686) and the saturated fatty acids (-0.679). When the olive varieties were studied individually, we showed that the axis 2 divides the varieties into two different groups: group one composed of "Chemlali" varieties (Chemlali Chaal, Chemlali Blettech, Chemlali Dokhane, Chemlali Sidi Bouzid, Chemlali Nabeul Chemlali Monastir, Chemlali Sousse) and Zalmati. On the other hand, group 2 combined Zarrazi, Oueslati and Chemlali Tataouine. For the varieties 'Zalmati' and 'Chemlali' their classification, by the PCA, in the same group along the axis 2 is essentially related to a remarkable similarity of their fruit size parameters, and the polyphenol content. Moreover, the classification results of the studied varieties are in agreement with previous publication (Khlif and Grati, 2001). This research study analyzed the chemical characteristics of 'Chemlali Tataouine' variety and showed that this variety has specific characteristics and differs from other 'Chemlali, in particular the oil of this variety has a lower percentage of linoleic acid against the other 'Chemlali' varieties. In fact, compared to 'Chemlali Sfax', it differs by the higher percentage of oleic acid and the lower percentage of linoleic acid. Besides, the studies of the oxidative stability of olive oils, whose induction time is up to 112 hours, demonstrate that the 'Chemlali Tataouine' variety is more stable than 'Chemlali' varieties originating from the Sfax and North regions of Tunisia. Therefore, distinguishing the 'Chemlali Tataouine' variety from other 'Chemlali' varieties shows that there is a matching between agro-morphological, chemical characters, genetic data based on AFLP and SSRs and the present work based on SNPs markers and SSR+SNP markers.



**Figure 3:** Classification of olive cultivars studied based on agro-morphological and chemical parameters using clustering the principal components analysis (PCA).

#### 4. Conclusion

In this study, we have shown that the compositional and DNA-based markers combined with bioinformatic tools (molecular phylogeny and statistical methodology analyses) are decisive, especially in estimating the genetic variation of Tunisian olive cultivars and resolving the denomination ambiguity.

#### 5. References

- Baldoni L, Cultrera NG, Mariotti R, Ricciolini C, Arcioni S, Vendramin GG, Buonamici A, Porceddu A, Sarri V, Ojeda M, Trujillo I, Rallo L, Belaj A, Perri E, Salimonti A, Muzzalupo I, Casagrande A, Lain O, Messina R, Testolin R (2009)** A consensus list of microsatellite markers for olive genotyping. *Molecular Breeding* 24: 213-231
- Barranco D, Cimato A, Fiorino P, Rallo L, Touzani A, Castaneda C, Sefarin F, Trujillo I (2000)** World catalogue of olive varieties. International Olive Oil Council, Madrid.
- Barranco D, Trujillo I, Rallo L (2005)** Libro I. Elaiografía Hispanica. In: *Varietades de olivo en España* (eds L. Rallo, D. Barranco, J.M. Caballero et al.). Junta de Andalucía, MAPA y Ediciones Mundi-Prensa, Madrid. pp. 45–229
- Belaj A, Satovic Z, Rallo L, Trujillo I (2002)** Genetic diversity and relationships in olive (*Olea europaea* L.) germplasm collections as determined by randomly amplified polymorphic DNA. *Theoretical and Applied Genetics* 105: 638–644
- Belaj A, Satovic Z, Trujillo I, Rallo L (2004)** Genetic relationships of Spanish olive cultivars using RAPD markers. *HortScience* 39(5): 948-951
- Belaj A, Trujillo I, De La Rosa R, Rallo L (2001)** Polymorphism and discrimination capacity of randomly amplified polymorphic markers in an olive germplasm bank. *Journal of American Society for Horticultural Sciences* 126: 64-71
- Ben Ayed R, Grati-kamoun N & Rebai A (2013)** An overview of the authentication of olive tree and oil. *Comprehensive Review in Food Science and Food Safety* 12(2): 218-227
- Ben Ayed R, Grati-Kamoun N, Moreau F & Rebai A (2009)** Comparative study of microsatellite profiles of DNA from oil and leaves of two Tunisian olive cultivars. *European Food Research and Technology* 229(5): 757-762
- Ben Ayed R, Grati-kamoun N, Sans-grout C, Moreau F & Rebai A (2012)** Characterization and authenticity of virgin olive oil (*Olea europaea* L.) cultivars by microsatellite markers. *European Food Research and Technology* 234(2): 263-271

- Ben Ayed R, Kallel I, Ben Hassen H & Rebai A (2014b)** SNP marker analysis for validating the authenticity of Tunisian olive oil. *Journal of Genetics* 93(3): 148-154
- Ben Ayed R, Sans-grout C, Moreau F, Grati-kamoun N & Rebai, A (2014a)** Genetic similarity among Tunisian olive cultivars and two unknown feral olive trees estimated through SSR markers. *Biochemical Genetics* 52(5-6): 258-268
- Besnard G, Baradat P, Chevalier D, Tagmout A, Bervillé A (2001)** Genetic differentiation in the olive complex (*Olea europaea* L.) revealed by RAPDs and RFLPs in the rRNA genes. *Genetic Resources and Crop Evolution* 48:165-182
- Besnard G, Khadari B, Baradat P, Bervillé A (2002)** *Olea europaea* (Oleaceae) phylogeography based on chloroplast DNA polymorphism. *Theoretical and Applied Genetics* 104:1353-1361
- Grati-Kamoun N, Lamy Mahmoud F, Rebai A, Gargouri A, Panaud O, Saar A (2006)** Genetic diversity of Tunisian olive tree (*Olea europaea* L.) cultivars assessed by AFLP markers. *Genetic Resources Crop Evolution* 53: 265-275
- Khlif M and Grati N (2001)** Caractérisation technologique des variétés d'olivier cultivées en Tunisie. *Ezzaitouna* ISSN : 0330-6828
- Martin GC, Sibbett GS (2005)** Botany of the olive. In: Sibbett G.S., Ferguson L., Coviello J.L., Lindstrand M. (Eds.), *Olive Production Manual*, University of California, Agriculture and Natural Resources, Oakland, California, pp. 15–19
- Martins-Lopes P, Gomes S, Lima-Brito J, Lopes J, Guedes-Pinto H (2009)** Assessment of clonal genetic variability in *Olea europaea* L. 'Cobrancosa' by molecular markers. *Scientia Horticulturae* 123(1): 82–89
- Muzzalupo I, Stefanizzi F, Perri E (2009)** Evaluation of olives cultivated in southern Italy by simple sequence repeat markers. *HortScience* 44(3): 582-588
- Rekik I, Salimonti A, Grati-Kamoun N, Muzzalupo I, Lepais O, Gerber S, Perri E & Rebai A (2008)** Characterization and identification of Tunisian olive tree varieties by microsatellite markers. *HortScience* 43(5): 1371-1376
- Rekik-Hakim I, Grati-Kamoun N, Makhloufi E & Rebaï A (2010)** Discovery and Potential of SNP Markers in Characterization of Tunisian Olive Germplasm. *Diversity* 2(1): 17-27
- Rugini E, De Pace C, Gutiérrez-Pesce P, Muleo R (2011) *Olea*. In: Kole C. (ed) *Wild crop relatives: genomic and breeding resources, temperate fruits*. Springer, Berlin/Heidelberg, pp 79–117
- Taamalli W, Geuna F, Banfi R, Bassi D, Daoud D, Zarrouk M (2007)** Using microsatellite markers to characterize the main Tunisian olive cultivars *Chemlali* and *Chetoui*. *J Hort Sci Biotechnol* 82: 25–28.
- Trujillo I, Rallo L, Arus P (1995)** Identification of olive cultivars by isozyme analysis. *Journal of the American Society for Horticultural Science* 120(2): 318–324