

Genetic Markers Analyses and Bioinformatic Approaches to Distinguish Between Olive Tree (*Olea europaea* L.) Cultivars

Rayda Ben Ayed¹ · Hanen Ben Hassen² · Karim Ennouri¹ · Ahmed Rebai¹

Received: 17 April 2015/Revised: 19 October 2015/Accepted: 23 October 2015/Published online: 29 February 2016
© International Association of Scientists in the Interdisciplinary Areas and Springer-Verlag Berlin Heidelberg 2016

Abstract The genetic diversity of 22 olive tree cultivars (*Olea europaea* L.) sampled from different Mediterranean countries was assessed using 5 SNP markers (FAD2.1; FAD2.3; CALC; SOD and ANTHO3) located in four different genes. The genotyping analysis of the 22 cultivars with 5 SNP loci revealed 11 alleles (average 2.2 per allele). The dendrogram based on cultivar genotypes revealed three clusters consistent with the cultivars classification. Besides, the results obtained with the five SNPs were compared to those obtained with the SSR markers using bioinformatic analyses and by computing a cophenetic correlation coefficient, indicating the usefulness of the UPGMA method for clustering plant genotypes. Based on principal coordinate analysis using a similarity matrix, the first two coordinates, revealed 54.94 % of the total variance. This work provides a more comprehensive explanation of the diversity available in Tunisia olive cultivars, and an important contribution for olive breeding and olive oil authenticity.

Keywords Markers · *Olea europaea* L. · Olive oil · Cultivars · Classification · PCoA · UPGMA · Bioinformatic approaches

1 Introduction

Olea europaea L. which is one of the oldest cultivated plants is an important oil-producing crop in the Mediterranean basin. Olive oil which is known for its beneficial effects on health can be consumed in crude form [1]. Moreover, the olive oil sector is a major component of the culture and socio-economy of many Mediterranean countries, including Tunisia. Given the high occurrence of mislabeling, homonyms and synonyms for olive [2], easy and accurate cultivar identification is an urgent necessity to manage its rich variability. Thus, important efforts are being made to obtain a unique and unequivocal genetic profile for every cultivar using powerful tools for assessing genetic variation in olive germplasm and to build a kind of identity card for every olive cultivar. The traditional genetic variation analyses allied to morphological and chemical markers is insufficient to study the relationship between cultivars due to the environmental effects on the phenotype and variation in chemical composition. Besides, these markers require expensive and numerous tests, which is considered as the major constraint [3, 4]. Recently, many molecular markers methods have been exploited to study the genetic diversity of olive oil cultivars such as random amplified polymorphic DNA (RAPDs) [4], amplified fragment length polymorphisms (AFLPs) [5, 6], simple sequence repeats (SSR) [3, 7, 8] and single-nucleotide polymorphism (SNP) [9–11]. These technical methods could allow the detection of DNA polymorphisms of some molecular markers, thus efficiently discerning between cultivars without any environmental influence. These methods could also help in solving the authenticity and the traceability of olive oil [12]. This study is interested in five SNPs (FAD2.1, FAD2.3, CALC, SOD, and ANTHO3) localized in four different genes: anthocyanidin synthase,

✉ Rayda Ben Ayed
raydabenayed@yahoo.fr

¹ Molecular and Cellular Screening Process Laboratory, Centre of Biotechnology of Sfax, Sidi Mansour Road, BP ‘1177’, 3018 Sfax, Tunisia

² Laboratory of Physics, Mathematics and Applications, Faculty of Sciences of Sfax, Sfax, Tunisia

Cu–Zn-superoxide dismutase, calcium binding protein and fatty acid desaturase [13]. In our previous work, we described the first efforts performed in collecting and characterizing Tunisian olive cultivars from different areas and their genetic diversity, relationships and traceability with olive cultivars using SSR markers [3, 7]. This paper reports the SNP markers employment to investigate genetic diversity and relationships among 16 olive cultivars from different Tunisian regions. Bioinformatic tools are used to compare these markers with SSR markers, previously studied and analyzed by our research group [3, 7, 12–14].

The present work is an assessment of the efficiency of such markers in the evaluation of the genetic diversity and relationship among the studied olive cultivars.

2 Materials and Methods

2.1 Plant Material

A total of sixteen Tunisian olive tree cultivars were chosen and used [15]. They were selected from different geographical regions of Tunisia. Six cultivars from four Mediterranean regions (French, Spain, Italy and Greece) were then added. For each cultivar, two trees were used, and from each tree, DNA was extracted from young leaves [7].

2.2 DNA Extraction

DNA was separately extracted from leaves using the CTAB methods described by Rekik et al. [8] and an additional purification was introduced, consisting in washing and eluting once with the QIAamp DNA stool (Qiagen) to eliminate contaminant molecules and generate a high-quality DNA for specific, reproducible and consistent amplifications [3].

2.3 Genotyping Analysis

SNP SOD (insertion/deletion type) was genotyped by a simple polymerase chain reaction and agarose gel electrophoresis. 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 (Table 1). The PCR product (171 bp) of the SNP (ANTHO3) was digested by MspI restriction enzyme (Fermentas, LIFE SCIENCES) at 37 °C overnight. This restriction enzyme recognizes the sequence AA/GG. The G-allele carrying PCR product is cleaved once by the enzyme generating two fragments (64 and 107 bp). The PCR product (476 bp) of the SNP (CALC) was digested by BstZI restriction enzyme (Promega) at 50 °C overnight. This

restriction enzyme recognizes the sequence CC/GG. The C-allele carrying PCR product is cleaved once by the enzyme leading to two fragments (316–160 bp). The two other SNPs (FAD2.1 and FAD2.3) were analyzed using PCR–RFLP. The PCR product (241 bp) of the SNPs (FAD2.1) and (240 bp) of the SNP (FAD2.3) were digested by BamHI restriction enzyme (Fermentas, LIFE SCIENCES) and Alw26I, respectively, at 37 °C overnight. The sizes of the restriction fragments of PCR product were 224 and 17, 130 and 110 bp for CC genotype of FAD2.1 SNP and FAD2.3 SNP, respectively. 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 were recorded and a binary data matrix was established with (1) for the presence of bands (each allele representing a band) and (0) for the absence of bands. Allele frequencies and heterozygosities (both observed and expected under Hardy–Weinberg equilibrium) were calculated using GDA program [16]. The power of discrimination (PD) was calculated for each SSR locus according to Brenner and Morris [17]:

$$PD = 1 - \sum_{i=1}^g p_i^2$$

where P_i is the frequency of i th genotype for the locus and the sum is overall genotypes.

The combined power of discrimination overall loci were then calculated as:

$$1 - \prod_{l=1}^L (1 - PD_l)$$

The data matrix was converted into a similarity matrix (S) values using Jaccard coefficient [18]. For a pair of two cultivars, i and j , this coefficient is calculated as follows:

$$S_{ij} = \frac{n_{ij}}{n_{ij} + n_i + n_j}$$

where n_i is the number of bands present in cultivar i and absent in cultivar j , n_j is the number of bands present in j and absent in i , and n_{ij} is the number of bands shared by both cultivars i and j

Similarity matrices were generated using SIMQUAL sub-program of NTSYS-PC software [19]. Similarity coefficients were used for the cultivars cluster analysis using SAHN sub-program of NTSYS-pc software and dendrograms were inferred using an arithmetic average (UPGMA) clustering algorithm. Principal coordinate analysis (PCoA) was performed using DCENTER and EIGEN provided in NTSYS-pc V.2.1. The graphical

Table 1 Characteristics of SNP markers used for DNA amplification in the present study

Gene name	SNPs name	SNP type	Tm ^a	Size of amplicon in bps	Ho	He	PD
Anthocyanidin synthase	ANTHO3	G/A	57	171	0.714	0.481	0.430
Cu–Zn-superoxide dismutase	SOD	I/D	57	190/176	0.681	0.490	0.475
Calcium binding protein	CALC	C/G	60	476	0.727	0.462	0.396
Fatty acid desaturase	FAD2.1	C/T	57	241	0.428	0.336	0.489
Fatty acid desaturase	FAD2.3	C/G	56	240	0.636	0.500	0.528
Total					3.188	2.272	2.321
Mean					0.637	0.454	0.464

For each locus, the type of polymorphism, the size of amplicon in base pairs, the frequency of heterozygotes found (Ho), the frequency of heterozygotes predict (He) and the power of discrimination (PD) are reported

I insertion, D deletion

^a Annealing temperature for PCR amplification

representation of the genetic relationships among olive cultivars were displayed in a two-dimensional plot using MXPLOT. Computations were achieved using the procedures in the NTSYS pc 2.1 software [19].

3 Results and Discussion

3.1 Genotyping Results and Characteristics of the Studied SNP Markers

The expected fragments for each SNP are reported in Sect. 2 and some of them are shown in Ben Ayed et al. [13].

While the observed heterozygosity for each marker ranges from 0.428 (FAD2.1) to 0.727 (CALC) (0.637 average), the expected heterozygosity varies between 0.336 and 0.5 with an average of 0.454 higher than the observed expected one (Table 1).

The discrimination power (DP) varies from 0.396 for the CALC marker to 0.528 for FAD2.3 marker with an average value of 0.464, which accords well with those of Reale et al. [20]. Although this value is lower than that shown by Rekik et al. with SSR markers (0.71), it is significantly higher than that reported by Cipriani et al. [21] in 12 Italian cultivars (0.44) and by Muzzalupo et al. [22] in 39 Italian cultivars (0.38), using SSR markers. The combined discriminating power is 0.95668938, indicating that the probability of finding two cultivars with the same genotypes combination for the 5 SNP markers is one per mile, showing that these markers' discriminating power is lower than SSR markers (0.99998716) of the same cultivars [7]. This result is expected since, unlike SSR markers, SNPs are biallelic.

For the allele frequencies of each studied SNP marker, a dominance of one allele over another is noted, except for

the FAD2.3 marker, in which both alleles are equal. Besides, most of the studied cultivars are heterozygous genotypes (Table 1), except for FAD2.1 markers, in which the frequency of heterozygous genotype is 42.85 % and that of homozygous genotype is 57.14 %.

3.2 Dendrogram Analysis

The SNP markers genotype data used in the dendrogram are shown in Fig. 2. Indeed, five cultivars groups could be defined by cutting the dendrogram at a distance of 0.73 (Fig. 2a). Group 1 consists of eight cultivars, 'Chemlali' (Chem_Chaàl, Chem_Blett, Chem_Dok, Chem_SB, Chem_Nab, Chem_Sous and Chem_Mon) and 'Zalmati'. The same group was found in the dendrogram based on SSR markers previously studied by Ben Ayed et al. [7]. The second group contains two foreign cultivars ('Koroneiki' and 'Picholine'). The third group comprises 'Arbequina', 'Zarrazi', three 'Chétoui' cultivars (Chetoui_Silia, Chet_Thibar and Chet_Nab) and 'Rkhaymi'. The fourth group constitutes three cultivars 'Chemch', 'Oueslati' and 'Chem_Tat'. The fifth group is made up of two foreign cultivars, 'Manzanilla' and 'Coratina'. Similar results were shown in the dendrogram generated by SSR markers previously studied by Ben Ayed et al. [7] and by the consensus dendrogram (Fig. 2b) generated in this paper using both SSR markers and SNP markers. The observation of the dendrogram does not reveal any correlations between genetic variability and geographical origin, which was also reported by Besnard et al. [23] and Grati-Kamoun et al. [5]. Nevertheless, Rao et al. [24] distinguishes between several olive cultivars located in Campania (southern Italy) by AFLPs markers to assess suspected cases of synonyms and homonyms, thus evaluating their potential relationships with morphological markers. They concluded that the morphological and molecular data yielded different hierarchical patterns.

3.3 Comparison Between SNP and SSR Markers

In this section, the five SNP markers are compared to other molecular markers previously used to study the genetic diversity of 22 olive cultivars as SSR markers [7].

3.3.1 Dendrograms Comparison

The similar matrices constructed with SSRs and SNPs were used to plot a dendrogram consensus for both markers (Fig. 1b). Four groups were obtained by cutting the dendrogram at a degree of similarity equal to 0.77. Group 1 consists of eight cultivars, ‘Chemlali’ (Chem_Chaàl, Chem_Blett, Chem_Dok, Chem_SB, Chem_Nab, Chem_Sous and Chem_Mon) and ‘Zalmati’. This group is common to all dendrograms SSR, SNP and (SNP + SSR). The second group includes four foreign cultivars: Picholine, Ascolana, Coratina and Manzanilla. The third group contains Chemch, Oueslati and Chem_Tat. The fourth group comprises Chetoui_Silia, Chet_Thibar, Chet_Nab and Rkhaymi cultivars. We also notice the existence of a separate group for the foreign cultivars (Picholine, Ascolana, Coratina and Manzanilla) and (Koroneiki and Picholine) for dendrograms based on SSR + SNP (Fig. 1b) and SNP markers (Fig. 1a), respectively. Hence, there is a partial homology between both SSR and SNP dendrograms. The dissimilarity is due to the biallelic nature of SNP markers. Despite the success of SNP markers to discern between Tunisian olive cultivars, consistent lower values of discrimination power than SSR markers previously used for the same cultivars [7] were observed. Indeed, the high number of alleles (47) revealed by the eight polymorphic SSR markers compared to the number of genotypes established by the SNP markers illustrates the difference between Tunisian olive cultivars in their efficiencies.

3.3.2 Cophenetic Correlation

The analysis of the hierarchical classification of Tunisian olive cultivars was performed using molecular markers SSR and SNP. To confirm their performance to differentiate the studied cultivars, we calculated the cophenetic correlation using NTSYSpc software (version 2.1). This correlation measures the degree of fidelity of the dendrogram obtained with the distance matrix or the agreement between two dendrograms and thus measures the clustering structure quality [25]. A value of cophenetic correlation coefficient near 1 indicates that the clustering solution reflects the data more accurately. The obtained profiles were compared using SSR markers (Fig. 2a), SNP markers (Fig. 2b) and both SNP and SSR markers (Fig. 2c). Using SSR markers ($r = 0.946$), the cophenetic correlation coefficient was found to be higher than SNP markers

($r = 0.833$). Moreover, the cophenetic correlation coefficient between SSR and SNP was relatively low ($r = 0.567$), indicating that it is better to use only one marker type to elucidate the relationship between the cultivars. Actually, SSR markers performed significantly better in the absence of such marker type and the clustering technique based on SNP markers are helpful and informative. This is because it generates a high value of cophenetic correlation coefficient ($0.833 > 0.7$), indicating a high agreement level between the clustering and the annotated natural cultivars classes. It is difficult to affirm that the obtained results from SSR markers are more informative since there exist several cluster analysis procedures such as Complete linkage (furthest neighbor), Single linkage (nearest neighbor), Unweighted pair-group average (UPGMA), Weighted pair-group average (WPGMA), Unweighted pair-group centroid (UPGMC), Weighted pair-group centroid (median). This implies that the results of the clustering can often be significantly affected by the appropriate cluster methods, whose selection for a specific application is the main motivation of researchers [26].

3.3.3 Principal Coordinate Analysis (PCoA)

Principal coordinate analysis used to explain genetic variation shows the variation pattern in a multidimensional case and gives a better understanding and interpretation of the relationship between individuals [27]. The relative variance of each coordinate indicates the importance of the related coordinate of total variance expressed in percentage. All the data obtained using 5 SNP and 8 SSR primers were used in principal coordinate analysis with simple matching coefficients of similarity. The first eight coordinates revealed 90.85 % of the total variance. The first two coordinates explained 54.94 % of the total variance, with 35.64 and 19.29 % for coordinates 1 and 2, respectively (Fig. 3), indicating that SNP and SSR markers are scattered over different parts of the genome. Olive cultivars were grouped into four main groups according to their similar characteristics (Fig. 3). Group I contains Chemlali olive cultivars and Zalmati, whereas group II includes the foreign olive cultivars: Coratina, Ascolana and Picholine. While group III comprises Chemchali and Oueslati cultivars, group IV contains only cultivars from northern Tunisia (Chetoui and Rkhaymi).

The obtained results are in accordance with those of Grati-Kamoun et al. [5] in their studies of the genetic diversity of the Tunisian olive oils using AFLP markers according to the fruit size (morphological character). However, Belaj et al. [28] studying the genetic variability of Spain olive cultivars using allozyme and RAPD markers, showed a moderate correlation with the fruit size.

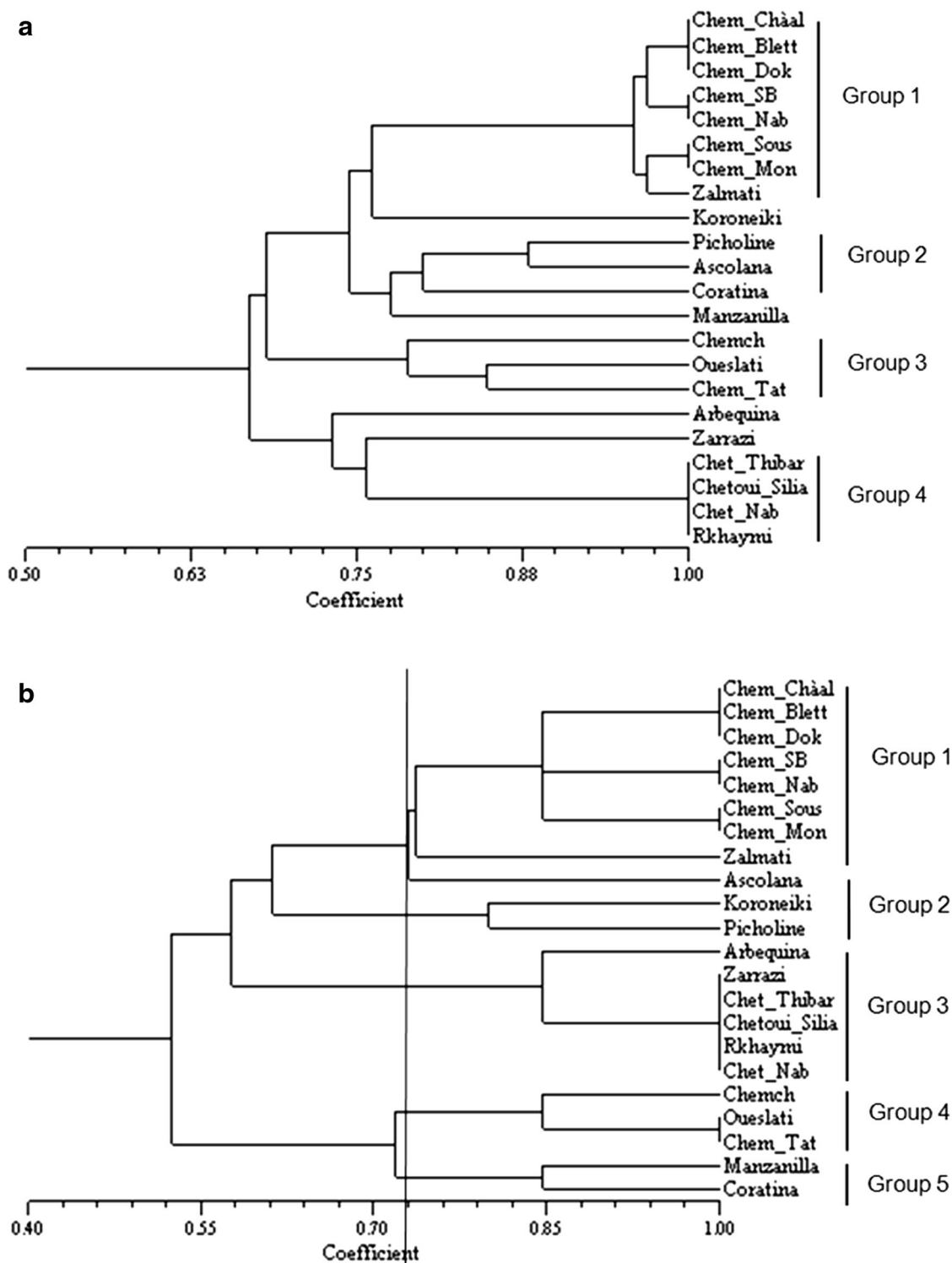


Fig. 1 Dendrogram produced by UPGMA clustering based on means of 5 SNP markers (**a**) and of both 5 SNP and 8 SSR markers (**b**) of 22 olive tree cultivars accessions from different Mediterranean countries

(16 cultivars from geographical regions from Tunisia, 2 cultivars from Spain, 2 cultivars from Italy, 1 cultivar from Greece and 1 cultivar from France)

Therefore, the difference between the morphological markers of the fruit and the chemical and molecular markers of the oil (mainly the previously studied SSR markers) is demonstrated. It could be related to the little

number of the studied genetic markers. Indeed, they do not cover all genomic regions responsible for the variability of agro-morphological and chemical characters, which can also be explained by the effect of environmental factors.

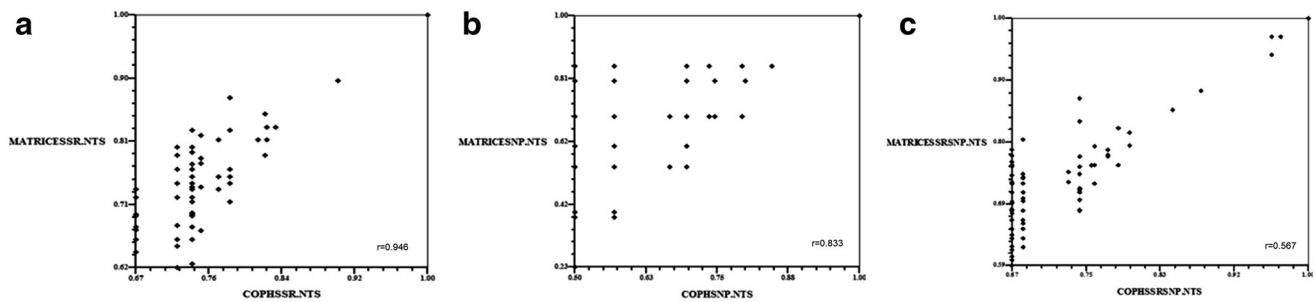


Fig. 2 Cophenetic correlation matrices using SSR (a), SNP (b) and SSR + SNP (c)

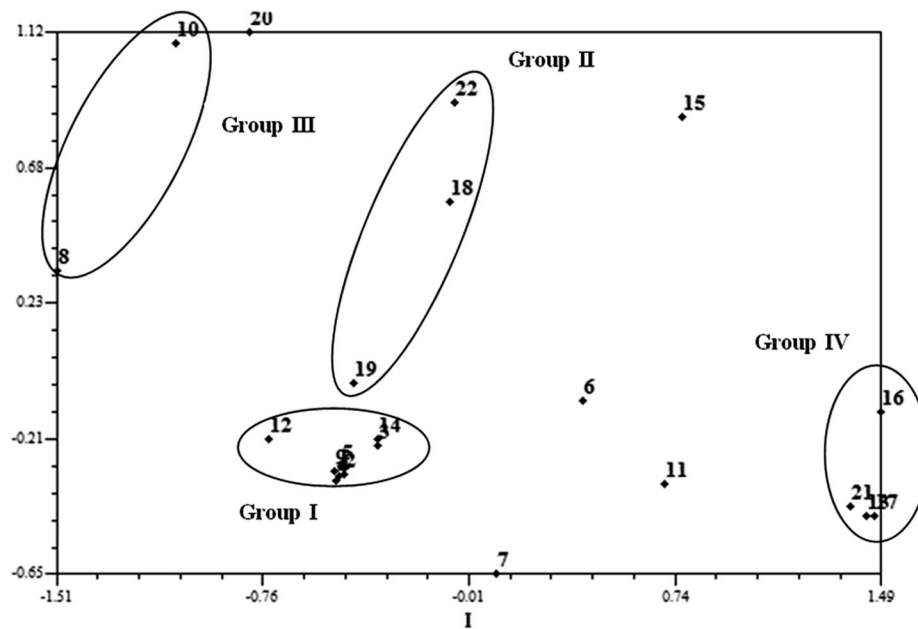


Fig. 3 Principal coordinates analysis (PCoA) plot of the olive cultivars based on the first two principal coordinates (coord1 = 35.646 % and coord2 = 19.298) for 22 cultivars and molecular markers (SSR and SNP; generated using NTSYS-pc) Rholf et al. [19] based on the genetic distance matrix and using DCENTER, EIGEN, and MXPLOT (provided in NTSYS-pc). 1 Chem_Chàal; 2

Chem_Blett; 3 Chem_SB; 4 Chem_Sous; 5 Chem_Dok; 6 Arbequina; 7 Koroneiki; 8 Chemch; 9 Chem_Mon; 10 Oueslati; 11 Zarrazi; 12 Zalmati; 13 Chet_Thibar; 14 Chem_Nab; 15 Chem_Tat; 16 Chetoui_Silia; 17 Rkhaymi; 18 Picholine; 19 Ascolana; 20 Manzanilla; 21 Chet_Nab; 22 Coratina

It is worthy to note that the studied Tunisian cultivars do not have a structure according to the geographical area, size and variety of olives. Indeed, in agreement with the results of Grati-Kamoun [5], Rekik et al. [8] and Taamalli et al. [29], we could show that the pomological and chemical criteria used for the varietal identification and determination of the genetic variability of Tunisian olive oil wealth may be insufficient [30]. Nonetheless, using the dendrograms obtained by genetic data (SSR and SNP), a classification of the studied cultivars according to the average weight of fruit was found. This confirms the hypothesis that the molecular markers such as SSRs and SNPs are powerful for distinguishing the cultivars of the *O. europaea* species and complementing the pomological and chemical analyses. Although sufficient variability exists to discriminate all the

Tunisian olive cultivars by morphological, agronomical and chemical data, the dendrogram based on SSR + SNP markers allowed the identification of clearer structure of tree grouping. This suggests that DNA markers are more informative in depicting genetic relationships. Each marker system measures different aspects of this genetic variability, explaining the lack of consistency in genetic diversity and relationship studies. Despite the efficiency of SSR markers to detect genetic variability and relationships, they should not be seen as a substitute of the traditional agro-morphological descriptors. All these markers systems should be considered as complementary tools to provide a more complete understanding of the diversity available in Tunisia olive cultivars and the way in which it can be best used for olive breeding and olive oil traceability and authenticity [30].

4 Conclusion

This study has revealed that the genetic diversity and distribution of Tunisian olive cultivars are strictly related to many markers. All these marker systems should be considered as complementary tools to provide a more complete understanding of the diversity of available Tunisian olive tree and oil cultivars and the way in which it can be best used for olive breeding and solving traceability and authenticity problems of monovarietal extra virgin olive oil. Therefore, at present, microsatellites (SSRs) are the most appropriate genetic markers used in olive cultivar characterization and olive oil authentication thanks to their advantages. Indeed, SSR markers are multiallelic, codominant, highly polymorphic, widely distributed along the plant genomes, easily amenable to PCR-based analyses and have great reproducibility [31]. Moreover, the SNP markers could give an idea about the characteristics of olive oil. The use of highly advanced genotyping techniques using other SNP markers in these genes will be suitable to confirm our findings and provide automated tools for olive cultivars identification and characterization.

Acknowledgments This work was supported by the Tunisian Ministry of Higher Education and Scientific Research.

Compliance with Ethical Standards

Conflict of interest We declare that we have no conflict of interest.

References

1. Elloumi J, Ben Ayed R, Aifa S (2012) An overview of olive oil biomolecules. *Curr Biotechnol* 1(2):115–124. doi:10.2174/2211550111201020115
2. Barranco D, Rallo L (2000) Olive cultivars in Spain. *HortTechnology* 10(1):107–110
3. 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. *Eur Food Res Technol* 229(5):757–762. doi:10.1007/s00217-009-1111-3
4. Busconi M, Foroni C, Corradi M, Bongiorno C, Cattapan F, Fogher C (2003) DNA extraction from olive oil and its use in the identification of the production cultivar. *Food Chem* 83(1):127–134. doi:10.1016/S0308-8146(03)00218-8
5. 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. *Genet Resour Crop Evol* 53:265–275. doi:10.1007/s10722-004-6130-0
6. Pafundo S, Agrimonti C, Marmiroli N (2005) Traceability of plant contribution in olive oil by amplified fragment length polymorphisms. *J Agric Food Chem* 53(18):6995–7002. doi:10.1021/jf050775x
7. 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. *Eur Food Res Technol* 234(2):263–271. doi:10.1007/s00217-011-1631-5
8. 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
9. Consolandi C, Palmieri L, Doveri S, Maestri E, Marmiroli N, Reale S, Lee D, Baldoni L, Tosti N, Severgnini M, De Bellis G, Castiglioni B (2007) Olive variety identification by ligation detection in a universal array format. *J Biotechnol* 129(3):565–574. doi:10.1016/j.jbiotec.2007.01.025
10. Consolandi C, Palmieri L, Severgnini M, Maestri E, Marmiroli N, Agrimonti C, Baldoni L, Donini P, De Bellis G, Castiglioni B (2008) A procedure for olive oil traceability and authenticity: DNA extraction, multiplex PCR and LDR-universal array analysis. *Eur Food Res Technol* 227(5):1429–1438. doi:10.1007/s00217-008-0863-5
11. Rekik-Hakim I, Grati-Kammoun N, Makhloufi E, Rebai A (2010) Discovery and potential of SNP markers in characterization of Tunisian olive germplasm. *Diversity* 2:17–27
12. Ben Ayed R, Grati-kamoun N, Rebai A (2013) An overview of the authentication of olive tree and oil. *Compr Rev Food Sci Food Saf* 12(2):218–227. doi:10.1111/1541-4337.12003
13. Ben Ayed R, Kallel I, Ben Hassen H, Rebai A (2014) SNP marker analysis for validating the authenticity of Tunisian olive oil. *J Genet* 93(3):148–154. doi:10.1007/s12041-014-0461-1
14. Ben Ayed R, Sans-grout C, Moreau F, Grati-kamoun N, Rebai A (2014) Genetic similarity among tunisian olive cultivars and two unknown feral olive trees estimated through SSR markers. *Biochem Genet* 52(5–6):258–268. doi:10.1007/s10528-014-9645-x
15. Trigui A, Msallem M (2002) Oliviers de Tunisie: catalogue des variétés autochtones et types locaux. Volume I (Identification variétale and caractérisation morpho-pomologique des ressources génétiques oléicoles de Tunisie). IRESA press, Tunis, p 159
16. Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer Associates, Sunderland, p 376
17. Brenner C, Morris JW (1990) Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. In: Madison WI (ed) International symposium on human identification. Promega Corporation, Madison, pp 21–53
18. Jaccard P (1908) Nouvelles recherches sur la distribution florale. *Bull Soc Vaud Sci Nat* 44:223–270
19. Rohlf FJ (1999) NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 2.1. Exeter Software, Setauket
20. Reale S, Doveri S, Díaz A, Angiolillo A, Lucentini L, Pilla F, Martin A, Donini P, Lee D (2006) SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars. *Genome* 49(9):1193–1205
21. Cipriani G, Marrazzo MT, Marconi R, Cimato A, Testolin R (2002) Microsatellite markers isolated in olive (*Olea europaea* L.) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. *Theor Appl Genet* 104(2–3):223–228
22. Muzzalupo I, Stefanizzi F, Perri E (2009) Evaluation of olives cultivated in southern Italy by simple sequence repeat markers. *HortScience* 44(3):582–588
23. Besnard G, Breton C, Baradat P, Khadari B, Bervillé A (2001) Cultivar identification in olive (*Olea europaea* L.) based on RAPD markers. *J Am Soc Hortic Sci* 126(6):668–675
24. Rao R, La Mura M, Corrado G, Ambrosino O, Foroni I, Perri E, Pugliano G (2009) Molecular diversity and genetic relationships of southern Italian olive cultivars as depicted by AFLP and morphological traits. *J Hortic Sci Biotechnol* 84(3):261–266
25. Farris JS (1969) On the cophenetic correlation coefficient. *Syst Zool* 18(3):279–285
26. Saraçlı S, Doğan N, Doğan I (2013) Comparison of hierarchical cluster analysis methods by cophenetic correlation. *J Inequal Appl* 2013:203. doi:10.1186/1029-242X-2013-203
27. Shiria MR, Choukanb R, Aliyevc RT (2014) Study of genetic diversity among maize hybrids using SSR markers and

- morphological traits under two different irrigation conditions. *Crop Breed J* 4(1):65–72
28. 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. *J Am Soc Hortic Sci* 126(1):64–71
 29. 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 Hortric Sci Biotechnol* 82(1):25–28
 30. Ben Ayed R, Ennouri K, Ben Hassen H, Triki MA, Rebai A (2015) Comparison between DNA-based, pomological and chemical markers accomplished by bioinformatic tools to distinguish within Tunisian olive cultivars. *J Fundam Appl Sci* 7(3): 408–421
 31. Ben Ayed R, Ben Hassen H, Ennouri K, Ben Marzoug K, Rebai A (2015) OGDD (Olive Genetic Diversity Database): a microsatellite markers’ genotypes database of worldwide olive trees for cultivar identification and virgin olive oil traceability. *Database: J Biol Databases Curation* 2015:1–9. doi:[10.1093/database/bav090](https://doi.org/10.1093/database/bav090)