CHILEAN JOURNAL of AGRICULTURAL RESEARCH

RESEARCH

Influence of cultivar and location on olive trees production revealed by geno-phenotypic profiling

Dina A. El-Alakmy[1], Ahmed B. EL-Mansy[2], Diaa Abd El Moneim[3*], Mohamed A. Awad[4], Hagar Tarek Elhefnawi[5], and Mahmoud Magdy[5, 6*]

[1]Arish University, Faculty of Environmental Agricultural Sciences, Department of Plant Production (Fruit Branch), El-Arish 45511, Egypt.
[2]Arish University, Faculty of Environmental Agricultural Sciences, Department of Plant Production (Vegetables Breeding and Biotechnology), El-Arish 45511, Egypt.
[3]Arish University, Faculty of Environmental Agricultural Sciences, Department of Plant Production (Genetic Branch), El-Arish 45511, Egypt.
[4]Arish University, Faculty of Environmental Agricultural Sciences, Department of Plant Production, El-Arish 45511, Egypt.
[5]Ain Shams University, Faculty of Agriculture, Department of Genetics, 11241, Cairo, Egypt.
[6]Murcia University, Faculty of Biology, Department of Plant Biology, 30100, Murcia, Spain.
[*]Corresponding authors (dabdelmoniem@aru.edu.eg; mahmoud.magdy@um.es)
Received: 21 March 2025; Accepted: 3 July 2025, doi:10.4067/S0718-58392025000500720

ABSTRACT

Olive (*Olea europaea* L.) is a key agricultural species in Mediterranean and semi-arid regions, where optimizing fruit and oil traits is crucial for both economic and nutritional value. This study investigates the impact of cultivar genetics and location on fruit characteristics, oil yield, quality parameters, and fatty acid composition in four olive cultivars—Picual, Sebhawy, Chemlali, and Manzanillo—cultivated in El-Sheikh Zuweid and El-Qantara. El-Sheikh Zuweid favored larger fruit sizes, higher oil content (21.58 ± 1.52%), and monounsaturated fatty acids (MUFA, 70.49 ± 1.56%), while El-Qantara exhibited higher polyunsaturated fatty acids (PUFA, 13.74 ± 2.79%) and linoleic acid (12.22 ± 2.80%). Among cultivars, 'Picual' excelled in fruit morphology, while 'Sebhawy' achieved the highest oil content (22.52 ± 1.29%), total soluble solids (23.05 ± 0.73%), and polyphenols (0.068 ± 0.010 g 100 g^{-1}). 'Chemlali' recorded the highest MUFA (69.57 ± 1.34%) and oleic acid (67.58 ± 2.87%), while 'Sebhawy' had the highest PUFA (14.64 ± 1.80%). ANOVA confirmed significant effects of location and cultivar on most traits, with interaction effects observed for total soluble solids and fatty acid composition. Genetic analyses using inter simple sequence repeat (ISSR) and start codon targeted (SCoT) markers showed moderate to high polymorphism, identifying 114 and 178 loci, respectively, with a combined Shannon diversity index of 7.33. Principal component analysis (PCA) and k-means clustering highlighted associations between specific molecular bands and desirable traits, such as fruit weight and oleic acid content. With 'Sebhawy' excelling in oil-related traits and 'Picual' demonstrating superior morphological adaptability in favorable conditions, these findings emphasize the role of genetic and environmental factors in olive trait development and the value of integrating morphological and molecular data for breeding programs.

Key words: Fatty acid composition, genetic diversity, genotype-environment interaction, ISSR markers, marker-assisted selection (MAS), morphological traits, olive cultivar selection, olive oil quality, principal component analysis (PCA), SCoT markers.

INTRODUCTION

The olive tree (*Olea europaea* L.) is a species of substantial agronomic and economic importance, especially in the Mediterranean region, where olive oil production represents a key sector in agriculture. Olive oil is valued for its unique fatty acid profile, rich in monounsaturated fats, and high levels of antioxidants, such as polyphenols, which contribute to both its health benefits and stability (Jimenez-Lopez et al., 2020). However, optimizing olive oil yield and quality is challenging due to the complex interaction between genetic, environmental, and management factors (Iqbal et al., 2023). Selection of appropriate cultivars and breeding strategies adapted to specific environmental conditions is therefore essential for enhancing production efficiency and oil quality.

The morphological characterization of olive cultivars has been traditionally used to assess fruit and oil traits, such as fruit size, weight, and oil content, which are key factors in determining market value and suitability for specific production environments (Navas-López et al., 2020). However, morphological traits alone are often insufficient to capture the genetic diversity within and among cultivars, as they can be influenced by environmental factors (Souabni et al., 2023). Molecular markers, such as inter simple sequence repeat (ISSR) and start codon targeted (SCoT) markers, have thus become valuable tools for genetic analysis in olives, providing stable and reliable insights into genetic diversity and aiding in the identification of genetic linkages to desirable traits (Alharbi et al., 2024).

The ISSR and SCoT markers have been widely used in olive research for assessing genetic diversity, cultivar identification, and marker-assisted selection (MAS) due to their ability to detect polymorphisms and reveal genetic relationships among cultivars (Galatali et al., 2021). These markers are particularly beneficial in breeding programs, where they can assist in selecting genotypes with favorable traits, such as high oil content or specific fatty acid profiles, which are critical for both yield and quality optimization (Alhaithloul et al., 2024; Alharbi et al., 2024). The integration of molecular and morphological data enables breeders to link specific molecular markers with agronomically valuable traits, facilitating the development of cultivars that are better adapted to local environmental conditions.

This study aims to analyze the combined effects of cultivar and location on olive fruit characteristics, oil yield, quality parameters, and fatty acid composition, using both morphological assessments and molecular markers. By integrating ISSR and SCoT marker data with morphological traits, we seek to identify specific molecular markers associated with desirable traits and assess the genetic diversity among four olive cultivars grown in two contrasting locations in Egypt, El-Sheikh Zuweid and El-Qantara. This approach provides insights into Genotype × Environment interactions, enabling targeted selection and breeding strategies to optimize olive oil production in variable environments.

MATERIALS AND METHODS

Study locations and ecological description

This study was conducted in two olive (*Olea europaea* L.)-growing regions in Egypt: El-Sheikh Zuweid and El-Qantara, both characterized by unique ecological conditions influencing olive growth and oil quality. El-Sheikh Zuweid, (31.21° N, 34.09° E; 88 m a.s.l.), North Sinai, this region has a semi-arid Mediterranean climate with hot, dry summers and mild, wet winters. Average annual rainfall is 175 mm, primarily during winter (December-February). Summers reach up to 25 °C in July and August, while winters drop to around 8 °C (FAO, 2021; EMA, 2022). Soil in this area is generally sandy to sandy loam, with moderate organic matter and good drainage, suitable for drought-tolerant olive trees.

El-Qantara, located in Ismailia Governorate along the Suez Canal (30.85° N, 32.31° E; 39 m a.s.l.), has a Mediterranean climate with higher humidity than El-Sheikh Zuweid. Rainfall averages 2.64 mm annually, concentrated during winter (FAO, 2021). Summers peak at approximately 36 °C, with winters averaging a minimum of 12 °C. The soil is predominantly clay-loam, with better fertility and moisture retention, supporting robust olive growth (EMA, 2022) (Figure 1).

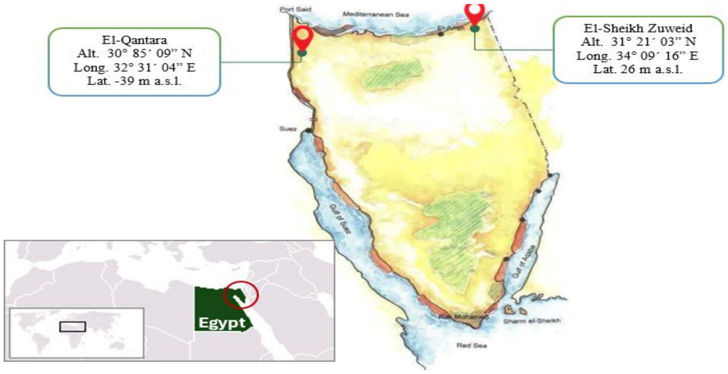

Figure 1. Location, altitude, longitude and latitude of the three study regions in Sinai (Northeast Egypt).

Plant material

The study utilized four olive cultivars: Picual, Sebhawy, Chemlali, and Manzanillo, selected based on commercial importance and adaptability to Mediterranean climates. Trees were managed under standard agricultural practices without supplemental irrigation, relying on seasonal rainfall. Fruits and fresh leaves were collected during the 2022 harvesting season at a consistent ripening stage from multiple trees per cultivar at both locations to ensure representative sampling.

Fruit characteristics

Fruit characteristics, including fruit length, fruit width, fruit volume, fruit weight, flesh weight, and stone weight, were measured using a digital caliper (Mitutoyo Corporation, Kawasaki, Japan) for dimensional parameters and an electronic balance (OHAUS Corporation, Parsippany, New Jersey, USA) for weights. Fruit volume (cm^3) was measured according to Rashidi et al. (2007) using the water displacement method. Twenty randomly selected fruits per cultivar per location were measured, and average values were recorded.

Oil yield and quality parameters

Oil content (OC) was determined using the Soxhlet extraction method with petroleum ether (Sigma-Aldrich, St. Louis, Missouri, USA) as the solvent, following standard protocols (AOAC, 2019). Total soluble solids (TSS) were measured as a direct index of oil accumulation during olive ripening with a digital refractometer (Atago, Tokyo, Japan) (Migliorini et al., 2011).

Oil quality parameters, including oil acidity (OA) and total polyphenols (TP), were measured as follows: OA was analyzed by titration using sodium hydroxide (NaOH) (Merck, Darmstadt, Germany), following the EEC/2568/92 method (European Community, 2013); TP was quantified using the Folin-Ciocalteu reagent (Sigma-Aldrich) following the protocol of Ainsworth and Gillespie (2007). Each parameter was measured in triplicate to ensure accuracy.

Fatty acid composition

Fatty acid composition, including palmitic (C16:0), stearic (C18:0), arachidic (C20:0), palmitoleic (C16:1), oleic (C18:1), linoleic (C18:2), and linolenic acids (C18:3), was determined using a gas chromatograph (GC-2010, Shimadzu Corporation, Kyoto, Japan) equipped with a flame ionization detector (FID). Fatty acid methyl esters (FAMEs) were prepared using methanol (Merck), following the method of IOOC (2001), and peaks were identified by comparing retention times with standards from Supelco (Bellefonte, Pennsylvania, USA). Results were expressed as percentages of the total fatty acids.

Molecular analysis: ISSR and SCoT techniques

Genetic diversity and cultivar differentiation were analyzed using inter simple sequence repeat (ISSR) and start codon targeted (SCoT) markers (Table 1). The DNA was extracted from fresh leaf samples using the cetyltrimethylammonium bromide (CTAB) method as described by Doyle and Doyle (1987), with chemicals obtained from Sigma-Aldrich. The DNA concentration and quality were verified with a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Table 1. List of the primer names and their nucleotide sequences used in the study for the inter simple sequence repeat (ISSR) and start codon targeted (SCoT).

N°	SCoT	Sequence (5→3)	ISSR		Sequence (5→3)
1	SCoT-1	ACGACATGGCGACCATCG	UBC811	ISSR-1	(GA)8C
2	SCoT-3	ACGACATGGCGACCACGC	UBC818	ISSR-2-I	(CA)8G
3	SCoT-4	CACCATGGCTACCACCAT	UBC849	ISSR-3-I	(GA)8T
4	SCoT-5	ACCATGGCTACCACCGTC	UBC827	ISSR-4-I	(TC)8RG
5	SCoT-6	CCATGGCTACCACCGCCT	UBC814	ISSR-5-I	(CT)8TG
6	SCoT-7	CCATGGCTACCACCGCAG	UBC844B	ISSR-10-I	(CT)8GC
7	SCoT-11	ACCATGGCTACCACCGCA	HB13	ISSR-25-I	(GAG)3GC
8	SCoT-22	ACAATGGCTACCACTGCA	HB15	ISSR-29	(GTG)3GC
9	SCoT-23	CAACAATGGCTACCACCG	TA2	ISSR-31-I	(CT)10G
10	SCoT-24	ACCATGGCTACCAGCGAG	TA3	ISSR-32-I	(AGG)6

The ISSR primers were selected based on their polymorphic potential and reproducibility. Each PCR reaction contained 50 ng DNA, 1X PCR buffer (Thermo Fisher Scientific), 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM primer, and 1 U Taq DNA polymerase (Thermo Fisher Scientific) in a total volume of 25 µL. The PCR profile included initial denaturation at 94 °C for 4 min, followed by 35 cycles of 30 s at 94 °C, 45 s at the primer-specific annealing temperature, and 2 min at 72 °C, with a final extension at 72 °C for 10 min (Pradeep Reddy et al., 2002). The SCoT primers targeting the start codon region were used following similar PCR conditions to ISSR, with primer-specific annealing temperatures. The reaction included 50 ng DNA, 0.4 µM SCoT primer, and 1 U Taq polymerase in a 25 µL volume (Collard and Mackill, 2009).

A 100 bp DNA ladder (Thermo Fisher Scientific) was used as a molecular weight marker to estimate fragment sizes. The PCR products from ISSR and SCoT markers were resolved by gel electrophoresis on a 1.5% agarose gel, stained with ethidium bromide (Sigma-Aldrich), and visualized under UV light using an imaging system (Gel Doc XR+, Bio-Rad, Hercules, California, USA).

Band scoring and data analysis

Band scoring for ISSR and SCoT markers was performed using TL120 software (TotalLab, Newcastle upon Tyne, UK). Bands were scored as present (1) or absent (0) to create a binary matrix for each cultivar across both locations. Only clear, reproducible bands were considered for analysis to ensure data accuracy. This binary matrix was subsequently used to calculate genetic diversity indices and genetic distances among cultivars (Rohlf, 2000).

ANOVA and molecular variance (AMOVA)

ANOVA was conducted to evaluate the effects of cultivar, location, and their interaction on all measured traits, including fruit characteristics, oil content, quality parameters, and fatty acid composition. The SPSS software (version 25) (IBM, Armonk, New York, USA) was used for ANOVA, with significance levels set at $P \leq 0.05$. Post hoc comparisons were performed using the least significant difference (LSD) test to identify significant differences between means. The AMOVA was conducted using FAMD software (Schlueter and Harris, 2006) among both locations based on ISSR and SCoT markers combined.

Data transformation and principal component analysis (PCA)

To standardize the data across varying scales, all morphological and biochemical traits were transformed using Z-scores prior to analysis. This standardization centered and scaled each trait to have a mean of 0 and a standard deviation of 1, allowing for consistent comparison across traits. Subsequently, a principal component analysis (PCA) was performed using R software (version 4.1.0) (R Foundation for Statistical Computing, Vienna, Austria). The PCA focused on fruit characteristics, oil quality, and fatty acid composition, and the biplot was generated to visualize the distribution of cultivars across the two locations based on the first two principal components, which captured the largest variance among the parameters (Jolliffe, 2002).

K-Means clustering for loci-trait associations

K-Means clustering was applied to identify associations between specific genetic loci and phenotypic traits. The binary data for ISSR and SCoT markers were standardized using Z-scores, and a heatmap was generated to highlight genetic loci potentially associated with key traits. This analysis, performed in orange software, facilitated the detection of clusters of loci associated with desirable traits, providing insights for marker-assisted selection (MacQueen, 1967; Demšar et al., 2013).

Multiple correlation heatmap analysis

To visualize relationships among all measured characteristics, a multiple correlation heatmap was constructed. Spearman correlation coefficients were calculated for each pair of variables, and a heatmap was generated using R software (version 4.1.0). This analysis allowed for the identification of highly correlated traits, providing insights into potential dependencies among morphological, chemical, and genetic parameters (Harrell, 2015).

RESULTS

Morphological assessment

The analysis of fruit characteristics, oil yield, quality parameters, and fatty acid composition was conducted for four olive cultivars—Picual, Sebhawy, Chemlali, and Manzanillo—across two distinct locations: El-Sheikh Zuweid and El-Qantara. Each parameter set reveals significant differences due to the effects of cultivar, location, and their interaction, as determined by the LSD test at $P \leq 0.05$.

Fruit characteristics

The results revealed significant variations in fruit morphological traits among the studied olive cultivars and between the two locations, El-Sheikh Zuweid and El-Qantara. El-Sheikh Zuweid demonstrated significantly higher averages across all traits compared to El-Qantara, with fruit length, width, volume, weight, flesh weight, and stone weight showing greater values ($P < 0.001$). Specifically, fruit weight was 4.385 ± 2.21 g in El-Sheikh Zuweid compared to 3.63 ± 1.93 g in El-Qantara, while fruit length was 23.00 ± 4.77 mm in El-Sheikh Zuweid *vs*. 20.44 ± 5.50 mm in El-Qantara.

Among cultivars, 'Picual' consistently exhibited the highest values for fruit morphology, with a fruit weight of 6.637 ± 0.72 g, fruit length of 26.47 ± 1.07 mm, and fruit width of 20.72 ± 1.05 mm. In contrast, 'Sebhawy' recorded the lowest values for all traits, with a fruit weight of 1.297 ± 0.27 g and fruit length of 13.72 ± 2.11 mm. 'Chemlali' and 'Manzanillo' showed intermediate performances, with 'Chemlali' excelling in stone weight (0.813 ± 0.11 g) and 'Manzanillo' demonstrating larger fruit volumes (49.83 ± 4.26 cm^3).

The ANOVA results confirmed significant effects of both location and cultivar on all fruit morphology traits ($P < 0.001$). However, the interaction between location and cultivar was nonsignificant, indicating that the observed variations were predominantly due to independent effects of these factors (Table 2).

Oil yield and quality parameters

Oil content, total soluble solids, oil acidity, and total polyphenols varied significantly between locations and among cultivars ($P < 0.001$). El-Sheikh Zuweid exhibited higher oil content (21.58 ± 1.52%) and total soluble solids (22.35 ± 1.75%) compared to El-Qantara (19.47 ± 1.74% and 19.83 ± 2.01%, respectively). Conversely, oil acidity was higher in El-Qantara (1.095 ± 0.28) than in El-Sheikh Zuweid (0.802 ± 0.16). Total polyphenol content was also greater in El-Sheikh Zuweid (0.066 ± 0.008 g 100 g^{-1}) compared to El-Qantara (0.055 ± 0.006 g 100 g^{-1}).

Among cultivars, 'Sebhawy' stood out with the highest oil content (22.52 ± 1.29%), total soluble solids (23.05 ± 0.73%), and total polyphenols (0.068 ± 0.010 g 100 g^{-1}), while also maintaining low oil acidity (0.820 ± 0.12). In contrast, 'Manzanillo' recorded the lowest oil content (18.52 ± 1.57%), total soluble solids (18.70 ± 1.05%), and total polyphenols (0.056 ± 0.006 g 100 g^{-1}). 'Picual' and 'Chemlali' exhibited intermediate values, with 'Picual' having higher oil acidity (1.150 ± 0.32) compared to other cultivars.

The interaction between location and cultivar was significant only for total soluble solids ($P < 0.001$), while other traits were influenced primarily by the independent effects of location and cultivar (Table 3).

Table 2. Fruit characteristics of assessed four olive cultivars in two locations. Data is presented as means ± standard deviation. The averages represent the main effect of analyzed factors (location and cultivars). Means labeled with different letters within the same factor were significantly different based on the LSD test ($P \leq 0.05$).

Studied factors		Fruit length	Fruit width	Fruit volume	Fruit weight	Flesh weight	Stone weight
		mm	mm	cm^{3}	g	g	g
Location							
El-Sheikh Zuweid		23.0 ± 4.7^{a}	17.5 ± 3.7^{a}	43.5 ± 7.2^{a}	4.3 ± 2.2^{a}	3.6 ± 2.0^{a}	0.7 ± 0.2^{a}
El-Qantara		20.4 ± 5.5^{b}	15.0 ± 3.9^{b}	38.6 ± 6.8^{b}	3.6 ± 1.9^{b}	3.0 ± 1.7^{b}	0.5 ± 0.2^{b}
Cultivar							
Picual		26.4 ± 1.0^{a}	20.7 ± 1.0^{a}	59.4 ± 4.3^{a}	6.6 ± 0.7^{a}	5.8 ± 0.6^{a}	0.7 ± 0.1^{a}
Sebhawy		13.7 ± 2.1^{c}	11.0 ± 1.4^{d}	24.2 ± 2.9^{d}	1.2 ± 0.2^{d}	1.0 ± 0.2^{d}	0.2 ± 0.07^{c}
Chemlali		22.9 ± 2.5^{b}	15.8 ± 1.5^{c}	30.8 ± 3.2^{c}	3.2 ± 0.4^{c}	2.4 ± 0.3^{c}	0.8 ± 0.1^{a}
Manzanillo		23.8 ± 1.3^{b}	17.6 ± 2.5^{b}	49.8 ± 4.2^{b}	4.8 ± 0.4^{b}	4.1 ± 0.3^{b}	0.6 ± 0.1^{b}
ANOVA	df	*P*-value					
Location (L)	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Cultivar (C)	3	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
L×C	3	0.233	0.346	0.898	0.127	0.176	0.803

Table 3. Oil yield and quality parameters of assessed four olive cultivars in two Egyptian locations. Data is presented as means ± standard deviation. The averages represent the main effect of analyzed factors (location and cultivars). Means labeled with different letters within the same factor were significantly different based on the LSD test ($P \leq 0.05$).

Studied factor		Oil content	Total soluble solids	Oil acidity	Total polyphenols
		%	°Brix		g 100 g^{-1}
Location					
El-Sheikh Zuweid		21.5 ± 1.5^{a}	22.3 ± 1.7^{a}	0.8 ± 0.1^{b}	0.06 ± 0.008^{a}
El-Qantara		19.4 ± 1.7^{b}	19.8 ± 2.01^{b}	1.09 ± 0.2^{a}	0.05 ± 0.006^{b}
Cultivar					
Picual		19.9 ± 1.2^{c}	22.1 ± 1.6^{b}	1.1 ± 0.3^{a}	0.05 ± 0.006^{c}
Sebhawy		22.5 ± 1.2^{a}	23.0 ± 0.7^{a}	0.8 ± 0.1^{b}	0.06 ± 0.01^{a}
Chemlali		21.1 ± 1.0^{b}	20.4 ± 2.3^{c}	0.7 ± 0.2^{b}	0.06 ± 0.006^{b}
Manzanillo		18.5 ± 1.5^{d}	18.7 ± 1.0^{c}	1.0 ± 0.2^{a}	0.05 ± 0.006^{c}
ANOVA	df	*P*-value			
Location (L)	1	< 0.001	< 0.001	< 0.001	< 0.001
Cultivar (C)	3	< 0.001	< 0.001	< 0.001	< 0.001
L×C	3	0.452	< 0.001	0.456	0.258

Fatty acid composition

Fatty acid composition analysis showed significant differences between locations and cultivars for several traits. El-Sheikh Zuweid samples demonstrated lower total saturated fatty acids (TSFA, 19.67 ± 1.75%) compared to El-Qantara (22.18 ± 2.73%), with corresponding lower levels of palmitic (C16:0), stearic (C18:0), and arachidic (C20:0) acids. In contrast, monounsaturated fatty acids (MUFA) were significantly higher in El-Sheikh Zuweid (70.49 ± 1.56%) compared to El-Qantara (64.02 ± 2.07%), with oleic acid (C18:1) showing a similar trend. Polyunsaturated fatty acids (PUFA), including linoleic (C18:2) and linolenic (C18:3) acids, were higher in El-Qantara (13.74 ± 2.79% and 1.526 ± 0.37%, respectively) than in El-Sheikh Zuweid (9.806 ± 2.37% and 0.889 ± 0.17%, respectively).

Among cultivars, 'Chemlali' exhibited the highest MUFA (69.57 ± 1.34%) and oleic acid (67.58 ± 2.87%), followed by 'Picual' (67.48 ± 2.57% and 65.18 ± 2.41%, respectively). 'Sebhawy' demonstrated the highest PUFA

(14.64 ± 1.80%) and linoleic acid (13.49 ± 1.50%), while ‘Manzanillo’ recorded intermediate levels across most traits. Saturated fatty acids, including palmitic, stearic, and arachidic acids, showed nonsignificant cultivar effects, although significant differences were observed for palmitoleic acid (C16:1), with higher levels in El-Sheikh Zuweid and ‘Chemlali’.

The ANOVA results confirmed significant effects of location and cultivars on all fatty acid traits ($P < 0.001$). Cultivar effects were highly significant for oleic acid, linoleic acid MUFA, and PUFA ($P < 0.05$), while the interaction between location and cultivar was highly significant for oleic acid (C18:1, $P < 0.001$) and MUFA ($P < 0.001$), while other traits showed moderate to low significance *p*-values, such as linoleic acid (C18:2, $P = 0.043$), and palmitic acid (C16:0, $P = 0.002$) (Table 4).

Table 4. Fatty acid composition of assessed four olive cultivars in two Egyptian locations. The data presented as means ± standard deviation. The averages represent the main effect of analyzed factors (location and cultivars). Means labeled with different letters within the same factor were significantly different based on the LSD test ($P \leq 0.05$).

Studied factor		Palmitic (C16:0)	Stearic (C18:0)	Arachidic (C20:0)	Total saturated fatty acids (TSFA)	Palmitoleic (C16:1)
Location						
El-Sheikh Zuweid		16.4 ± 1.1^{b}	2.6 ± 0.2^{b}	0.5 ± 0.8^{b}	19.6 ± 1.7^{b}	2.4 ± 0.4^{a}
El-Qantara		18.1 ± 1.1^{a}	3.1 ± 0.4^{a}	0.9 ± 1.6^{a}	22.1 ± 2.7^{a}	1.7 ± 0.4^{b}
Cultivar						
Picual		17.3 ± 0.4^{ab}	3.0 ± 0.5^{a}	0.6 ± 0.2^{b}	21.1 ± 0.4^{a}	2.3 ± 0.3^{ab}
Sebhawy		16.5 ± 1.6^{b}	2.8 ± 0.5^{ab}	0.7 ± 0.3^{ab}	20.0 ± 2.3^{b}	1.6 ± 0.5^{c}
Chemlali		17.4 ± 2.0^{a}	2.8 ± 0.3^{ab}	0.7 ± 0.3^{ab}	20.9 ± 2.3^{ab}	1.9 ± 0.5^{bc}
Manzanillo		17.8 ± 0.9^{a}	2.7 ± 0.2^{b}	0.9 ± 1.7^{a}	21.4 ± 3.3^{a}	2.4 ± 0.6^{a}
ANOVA	df	*P*-value				
Location (L)	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Cultivar (C)	3	0.042	0.031	0.037	0.037	0.006
L×C	3	0.002	0.039	0.002	0.003	0.047

Studied factor		Oleic (C18:1)	Monounsaturated fatty acids (MUFA)	Linoleic (C18:2)	Linolenic (C18:3)	Polyunsaturated fatty acids (PUFA)
Location						
El-Sheikh Zuweid		68.0 ± 1.4^{a}	70.4 ± 1.5^{a}	8.9 ± 2.3^{b}	0.8 ± 0.1^{b}	9.8 ± 2.3^{b}
El-Qantara		62.2 ± 2.0^{b}	64.0 ± 2.0^{b}	12.2 ± 2.8^{a}	1.5 ± 0.3^{a}	13.7 ± 2.7^{a}
Cultivar						
Picual		65.1 ± 2.4^{b}	67.4 ± 2.5^{b}	9.9 ± 1.8^{b}	1.4 ± 0.4^{a}	11.3 ± 2.2^{b}
Sebhawy		63.5 ± 3.3^{d}	65.2 ± 2.8^{d}	13.4 ± 1.5^{a}	1.1 ± 0.4^{ab}	14.6 ± 1.8^{a}
Chemlali		67.5 ± 2.8^{a}	69.5 ± 1.3^{a}	8.3 ± 0.9^{c}	1.0 ± 0.4^{b}	9.3 ± 1.1^{c}
Manzanillo		64.3 ± 4.0^{c}	66.7 ± 2.5^{c}	10.5 ± 1.4^{b}	1.2 ± 0.4^{ab}	11.7 ± 1.6^{b}
ANOVA	df	*P*-value				
Location (L)	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Cultivar (C)	3	< 0.001	< 0.001	< 0.001	0.011	< 0.001
L×C	3	< 0.001	< 0.001	0.043	0.764	0.091

Variability and principal component analysis of morphological traits

Panel A presents box plots for each morphological trait, revealing substantial variability across samples from the two locations, El-Sheikh Zuweid and El-Qantara. For example, fruit length has a mean Z-score of 0.35 with a standard deviation of 1.12, with El-Sheikh Zuweid samples showing higher median values. The range of fruit length values spans from-1.4 to 1.8 Z-score units, with El-Sheikh Zuweid cultivars clustered around the upper range. This difference indicates that environmental conditions in El-Sheikh Zuweid may favor larger fruit development (Figure 2A).

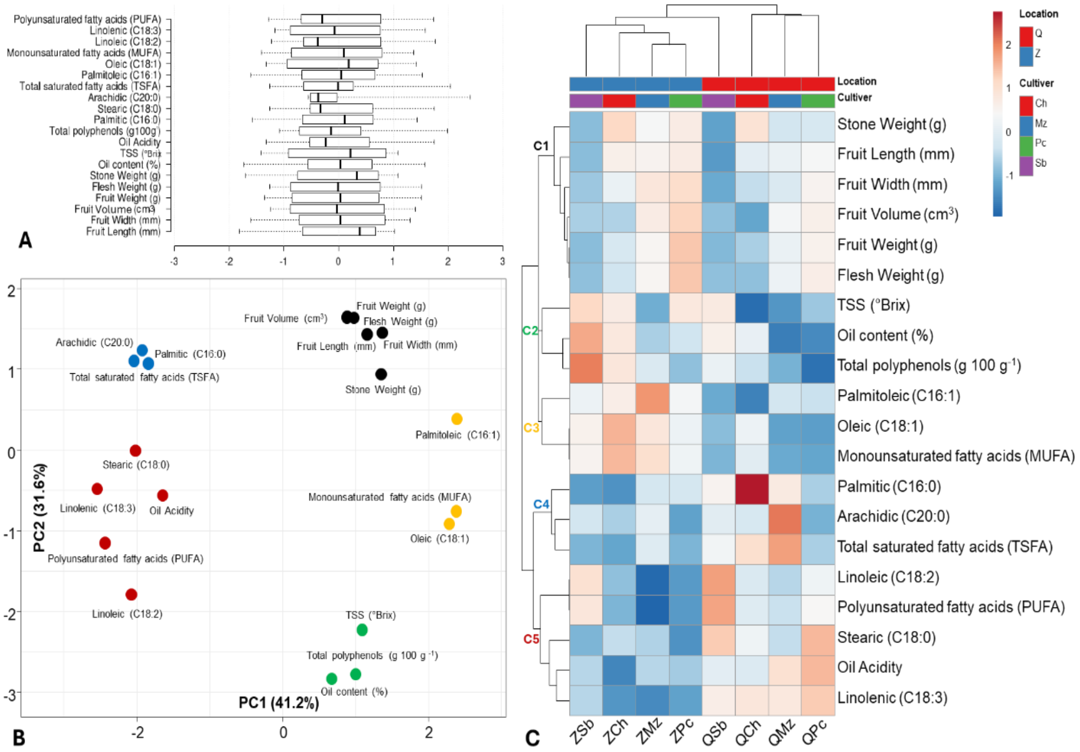

Figure 2. Analysis of morphological and biochemical traits in four olive cultivars across two Egypt locations. (A) Box plots showing variability in key traits across all samples, including fruit size, oil content, and polyphenols. (B) Principal component analysis (PCA) biplot highlighting five trait clusters (C1 to C5) based on principal components, with each cluster representing related traits such as fruit morphology (C1), oil quality (C2 and C3), and fatty acid composition (C4 and C5). (C) Correlation heatmap with hierarchical clustering of traits and samples, illustrating trait associations and grouping patterns influenced by cultivar and location. TSS: Total soluble solids; Sb: 'Sebhawy'; Mz: 'Manzanillo'; Ch: 'Chemlali'; Pc: 'Picual'; Z: El-Sheikh Zuweid; Q: El-Qantara. For example, ZSb indicates 'Sebhawy' from El-Sheikh Zuweid, while QPc refers to 'Picual' from El-Qantara.

Fruit width also displays notable variability, with a mean Z-score of 0.3 and a standard deviation of 1.05. The majority of samples from El-Sheikh Zuweid show fruit width Z-scores above 0.5, while El-Qantara samples are centered closer to the mean. In terms of oil content, the mean Z-score across all samples is 0.4 (*SD* = 1.24), with values ranging from-1.3 to 2.1. El-Qantara samples, particularly from the 'Picual' and 'Chemlali' cultivars, show consistently higher oil content, with median Z-scores around 1.5. This higher oil content suggests that environmental conditions in El-Qantara may support oil accumulation, possibly due to soil or climate factors that enhance oil yield (Figure 2A). For polyphenol content, the variability is particularly pronounced in El-Qantara samples, with Z-scores ranging from-1.1 to 1.8. The elevated polyphenol levels, especially in 'Picual', align with high oil content values, suggesting a correlation between polyphenol and oil production in these samples. This pattern supports the findings from the correlation analysis, where polyphenols and oil content were clustered together as part of C3 (Figure 2A).

The PCA biplot in Panel B provides insights into the clustering of traits, with the first two principal components (PC1 and PC2) explaining a combined variance of 36.6% (PC1 = 19.8%, PC2 = 16.8%). The traits are grouped into five distinct clusters. Cluster C1 (black, Figures 2B, 2C) includes stone weight, fruit length, fruit width, fruit volume, fruit weight, and flesh weight—all of which are related to the physical size and mass of the fruit. These traits showed high loadings on the primary components in the PCA, reflecting a strong correlation among them. This cluster underscores the genetic and environmental impact on fruit morphology, with samples from El-Sheikh Zuweid generally displaying higher Z-scores for these traits. The larger fruit size and weight in this location may be due to local environmental factors that favor fruit growth, making these traits essential for assessing yield potential.

Cluster C2 (green, Figures 2B, 2C) encompasses total soluble solids (TSS), oil content, and total polyphenols. These traits are indicators of fruit and oil quality, as TSS influences flavor and sugar content, while oil content and polyphenols contribute to the nutritional and antioxidant properties of the oil. This cluster is particularly prominent in cultivars from El-Qantara, where high oil content and polyphenol levels were observed, especially in 'Picual'. The strong association within this cluster suggests that both genetic predisposition and environmental factors in El-Qantara enhance these quality attributes, supporting the production of high-value olive oil.

Cluster C3 (yellow, Figures 2B, 2C) includes palmitoleic acid (C16:1), oleic acid (C18:1), and monounsaturated fatty acids (MUFA). These fatty acids are valuable for the nutritional profile and stability of olive oil, with oleic acid being particularly beneficial for health and oil longevity. The close grouping of these traits suggests a genetic basis for fatty acid composition, with certain cultivars, such as 'Chemlali' from El-Qantara, showing elevated oleic acid levels. This cluster highlights the importance of monounsaturated fats in olive oil quality and potential cultivar selection for high-oleic content oils.

Cluster C4 (blue, Figures 2B, 2C) consists of palmitic acid (C16:0), arachidic acid (C20:0), and total saturated fatty acids (TSFA). This cluster represents the saturated fatty acids in olive oil, which contribute to oil texture and stability. While saturated fats are less prominent in olive oil compared to monounsaturated and polyunsaturated fats, their presence influences overall oil quality and shelf life. These traits are relatively consistent across locations, indicating that saturated fat content may be more genetically controlled with limited environmental impact.

Cluster C5 (red, Figures 2B, 2C) includes linoleic acid (C18:2), polyunsaturated fatty acids (PUFA), stearic acid (C18:0), oil acidity, and linolenic acid (C18:3). These traits are associated with the PUFA content and stability aspects of olive oil. Higher linoleic and linolenic acid levels contribute to the health benefits of olive oil but can also affect its stability and shelf life. Oil acidity, a key quality indicator, is included in this cluster and shows significant variation across samples, with El-Sheikh Zuweid cultivars exhibiting higher acidity levels. This clustering suggests that while polyunsaturated fats enhance nutritional value, their levels and the oil's acidity are influenced by both genetic factors and environmental conditions, particularly in El-Sheikh Zuweid.

In the hierarchical clustering (Figure 2C), samples group by location, with El-Sheikh Zuweid samples clustering together based on higher values for C1 and C2 traits, while El-Qantara samples are grouped by elevated C3 and C4 traits. El-Sheikh Zuweid samples exhibit higher Z-scores for fruit length, width, and flesh weight, reinforcing the idea that this location promotes the co-expression of fruit size traits. Conversely, El-Qantara samples, particularly 'Picual' and 'Chemlali', show clustering driven by high oil content and polyphenols, confirming the influence of location-based conditions on oil quality traits.

The clustering patterns observed in the PCA and heatmap analysis suggest that environmental factors and genetic predisposition influence the expression of key traits differently across locations. The traits in Cluster C1 highlight El-Sheikh Zuweid's potential for cultivating olives with larger fruits, while Clusters C2 and C3 underscore El-Qantara's suitability for high-quality oil production, particularly in terms of oil content, polyphenols, and monounsaturated fatty acids. These insights provide a basis for targeted selection in breeding programs, allowing for location-based optimization of both yield and quality traits in olive production.

DNA fingerprinting and molecular assessment

The molecular diversity and genetic structure of four olive cultivars grown in two distinct locations—El-Sheikh Zuweid (Z group) and El-Qantara (Q group)—were assessed using ISSR and SCoT markers. Each cultivar-location combination is identified by a sample code where "Z" denotes El-Sheikh Zuweid and "Q" represents El-Qantara. The following letters denote the specific cultivar: "Sb" for ‘Sebhawy’, "Mz" for ‘Manzanillo’, "Ch" for ‘Chemlali’, and "Pc" for ‘Picual’. For example, "ZSb" indicates ‘Sebhawy’ from El-Sheikh Zuweid, while "QPc" refers to ‘Picual’ from El-Qantara.

DNA polymorphism and diversity

Using inter simple sequence repeat (ISSR) and start codon targeted (SCoT) markers, we assessed DNA polymorphism and genetic diversity across cultivars. The combined analysis of ISSR and SCoT markers provided a comprehensive profile of genetic variability in the studied cultivars.

A total of 10 ISSR and 10 SCoT primers were initially tested across four olive cultivars from two locations. Of these, seven ISSR primers successfully amplified, with the highest polymorphism information content (PIC) observed for primer ISSR-1 (PIC = 0.350) and the lowest for primer ISSR-2-I (PIC = 0.306), resulting in an average PIC of 0.334 across the effective ISSR primers. In contrast, all 10 SCoT primers amplified successfully, with the highest PIC observed for primer SCoT-1 (PIC = 0.349) and the lowest for primer SCoT-4 (PIC = 0.261), yielding an average PIC of 0.307. These values underscore the polymorphic efficiency of each primer set in assessing genetic diversity.

For ISSR markers, a total of 114 loci were analyzed, with no missing data across the samples. The mean molecular weight for ISSR bands was 712 bp (Table 5). Across all individuals, 38 bands were detected, with an average band frequency of 0.3 per individual. The cv. ZMz (‘Manzanillo’ from El-Sheikh Zuweid) exhibited the highest diversity within ISSR, with 40 bands, whereas QPc (‘Picual’ from El-Qantara) showed the lowest, with 25 bands. This variation indicates that ZMz has a more diverse ISSR profile compared to QPc. The Shannon diversity index for ISSR was 6.07, reflecting moderate genetic diversity within the cultivars across both locations.

The SCoT markers produced a total of 178 loci with a mean molecular weight of 818 bp (Table 5). Across all individuals, 59 bands were observed, with an average band frequency of 0.4 per individual, slightly higher than that observed with ISSR markers. The cv. ZSb (‘Sebhawy’ from El-Sheikh Zuweid) exhibited the highest number of bands at 73, indicating higher genetic diversity, whereas QPc had the fewest bands, with 48. Shannon’s diversity index for SCoT was 6.57, indicating greater genetic variation captured by SCoT markers compared to ISSR.

Table 5. Inter simple sequence repeat (ISSR) and start codon targeted (SCoT) marker outputs showing band diversity across cultivars and locations. Each lane represents a unique cultivar-location combination. Bands indicate the presence of specific loci used to assess genetic diversity. Z: El-Sheikh Zuweid; Q: El-Qantara Sb: ‘Sebhawy’; Mz: ‘Manzanillo’; Ch: ‘Chemlali’; Pc: ‘Picual’.; MW: molecular weight; *SD*: standard deviation.

Per Individual		MW (bp)	ZSb	QSb	ZMz	QMz	ZCh	QCh	ZPc	QPc	Mean	*SD*
ISSR	Total	-	29.0	22.0	40.0	45.0	45.0	37.0	35.0	25.0	38.0	7.0
	Average	712	0.3	0.2	0.4	0.4	0.4	0.3	0.3	0.2	-	-
SCoT	Total	-	73.0	69.0	68.0	73.0	60.0	50.0	66.0	48.0	59.0	11.0
	Average	818	0.4	0.4	0.4	0.4	0.3	0.3	0.4	0.3	-	-
Combined	Total	-	102.0	91.0	108.0	118.0	105.0	87.0	101.0	73.0	97.0	17.0
	Average	777	0.3	0.3	0.4	0.4	0.4	0.3	0.3	0.3	-	-

When ISSR and SCoT markers were combined, 292 loci were analyzed, yielding a mean molecular weight of 777 bp. Across all samples, 97 bands were observed, with an average band frequency of 0.3 per individual. The cv. ZCh (‘Chemlali’ from El-Sheikh Zuweid) had the highest diversity, with 118 bands, while QPc had the lowest diversity with 73 bands. Shannon’s diversity index for the combined markers was the highest at 7.33, indicating robust genetic diversity across all cultivars and locations. These results suggest that combining ISSR and SCoT markers provides a more comprehensive view of genetic variability, capturing both moderate and high levels of diversity within and between locations.

Genetic structure and differentiation

Principal component analysis (PCA). The principal component analysis (PCA) plot (left panel of Figure 3) illustrates the genetic relationships among olive cultivars from two distinct locations: El-Sheikh Zuweid (Z group, black squares) and El-Qantara (Q group, red circles). Each point represents a sample based on combined ISSR and SCoT marker data, plotted along the first two principal components (PC1 and PC2), which explain 19.8% and 16.8% of the variance, respectively. In the PCA plot, samples from El-Sheikh Zuweid, such as ZSb, ZCh, and ZMz, are generally closer to each other, suggesting high genetic similarity within this location. This clustering pattern implies a degree of genetic cohesion among these samples, potentially due to shared environmental influences or selective breeding practices. On the other hand, samples from El-Qantara, such as QPc, QCh, and QSb, show more dispersed positioning. This dispersion indicates greater genetic differentiation within the Q group, possibly due to distinct environmental conditions or selection-based genetic drift specific to the El-Qantara region. The cultivars from each location (e.g., ZMz and QMz) indicate that while genetic profiles may differ between locations, cultivars within each location retain some genetic characteristics that distinguish them.

The Euclidean distance-based heatmap, generated from binary data (presence/absence of ISSR and SCoT markers), illustrates the genetic distances among samples, with hierarchical clustering based on Euclidean distance. In the heatmap, cultivars from the same location tend to cluster together, supporting the PCA findings. For instance, samples ZSb and ZCh from El-Sheikh Zuweid are closely positioned, forming a tight cluster indicative of high genetic similarity. Similarly, samples within the El-Qantara group, such as QPc and QSb, also show clustering, though with a broader range of genetic variation compared to El-Sheikh Zuweid. This Euclidean distance-based clustering reaffirms the distinct genetic characteristics associated with each location, as seen in the PCA plot, and highlights subtle intra-group genetic variations.

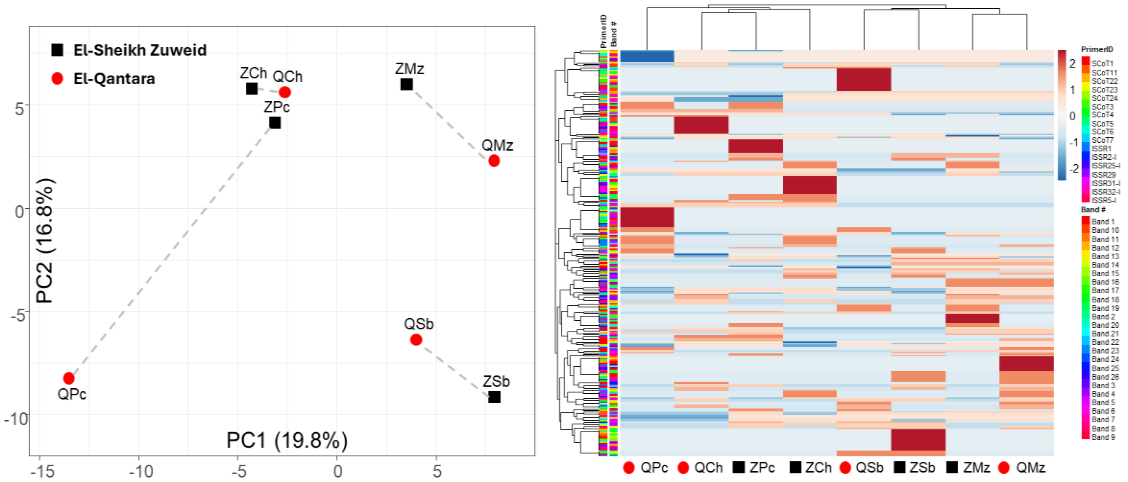

Figure 3. Principal component analysis (PCA) biplot of combined ISSR and SCoT markers. Left panel- PCA plot based on combined inter simple sequence repeat (ISSR) and start codon targeted (SCoT) marker data, with black squares representing cultivars from El-Sheikh Zuweid (Z) and red circles representing cultivars from El-Qantara (Q). The axes display PC1 (19.8% variance) and PC2 (16.8% variance). Right panel- Euclidean distance heatmap based on binary marker data, showing clustering patterns among samples. Red colors indicate high similarity, while blue colors represent greater genetic divergence. Sb: 'Sebhawy'; Mz: 'Manzanillo'; Ch: 'Chemlali'; Pc: 'Picual'. For example, ZSb indicates 'Sebhawy' from El-Sheikh Zuweid, while QPc refers to 'Picual' from El-Qantara.

Hybridization analysis using ML index

Hybridization levels between the Z and Q parental groups were analyzed using a maximum likelihood (ML) index. This index evaluates the genetic similarity of each individual to the two parental groups (Z and Q), with values ranging from 0 to 1. An ML index close to 1 suggests high similarity to the Z parental group, while a value close to 0 indicates high similarity to the Q parental group. Intermediate values represent varying degrees of hybridization, with fractions reflecting the genetic contribution from each parental group.

The ML index analysis (Figure 4) revealed distinct patterns of hybridization across samples. Individuals from El-Sheikh Zuweid, such as ZSb, ZMz, and ZCh, had ML index values closer to 1, indicating a strong genetic similarity to the Z group. These individuals showed minimal evidence of hybridization, suggesting that they have retained a genetic profile closely aligned with the Z parental group.

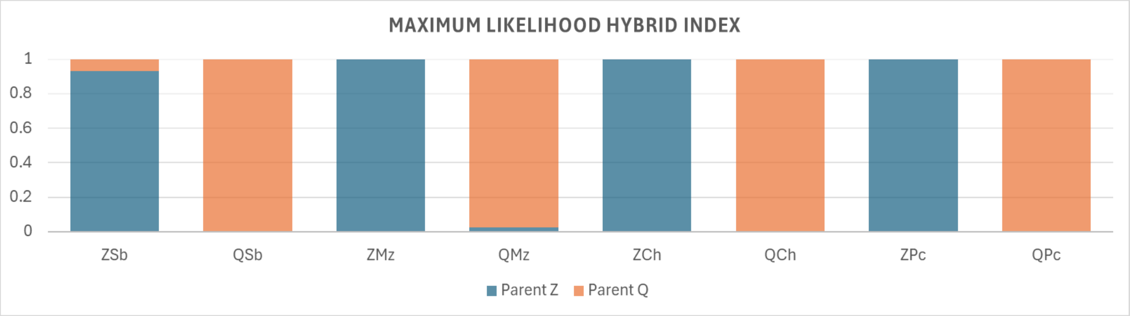

Figure 4. Hybridization analysis using maximum likelihood (ML) index values for each sample, with values ranging from 0 to 1. An ML index close to 1 indicates high genetic similarity to the Z parental group (El-Sheikh Zuweid), while values closer to 0 indicate similarity to the Q parental group (El-Qantara). Intermediate values suggest partial hybridization between the two groups. Q values is represented as 1 – Z values. Sb: 'Sebhawy'; Mz: 'Manzanillo'; Ch: 'Chemlali'; Pc: 'Picual'.

In contrast, individuals from El-Qantara, particularly QSb and QPc, exhibited ML index values closer to 0. This indicates a high degree of genetic similarity to the Q parental group and low hybridization with the Z group. However, some individuals, such as QCh and ZPc, displayed intermediate ML index values, reflecting partial hybridization and genetic contributions from both parental groups. This partial hybridization indicates some level of gene flow between the two groups, potentially due to environmental adaptation or selective breeding practices that have introduced genetic material from one location to the other.

Overall, the ML index results demonstrate that while most individuals align strongly with their respective parental groups (Z or Q), a subset exhibits genetic characteristics intermediate between the two groups, suggesting the presence of hybrids. These hybrids show a mix of genetic traits from both Z and Q groups, providing insights into the extent of gene flow and genetic exchange across locations.

Analysis of molecular variance (AMOVA)

The genetic differentiation among and within populations was analyzed through analysis of molecular variance (AMOVA) based on the combined ISSR and SCoT data. The AMOVA results (Table 3) showed that most genetic variation was within populations, accounting for 93.26% of the total variance, while only 6.74% of the genetic variation was attributed to differences among populations. These results suggest that most of the genetic diversity is contained within individual cultivars across both locations, with minimal differentiation between the two geographical locations (El-Sheikh Zuweid and El-Qantara). This finding aligns with the hypothesis that genetic variation is largely cultivar-specific rather than location-dependent (Table 6).

Table 6. AMOVA results based on combined Inter simple sequence repeat (ISSR) and start codon targeted (SCoT) markers for the four olive cultivars across two locations. The analysis shows that most of the genetic variation is within populations.

Source of variation	Sum of squares (SS)	Degrees of freedom (df)	Variance components	Percentage of variation
Among locations	0.304	1	0.024	6.74
Within location	1.417	6	0.236	93.26
Total	1.721	7		100.00

Clustering and association of genetic loci with morphological traits

Fruit characteristics. The k-means clustering analysis identified associations between specific genetic loci and key fruit characteristics, indicating possible genetic markers for traits related to fruit size, weight, and volume. Within this category, one subcluster showed loci linked to fruit length and width, where samples with high Z-scores for these traits consistently displayed the presence of Band B234 (SCoT-1, Band 19). This locus may serve as a potential marker for larger fruit size, which is a valuable trait for olive production and breeding. Another subcluster related to fruit characteristics included loci associated with fruit volume, fruit weight, and flesh weight. Samples that scored high for these traits exhibited loci such as Band B157 (SCoT-1, Band 8) and Band B201 (SCoT-22, Band 11). These loci could be used as genetic markers for selecting cultivars with desirable fruit mass and quality traits, as increased flesh weight and volume contribute to better fruit quality and yield (Figure 5).

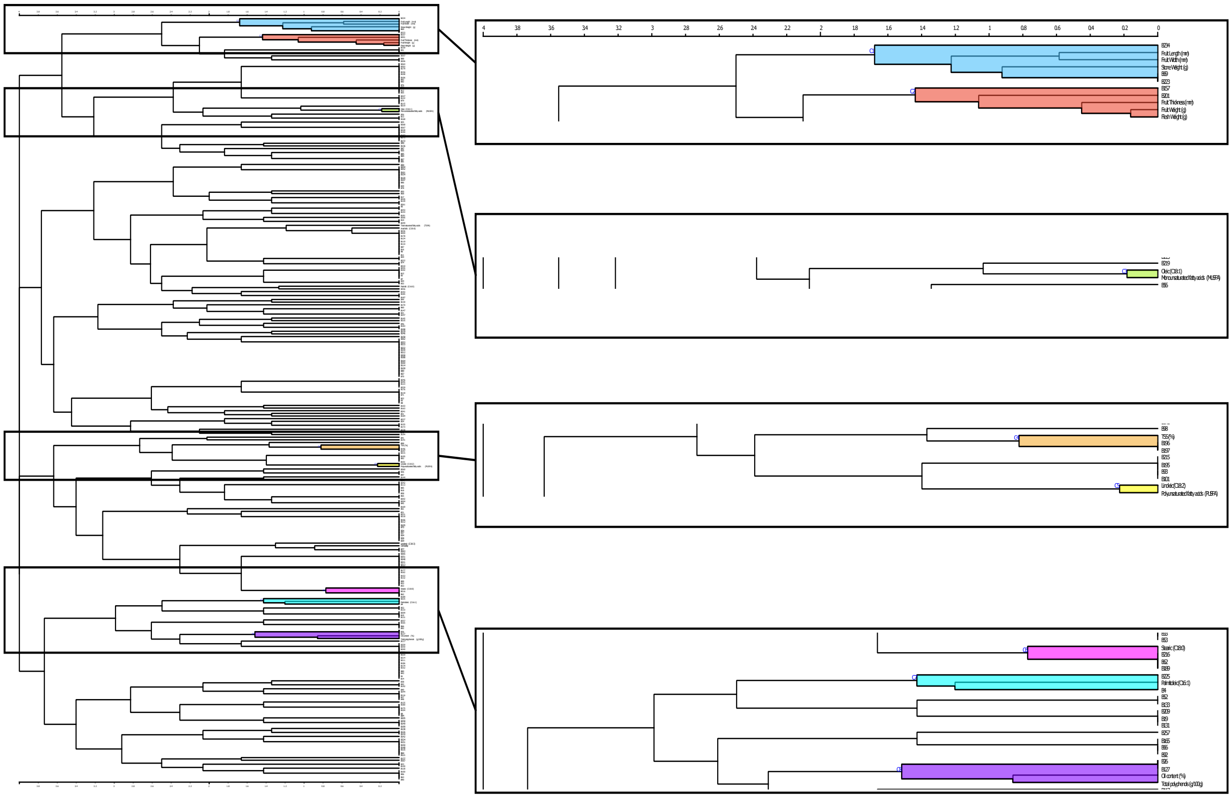

Figure 5. Hierarchical clustering analysis integrating morphological data (transformed z-scores) and binary molecular data from start codon targeted (SCoT) and inter simple sequence repeat (ISSR) markers. The k-means clustering identified associations between genetic loci and key traits, including fruit morphology, oil quality, and fatty acid composition. Highlighted clusters represent loci linked to specific morphological traits, oil yield, and fatty acid characteristics.

Oil yield and quality parameters

The analysis also identified loci associated with oil yield and quality parameters, particularly for traits like oil content and total soluble solids (TSS). In one subcluster, loci correlated with high oil content and total polyphenols, essential for oil quality and stability, were observed. Samples with elevated oil content and polyphenol levels displayed the presence of Band B127 (SCoT-3, Band 4). This locus may be significant for selecting cultivars that produce high-quality oil with rich antioxidant properties. Another subcluster revealed loci associated with total soluble solids (TSS), a parameter related to fruit ripeness and sugar accumulation. Samples with high TSS scores consistently exhibited Band B196 (SCoT-6, Band 4) and Band B197 (SCoT-23, Band 5), suggesting that these loci could be linked to traits influencing fruit maturity and sugar content, valuable for breeding improved ripening and sweetness profiles (Figure 5).

Fatty acid composition

The clustering analysis further identified loci associated with the fatty acid composition of olive oil, a critical quality trait for nutritional and commercial value. One subcluster included loci linked to oleic acid (C18:1) and monounsaturated fatty acids (MUFA), both highly desirable in olive oil due to their health benefits. Samples with higher Z-scores for oleic acid content consistently displayed Band B223 (SCoT-5, Band 16), suggesting that this locus may serve as a marker for cultivars with superior oil quality. Another subcluster showed loci associated with linoleic acid (C18:2) and PUFA, with samples high in these fatty acids displaying Band B69 (ISSR-25-I, Band 11). Given that linoleic acid affects the oxidative stability and shelf life of olive oil, this locus may be helpful in selecting cultivars with balanced fatty acid profiles. Additional subclusters were associated with other fatty acids, such as stearic acid (C18:0) and palmitoleic acid (C16:1). Loci linked to stearic acid, such as Band B216 (SCoT-3, Band 17), Band B62 (ISSR-5-I, Band 11), and Band B189 (SCoT-22, Band 10), were consistently present in samples with higher Z-scores for this saturated fatty acid. Although less desirable in high amounts, stearic acid is an integral part of oil composition, and these markers may help manage fatty acid profiles. For palmitoleic acid, a subcluster associated loci such as Band B225 (SCoT-24, Band 9), Band B4 (ISSR-1, Band 3), and Band B26 (ISSR-25-I, Band 5) with higher concentrations of this MUFA, indicating their potential role as markers for cultivars with specific fatty acid characteristics. The detailed identification of these loci associated with each morphological and oil quality trait provides a valuable resource for further validation (Figure 5).

DISCUSSION

The current study highlights the critical influence of both location and cultivar on olive fruit characteristics, oil yield, quality parameters, and fatty acid composition, aligning with findings from previous studies that underscore the significance of Genotype × Environment interactions in olive production (Navas-López et al., 2020; Yılmaz-Düzyaman et al., 2023). By integrating morphological and molecular analyses, we obtained a comprehensive view of how genetic and environmental factors interact to shape key agronomic traits in olive cultivars.

Environmental conditions such as soil type, temperature, and water availability were found to significantly affect fruit morphology and oil quality between the two study locations. Larger fruit sizes and higher oil yields in El-Sheikh Zuweid, compared to El-Qantara, suggest that local environmental conditions in El-Sheikh Zuweid may be more conducive to promoting fruit growth and oil accumulation (Benlloch-González et al., 2019; Volakaki et al., 2023). These results support earlier work by Bortoluzzi et al. (2023), who demonstrated that specific environmental factors could enhance yield and stability in olive oil, providing a basis for selecting optimal production sites based on desired fruit and oil characteristics.

The significant differences among cultivars in terms of oil yield, oil acidity (OA), and polyphenol content confirm the importance of genetic factors in determining olive oil characteristics. 'Sebhawy' emerged as a high oil-yielding cultivar with favorable oil quality attributes, such as high polyphenol content and low OA levels, aligning with findings from other studies that identify 'Sebhawy' as a premium cultivar for oil production additionally, 'Picual' ranks highest for fruit weight, along with acceptable oil yield and high-quality traits. In contrast, 'Chemlali' displayed higher oleic acid content, particularly in El-Sheikh Zuweid, which could be advantageous for nutritional purposes due to the health benefits associated with PUFA (Vossen, 2007; Jimenez-Lopez et al., 2020; Ali et al., 2022).

The fatty acid composition, particularly the balance of oleic and linoleic acids, was notably influenced by both cultivar and location, confirming findings by Rey-Giménez and Sánchez (2024) and Iqbal et al. (2023) on the genetic and environmental effects on olive oil's fatty acid profile. Higher oleic acid levels in cultivars grown in El-Sheikh Zuweid, especially in 'Picual' and 'Chemlali', suggest a more stable and potentially healthier oil profile due to the antioxidant properties of monounsaturated fats. This is significant as higher oleic acid content is linked to greater oxidative stability and health benefits, factors that are highly desirable in olive oil (Montaño et al., 2016).

The interaction between location and cultivar observed in this study underscores the need to consider both environmental and genetic factors in olive oil production. For example, the increase in polyphenol content in 'Chemlali' grown in El-Sheikh Zuweid suggests that this cultivar may optimize its antioxidant potential under specific environmental conditions, a finding supported by Yılmaz-Düzyaman et al. (2023), who reported enhanced polyphenol levels under environmental factors, and location in certain cultivars. Similarly, 'Picual' exhibited higher oil yield and oleic acid content in El-Sheikh Zuweid, indicating that it responds favorably to this location's conditions, consistent with the findings of Navas-López et al. (2020), who noted that 'Picual' often performs better in warmer climates with higher water availability. Ben Hmida et al. (2022) noted that oil extracted from the olive cultivars grown in the hottest grove contained less oleic acid than oil from those grown in a milder environment.

The observed decrease in linoleic acid in El-Sheikh Zuweid for 'Sebhawy' aligns with research suggesting that cooler and more stable temperatures, such as those found in El-Qantara, promote higher PUFA content in olive oil (Ben Hmida et al., 2022). This finding implies that 'Sebhawy', with its high linoleic acid content in El-Qantara, may be more suited to environments that enhance its PUFA profile, benefiting consumers interested in oils with a higher proportion of polyunsaturated fats.

The results of this study have practical implications for olive oil producers, particularly in regions with diverse environmental conditions. Selecting cultivars like 'Picual' and 'Sebhway' for locations similar to El-Sheikh Zuweid may maximize oil yield and quality due to their adaptability to these environmental conditions, while cultivars such as 'Sebhawy' may be better suited to locations like El-Sheikh Zuweid for producing oils with specific fatty acid profiles. This study reinforces the importance of Genotype × Environment interaction in olive production, as previously highlighted by Rey-Giménez and Sánchez (2024). Optimizing cultivar selection for specific environments not only enhances yield but also tailors the nutritional and sensory attributes of olive oil, which can increase the market value and consumer appeal of the product.

The use of ISSR and SCoT markers enabled us to assess the genetic diversity and structure within and among the olive cultivars. The combined Shannon diversity index of 7.33 indicates a moderate to high level of polymorphism, which is critical for maintaining genetic diversity in breeding populations (Alhaithloul et al., 2024). The genetic differentiation observed among cultivars and locations highlights the value of molecular markers in revealing genetic variability that may not be captured through morphological traits alone. These findings align with prior research emphasizing the role of molecular markers in elucidating genetic relationships and aiding in cultivar identification (Galatali et al., 2021).

One of the key outcomes of this study is the identification of associations between specific ISSR and SCoT marker bands and desirable traits. For instance, the presence of Band B234 (SCoT-1, Band 19) was consistently associated with larger fruit sizes, while Band B127 (SCoT-3, Band 4) correlated with high oil content and polyphenol levels. This integration of molecular and morphological data provides valuable insights into marker-trait associations that can facilitate the selection of cultivars with specific attributes, supporting findings by Alhaithloul et al. (2024) on the value of molecular markers in enhancing selection precision.

The clustering analysis further identified markers linked to fatty acid composition, such as Band B223 (SCoT-5, Band 16) for high oleic acid levels. These molecular markers are essential for breeding programs focused on improving oil quality traits, especially in regions where environmental conditions may vary. The association between specific bands and traits demonstrates the potential of molecular markers to support marker-assisted selection (MAS) by enabling the early identification of desirable genotypes, even before phenotypic expression (De Mori and Cipriani, 2023).

The integration of morphological and molecular data offers a powerful approach to olive breeding, as it combines phenotypic assessment with genetic insights to support more effective cultivar selection. The observed clustering patterns in both the PCA and k-means analysis underscore the importance of combining genetic and morphological data for a more comprehensive understanding of trait relationships and environmental adaptability (Zhu et al., 2019; Vogt, 2022; Sales et al., 2024). For example, traits in Cluster C1 (e.g., fruit size and weight) were highly responsive to El-Sheikh Zuweid conditions, while oil quality traits in Cluster C2 were more pronounced in El-Qantara, suggesting that these clusters can guide location-based breeding strategies.

The ability to associate specific molecular markers with distinct morphological traits also enables breeders to make more targeted selections, aligning with the objectives of precision agriculture and sustainable crop production. This integrative approach is increasingly relevant as climate change and soil degradation continue to challenge traditional agriculture, emphasizing the need for cultivars that are both high-yielding and resilient (Benlloch-González et al., 2019). Additionally, the observed associations between genetic loci and fatty acid composition can be used to develop cultivars with optimized nutritional profiles, meeting growing consumer demand for health-beneficial oils (Jimenez-Lopez et al., 2020).

CONCLUSIONS

This study underscores the importance of cultivar and location in determining olive fruit characteristics, oil yield, and quality, with molecular markers providing valuable insights into genetic diversity and trait associations. Our findings support the idea that cultivar selection should be tailored to specific environmental conditions to optimize desirable oil characteristics, whether for yield, stability, or nutritional profile. Future research could further explore the mechanisms by which specific environmental factors influence oil composition, offering a more comprehensive understanding that can guide olive cultivation strategies in diverse regions. By integrating morphological and molecular data, we identified specific markers linked to desirable traits, which can enhance breeding programs aimed at improving both yield and quality in olives. The findings highlight the potential of molecular markers for optimizing olive oil production through targeted cultivar selection, particularly in diverse and challenging environments, contributing to sustainable agriculture and the production of high-value olive oil.

Author contributions

Conceptualization: D.A., M.B., D.A.E., M.M. Methodology: D.A., M.A., H.T., M.M. Software: D.A.E., M.M. Validation: M.A., H.T. Formal analysis: M.M. Investigation: D.A.E., D.A., M.B., M.M. Resources: D.A., M.B., M.A. Data curation: D.A.E. Writing-original draft: D.A.E., M.M. Writing-review & editing: D.A., M.A., D.A.E., M.M. Visualization: M.M. Supervision: D.A.E., M.M. Project administration: D.A.E. Funding acquisition: D.A.E. All co-authors reviewed the final version and approved the manuscript before submission.

Acknowledgements

Mahmoud Magdy acknowledges the financial support provided by the Saavedra Fajardo postdoctoral fellowship (Ref. 20245/SF/24)- Fundación Séneca, the Science and Technology Agency of the Region of Murcia, Spain.

References

Ainsworth, E.A., Gillespie, K.M. 2007. Estimation of total phenolic contents and other oxidation substrates in plant tissue using Folin-Ciocalteu reagent. Nature Protocols 2:875-877. doi:10.1038/nprot.2007.102.

Alhaithloul, H.A.S., Awad, N.S., Qari, S.H., El-Homosy, R.F., El-Sayed, M.Q., Alqahtani, M.M., et al. 2024. Genetic diversity, chemical constituents, and anatomical analysis of eight popular olive (*Olea europaea* L.) cultivars in Al-Jouf region, Saudi Arabia. Scientific Reports 14:14688. doi:10.1038/s41598-024-65542-y.

Alharbi, R.M., Gomah, N.H., Ragab, W.S.M., Awad, N.S., Abdein, M.A. 2024. Molecular characterization of elite olive cultivars grown in the North region of Saudi Arabia using SCoT, ISSR, and SRAP techniques. Biocatalysis and Agricultural Biotechnology 60:103283. doi:10.1016/j.bcab.2024.103283.

Ali, H.A.M., Gowda, A.M., Farrag, H.M., Radwan, E.M.A. 2022. Performance of some olive (*Olea europaea* L.) cultivars grown under saline stress conditions in newly reclaimed soils. Assiut Journal of Agricultural Sciences 53(4):39-54. doi:10.21608/ajas.2022.144378.1152.

AOAC. 2019. Official methods of analysis. 21st ed. Association of Official Agricultural Chemists (AOAC), Washington D.C., USA.

Ben Hmida, R., Gargouri, B., Chtourou, F., Abichou, M., Sevim, D., Bouaziz, M. 2022. Study on the effect of climate changes on the composition and quality parameters of virgin olive oil "Zalmati" harvested at three consecutive crop seasons: Chemometric discrimination. ACS Omega 7(44):40078-40090. doi:10.1021/acsomega.2c04813.

Benlloch-González, M., Sánchez-Lucas, R., Bejaoui, M.A., Benlloch, M., Fernández-Escobar, R. 2019. Global warming effects on yield and fruit maturation of olive trees growing under field conditions. Scientia Horticulturae 249:162-167. doi:10.1016/j.scienta.2019.01.046.

Bortoluzzi, L., Casal, S., Cruz, R., Peres, A.M., Baptista, P., Rodrigues, N. 2023. Influence of interannual climate conditions on the composition of olive oil from centenarian olive trees. Agronomy 13(12):2884. doi:10.3390/agronomy13122884.

Collard, B.C.Y., Mackill, D.J. 2009. Start codon targeted (SCoT) polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27(1):86-93. doi:10.1007/s11105-008-0060-5.

De Mori, G., Cipriani, G. 2023. Marker-assisted selection in breeding for fruit trait improvement: A review. International Journal of Molecular Sciences 24(10):8984. doi:10.3390/ijms24108984.

Demšar, J., Curk, T., Erjavec, A., Hočevar, T., Milutinović, M., Možina, M., et al. 2013. Orange: Data mining toolbox in Python. Journal of Machine Learning Research 14:2349-2353. doi:10.5555/2567709.2567736.

Doyle, J.J., Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15.

EMA. 2022. Climatic and soil data report for North Sinai and El-Qantara (Ismailia Governorate) [2000-2020 Data report generated upon request]. Egyptian Meteorological Authority, Cairo, Egypt.

European Community. 2013. Commission Regulation (EU) N° 1348/2013 of 16 December amending Regulation (EEC) N° 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. Official Journal of the European Union L 338:31-67.

FAO. 2021. Climate data for agricultural planning. Food and Agriculture Organization (FAO), Rome, Italy. Available at https://www.fao.org/climate-change/resources/data-tools/en (accessed March 2025).

Galatali, S., Ghafoor, N., Kaya, E. 2021. Characterization of olive (*Olea europaea* L.) genetic resources via PCR-based molecular marker systems. European Journal of Biology and Biotechnology 2(1):26-33. doi:10.24018/ejbio.2021.2.1.146.

Harrell, F.E. 2015. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. Springer-Verlag, New York, USA. doi:10.1007/978-3-319-19425-7.

IOOC. 2001. Preparation of the fatty acid methyl esters from olive oil and olive-pomace oil. COI/T.20/Doc. N°24. International Olive Oil Council (IOOC), Madrid, Spain.

Iqbal, J., Ashraf, M.A., Cai, Q., Riaz, A., Li, Y., Shahid, I., et al. 2023. Simple sequence repeats-based DNA fingerprinting and varietal identification of mango cultivars. Preprints. doi:10.20944/preprints202309.1723.v1.

Jimenez-Lopez, C., Carpena, M., Lourenço-Lopes, C., Gallardo-Gomez, M., Lorenzo, J.M., Barba, F.J., et al. 2020. Bioactive compounds and quality of extra virgin olive oil. Foods 9(8):1014. doi:10.3390/foods9081014.

Jolliffe, I.T. 2002. Principal component analysis. 2nd ed. Springer Series in Statistics. Springer-Verlag, New York, USA.

MacQueen, J.B. 1967. Some methods for classification and analysis of multivariate observations. p. 281-297. In Le Cam, L.M., and Neyman, J. (eds.) Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. University of California Press, Berkeley, California, USA.

Migliorini, M., Cherubini, C., Mugelli, M., Gianni, G., Trapani, S., Zanoni, B. 2011. Relationship between the oil and sugar content in olive oil fruits from Moraiolo and Leccino cultivars during ripening. Scientia Horticulturae 129(4):919-921. doi:10.1016/j.scienta.2011.05.023.

Montaño, A., Hernández, M., Garrido, I., Llerena, J.L., Espinosa, F. 2016. Fatty acid and phenolic compound concentrations in eight different monovarietal virgin olive oils from Extremadura and the relationship with oxidative stability. International Journal of Molecular Sciences 17(11):1960. doi:10.3390/ijms17111960.

Navas-López, J.F., Cano, J., de la Rosa, R., Velasco, L., León, L. 2020. Genotype by environment interaction for oil quality components in olive tree. European Journal of Agronomy 119:126115. doi:10.1016/j.eja.2020.126115.

Pradeep Reddy, M., Sarla, N., Siddiq, E. 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128:9-17. doi:10.1023/A:1020691618797.

Rashidi, M., Seyfi, K., Gholami-Parashkouhi, M. 2007. Determination of kiwifruit volume using image processing. ARPN Journal of Agricultural and Biological Science 2(6):17-22.

Rey-Giménez, R., Sánchez, A.G. 2024. Effect of cultivar and environment on chemical composition and geographical traceability of Spanish olive oils. Journal of the American Oil Chemists' Society 101:371-382. doi:10.1002/aocs.12774.

Rohlf, F.J. 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket, New York, USA.

Sales, H., Alves, M.L., Sampaio, A.M., Nunes, J., Vaz Patto, M.C. 2024. Understanding the phenotypic variability of the olive 'Galega vulgar' fruits and oil-related traits across environments. Scientia Horticulturae 326:112738. doi:10.1016/j.scienta.2023.112738.

Schlueter, P.M., Harris, S.A. 2006. Analysis of multilocus fingerprinting data sets containing missing data. Molecular Ecology Notes 6(2):569-572. doi:10.1111/j.1471-8286.2006.01225.x.

Souabni, H., Amar, F., Gogorcena, Y., Fendri, M., Mnasri, S., Triki, M., et al. 2023. Morphological and molecular characterization of minor olive (*Olea europaea* L.) cultivars in Tunisia. Journal of Plant Biochemistry and Biotechnology 32:27-37. doi:10.1007/s13562-022-00726-4.

Vogt, G. 2022. Environmental adaptation of genetically uniform organisms with the help of epigenetic mechanisms—An insightful perspective on ecoepigenetics. Epigenomes 7(1):1. doi:10.3390/epigenomes7010001.

Volakaki, A.-M., Ioanna, M., Georgios, K. 2023. Genotype, environment, and interactive effects on olive tree photosynthesis and fruit ripening. Acta Horticulturae 1384:299-306. doi:10.17660/ActaHortic.2023.1384.39.

Vossen, P. 2007. Olive oil: History, production, and characteristics of the world's classic oils. HortScience 42(5):1093-1100.

Yılmaz-Düzyaman, H., Medina-Alonso, M.G., Sanz, C., Pérez, A.G., de la Rosa, R., León, L. 2023. Influence of genotype and environment on fruit phenolic composition of olive. Horticulturae 9(10):1087. doi:10.3390/horticulturae9101087.

Zhu, S., Niu, E., Shi, A., Mou, B. 2019. Genetic diversity analysis of olive germplasm (*Olea europaea* L.) with genotyping-by-sequencing technology. Frontiers in Genetics 10:755. doi:10.3389/fgene.2019.00755.