ONLINE RESOURCES

SNP marker analysis for validating the authenticity of Tunisian olive oil

RAYDA BEN AYED[1]*, IMEN KALLEL[1], HANEN BEN HASSEN[2] and AHMED REBAI[1]

[1]*Centre of Biotechnology of Sfax, P.B. '1177', 3018 Sfax, Tunisia*
[2]*Laboratory of Physics-Mathematics and Applications, Faculty of Sciences of Sfax, BP 1177, 3000 Sfax, Tunisia*

[Ben Ayed R., Kallel I., Ben Hassen H. and Rebai A. 2014 SNP marker analysis for validating the authenticity of Tunisian olive oil. *J. Genet.* **93**, e148–e154. Online only: http://www.ias.ac.in/jgenet/OnlineResources/93/e148.pdf]

Introduction

Olive (*Olea europaea* L.), which is an important oil-producing crop, is one of the oldest agricultural plant in the Mediterranean basin. The oil obtained is known for its nutritional and healthy benefits compared to other vegetable oils, and can be consumed in its crude form (Roche *et al.* 2000; Elloumi *et al.* 2012). Moreover, the olive oil sector plays an important role in the culture and socio-economy of many Mediterranean countries, including Tunisia.

Traditionally, genetic variation analyses relying on morphological and chemical markers are insufficient to study the relationship and traceability between cultivars due to the environmental effect on the possibly large phenotype and the chemical composition, thus making it expensive (Busconi *et al.* 2003; Ben-Ayed *et al*. 2009, 2013). Recently, several molecular marker types, such as random amplified polymorphic DNA (RAPDs) (Busconi *et al.* 2003), amplified fragment length polymorphisms (AFLPs) (Pafundo *et al.* 2005; Grati Kamoun *et al.* 2006), simple sequence repeats (SSR) (Testolin and Lain 2005; Rekik *et al.* 2008; Ben-Ayed *et al.* 2009, 2012, 2014) and single nucleotide polymorphism (SNP) (Reale *et al.* 2006; Consolandi *et al.* 2008; Rekik-Hakim *et al.* 2010) have been developed. These can be used as both detection of DNA polymorphisms and for effective distinction between different cultivars, thus solving traceability without any environmental influence.

Despite the potential advantages of using SNPs for the authentication of major crop species as coffee (Spaniolas *et al.* 2006), to the best of our knowledge, the identification of SNP markers has not yet been documented in olive oil. Compared with other genetic markers, SNPs are beneficial from a technological viewpoint change in a single nucleotide allows the distinction of very similar cultivars. Moreover, these molecular markers requiring short DNA amplicons for genotyping and are genetically stable; their high density in genomes usually allows several SNPs to occur in a single locus of a few hundred base pairs. They can also occur in coding regions, occasionally leading to amino acid changes in the encoded proteins. They are, generally, biallelic, codominant and open to high throughput genotyping and automation (Brookes 1999).

The ultimate goal of this study was to evaluate the associations between SNP and olive oil quality parameters and their effectiveness in authenticity. Indeed, it highlights that five SNPs localized in four different genes: anthocyanidin synthase, Cu–Zn superoxide dismutase, calcium-binding protein and fatty acid desaturase. The two first SNPs (FAD2.1 and FAD2.3) are localized in the *FAD2* gene, involved in the biosynthesis of highly unsaturated fatty acids (HUFA) from the precursor polyunsaturated fatty acids (PUFA) (Vance and Vaucheret 2001). The third studied SNP, named CALC, is situated in the gene encoding the calcium-binding protein which plays a crucial role in plant defense against stress conditions (cold, drought or salinity) (Hirschi 1999). The other SNP, called SOD, is an insertion/deletion polymorphism type localized in Cu–Zn superoxide dismutase gene which is involved in the oxidative stress, one of the unfavourable conditions (Apel and Hirt 2004; Flexas *et al.* 2006). The last SNP is the ANTHO3 localized in the anthocyanidin synthase. This study particularly aims at evaluating the potential relationship between olive oil cultivars using the five SNPs and their association with the quantitative and qualitative parameters. It also aims at identifying the SNPs usefulness in the traceability of Tunisian olive oil.

Materials and methods

Plant material

A total of 16 Tunisian olive tree cultivars were selected from different geographical regions from north to south of Tunisia. For each cultivar two trees were used and from each tree DNA was extracted from monovarietal olive oil samples.

*For correspondence. E-mail: raydabenayed@yahoo.fr.

Keywords. olive oil; SOD; CALC; FAD; ANTHO; authentification.

DNA extraction

DNA was extracted from oil using QIAamp DNA Stool Mini kit (Qiagen, France). It was extracted from 200 μL of each olive oil sample based on pathogen detection protocol with some modifications (Ben-Ayed *et al.* 2009). DNA was quantified by incorporating Hoechst H33258 dye and detected by spectrofluorometer (Tecan GENIOS Plus, France). Dilution series of Lamda DNA (D150A Promega, France) were used as standard calibration. Genomic DNA thus obtained was dissolved in TE buffer (10 mM Tris-HCl pH 8.1 mM EDTA pH 8) at −20°C.

Genotyping

SNPs recently found in olive genome (Reale *et al.* 2006; Consolandi *et al.* 2008; Rekik-Hakim *et al.* 2010), were exploited as molecular genetic DNA markers since they could discriminate very similar cultivars with a single nucleotide change. The five SNPs under study were selected in the coding regions of genes involved in fruit characteristics and oil compositions.

The SNP SOD (insertion/deletion type) was genotyped by a simple polymerase chain reaction (PCR) 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; figure 1). The PCR product (171 bp) of the SNP (ANTHO3) was digested with *Msp*I restriction enzyme (Fermentas, Life Sciences, France) at 37°C overnight. This restriction enzyme recognizes the sequence AA/GG. The G-allele carrying the PCR product was cleaved once by the enzyme generating two fragments (64 and 107 bp). The PCR product (476 bp) of SNP (CALC) was digested by *Bst*ZI restriction enzyme (Promega) at 50°C overnight. This restriction enzyme recognizes the sequence CC/GG. The C-allele carrying the PCR product was cleaved once by the enzyme producing two fragments (316 and 160 bp). The two other SNPs (FAD2.1 and FAD2.3) were analysed using PCR-RFLP. The PCR products (241 and 240 bp) of the SNPs (FAD2.1) and (FAD2.3) were digested using *Bam*HI restriction enzymes (Fermentas) and *Alw26I*, respectively, at 37°C overnight. The sizes of the restriction fragments of PCR products were 224 and 17 bp, 130 and 110 bp for CC genotype of FAD2.1 SNP and FAD2.3 SNP, respectively. All digestion products were separated by electrophoresis an ethidium bromide-stained agarose (3% NuSieve) gel and visualized under UV light.

Statistical analysis

The analysis of the relationship between SNP markers, and the quantitative and qualitative parameters was performed in many steps using several statistical techniques. Firstly, the differences between the classes of qualitative traits in allele and genotype frequencies were evaluated by the chi-square test. For quantitative traits, a Student's *t*-test was employed to assess the significant difference between the means of genotype groups for each SNP. The SPSS program ver. 13.0 (SPSS, Chicago, USA) was applied to study the correlation between quantitative and qualitative parameters. All the tests were declared statistically significant when P values were <0.05. Finally, to test the association of all SNPs with quantitative traits, a multi-way analysis of variance (only main effects were included) was performed. Multinomial logistic regression was also used to test the associations of the five SNPs with qualitative traits separately.

Results and discussion

Genotyping results and characteristics of the studied SNP markers

The expected fragments for each SNP are detailed in the material and method section. The observed heterozygosity

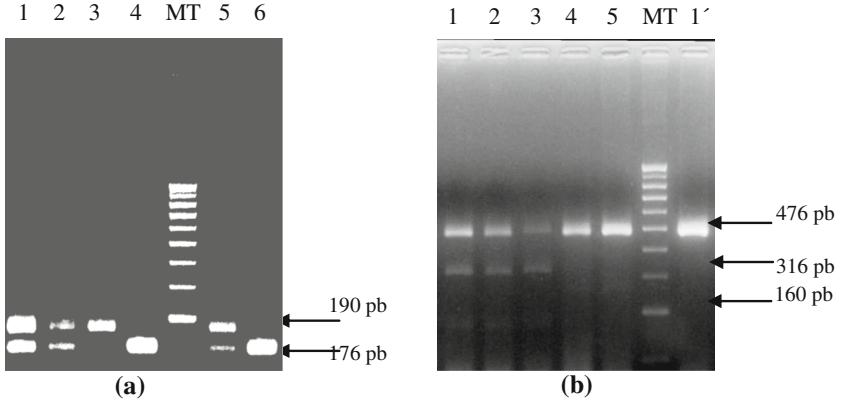

Figure 1. (a) PCR amplification of DNA sequences specific SNPs type insertion/deletion (SOD) on 2% agarose gel; (b) digestion products of SNPs type CALC tested on 3% agarose gel. 1, Chemlali_Sousse; 2, Zalmati; 3, Oueslati; 4, Koroneiki; 5, Chetoui_Nabeul; 6, Picholine; 1′, PCR product of SNP CALC for the sample Chemlali_Sousse; MT, size marker (100 bp DNA step ladder, Promega).

Table 1. Characteristics of SNP markers used for DNA amplification.

Gene name	SNPs	SNP type	Tm^a	Primer sequence (5′→3′)	Size of amplicon (bps)	Enzyme	References	H_o	H_e	PD
Anthocyanidine synthase	ANTHO3	G/A	57	5′-gttggtggcatggaagaact-3′ 5′-gacccattttccctcacaga-3′	171	*Msp*I	*	0.714	0.481	0.430
Cu–Zn superoxide dismutase	SOD	I/D	57	5′-ttcccaccgatattgttaaggg-3′ 5′-catatagcatagcacccgcaac-3′	190/176	–	Reale *et al.* (2006)	0.681	0.490	0.475
Calcium-binding protein	CALC	C/G	60	5′-cacggacaggaattccaagccttca-3′ 5′-tgccgcttttgttcgtcatcattttct-3′	476	*Bst*ZI	*	0.727	0.462	0.396
Fatty acid desaturase	FAD2.1	C/T	57	5′-gccccattgtttccagggatc-3′ 5′-ggtgtcgtctagccactggt-3′	241	*Bam*HI	*	0.428	0.336	0.489
Fatty acid desaturase	FAD2.3	C/G	56	5′-gggagacattacgaccggtt-3′ 5′-cagggagggatgtgtatgct-3′	240	*Alw*26I	*	0.636	0.500	0.528
Total								3.188	2.272	2.321
Mean								0.637	0.454	0.464

*Primers designed in this study. I, insertion; D, deletion. For each locus, the type of polymorphism, size of amplicon in base pairs, frequency of heterozygotes observed (H_o), frequency of heterozygotes expected (H_e) and the power of discrimination (PD) are reported. [a]Annealing temperature for PCR amplification.

of each marker ranged from 0.428 (FAD2.1) to 0.727 (CALC) (0.637 being the average), while the expected heterozygosity was from 0.336 to 0.5 with an average of 0.454, which indicates an excess of heterozygosity for all markers (table 1).

The discriminating power (DP) varies from 0.396 for the CALC marker to 0.528 for FAD2.3 marker with an average value of 0.464. Similar values were found by Reale *et al.* (2006) in DNA samples extracted from 65 Tunisian olive cultivars. Although the average values are lower than those of Rekik *et al.* (2008) with SSR markers (0.71), they are significantly higher than those reported by Cipriani *et al.* (2002) in 12 Italian cultivars (0.44) and Muzzalupo *et al.* (2009) in 39 Italian cultivars (0.38) using SSR markers. The combined DP was 0.956, indicating that the probability of finding two cultivars with the same genotypes combination for the five SNP markers was about 4%. This showed that the set of the DP of these markers was lower than SSR markers (0.999) in the same cultivars set (Ben-Ayed *et al.* 2012), which was expected since the SNPs were biallelic, whereas SSR markers had an average of four alleles.

The allele frequencies of the five studied SNPs showed that there is a dominance of all markers, except for FAD2.3 in which the two alleles have similar frequencies. Most of the studied cultivars have heterozygous genotypes (table 1) except FAD2.1 marker in which the frequency of heterozygous genotype is about 43%.

Bivariate analysis

Relationship between SNPs polymorphisms and qualitative parameters: Table 2 shows the absence of any significant associations between the three SNP (SOD, ANTHO3 and FAD2.3) genotypes and none of the qualitative traits was considered in this work. However, significant associations were observed for CALC (C/G) SNP with the fruit maturation level ($P = 0.002$) and with the fruit oil content ($P = 0.006$). Other considerable associations were also observed between the CALC marker and tree tolerance capacity and productivity. Indeed, the late maturation step of the olive was essentially observed for cultivars having CALC-CG genotype representing 62.5% of the sample, 80% of which are Chemlali cultivars. Nevertheless, only 25% of the studied varieties have a middling maturation and GG genotype, and none of the varieties carrying the GG genotype has a late maturation. Therefore, for this marker, the C allele is strictly linked to the late maturation of the olive fruit. Further, it can be noted that the fruit oil content of olive depend on this marker genotype variation. All varieties carrying CG genotype have high or very high fruit oil content, and represented about 70% of the total samples, including Chemlali, Oueslati and Zalmati cultivars. None of the cultivars with homozygote GG genotype showed a high fruit oil content. In the same context, the productivity of the olive trees depend on the genotype variation of CALC markers. Actually, 50% of the studied varieties which carried the CG genotype (including the eight Chemlali) have a better alternating productivity than the other varieties. These results suggest that the C allele of CALC marker may be related to the higher fruit oil content and higher productivity level.

A highly significant association of FAD2.1 genotype with the maturation period can be observed ($P = 0.008$) (table 2). All varieties carrying the homozygous TT genotype exhibited late fruit maturation. They were mainly Chemlali cultivar representing 50% of the samples. While 31.25% of the studied varieties carrying the heterozygous CT genotype (majority Chetoui cultivar) require an average fruit maturation, only one variety (Zarrazi) an early maturation and carry CT genotype. A second interesting association of FAD2.1 genotype variation was found with the rate of oil per fruit ($P = 0.023$). The varieties carrying the homozygous TT genotype (43.75% of all samples, whose majority being

Table 2. Association between CALC and FAD2.1 genotypes and qualitative parameters.

Polymorphisms / Qualitative parameters		CALC			FAD2.1		
		CG (%)	GG (%)	Chi-square (*P* value)	TT (%)	CT (%)	Chi-square (*P* value)
Fruit maturation (FM)	Early (premature)	0	6.25	12.27	0	6.25	9.6
	Middling	6.25	25	**0.002**	0	31.25	**0.008**
	Lately	62.5	0		50	12.5	
Fruit oil content (FOC)	Weak	0	6.25	12.50	0	6.25	9.5
	Middling	0	18.75	**0.006**	0	18.75	**0.023**
	High	50	0		43.75	6.25	
	Very high	18.75	6.25		6.25	18.75	
Productivity	Weak	0	12.5	12.89	0	12.5	10.5
	Weak and alternating	0	6.25	**0.012**	0	6.25	**0.033**
	Middling and alternating	6.25	12.5		0	18.75	
	High and alternating	50	0		43.75	6.25	
	Very high	12.5	0		6.25	6.25	
Tolerance	Cold	0	6.25	12.89	0	6.25	9.833
	Cold and salinity	0	18.75	**0.012**	0	18.75	**0.043**
	Drought	12.5	6.25		6.25	12.5	
	Drought and salinity	50	0		43.75	6.25	

Variable that has statistical significance for all tests was declared when *P* values are <0.05 are in bold.

Chemlali cultivar) provide higher content in oil per fruit than that in the varieties with CT genotypes (6.25% of the samples, whose majority was Chetoui cultivar). Likewise, a significant association of the FAD2.1 marker with productivity parameter was observed ($P = 0.033$) (table 2). Therefore, the varieties that carried the homozygous TT genotype, gave a high and alternating productivity (43.75% of the samples), whereas, only one variety (Zarrazi) with heterozygous genotype CT gave a high and alternating productivity.

Relationship between SNPs polymorphisms and quantitative parameters: Table 3 shows that high significant associations were found between CALC SNP and the three parameters, namely carotenes, weight of fruit at maturation (WFM) and cholesterol. However, a significant difference of the average rate in carotene pigment, between the heterozygote varieties with CG-CALC genotype and GG-CALC genotypes ($P = 0.008$) was observed. A positive relationship between the rate of cholesterol and the genotype variation for this marker ($P = 0.025$) also be noted. In that case, the varieties with CG genotypes have the highest cholesterol rate. The third significant association of CG-CALC variant was shown with the polyunsaturated fatty acid rate (C18:3) ($P = 0.037$). These three positive associations between the CALC polymorphisms and the parameters carotene, oil content cholesterol and linoleic acid suggest that the heterozygote varieties with CG genotypes produce, on average, higher levels of these compounds than the GG genotypes varieties.

With respect to SNP (FAD2.1), five highly significant associations with some quantitative parameters are proved. These correlations are observed with the accumulation of both linoleic (C18:2; $P = 0.008$) and linolenic (C18:3; $P = 0.001$) unsaturated fatty acids, with cholesterol rate ($P < 0.001$), carotene rate ($P = 0.002$) and finally with WFM parameter ($P = 0.003$) (table 3). All these positive associations concern essentially the homozygous varieties (TT- FAD2.1) except the cholesterol content and WFM parameters, which are mostly linked to the heterozygote cultivars (CT- FAD2.1). It has been described that this gene is involved in the synthesis of the unsaturated fatty acid (Vance and Vaucheret 2001), which suggests a direct effect of FAD2.1 genotypic variations on the rate of polyunsaturated fatty acid (such as C18:2 and C18:3) for each variety. Substantial associations were found with the polyunsaturated fatty acid C18:3 ($P = 0.002$) content, carotene content ($P < 0.001$) and, at a lower degree, to the cholesterol rate ($P = 0.015$). In fact, the homozygous varieties GG (all Chemlali cultivar) have a higher content of polyunsaturated fatty acids (particularly C18:3) than the homozygous CC and heterozygous CG varieties (50% of which are Chetoui). Moreover, GG varieties have more carotene pigment than the two other genotypes. The cholesterol rate was significantly higher in the varieties carrying the heterozygous genotype CG-FAD2.3 than the two other homozygous genotypes (tables 3, a&b).

Regarding SOD, this marker is an insertion/deletion polymorphism, located in the Cu–Zn superoxide dismutase gene involved in the defense mechanism of the plant against the oxidative stress (Apel and Hirt 2004) and the hydric stress (Flexas *et al.* 2006). The analysis of genotype association of SOD marker with quantitative traits has not shown any significant association, except for the rate of unsaturated fatty acid type C18:2. This relationship is highly

Table 3(a). Association between SOD, CALC, FAD2.1, FAD3.2 and ANTHO3 genotypes and quantitative parameters of olive oil.

Parameters SNP		Cholesterol		β-sitosterol		Chlorophyll		Carotene		FWM	
		Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*
SOD	ID	0.0375 ± 0.03	**0.045**	84.10 ± 0.64	0.9	2.313 ± 1.08	0.519	5.239 ± 1.53	0.598	1.166 ± 0.83	0.933
	II	0.083 ± 0.01		84.03 ± 1.56		2.767 ± 0.85		4.733 ± 0.90		1.616 ± 0.75	
CALC	CG	0.034 ± 0.03	**0.025**	83.95 ± 0.92	0.322	2.777 ± 1.10	0.070	5.736 ± 1.35	**0.008**	1.238 ± 0.42	<**0.001**
	GG	0.080 ± 0.0		84.45 ± 0.30		1.749 ± 0.37		3.843 ± 0.08		2.566 ± 0.63	
FAD 2.1	CT	0.081 ± 0.0	<**0.001**	84.27 ± 0.95	0.438	2.131 ± 0.75	0.255	4.177 ± 0.67	**0.003**	2.21 ± 0.79	**0.001**
	TT	0.016 ± 0.17		83.92 ± 0.72		2.782 ± 1.28		6.113 ± 1.33		1.096 ± 0.08	
FAD3.2	CC	0.042 ± 0.03	**0.015**	83.675 ± 0.12	0.132	2.300 ± 0.07	0.212	5.650 ± 1.65	<**0.001**	1.460 ± 0.63	0.110
	CG	0.07 ± 0.02		84.543 ± 1.06		2.056 ± 0.70		4.064 ± 0.40		2.041 ± 0.88	
	GG	0.01 ± 0.0		83.70 ± 0.0		3.174 ± 1.45		6.800 ± 0.0		1.070 ± 0.00	
ANTHO3	GA	0.035 ± 0.03	**0.041**	84.15 ± 0.65	0.616	2.364 ± 1.12	0.802	5.401 ± 1.48	0.225	1.691 ± 0.86	0.749
	AA	0.077 ± 0.01		83.90 ± 1.30		2.525 ± 0.84		4.375 ± 1.03		1.537 ± 0.63	

P, *P* value of student's test; SD, standard deviation; FWM, fruit weight at maturation.
Variable that has statistical significance for all tests was declared when *P* values are <0.05 are in bold.

important, and the varieties having the ID-SOD genotype (81.25% of all samples) produce more linoleic acid ($P = 0.008$) than the other II-SOD cultivars which regroup tree varieties: Chemlali Tataouine, Chemcheli and Oueslati) (table 3b).

For ANTHO3 marker analysis, only two genotypes for ANTHO3 SNP, AA and AG, were used. About 75% of the varieties were heterozygous AG-ANTHO3, including Chetoui and Chemlali cultivars. A significant association was established with cholesterol rate ($P = 0.041$), with a higher average for homozygous cultivars (AA) (25% of all samples) (tables 3, a&b).

Multivariate analysis using variance and logistic regression

This section focusses on assessing the relationship between the oil parameters and the studied SNP markers using the analysis of variance. The results (table 4) have shown that a significant association (after Bonferroni correction for multiple testing) exists between the four SNPs (CALC, ANTHO3, FAD2.1 and FAD2.3) with the content of C18:3 and carotenes. The associations of FAD2.1 with the content of C18:3 ($P = 0.014$), carotenes ($P = 0.011$) and cholesterol ($P = 0.001$) were confirmed by multinomial logistic regression analysis using genotype markers as dependent variables

Table 3(b). Association between SOD, CALC, FAD2.1, FAD3.2 and ANTHO3 genotypes and quantitative parameters of olive oil.

Parameters SNP		Oleic acid C18:1		Linoleic acid C18:2		Linolenic acid C18:3		Unsaponifiable		Oxidative stability	
		Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*	Mean ± SD	*P*
SOD	ID	0.226 ± 0.03	0.230	16.31 ± 2.94	**0.008**	0.638 ± 0.62	0.312	1.259 ± 0.09	0.235	45.24 ± 15.25	0.191
	II	0.250 ± 0.10		10.87 ± 0.96		0.600 ± 0.00		1.027 ± 0.59		57.73 ± 4.24	
CALC	CG	0.238 ± 0.02	1.143	15.59 ± 3.33	0.615	0.651 ± 0.06	**0.037**	1.178 ± 0.34	0.706	45.62 ± 8.66	0.445
	GG	0.214 ± 0.04		14.62 ± 3.98		0.588 ± 0.03		1.248 ± 0.10		51.90 ± 24.10	
FAD 2.1	CT	0.227 ± 0.03	0.694	13.21 ± 3.62	**0.010**	0.593 ± 0.02	**0.003**	1.153 ± 0.36	0.513	54.09 ± 18.60	0.073
	TT	0.233 ± 0.02		17.37 ± 1.55		0.670 ± 0.06		1.268 ± 0.10		41.08 ± 4.00	
FAD3.2	CC	0.250 ± 0.01	0.224	15.80 ± 2.60	0.081	0.64 ± 0.07	**0.002**	1.141 ± 0.04	0.542	44.75 ± 5.50	0.215
	CG	0.230 ± 0.04		13.60 ± 3.62		0.593 ± 0.02		1.14 ± 0.36		53.66 ± 18.66	
	GG	0.212 ± 0.005		18.16 ± 1.68		0.700 ± 0.00		1.20 ± 0.00		38.25 ± 0.50	
ANTHO3	GA	0.224 ± 0.03	0.144	16.23 ± 3.05	0.055	0.645 ± 0.06	0.098	1.236 ± 0.07	0.565	45.51 ± 15.90	0.343
	AA	0.250 ± 0.01		12.48 ± 3.31		0.590 ± 0.02		1.130 ± 0.52		53.80 ± 8.60	

P, *P* value of student's test; SD, standard deviation.
Variable that has statistical significance for all tests was declared when *P* values are <0.05 are in bold.

Table 4. *P* values of Fisher tests for the association study between SNP markers and oil quality characteristics.

Model	SOD	FAD2.3	FAD2.1	CALC	ANTHO3
WFM	–	0.384	0.271	0.098	0.948
Acidity	–	0.143	0.853	0.275	0.920
Total polyphenol content (mg/kg)	–	0.333	0.053	0.346	0.408
Rate unsaponifiable (%)	–	0.632	0.435	0.262	0.669
Oxidative stability (h at 100°C)	–	0.747	0.594	0.696	0.996
Chlorophyll content (ppm)	–	0.580	0.241	0.081	0.142
C18:2	–	0.526	0.161	0.228	0.909
C18:3	–	**0.047***	**0.046***	**0.006**	**0.002**
Cholesterol	–	0.513	**0.002**	**3.10^{-5}**	**7.10^{-6}**
Carotene content (ppm)	–	**10^{-5}**	**10^{-6}**	**7.10^{-8}**	**4.10^{-8}***
β-sitosterol	–	0.165	0.808	0.700	0.618

Variable that has statistical significance for all tests was declared when *P* values are <0.05 are in bold. *Variable that has statistical significance for all tests was declared when *P* values are <0.05 and has biological relevant.

and oil characteristics as explanatory variables (table 5). Nonetheless, only the plant tolerance showed a significant association with the CALC marker. These analyses have also confirmed the association already identified in the bivariate analyses as those of FAD2.3 and FAD2.1 with polyunsaturated fatty acids (C18:2 and C18:3). Indeed, most of the results are biologically relevant since these two markers are located in the FAD2 gene involved in the synthesis of polyunsaturated fatty acids. Besides, the ANTHO3 marker, located in anthocyanidin synthase gene is involved in the fruit pigmentation during maturation, which may explain the high association between these markers and the carotene content ($P = 4.10^{-8}$).

This study has revealed that the genetic diversity and distribution of Tunisian olive cultivars is related to many qualitative and quantitative parameters. Numerous and highly important associations were established, especially with the content in polyunsaturated fatty acid (C18:2 and C18:3), cholesterol rate, high and alternating productivity. Therefore, the SNP markers could help in the determination of olive oil characteristics, quality and to confirm its authenticity.

Table 5. *P* values given by the binary logistic regression analysis.

Model	SOD	FAD2.3	FAD2.1	CALC	ANTHO3
WFM	0.981	–	**0.004**	**0.002**	0.804
Acidity	0.289	–	0.803	0.584	0.175
Total polyphenol content (mg/kg)	0.450	–	**0.001**	**0.002**	0.874
Rate unsaponifiable (%)	0.069	–	0.347	0.705	**0.004**
Oxidative stability (h at 100°C)	0.273	–	0.001	**0.004**	0.635
Chlorophyll content (ppm)	0.676	–	0.065	**0.090**	0.374
C18:2	**0.007**	–	**0.044***	0.903	**0.045**
C18:3	0.248	–	**0.014***	**0.074**	**0.049**
Cholesterol	0.103	–	**0.001**	**0.012**	0.089
Carotene content (ppm)	0.607	–	**0.011**	**0.017**	0.179
β-sitosterol	0.090	–	0.773	0.103	0.059
Fruit maturation	0.861	–	**0.003**	**0.002**	0.765
Fruit shape	0.672	–	0.231	0.101	0.588
Pollinization	0.672	–	0.231	0.101	0.588
Tolerance	0.425	–	**0.024**	**0.002***	0.308
FOC	0.243	–	0.409	0.071	**0.047**
Productivity of the tree	0.274	–	0.288	**0.044**	0.071

Variable that has statistical significance for all tests was declared when *P* values are <0.05 are in bold. *Variable that has statistical significance for all tests was declared when *P* values are <0.05 and have biological relevant.

Acknowledgements

This work was supported by the Tunisian Ministry of Higher Education and Scientific Research. The authors would like to thank Mrs Leila Mahfoudhi, the English teacher in the Sfax Faculty of Science, for proofreading the manuscript.

References

Apel K. and Hirt H. 2004 Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Ann. Rev. Plant Biol.* **55**, 373–399.

Ben-Ayed R., Grati-Kamoun N., Moreau F. and Rebaï A. 2009 Comparative study of microsatellite profiles of DNA from oil and leaves of two Tunisian olive cultivars. *Eur. Food Res. Tech.* **229**, 757–762.

Ben-Ayed R., Grati-Kamoun N., Sans-grout C., Moreau F. and Rebai A. 2012 Characterization and authenticity of virgin olive oil (*Olea europaea* L.) cultivars by microsatellite markers. *Eur. Food Res. Tech.* **234**, 263–271.

Ben-Ayed R., Grati-Kamoun N. and Rebai A. 2013 An overview of the authentication of olive tree and oil. *Compr. Rev. Food Sci. Food Safety* **12**, 218–227.

Ben-Ayed R., Sans-Grout C., Moreau F., Grati-Kamoun N. and Rebai A. 2014 Genetic similarity among Tunisian olive cultivars and two unknown feral olive trees estimated through SSR markers. *Biochem. Genet.* **52**, 258–268.

Brookes A. J. 1999 The essence of SNPs. *Gene* **234**, 177–186.

Busconi M., Foroni C., Corradi M., Bongiorni C., Cattapan F. and Fogher C. 2003 DNA extraction from olive oil and its use in the identification of the production cultivar. *Food Chem.* **83**, 127–134.

Cipriani G., Marrazzo M. T., Marconi R., Cimato A. and 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**, 223–228.

Consolandi C., Palmieri L., Severgnini M., Maestri E., Marmiroli N., Agrimonti C. *et al.* 2008 A procedure for olive oil traceability and authenticity: DNA extraction, multiplex PCR and LDR-universal array analysis. *Eur. Food Res. Technol.* **227**, 1429–1438.

Elloumi J., Ben-Ayed R. and Aifa S. 2012 An overview of olive oil biomolecules. *Curr. Biotechnol.* **1**, 115–124.

Flexas J., Bota J., Galmés J., Medrano H. and Ribas-Carbó M. 2006 Keeping a positive carbon balance under adverse conditions: responses of photosynthesis and respiration to water stress. *Physiol. Plant.* **127**, 343–352.

Grati Kamoun N., Lamy Mahmoud F., Rebai A., Gargouri A., Panaud O. and Saar A. 2006 Genetic diversity of Tunisian olive tree (*Olea europaea* L.) cultivars assessed by AFLP markers. *Genet. Res. Crop Evol.* **53**, 265–275.

Hirschi K. D. 1999 Expression of Arabidopsis CAX1 in tobacco: altered calcium homeostasis and increase stress sensitivity. *Plant Cell* **11**, 2113–2122.

Muzzalupo I., Stefanizzi F. and Perri E. 2009 Evaluation of olives cultivated in Southern Italy by simple sequence repeat markers. *Hort. Sci.* **44**, 582–588.

Pafundo S., Agrimonti C. and Marmiroli N. 2005 Traceability of plant contribution in olive oil by amplified fragment length polymorphisms. *J. Agr. Food Chem.* **53**, 6995–7002.

Reale S., Doveri S., Díaz A., Angiolillo A., Lucentini L., Pilla F. *et al.* 2006 SNP-based markers for discriminating olive (*Olea europaea L.*) cultivars. *Genome* **15**, 1193–1209.

Rekik I., Salimonti A., Grati-Kamoun N., Muzzalupo I., Lepais O., Gerber S. *et al.* 2008 Characterization and identification of Tunisian olive tree varieties by microsatellite markers. *Hort. Sci.* **43**, 1371–1376.

Rekik-Hakim I., Grati-Kammoun N., Makhloufi E. and Rebaï A. 2010 Discovery and potential of SNP Markers in characterization of Tunisian olive germplasm. *Diversity* **2**, 17–27.

Roche H. M., Gibney M. J., Kafatos A., Zampelas A. and Williams C. M. 2000 Beneficial properties of olive oil. *Food Res. Int.* **33**, 227–231.

Spaniolas S., May S. T., Bennett M. J. and Tucker G. A. 2006 Authentication of coffee by means of PCR–RFLP analysis and lab-on-a-chip capillary electrophoresis. *J. Agric. Food Chem.* **54**, 7466–7470.

Testolin R. and Lain O. 2005 DNA extraction from olive oil and PCR amplification of microsatellite markers. *Food Sci.* **70**, 108–112.

Vance V. and Vaucheret H. 2001 RNA silencing in plants-defense and counter defense. *Science* **292**, 2277–2280.

Received 24 June 2014, in revised form 2 August 2014; accepted 13 August 2014
Unedited version published online: 19 August 2014
Final version published online:19 December 2014