DOI: http://doi.org/10.1590/1678-992X-2020-0308

ISSN 1678-992X

SCIENTIA AGRICOLA

Research Article

Molecular diversity and ecogeographic distribution of Algerian wild olives (*Olea europaea* subsp. *europaea* var. *sylvestris*)

Wahiba Falek[1], Sara Sion[2], Cinzia Montemurro[2], Isabella Mascio[2]*, Susanna Gadaleta[2], Valentina Fanelli[2], Michele Antonio Savoia[2], Luciana Piarulli[3], Sakina Bechkri[4], Douadi Khelifi[1], Monica Marilena Miazzi[2]

[1]Ecole Nationale Supérieure de Biotechnologie, Ville universitaire Ali Mendjeli, BP E66 – 251000 – Constantine – Algeria.

[2]Università degli Studi di Bari Aldo Moro – Dipartimento di Scienze del Suolo, della Pianta e degli Alimenti (DiSSPA), Via Amendola, 165/A – 70126 – Bari – Italia.

[3]Spin Off dell'Università degli Studi di Bari Aldo Moro/ SINAGRI S.r.l. – Bari – Italia.

[4]University Frères Mentouri Constantine – Lab. of Genetic Biochemistry and Plant Biotechnology, P.O. Box, 325 Ain El Bey Way – 25017 – Constantine – Algeria.

*Corresponding author <mascioisa@gmail.com>

Edited by: Leonardo Oliveira Medici

Received September 25, 2020

Accepted January 08, 2021

ABSTRACT: Olive is one of the most important crops in the Mediterranean Basin, because of the olive oil economic value and its role in characterization of the rural landscape. The strong influence of climatic changes on the modern agriculture and the availability of a large source of genetic variability pose as crucial future challenges. Therefore, safeguarding olive genetic resources becomes fundamental, not only in cultivated forms in *ex situ* collections, but also in terms of wild trees in their natural habitat. In this study, 174 samples of oleaster collected in different parts of Algeria were analyzed by 16 nuclear Simple Sequence Repeats (SSRs). The analysis showed a huge genetic variability in the oleaster, and the STRUCTURE and Principal Coordinate Analyses (PCoA) highlighted clusterization of genotypes according to their geographic origin and bioclimatic conditions. Genotypes adapted to harsh climatic conditions were identified, which could be useful to enrich the panel of olive genotypes for breeding purposes and preserve genetic diversity of this species from erosion risks.

Keywords: SSR, Algeria, Oleaster, biodiversity, bioclimate

Introduction

Olive (*Olea europaea* L.) is one of the most ancient and socio-economically important trees of the Mediterranean Basin. The subspecies *europaea* has two botanical varieties: var. *europaea,* which corresponds to cultivated olive, and var. *sylvestris,* the wild form, also known as Oleaster (Green, 2002). Olive domestication dates back to approximately 6000 years ago in the Middle East of the Mediterranean Basin (Zohary and Hopf, 1994; Vossen, 2007; Besnard et al., 2018), where genetic studies support a major domestication event followed by the spread and secondary diversification of the crop in westernmost regions of the Mediterranean Basin (Díez et al., 2015; Besnard et al., 2001; D'Agostino et al., 2018). Even if selection of cultivars were associated to limited erosion of genetic diversity due to admixture between different genepools, wild forms can still be considered an important reserve of genes for adaptability and favorable agronomic traits (Hannachi et al., 2009; Besnard et al., 2013; Miazzi et al., 2020). Due to the great resistance to wind, drought, and salinity, wild forms play an important role in the preservation of Mediterranean ecosystems (Belaj et al., 2007). In addition, although they have low weight and low oil content fruit, some favorable features could be introduced into the cultivated varieties (Díaz-Rueda et al., 2020; León et al., 2018). For instance, wild olive can represent an interesting source of genes for resistance to diseases, such as the Olive Quick Decline Syndrome (OQDS) (Saponari et al., 2019), for which resistant or molecular targets and tolerant genotypes have been identified (Novelli et al., 2019). Nowadays, most olive wild-looking trees are actually feral forms from hybridization events between oleaster and cultivars, while the populations of genuine wild olives remain limited to isolated areas, such as remote areas of North Africa (Lumaret et al., 2004). In Algeria, olive is an important crop cultivated mainly in the northern part of the country, where it is represented by old cultivars poorly characterized (Boucheffa et al., 2017). Studies on Algerian olive genetic diversity focused on the evaluation of the variability distribution in cultivated and wild olives (Abdessemed et al., 2015; Boucheffa et al., 2017) or on the establishment of relationships between *O. europaea* subspecies (Rubio de Casas et al., 2006). Collected in undisturbed areas of Algeria, the genetic variability of wild olive germplasm has been analyzed using SSR markers to preserve it from the erosion risk and identify enhanced traits for the improvements of cultivated olive.

Materials and Methods

Plants material

We collected 174 wild olive types from their undisturbed natural habitat in 33 provinces of northern Algeria. The sampling sites were selected to represent different microclimatic and soil conditions where the olive tree can grow, from the coastal area to the inner regions of northeastern Algeria. In each sampling site, individual plants were collected and geo-referenced using a Global Positioning System Tracker (Figure 1; Table 1).

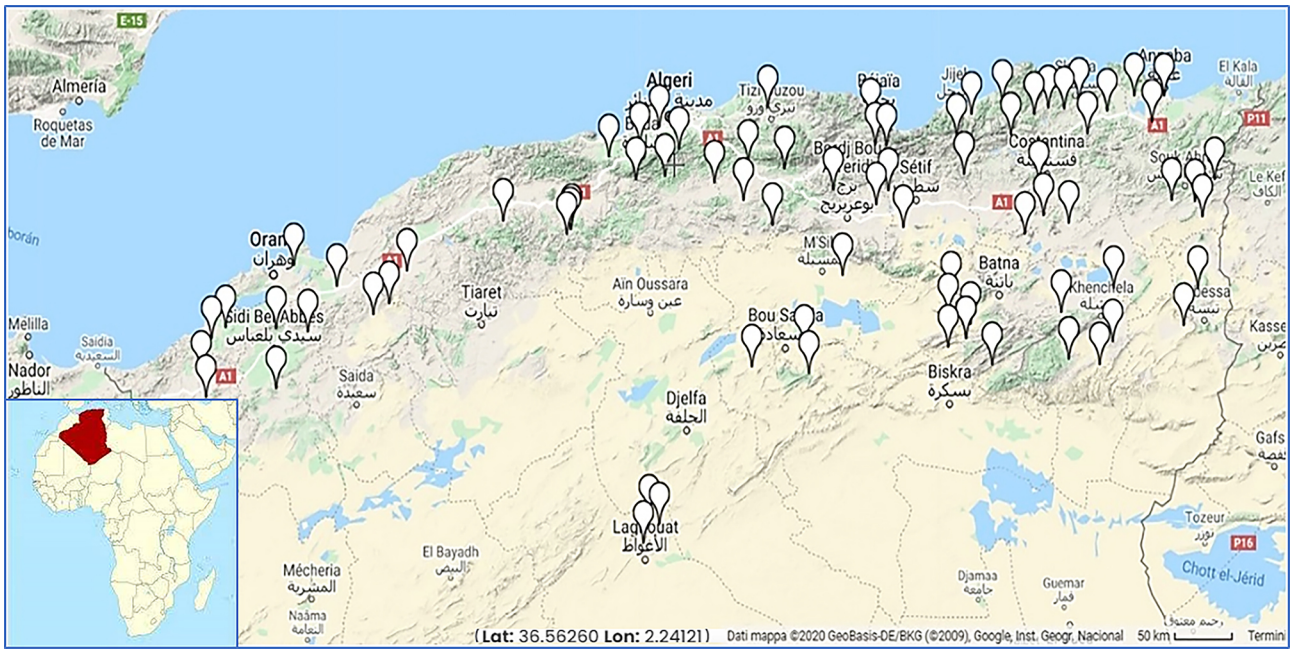


Figure 1 – Map of wild olive sampling sites in 33 provinces of northern Algeria in North Africa. The geographic coordinates of the sampling sites are reported in Table 1. Source:https://www.scribblemaps.com/create/#/lat=36.87962061&lng=40.78125&z=3&t=hybrid).

For each sample, the following climatic parameters of the collection sites were gathered: altitude, annual rainfall expressed in millimeter (P mm), average of the maximum temperature of the hottest month (M °C), average of the minimum temperature of the coldest month (m °C), Emberger coefficient (Q2) (Emberger, 1930) (National Office of Meteorology, Algeria http://www.aps.dz/en/; http://fr.climate-data.org/). Precipitation and temperature data were corrected according to different altitudes and the pluvio-thermic Emberger coefficient (Q2) was determined according to Bechkri and Khelifi (2017) (Table 1).

Molecular characterization

We collected 20 young leaves from different parts of the canopy of 174 olive trees and the leaves were immediately frozen after collection. For genomic DNA extraction, 150 mg of leaves were used following the protocol described by Spadoni et al. (2019). DNA quality and quantity were assessed spectrophotometrically and normalized to 50 ng μL^{-1}.

Genotypes were analyzed by 16 microsatellite markers (Carriero et al., 2002; Cipriani et al., 2002; De La Rosa et al., 2002; Sefc et al., 2000), considered informative and effective to discriminate olive cultivar (Baldoni et al., 2009). Amplifications were conducted according to Spadoni et al. (2019).

The amplification of products was detected using the automatic sequencer ABI PRISM 3100 Avant Genetic Analyzer and data were collected with Gene Mapper genotyping software v.5.0, using a size standard.

Genetic analysis

The genetic diversity of genotypes was estimated through the following indices: number of alleles (Na), effective number of alleles (Ne), Shannon information index (I), observed (Ho) and expected (He) heterozygosity, and fixation index (F) using the GenAlex software v.6.5. The polymorphic information content (PIC) was calculated by using Cervus v 3.0 to describe the informativeness of each marker. Moreover, Cervus v 3.0 software was used to evaluate the significance of estimates per locus tested by permutations (9,999 replicates). The frequency of null alleles F (null) and departure from Hardy-Weinberg equilibrium (HWE) were tested, applying sequential Bonferroni correction.

GenAlex v.6.5 was also used to perform the principal coordinate analysis (PCoA) using the Nei's unbiased genetic distance pairwise population matrix. The inter-individual relationship was calculated for the partition of olive samples into specific groupings. The Bayesian clustering algorithm implemented in the STRUCTURE software version 2.3.4 was used to infer the structure of the studied germplasm, assuming 10 genetic clusters (K) and performing 10 independent runs with 100,000 Markov Chain Monte Carlo (MCMC) iterations for each K following a burn-in of 10,000 iterations. The optimal value of K was determined based on the δK test (Evanno et al., 2005) using the STRUCTURE HARVESTER software. Accessions were assigned to defined populations if the value of the corresponding membership coefficient (qi) was higher than 0.6, otherwise, they were considered admixed

Table 1 – Estimation of ecogeographic parameters. For each sample coordinates, altitude, annual rainfall (Pmm) expressed in millimeter, average of maximum temperature of the hottest month (M °C), average of minimum temperature of the coldest month (m °C), Emberger coefficient (Q2) and the group clustering based on the Structure and dendrogram analysis are indicated.

N°	Code	Latitude	Longitude	Altitude	P mm	m °C	M °C	Q2	Structure analysis K = 3	Dendrogram
1	AinDefla–BOK	36°1'8.72" N	2°10'2.20" E	529	636.8	4.57	33	76.83	Gp1	Cluster 1
2	AinDefla–DTBZ_1	35°58'3.02" N	2°07'0.59" E	832	758	3.36	30.9	94.47	Gp1	Cluster 1
3	AinDefla–DTBZ_2	36°0'0.41" N	2°09'0.72" E	634	678.8	4.15	32.3	82.82	Gp1	Cluster 1
4	AinTémouchent–AET	35°17'2.83" N	1°13'2.00" W	245	493	6.76	30.1	72.36	Gp2.1	Cluster 3
5	AinTémouchent–CE	35°21'9.89" N	1°05'7.31" W	124	396.2	7.24	31	57.26	Gp1	Cluster 2
6	Alger–D	36°40'1.39" N	2°59'8.81" E	189	741.8	7.75	28.9	120.4	Gp1	Cluster 2
7	Annaba–AB	36°42'4.81" N	7°37'4.32" E	73	712.52	6.42	31	99.39	Gp1	Cluster 1
8	Annaba–DEB_1	36°52'0.89" N	7°44'8.51" E	40	686.12	6.55	31.2	95.32	Gp1	Cluster 1
9	Annaba–DEB_2	36°52'0.38" N	7°44'9.69" E	46	690.92	6.53	31.2	96.06	Gp1	Cluster 1
10	Annaba–DOEA	36°52'7.25" N	7°27'5.18" E	66	706.92	6.45	31.1	98.52	Gp1	Cluster 3
11	Batna–ATi	35°34'34.80" N	5°43'46.95" E	918	368.2	0.02	35.7	35.36	Gp2.2	Cluster 3
12	Batna–ATM_1	35°22'34.35" N	5°55'8.21" E	808	324.2	0.46	36.5	30.85	Gp2.2	Cluster 1
13	Batna–ATM_2	35°17'0.09" N	5°52'05.96" E	809	324.6	0.45	36.5	30.89	Gp2.2	Cluster 3
14	Batna–ATo	35°17'0.09" N	5°52'05.96" E	942	377.8	–0.08	35.6	36.36	Gp2.2	Cluster 3
15	Batna–DM_1	35°17'1.03" N	5°52'4.42" E	1000	401	–0.31	35.2	38.78	Gp1	Cluster 1
16	Batna–DM_2	35°17'1.03" N	5°52'4.42" E	815	327	0.43	36.5	31.14	Gp2.1	Cluster 3
17	Batna–DM_3	35°13'34.14" N	5°42'25.45" E	539	216.6	1.53	38.4	20.16	Gp2.1	Cluster 3
18	Batna–G	35° 6'51.25" N	6° 07'20.63" E	900	361	0.09	35.9	34.62	Gp2.2	Cluster 3
19	Batna–KF	35°26'00.25" N	5°42'02.92" E	977	391.8	–0.22	35.3	37.82	Gp2.2	Cluster 3
20	Batna–MEAS_1	35°15'6.57" N	5°53'6.60" E	890	357	0.13	35.9	34.21	Gp2.2	Cluster 1
21	Batna–MEAS_2	35°15'6.57" N	5°53'6.60" E	101	405.8	–0.36	35.1	39.29	Gp2.2	Cluster 1
22	Batna–OA	35°26'51.98" N	5°44'56.93" E	1270	509	–1.39	33.3	50.38	Gp2.2	Cluster 3
23	Batna–OSS	35°34'43.30" N	5°40'17.00" E	794	318.6	0.51	36.6	30.28	Gp2.2	Cluster 3
24	Batna–S	35°27'05.83" N	5°33'09.23" E	696	279.4	0.9	37.3	26.34	Gp2.2	Cluster 3
25	Batna–SAS	35°27'05.83" N	5°33'09.23" E	726	291.4	0.78	37.1	27.54	admixed	Cluster 4
26	Batna–SDEH	35°27'03.41" N	5°34'01.69" E	710	285	0.85	37.2	26.9	Gp2.2	Cluster 3
27	Batna–SDES	35°26'05.32" N	5°34'06.00" E	729	292.6	0.77	37.1	27.66	Gp2.2	Cluster 3
28	Batna–SK	35°26'09.14" N	5°34'04.70" E	701	281.4	0.88	37.3	26.54	Gp2.2	Cluster 3
29	Batna–SS	35°20'49.88" N	5°38'02.48" E	717	287.8	0.82	37.1	27.18	Gp2.1	Cluster 2
30	Batna–T	35°25'09.65" N	5°42'05.63" E	1232	493.8	–1.24	33.5	48.71	Gp2.2	Cluster 3
31	Batna–TAT	35°36'44.41" N	5°48'41.47" E	967	387.8	–0.18	35.4	37.4	Gp2.2	Cluster 1
32	Bejaia–B	36°48'6.75" N	4°59'5.26" E	67	826.2	7.14	30.3	122.7	admixed	Cluster 3
33	Bejaia–K	36°39'9.46" N	4°50'4.79" E	94	837	7.03	30.1	124.7	admixed	Cluster 3
34	Bejaia–OD	36°51'6.65" N	4°48'5.80" E	36	813.8	7.26	30.5	120.3	Gp1	Cluster 1
35	Bejaia–OG	36°42'4.63" N	4°56'9.61" E	67	826.2	7.14	30.3	122.7	Gp2.1	Cluster 2
36	Biskra–AZREC	35°09'1.15" N	5°50'02.89" E	981	492.4	3.3	35.1	53.11	Gp2.2	Cluster 3
37	Biskra–D_1	34°37'1.52" N	5° 05'55.24" E	190	176	6.47	40.6	17.66	Gp2.2	Cluster 3
38	Biskra–D_2	34°36'0.48" N	5° 04'03.96" E	165	166	6.57	40.8	16.62	Gp2.2	Cluster 3
39	Biskra–D_3	34°36'54.07" N	5° 05'55.55" E	191	176.4	6.46	40.6	17.71	Gp2.2	Cluster 3
40	Biskra–EG	34°42'21.84" N	5°15'49.49" E	160	164	6.59	40.9	16.42	Gp1	Cluster 1
41	Biskra–F	34°42'47.89" N	5°19'06.99" E	147	158.8	6.64	41	15.88	Gp2.2	Cluster 1
42	Biskra–L	34°37'33.83" N	5°23'07.92" E	1047	518.8	3.04	34.7	56.3	Gp2.2	Cluster 3
43	Biskra–T	34°41'44.90" N	5°22'35.00" E	142	156.8	6.66	41	15.67	Gp2.2	Cluster 3
44	Blida–DC_1	36°27'3.16" N	2°52'1.19" E	763	995.4	4.96	28.1	147.4	Gp1	Cluster 1
45	Blida–DC_2	36°26'9.22" N	2°52'1.28" E	993	1087.4	4.04	26.5	165.9	Gp1	Cluster 1
46	Blida–DC_3	36°27'2.59" N	2°51'6.48" E	892	1047	4.44	27.2	157.7	Gp1	Cluster 1
47	BordjBouArreridj–OHEB	36°12'4.31" N	4°22'4.46" E	577	240.82	3.2	39.2	22.97	Gp1	Cluster 1
48	BordjBourArreridj–BL	36°16'3.39" N	4°47'4.15" E	731	302.42	2.59	38.1	29.23	Gp2.1	Cluster 3
49	BordjBourArreridj–DEZ	36°14'6.48" N	4°49'4.03" E	730	302.02	2.59	38.1	29.19	Gp2.2	Cluster 3
50	Bouira–AEA	36°25'1.39" N	3°54'2.60" E	570	680.6	3.28	31.8	81.8	Gp1	Cluster 1
51	Bouira–DI	36°26'8.49" N	3°56'8.15" E	928	823.8	1.85	29.3	102.9	Gp1	Cluster 1
52	Bouira–DT	36°25'6.63" N	4°05'1.24" E	1050	872.6	1.36	28.5	110.5	Gp1	Cluster 1
53	Bouira–OEZ	36°14'5.87" N	3°57'2.31"E	525	662.6	3.46	32.1	79.26	Gp1	Cluster 1
54	Boumerdes–C	36°43'7.74" N	3°26'4.94" E	34	736.2	6.83	32	100.5	Gp1	Cluster 1
55	Boumerdes–L	36°39'9.73" N	3°20'7.16" E	81	755	6.64	31.6	103.7	Gp1	Cluster 1
56	Boumerdes–T	36°44'6.67" N	3°30'5.70" E	50	742.6	6.76	31.8	101.6	Gp1	Cluster 1

Continue...

Table 1 – Continuation.

57	Chlef–S_1	35°59'0.20" N	1°30'2.04" E	536	730	4.82	30.5	97.66	Gp1	Cluster 3
58	Chlef–S_2	35°59'8.58" N	1°27'6.98" E	582	766.8	4.64	30.1	103.1	Gp1	Cluster 1
59	Constantine–AS	36°16'22.85'' N	6°31'18.10'' E	626	467.72	2.64	41.9	40.84	Gp2.1	Cluster 2
60	Constantine–AZ	36°13'7.21" N	6°31'4.03 "E	720	505.32	2.26	41.3	44.45	Gp1	Cluster 1
61	Constantine–BH_1	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	admixed	Cluster 3
62	Constantine–BH_2	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	Gp2.2	Cluster 3
63	Constantine–BH_3	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	admixed	Cluster 3
64	Constantine–BH_4	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	Gp2.2	Cluster 3
65	Constantine–BH_5	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	Gp2.2	Cluster 3
66	Constantine–BH_6	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	Gp2.2	Cluster 3
67	Constantine–BH_7	36°30'20.38" N	6°33'07.98" E	435	391.32	3.4	43.3	33.68	Gp2.2	Cluster 3
68	Constantine–DEO_1	36°24'4.13" N	6°41'1.84" E	969	604.92	1.26	39.5	54.25	Gp1	Cluster 1
69	Constantine–DEO_2	36°23'2.85" N	6°40'6.19" E	787	532.12	1.99	40.8	47.05	Gp2.1	Cluster 3
70	Constantine–OAM_1	36°18'17.48'' N	6°34'24.46" E	563	442.52	2.89	42.4	38.46	admixed	Cluster 1
71	Constantine–OAM_2	36°18'17.48'' N	6°34'24.46'' E	563	442.52	2.89	42.4	38.46	Gp2.1	Cluster 2
72	Constantine–OAM_3	36°18'17.48'' N	6°34'24.46" E	563	442.52	2.89	42.4	38.46	Gp2.1	Cluster 2
73	Guelma–A	36°30'15.59" N	7°30'30.10" E	500	588.8	3.51	34.9	679.1	Gp1	Cluster 4
74	Guelma–B_1	36°28'30.99" N	7°28'58.42" E	173	611.6	4.82	37.2	644.7	Gp2.1	Cluster 4
75	Guelma–B_2	36°32'41.20" N	7°28'33.43" E	350	589.2	4.11	35.9	663	Gp2.1	Cluster 3
76	Guelma–BM	36°28'37.85" N	7°36'18.10" E	261	719.6	4.46	36.6	653.7	admixed	Cluster 2
77	Guelma–GBS	36°32'32.16" N	7°30'13.23" E	230	624	4.59	36.8	650.5	Gp2.1	Cluster 3
78	Guelma–H	36°30'20.68" N	7°27'2.55" E	174	659.6	4.81	37.2	644.8	admixed	Cluster 1
79	Jijel–M	36°42'9.64" N	6°16'1.82" E	52	1090.6	6.52	31.1	152.3	Gp1	Cluster 2
80	Jijel–S	36°33'5.43" N	6°16'8.47" E	250	1249	5.73	29.7	178.7	admixed	Cluster 1
81	Jijel–SM	36°37'1.25" N	6°16'3.35" E	93	1123.4	6.36	30.8	157.6	admixed	Cluster 1
82	Jijel–TOK_1	36°39'53.68'' N	5°45'03.77'' E	573	1272.4	4.44	27.5	224.8	Gp2.1	Cluster 1
83	Jijel–TOK_2	36°39'53.68'' N	5°45'03.77'' E	573	1272.4	4.44	27.5	224.8	Gp2.1	Cluster 2
84	Jijel–TZ_1	36°39'53.68'' N	5°45'03.77'' E	573	1272.4	4.44	27.5	224.8	Gp2.1	Cluster 2
85	Jijel–TZ_2	36°39'53.68'' N	5°45'03.77'' E	573	1272.4	4.44	27.5	224.8	Gp2.1	Cluster 2
86	Khenchela–AS_1	35°26'1.28" N	7°05'2.39" E	1133	582.39	1.4	33.9	61.56	Gp1	Cluster 1
87	Khenchela–AS_2	35°26'1.37" N	7°05'2.34" E	1130	581.19	1.41	33.9	61.42	Gp1	Cluster 1
88	Khenchela–DK	35°25'08.30" N	7°04'08.97" E	1049	548.79	1.74	34.4	57.56	Gp2.2	Cluster 3
89	Khenchela–DL_1	35°29'1.57" N	7°15'2.15" E	1012	533.99	1.88	34.7	55.82	Gp1	Cluster 1
90	Khenchela–DL_2	35°29'1.57" N	7°15'2.15" E	1013	534.39	1.88	34.7	55.87	admixed	Cluster 4
91	Khenchela–DL_3	35°29'1.57" N	7°15'2.15" E	1020	537.19	1.85	34.6	56.19	admixed	Cluster 4
92	Khenchela–DT_1	34°58'4.79" N	7°02'6.36" E	824	458.79	2.64	36	47.15	admixed	Cluster 3
93	Khenchela–DT_2	34°58'4.79" N	7°02'6.36" E	900	489.19	2.33	35.5	50.62	Gp2.2	Cluster 3
94	Khenchela–DT_3	34°58'5.79" N	7° 03'6.36" E	1000	529.19	1.93	34.8	55.26	Gp2.2	Cluster 3
95	Khenchela–HES	35°26'8.30" N	7°04'9.02" E	1064	554.79	1.68	34.3	58.27	Gp1	Cluster 1
96	Khenchela–Z_1	34°57'7.34" N	7°02'2.09" E	788	444.39	2.78	36.3	45.52	Gp2.2	Cluster 3
97	Khenchela–Z_2	34°57'3.84" N	7°02'2.64" E	770	437.19	2.85	36.4	44.71	Gp2.2	Cluster 3
98	Laghouat–DGES	34°11'9.82" N	3°03'4.27" E	975	259.2	1.87	34.8	26.96	Gp1	Cluster 1
99	Laghouat–DUAD_1	33°45'0.17" N	2°39'5.22" E	900	229.2	2.17	35.4	23.68	Gp1	Cluster 2
100	Laghouat–DUAD_2	33°45'0.11" N	2°39'5.37" E	877	220	2.26	35.5	22.68	Gp1	Cluster 1
101	Mascara–BO	35°32'5.84" N	0°21'8.39" W	178	267.9	5.53	37.6	28.63	Gp1	Cluster 1
102	Mascara–O	35°32'2.73" N	0°21'9.51" W	196	275.1	5.46	37.5	29.45	Gp1	Cluster 1
103	Mascara–S	35°33'1.62" N	0°15'0.32" W	74	226.3	5.95	38.4	23.95	Gp1	Cluster 1
104	Médea–B	36°08'1.67" N	2°52'3.10" E	826	698.4	4.32	33.9	80.9	Gp1	Cluster 1
105	Médea–OEH	36°09'4.40" N	2°57'6.18 "E	818	695.2	4.35	34	80.46	Gp1	Cluster 1
106	Médea–OEM	36°12'0.00" N	3°08'6.59" E	615	614	5.16	35.4	69.63	Gp1	Cluster 1
107	Médea–OEZ	36°10'6.57" N	2°59'9.58" E	777	678.8	4.51	34.3	78.24	Gp1	Cluster 1
108	Mila–AA	36°30'0.95" N	6°05'9.28" E	264	236.09	6.02	37.6	25.61	Gp1	Cluster 1
109	Mila–DB_1	36°27'4.26" N	6°00'2.82" E	755	432.49	4.06	34.2	49.21	Gp1	Cluster 1
110	Mila–DB_2	36°27'1.76" N	6°00'5.06" E	776	440.89	3.98	34.1	50.27	Gp2.1	Cluster 2
111	Mila–DEK	36°32'2.49" N	5°59'7.66" E	603	371.69	4.67	35.3	41.66	Gp1	Cluster 1
112	Mila–DER	36°27'8.94" N	6°06'2.59" E	260	234.49	6.04	37.7	25.43	Gp1	Cluster 1
113	Mila–HB	36°32'0.29" N	6°00'8.43" E	603	371.69	4.67	35.3	41.66	Gp1	Cluster 1
114	Mila–SK	36°22'5.88" N	6°19'4.34" E	643	387.69	4.51	35	43.63	Gp1	Cluster 4
115	Mila–ZDBH	36°28'0.77" N	6°09'6.56" E	356	272.89	5.66	37	29.86	Gp1	Cluster 1

Continue...

Table 1 – Continuation.

116	M'Sila–D	35°02'9.93" N	4°06'5.58" E	958	388.6	2.5	27.9	52.47	Gp1	Cluster 3
117	M'Sila–DA	35°02'9.75" N	4°06'5.84" E	591	212	3.2	31.2	25.97	Gp1	Cluster 1
118	M'Sila–DM_1	35°14'8.62" N	4°09'9.95" E	1005	377.6	1.54	28.3	48.4	Gp1	Cluster 1
119	M'Sila–DM_2	35°59'3.99" N	4°11'3.15" E	999	375.2	1.57	28.3	48.06	Gp2.1	Cluster 3
120	M'Sila–DOH	35°58'1.24" N	4°26'1.25" E	1080	437.4	2.02	27.1	59.92	admixed	Cluster 1
121	M'Sila–HD	35°58'6.18" N	4°23'8.18" E	943	382.6	2.56	28	51.57	Gp2.2	Cluster 1
122	Oran–FM	35°40'2.48" N	0°47'15.2" W	436	660	7.28	24.9	128.4	Gp1	Cluster 1
123	Oran–M	35°38'6.32" N	0°45'2.89" W	217	484.8	8.16	26.5	90.91	Gp1	Cluster 1
124	Oran–S	35°38'6.98" N	0°42'0.27" W	212	480.8	8.18	26.5	90.08	Gp2.2	Cluster 3
125	OumElBouaghi–DL	35°42'9.79" N	7°00'6.44" E	875	392.32	1.26	35.1	39.76	Gp1	Cluster 1
126	OumElBouaghi–DSR	35°54'1.61" N	7°07'5.34" E	1214	527.92	–0.10	32.7	55.16	admixed	Cluster 3
127	OumElBouaghi–F	35°52'9.90" N	7°07'3.97" E	1003	443.52	0.74	34.2	45.47	admixed	Cluster 1
128	Relizane–AER	35°37'4.62" N	0°23'7.65" E	330	446.5	4.36	37	46.99	Gp1	Cluster 1
129	Relizane–EK_1	35°35'7.77" N	0°20'2.65" E	471	502.9	3.8	36	53.62	admixed	Cluster 1
130	Relizane–EK_2	35°35'7.77" N	0°20'2.65" E	471	502.9	3.8	36	53.62	Gp1	Cluster 1
131	Relizane–Z	35°42'8.50" N	0°46'2.47" E	382	467.3	4.15	36.6	49.41	Gp2.1	Cluster 2
132	Setif–BO	36°25'7.73" N	4°53'1.84" E	1109	435.94	–0.31	34.2	43.35	admixed	Cluster 4
133	Setif–DB	35°41'4.09" N	5°10'4.80" E	854	333.94	0.71	36	32.49	Gp1	Cluster 1
134	Setif–DL	36°23'8.86" N	4°57'8.85" E	776	302.74	1.02	36.5	29.26	admixed	Cluster 1
135	Setif–EN	36°23'2.18" N	4°57'9.75" E	641	248.74	1.56	37.5	23.77	admixed	Cluster 1
136	SidiBelAbbes–AEB	35°22'7.93" N	0°29'8.04" W	516	391.5	2.64	35.7	40.6	Gp1	Cluster 1
137	SidiBelAbbes–SB	35°14'8.23" N	0°35'1.11" W	435	359.1	2.96	36.3	36.97	admixed	Cluster 3
138	SidiBelAbbes–SH	35°17'8.34" N	0°33'1.97" W	419	352.7	3.02	36.4	36.26	admixed	Cluster 3
139	Skikda–BT	36°42'9.43" N	7°18'8.63" E	39	822.19	8.75	28.8	140.4	Gp2.1	Cluster 2
140	Skikda–DSZEA_1	36°41'6.65" N	7°19'3.94" E	46	827.79	8.72	28.8	141.5	Gp2.1	Cluster 3
141	Skikda–DSZEA_2	36°41'6.65" N	7°19'3.94" E	46	827.79	8.72	28.8	141.5	Gp2.1	Cluster 2
142	Skikda–DSZEA_3	36°41'6.65" N	7°19'3.94" E	46	827.79	8.72	28.8	141.5	Gp2.1	Cluster 2
143	Skikda–DSZEA_4	36°41'6.65" N	7°19'3.94" E	46	827.79	8.72	28.8	141.5	Gp2.1	Cluster 2
144	Skikda–SS_1	36°42'09.37" N	7°17'07.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
145	Skikda–SS_10	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
146	Skikda–SS_2	36°42'09.37" N	7°17'07.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 4
147	Skikda–SS_3	36°42'09.37" N	7°17'07.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
148	Skikda–SS_4	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
149	Skikda–SS_5	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
150	Skikda–SS_6	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
151	Skikda–SS_7	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 1
152	Skikda–SS_8	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 2
153	Skikda–SS_9	36°42'09.37" N	7°17'7.07" E	43	825.39	8.74	28.8	141	Gp2.1	Cluster 1
154	Skikda–TN_1	36°42'3.11" N	7°19'1.36" E	45	826.99	8.73	28.8	141.3	Gp2.1	Cluster 2
155	Skikda–TN_2	36°42'3.11" N	7°19'1.36" E	45	826.99	8.73	28.8	141.3	Gp2.1	Cluster 2
156	Skikda–TN_3	36°42'3.11" N	7°19'1.36" E	45	826.99	8.73	28.8	141.3	Gp2.1	Cluster 3
157	Skikda–ZSZEA_1	36°40'06.42" N	7°17'4.97" E	118	885.39	8.44	28.3	153	Gp2.1	Cluster 2
158	Skikda–ZSZEA_2	36°40'06.42" N	7°17'4.97" E	118	885.39	8.44	28.3	153	Gp2.1	Cluster 2
159	Skikda–ZSZEA_3	36°40'06.42" N	7°17'4.97" E	118	885.39	8.44	28.3	153	Gp2.1	Cluster 2
160	SoukAhras–C	36°14'7.99" N	7°57'2.37" E	546	679.8	1.95	33.3	74.46	Gp2.1	Cluster 3
161	SoukAhras–Z_1	36°11'8.49" N	7°57'1.60" E	895	819.4	0.56	30.8	92.86	Gp2.1	Cluster 3
162	SoukAhras–Z_2	36°11'8.54" N	7°57'2.31" E	887	816.2	0.59	30.9	92.42	Gp2.2	Cluster 3
163	Tébessa–DT	35°00'6.60" N	7°39'6.99" E	1004	415	1.08	34	43.25	admixed	Cluster 1
164	Tébessa–L	35°49'6.20" N	7°52'3.07" E	691	289.8	2.33	36.2	29.37	admixed	Cluster 4
165	Tipaza–PC	36°41'2.34" N	2°47'9.37" E	45	640.2	8.31	30.6	98.33	Gp1	Cluster 2
166	Tipaza–PM	36°37'2.42" N	2°24'4.38" E	18	629.4	8.42	30.8	96.33	Gp2.1	Cluster 2
167	Tipaza–SR	36°34'8.00" N	2°29'4.85" E	146	680.6	7.9	29.9	106	Gp1	Cluster 1
168	Tissemsilt–TEH	35°54'7.61" N	2°04'0.32" E	931	455.8	1.97	31.1	53.67	Gp1	Cluster 1
169	TiziOuzou–BY	36°33'7.42" N	4°12'9.04" E	565	1024.5	5.09	33.4	124.3	Gp1	Cluster 1
170	TiziOuzou–DD	36°29'9.96" N	4°14'7.96" E	1001	1198.9	3.35	30.3	152.5	Gp1	Cluster 1
171	TiziOuzou–M	36°42'2.15" N	4°00'3.36" E	87	833.3	7	36.7	96.23	Gp1	Cluster 1
172	TiziOuzou–O	36°35'0.32" N	4°09'5.67" E	232	891.3	6.42	35.7	104.5	Gp1	Cluster 1
173	Tlemcen–H	34°58'0.96" N	1°22'00.5" W	375	463	5.08	33	56.89	admixed	Cluster 3
174	Tlemcen–R	35°08'0.68" N	1°26'8.65" W	119	258.2	6.11	34.8	30.88	Gp1	Cluster 2

ancestry. An Unweighted Neighbor–Joining dendrogram was generated in the DARWIN software version 6.0.010 with 1000 bootstraps value for tree construction and the tree was viewed using FigTree 2016–10–04–v1.4.4. Finally, the XLSTAT 2020.5.1 software was used to calculate the Spearman correlation coefficient to assess the correlation between clusterization, obtained by the software STRUCTURE, and ecogeographic parameters.

Results

Genetic diversity

Each of the 174 samples was successfully amplified at 16 SSR loci. We obtained 173 alleles (Na) (average 10.81 alleles per locus), ranging from three for DCA15 to 18 for UDO43. The number of effective alleles (Ne) varied between 1.79 (GAPU45) and 12.74 (DCA16) with a mean value of 6.15. The Shannon information Index (I) ranged from 0.86 (GAPU45) to 2.78 (DCA16). Observed heterozygosity (Ho) ranged from 0.32 (EMOL) to 0.89 (DCA05) while expected heterozygosity (He) ranged from 0.44 (GAPU45) to 0.92 (DCA16), with a mean Fixation Index (F) of 0.11. All microsatellite markers were confirmed as highly polymorphic, with PIC values higher than 0.50 except for DCA15 and GAPU45. Null alleles were detected at a frequency higher than 0.2 only in EMOL and DCA17 loci. The departure from Hardy–Weinberg equilibrium was significant for five of the 16 loci analyzed (Table 2).

Table 2 – Genetic diversity indices of 16 SSR markers detected in 174 wild olive samples.

Locus	Na	Ne	I	Ho	He	F	PIC	HW	F(null)
DCA03	8	3.25	1.51	0.72	0.69	–0.04	0.65	NS	–0.0219
DCA04	12	6.75	2.14	0.75	0.85	0.12	0.84	NS	0.0642
DCA05	13	7.03	2.11	0.89	0.86	–0.04	0.84	NS	–0.0194
DCA09	13	8.62	2.31	0.75	0.88	0.15	0.87	NS	0.0832
DCA13	9	5.49	1.89	0.84	0.82	–0.03	0.80	***	–0.0232
DCA15	3	2.37	0.94	0.40	0.58	0.31	0.49	***	0.1837
DCA16	22	12.74	2.78	0.86	0.92	0.07	0.92	ND	0.0343
DCA17	12	7.07	2.14	0.38	0.86	0.55	0.84	***	0.3809
DCA18	14	8.59	2.33	0.87	0.88	0.02	0.87	NS	0.0101
GAPU45	5	1.79	0.86	0.34	0.44	0.23	0.40	NS	0.1065
GAPU71b	7	3.56	1.53	0.68	0.72	0.05	0.68	*	0.0058
GAPU101	7	4.35	1.62	0.84	0.77	–0.08	0.73	NS	–0.043
EMO90	6	2.29	1.12	0.61	0.56	–0.08	0.52	**	–0.0732
EMOL	7	2.13	1.17	0.32	0.53	0.41	0.51	***	0.2617
UDO28	17	10.94	2.60	0.79	0.91	0.13	0.90	NS	0.0694
UDO43	18	11.42	2.60	0.86	0.91	0.05	0.91	ND	0.0264
Mean	10.81	6.15	1.85	0.68	0.76	0.11	0.74		0.0653
Total	173.00								

Na = number of alleles; Ne = number of effective alleles; I = Shannon information index; Ho = observed heterozygosity; He = expected heterozygosity; F = fixation index; PIC = polymorphic information content; HW = Hardy–Weinberg equilibrium test (NS = Not significant; ND = not determined; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$ after Bonferroni correction); F(null), estimated frequency of null alleles.

Genetic relationships

The PCoA explained 13.47 % and 9.13 % of the total variance for the first (PCo1) and the second (PCo2) principal coordinates, respectively. The PCo1 separated genotypes collected in western Algeria, from the eastern genotypes collected in the provinces of Jijel, Mila and Batna, near the border with Tunisia. The PCo2 divided, further, the Eastern samples collected in the temperate north–eastern provinces of Skikda, Guelma, Constantine, Souk Ahras and Jijel, from genotypes collected in the inner arid provinces of Batna, Biskra, Khenchela, Oum El Bouaghi and Tebessa (Figure 2).

Population structure

The population structure indicated two populations (K = 2) as the best model that fits genotype distributions, followed by K = 3 (Figure 3). At K = 2, two genetic clusters are distinguished. The first gene pool (Gp1) consisted of 78 genotypes, while the second (Gp2) included 86 samples. The Gp1 cluster included samples collected in the Northwest of the country, in particular in a region between Tizi Ouzou and Mascara provinces, although it also included several genotypes from the eastern provinces of Annaba and Mila. However, the Gp2 cluster contained genotypes mainly collected in northeastern Algeria with the exception of few genotypes from Oran, Sidi Bel Abbes, Ain Témouchent, and Tlemcen provinces located near the border with Morocco.

At K = 3, the genetic cluster Gp1 had approximately the same composition as at K = 2, while the Gp2 cluster splits up into two sub-groups: Gp2.1, which included mainly genotypes from the northeastern provinces of Skikda, Guelma, Souk Ahras, and Jijel (except for the samples MILA–DB_2, BEJAIA–OG, AIN TÉMOUCHENT–AET and RELIZANE–Z); Gp2.2, which encompassed the genotypes collected in the inner mountainous provinces of Batna, Biskra and Khenchela (except for the samples M'SILA–HD, BORDJBOURARRERIDJ–DEZ and ORAN–S).

The admixed group included 25 genotypes collected in the inner provinces of the northern belt of the country (RELIZANE–EK_1, M'SILA–D, BEJAIA–K, SETIF–BO BATNA–SDES, OUMELBOUAGHI–DL, CONSTANTINE–BH_1, etc.) (Table 1).

The Unweighted Neighbor–Joining dendrogram showed four clusters. Three were compatible with clusters obtained by the population structure analysis and the fourth group included only nine samples (Figure 4). In particular, Cluster I contained 76 genotypes and reflected cluster Gp1 composition obtained from Structure. This group included main samples collected in the central and northwestern provinces with the exception of Annaba (Far East). Cluster II contained 34 samples, matching Gp2.1 (Northeast) except for six genotypes (AINTEMOUCHENT–CE, RELIZANE–Z, TIPAZA–PM, TIPAZA–PC, LAGHOUAT–DUAD_1 and

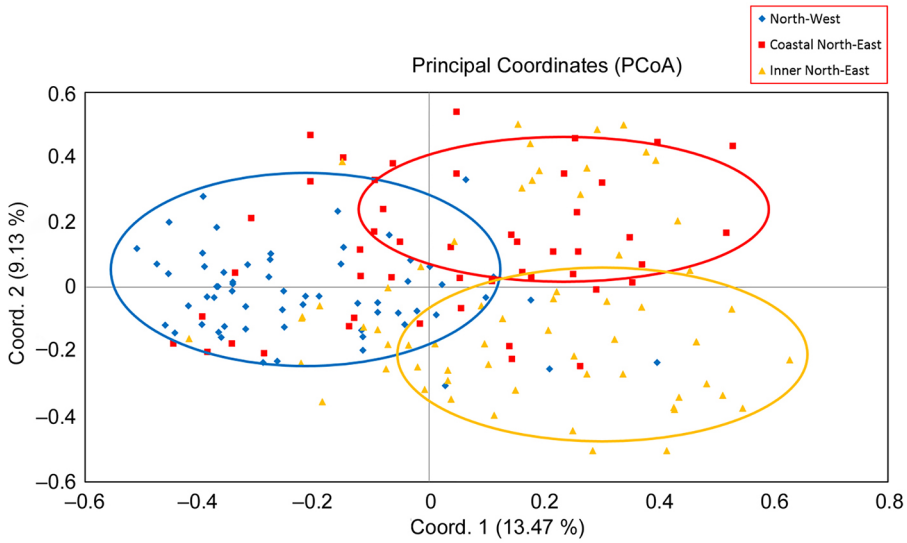


Figure 2 – Principal coordinates analysis (PCoA). Differentiation between 174 Algerian wild olive genotypes based on 16 polymorphic microsatellite markers.

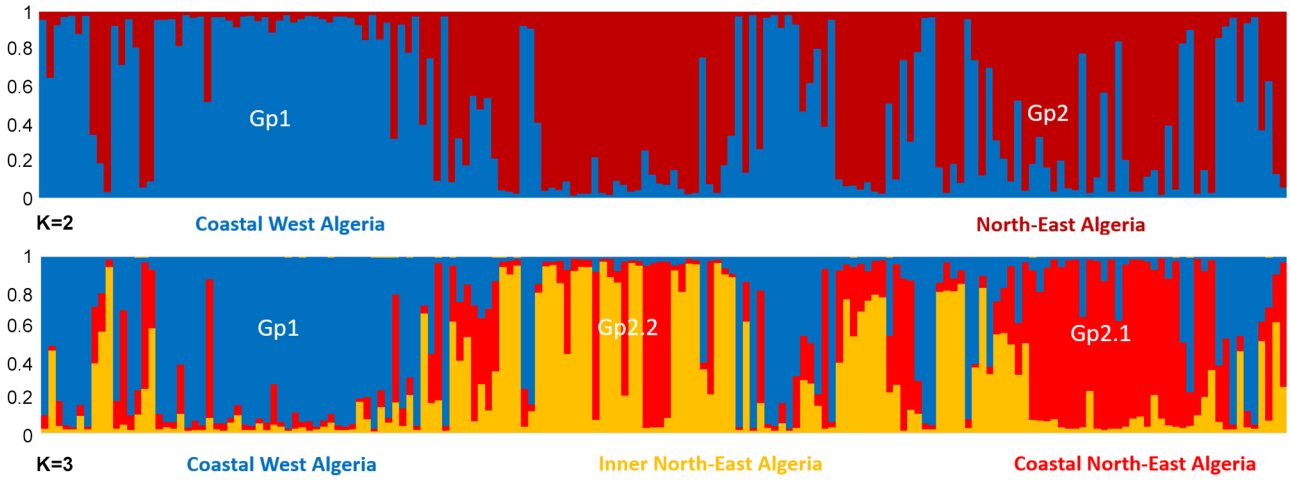


Figure 3 – Analysis of the genetic structure of 174 wild olive accessions. Bar plot showing clusters inferred by STRUCTURE at K = 2 and K = 3. Each vertical line stands for a single accession assigned to defined populations if the value of the corresponding membership coefficient (qi) was higher than 0.6. The samples were sorted according to their geographic origin from West to East.

TLEMCEN-R). Cluster III (55 genotypes) included the samples collected in the southeastern provinces of Batna, Biskra and Khenchela, matching Gp2.2. Finally, Cluster IV contained nine samples, all from the northeastern part of Algeria (Table 1).

To analyze the possible correlation between the genetic structure of 174 wild olive genotypes and bioclimatic conditions, five ecogeographic parameters were used as described in the Materials and Methods section (Table 1).

The results revealed high heterogeneity of samples within the subpopulation Gp1, while the subpopulation Gp2.1 included mostly samples growing at altitude < 600 m, and subject to abundant rains (p > 600 mm), with a temperature range between 4 °C and 30 °C, and an Emberger coefficient > 100 for 66 % of the samples. Conversely, the subpopulation Gp2.2 included genotypes mostly collected at altitude > 600 m, p < 600 mm, M > 30 °C, m < 4 °C. Moreover, the Q2 value was < 100 for all the samples (Figure 5A). The Spearman correlation coefficient was calculated to statistically support the correlation between the ecogeographic parameters and clusterization obtained by the software STRUCTURE. The results showed a significant correlation (p < 0.001) with all the variables considered. The highest positive correlation was observed between the values that support the membership to the Gp2.1 group in the software STRUCTURE and the Q2 values (0.519), while the highest negative correlation was observed between the Gp2.1 group and the Altitude values (−0.383) (Figure 5B).

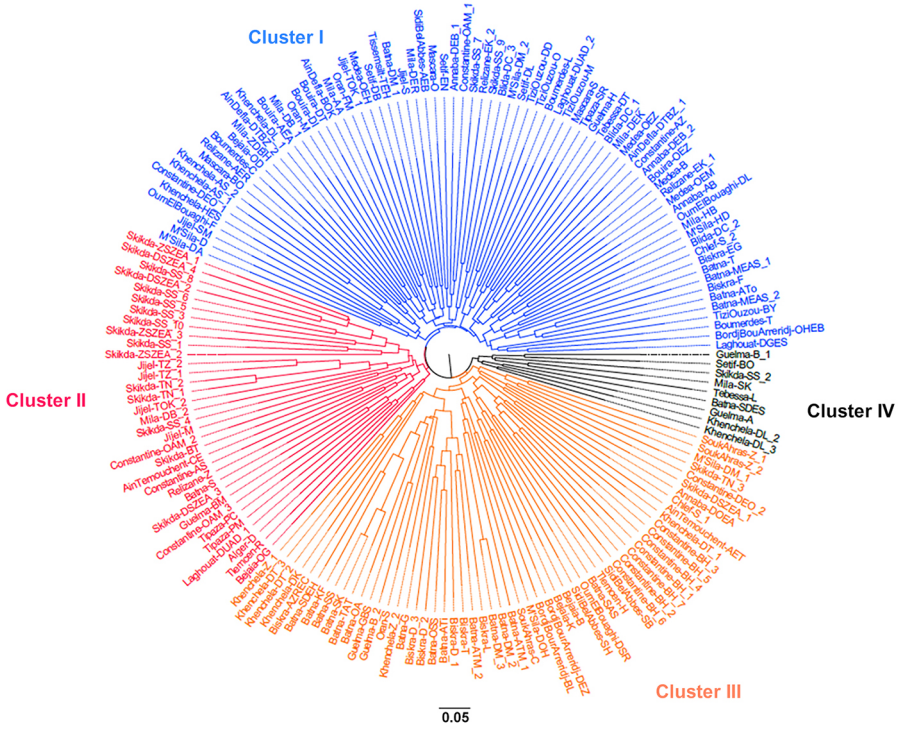


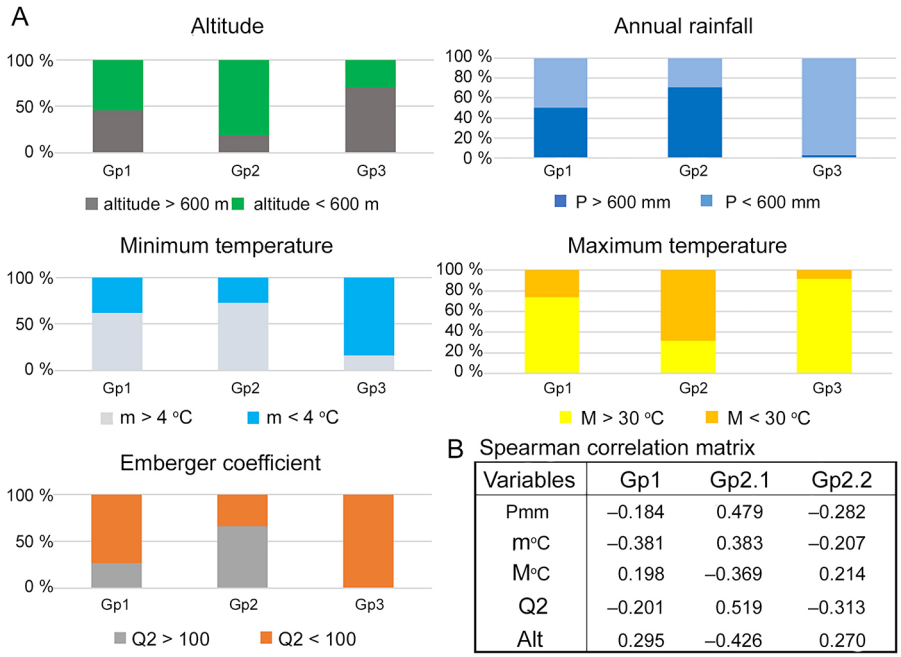
Figure 4 – Unweighted Neighbor–Joining dendrogram generated in DARWIN software version 6.0.010 showing clusterization of the samples analyzed.

Discussion

Biological diversity is a crucial factor to increase and improve productivity in agriculture. Algeria is characterized by low population density and the presence of olive cultivations restricted to the northern coastal region. In Algeria, wild-looking forms of olive trees are preserved in natural areas where they can survive in small grove or scattered plants, due to the isolation from cultivated orchards. Prospections were conducted in 33 provinces in the northern region of Algeria, allowing the collection of 174 samples. The genetic analysis was carried out with 16 microsatellite markers, which were highly informative (Pasqualone et al., 2015; Sabetta et al., 2017; Saddoud et al., 2020). Other authors have reported that the results confirmed a high genetic diversity of the Algerian wild olive (Baldoni et al., 2006; Besnard and Bervillé, 2002; Lumaret et al., 2004; Mousavi et al., 2017; Mulas et al., 2004). Deviations from HW equilibrium and positive values of the inbreeding coefficient in some locus were observed, despite the width of the sampling in the analyzed loci. As already observed in cultivated olive (Di Rienzo et al., 2018; Muzzalupo and Perri, 2009), a certain degree of inbreeding can be favored by the geographic isolation of plants, which promote self-cross reproduction instead of open-pollinated reproduction, as reported in previous works on isolated olive trees (Besnard et al., 2007; Diaz et al., 2006). Indeed, the PCoA underlined the separation of the genotypes analyzed in three groups, according to their growing geographic areas: northwestern coastal area, northeastern coastal area, and northeastern mountainous areas. These three Regions are well separated by physical barriers, such as the Atlas Mountains in the West and the Aurès Massif, in the East (Figure 1).

The STRUCTURE Bayesian-based analysis detected two main populations: Gp1 and Gp2. One cluster could be composed of genuine wild olives while the other one could include feral olives. However, a further subdivision was revealed within the Gp2 cluster. Many hypotheses about the origin of this clusterization may be raised. Based on previous studies, three principal gene pools were identified for domesticated olive, corresponding to three main geographical areas: western (Q1), central Mediterranean (Q2) and eastern Mediterranean (Q3) (Díez et al., 2015; Besnard et al., 2013). Therefore, the distinction of two subgroups within Gp2 could result from a further differentiation within the local oleaster or from the presence of feral forms derived from different cultivated gene pools.

The clusterization obtained by structure seems also to be related to the growing climatic conditions of samples. Indeed, Gp1 collects the genotypes from the large coastal planes of Algeria, from the border with Morocco through the central Bejaia province, toward the plains of Oran and Annaba, characterized by intensive cultivation of olive. The Gp2.1 mostly



B Spearman correlation matrix

Variables	Gp1	Gp2.1	Gp2.2
Pmm	–0.184	0.479	–0.282
m°C	–0.381	0.383	–0.207
M°C	0.198	–0.369	0.214
Q2	–0.201	0.519	–0.313
Alt	0.295	–0.426	0.270

Figure 5 – A Stacked bar–plots illustrating, for each of the three clusters identified by STRUCTURE for K = 3, the percentages of genotypes collected in areas characterized by different ecogeographic parameters (altitude, annual rainfall (P), Minimum (m) –maximum (M) temperature, and Emberger coefficient (Q2). B Spearman correlation coefficient, significant at $p < 0.001$, indicating the correlation between the ecogeographic parameters and the cluster obtained by STRUCTURE software.

included genotypes collected in northeastern Algeria, in the provinces of Skikda, Guelma, Souk Ahras, and Jijel, where the coast is predominantly mountainous with small plains characterized by mild temperature, high rainfall and moderate altitude. The Gp2.2 group included the genotypes collected in the inner eastern part of the country, in the regions of Batna, Biskra and Khenchela, characterized mainly by extreme temperature changes, lack of rain and high altitudes (Figure 5A). In regions near the desert, higher temperature and low precipitations hinder olive cultivation; thus, the genotypes belonging to this group could be particularly adapted to the harsh climatic conditions, such as aridity and thermic excursions. These genotypes could be useful in breeding programs for tolerance to drought, as well as resources for the introduction of olive populations in habitats where adverse conditions endanger this species.

Many samples fall into the admixed group and were mostly collected in the plains of Sétif, Constantine. and Oum El Bouaghi provinces, main centers of grain cultivation during the French colonial period. This area, characterized by fertile soils and Mediterranean climate, is traditionally devoted to agriculture where olive cultivation and human selection could have contributed to an admixture between cultivars and wild populations, as previously reported in other countries (Belaj et al., 2007; Boucheffa et al., 2019; García–Verdugo et al., 2009).

The Unweighted Neighbor–Joining dendrogram confirmed the three main groups outlined by STRUCTURE; nevertheless, it also revealed a fourth group that included nine samples collected in the area between the North and South East, probably the result of mixed pollination between wild and domesticated. Olive is generally considered a wind–pollinated species and its pollen spreads in a range of about 100 m; however, evidences have been found that it can move across kilometers, at low concentrations (Pinillos and Cuevas, 2009). Pollen dispersion and geographical barriers are probably the basis of the complex genetic structure observed in the Algerian wild germplasm and of the different gene pools found in the different geo–climatic conditions. The detection of three main gene pools shows how the geographical barriers can determine partial genetic isolation even on local scale, according to other studies (Belaj et al., 2007; Boucheffa et al., 2019; Breton et al., 2006; Sion et al., 2019). This wide genetic variability deserves further investigation to better understand the relationship between wild and feral forms spread in these areas. Indeed, wild genotypes constitute a priceless resource of genes that need to be preserved and conserved. On the other hand, feral forms represent a variability source useful for olive breeding programs. Future studies should compare the large genetic variability of wild olive with the variability in varieties cultivated in Algeria to investigate their relationships.

Conclusion

Our study provides a genetic characterization of 174 samples of oleaster collected in different regions of northern Algeria. The accessions were clustered according to geographic origin and consequently to their characteristic climatic conditions, which allows the identification of samples from an area characterized by higher temperatures and low precipitation, making them a good source of genes for tolerance to harsh climatic conditions, which is crucial to face challenges posed by climate change.

Acknowledgments

This research was funded by the Exceptional National program 2019/2020. W.F. acknowledges the financial support of Ecole Nationale Supérieure de Biotechnologie Constantine, Algeria. Experiments were also supported by statutory funds of the Department of Soil, Plant and Food Sciences (DISSPA), University of Bari, Italy. The 6th. author has been supported by MIUR – PON Ricerca e Innovazione 2014 – 2020 (project AIM1809249 – attività 2, linea 1). The 2nd. and 5th. authors were supported by PhD program of University of Bari Aldo Moro "Biodiversity, agriculture and environment" curriculum of "Genetics and plant biotechnology". The 7th. author was supported by ADISU and University of Bari Aldo Moro with post-graduate fellowship.

Authors' Contributions

Wahiba Falek, Cinzia Montemurro and Sara Sion equally contributed to the work. **Conceptualization:** Montemurro, C.; Mascio, I.; Miazzi, M.M. **Data acquisition:** Falek, W.; Bechkri, S.; Khelifi, D. **Data analysis:** Falek, W.; Gadaleta, S.; Mascio, I.; Sion, S.; Fanelli, V.; Savoia, M.A.; Piarulli, L. **Design of methodology:** Montemurro, C.; Miazzi, M.M.; Mascio, I.; Falek, W. **Writing and editing:** Mascio, I.; Falek, W.; Montemurro, C.; Fanelli, V.; Sion, S.; Miazzi, M.M.

References

Abdessemed, S.; Muzzalupo, I.; Benbouza, H. 2015. Assessment of genetic diversity among Algerian olive (*Olea europaea* L.) cultivars using SSR marker. Scientia Horticulturae 192: 10–20. https://doi.org/10.1016/j.scienta.2015.05.015

Baldoni, L.; Cultrera, N.G.; Mariotti, R.; Ricciolini, C.; Arcioni, S.; Vendramin, G.G.; Buonamici, A.; Porceddu, A.; Sarri, V.; Ojeda, M.A.; Trujillo, I.; Rallo, L.; Belaj, A.; Perri, E.; Salimonti, A.; Muzzalupo, I.; Casagrande, A.; Lain, O.; Messina, R.; Testolin, R. 2009. A consensus list of microsatellite markers for olive genotyping. Molecular Breeding 24: 213–231. https://doi.org/10.1007/s11032-009-9285-8

Baldoni, L.; Tosti, N.; Ricciolini, C.; Belaj, A.; Arcioni, S.; Pannelli, G.; Germanà, M.A.; Mulas, M.; Porceddu, A. 2006. Genetic structure of wild and cultivated olives in the central Mediterranean basin. Annals of Botany 98: 935–942. https://doi.org/10.1093/aob/mcl178

Bechkri, S.; Khelifi, D. 2017. Variation in *Vicia sativa* s.l. from Algeria based on morphological characters and ecogeographic parameters. Genetic Resources and Crop Evolution 64: 815-832. https://doi.org/10.1007/s10722-016-0404-1

Belaj, A.; Munoz-Diez, C.; Baldoni, L.; Porceddu, A.; Barranco, D.; Satovic, Z. 2007. Genetic diversity and population structure of wild olives from the North-Western Mediterranean assessed by SSR markers. Annals of Botany 100: 449–458. https://doi.org/10.1093/aob/mcm132

Besnard, G.; Bervillé, A. 2002. On chloroplast DNA variations in the olive (*Olea europaea* L.) complex: comparison of RFLP and PCR polymorphisms. Theoretical and Applied Genetics 104: 1157–1163. https://doi.org/10.1007/s00122-001-0834-8

Besnard, G., Baradat, P., Breton, C., Khadari. B., Bervillé, A. 2001. Olive domestication from structure of oleasters and cultivars using RAPDs and mitochondrial RFLP. Genetics Selection Evolution 13: S251–S268. https://doi.org/10.1186/BF03500883

Besnard, G.; El Bakkali, A.; Haouane, H.; Baali-Cherif, D.; Moukhli, A.; Khadari, B. 2013. Population genetics of Mediterranean and Saharan olives: geographic patterns of differentiation and evidence for early generations of admixture. Annals of Botany 112: 1293-1302. https://doi.org/10.1093/aob/mct196

Besnard, G.; Rubio de Casas, R.; Vargas, P. 2007. Plastid and nuclear DNA polymorphism reveals historical proceses of isolation and reticulation in the olive tree complex (*Olea europaea*). Journal of Biogeography 34: 736–752. https://doi.org/10.1111/j.1365-2699.2006.01653.x

Besnard, G.; Terral, J.F.; Cornille, A. 2018. On the origins and domestication of the olive: A review and perspectives. Annals of Botany 121: 385-403. https://doi.org/10.1093/aob/mcy002

Boucheffa, S.; Miazzi, M.M.; Di Rienzo, V.; Mangini, G.; Fanelli, V.; Tamendjari, A.; Pignone, D.; Montemurro, C. 2017. The coexistence of oleaster and traditional varieties affects genetic diversity and population structure in Algerian olive (*Olea europaea*) germplasm. Genetic Resources and Crop Evolution 64: 379–390. https://:doi.org/10.1007/s10722-016-0365-4

Boucheffa, S.; Tamendjari, A.; Sanchez-Gimeno, A.C.; Rovellini, P.; Venturini, S.; Di Rienzo, V.; Miazzi, M.M.; Montemurro, C. 2019. Diversity assessment of algerian wild and cultivated olives (*Olea europaea* L.) by molecular, morphological, and chemical traits. European Journal of Lipid Science and Technology 121: 1800302. https://doi.org/10.1002/ejlt.201800302

Breton, C.; Tersac, M.; Bervillé, A. 2006. Genetic diversity and gene flow between the wild olive (oleaster, *Olea europaea* L.) and the olive: several Plio-Pleistocene refuge zones in the Mediterranean basin suggested by simple sequence repeats analysis. Journal of Biogeography 33: 1916–1928. https://doi.org/10.1111/j.1365-2699.2006.01544.x

Carriero, F.; Fontanazza, G.; Cellini, F.; Giorio, G. 2002. Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.). Theoretical and Applied Genetics 104: 301–307. https://doi.org/10.1007/s001220100691

Cipriani, G.; Marrazzo, M.T.; Marconi, R.; Cimato, A.; Testolin, R. 2002. Microsatellite markers isolated in olive (*Olea europaea* L.) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. Theoretical and Applied Genetics 104: 223–228. https://doi.org/101007/s001220100685

D'Agostino, N.; Taranto, F.; Camposeo, S.; Mangini, G.; Fanelli, V.; Gadaleta, S.; Miazzi, M.M.; Pavan, S.; Di Rienzo, V.; Sabetta, W. 2018. GBS-derived SNP catalogue unveiled wide genetic variability and geographical relationships of Italian olive cultivars. Scientific Reports 8: 15877.

De La Rosa, R.; James, C.M.; Tobutt, K.R. 2002. Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae. Molecular Ecology Notes 2: 265–267. https://doi.org/10.1046/j.1471-8286.2002.00217.x

Diaz, A.; Martin, A.; Rallo, P.; Barranco, D.; Rosa, R.D.L. 2006. Self-incompatibility of 'Arbequina' and 'Picual' olive assessed by SSR markers. Journal of American Society for Horticultural Science 131: 250–255. https://doi.org/10.21273/JASHS.131.2.250

Díaz-Rueda, P.; Franco-Navarro, J.D.; Messora, R.; Espartero, J.; Rivero-Núñez, C.M.; Aleza, P.; Capote, N.; Cantos, M.; García-Fernández, J.L.; Cires, A.; Belaj, A.; León, L.; Besnard, G.; Colmenero-Flores, J.M. 2020. SILVOLIVE, a germplasm collection of wild subspecies with high genetic variability as a source of rootstocks and resistance genes for olive breeding. Frontiers in Plant Science 11: 629. https://doi.org/10.3389/fpls.2020.00629

Díez, C.M.; Trujillo, I.; Martinez-Urdiroz, N.; Barranco, D.; Rallo, L.; Marfil, P.; Brandon, S.G. 2015. Olive domestication and diversification in the Mediterranean Basin. New Phytologist 206: 436–447.

Di Rienzo, V.; Miazzi, M.M.; Fanelli, V.; Sabetta, W.; Montemurro, C. 2018. The preservation and characterization of Apulian olive germplasm biodiversity Acta Horticulturae 1199: 1-6. https://doi.org/10.17660/ActaHortic.2018.1199.1

Emberger, L. 1930. The vegetation of the Mediterranean region, an attempt at a classification of plant groups = La vègètation de la region mediterranneenne, essai d'une classification des groupements vègètaux. Revue Générale de Botanique 42: 641-662, 705-721 (in French).

Evanno, G.; Regnaut, S.; Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

García-Verdugo, C.; Fay, M.F.; Granado-Yela, C.; Rubio de Casas, R.; Balaguer, L.; Besnard, G.; Vargas, P. 2009. Genetic diversity and differentiation processes in the ploidy series of *Olea europaea* L.: a multiscale approach from subsp. to insular populations. Molecular Ecology 18: 454–467. https://doi.org/10.1111/j.1365-294X.2008.04027.x

Green, P.S. 2002. A Revision of Olea L. (*Oleaceae*). Kew Bulletin 57: 91-140. https://doi.org/10.2307/4110824

Hannachi, H.; Sommerlatte, H.; Breton, C.; Msallem, M.; El Gazzah, M.; El Hadj, S.B.; Berville, A. 2009. Oleaster (var. *sylvestris*) and subsp. *cuspidata* are suitable genetic resources for improvement of the olive (*Olea europaea* subsp. *europaea* var. *europaea*). Genetic Resources and Crop Evolution 56: 393–403. https://doi.org/10.1007/s10722-008-9374-2

León, L.; De la Rosa, R.; Velasco, L.; Belaj, A. 2018. Using wild olives in breeding programs: implications on oil quality composition. Frontiers in Plant Science 9: 232. https://doi.org/10.3389/fpls.2018.00232

Lumaret, R.; Ouazzani, N.; Michuad, H.; Vivier, G.; Deguilloux, M.F.; Di Giusto, F. 2004. Allozyme variation of oleaster populations (wild olive tree) (*Olea europaea* L.) in the Mediterranean basin. Heredity 92: 343–351. https://doi.org/10.1038/ sj.hdy.6800430

Miazzi, M.M.; Di Rienzo, V.; Mascio, I.; Montemurro, C.; Sion, S.; Sabetta, W.; Vivaldi, G.A.; Camposeo, S.; Caponio, F.; Squeo, G.; Difonzo, G.; Loconsole, G.; Bottalico, G.; Venerito, P.; Montilon, V.; Saponari, A.; Altamura, G.; Mita, G.; Petrontino, A.; Fucilli, V.; Bozzo, F. 2020. Re.Ger.O.P.: an integrated project for the recovery of ancient and rare olive germplasm. Frontiers in Plant Science 11: 73. https://doi.org/10.3389/fpls.2020.00073

Mousavi, S.; Mariotti, R.; Bagnoli, F.; Costantini, L.; Cultrera, N.G.M.; Arzani, K.; Pandolfi, S.; Vendramin, G.G.; Torkzaban, B.; Hosseini-Mazinani, M.; Baldoni, L. 2017. The eastern part of the Fertile Crescent concealed an unexpected route of olive (*Olea europaea L.*) differentiation. Annals of Botany 119: 1305-1318. https://doi.org/10.1093/aob/mcx027

Mulas, M.; Fadda, A.; Cauli, E. 2004. First observations on wild olive clones (*Olea europaea var. sylvestris* Hoff. E Link) selected for forest use = Prime osservazioni su cloni di oleastro (*Olea europaea* var. *sylvestris* Hoff. E Link) selezionati per l'utilizzo forestale. Italus Hortus 4: 214–217 (in Italian).

Muzzalupo, I.; Perri, E. 2009. Genetic diversity in olive tree cultivars from Italy and other countries of the Mediterranean basin as revealed by RAPD and SSR molecular marker. Advances in Horticultural Science 23: 263-275. https://doi.org/10.1400/121244

Novelli, S.; Gismondi, A.; Di Marco, G.; Canuti, L.; Nanni, V.; Canini, A. 2019. Plant defense factors involved in *Olea europaea* resistance against *Xylella fastidiosa* infection International Journal of Plant Research 132: 439-455. https://doi.org/10.1007/s10265-019-01108-8

Pasqualone, A.; Di Rienzo, V.; Miazzi, M.M.; Fanelli, V.; Caponio, F.; Montemurro, C. 2015. High resolution melting analysis of DNA microsatellites in olive pastes and virgin olive oils obtained by talc addition. European Journal of Lipid Science and Technology 117: 2044-2048. https://doi.org/10.1002/ejlt.201400654

Pinillos, V.; Cuevas, J. 2009. Open pollination provides sufficient level of cross-pollen in Spanish monovarietal olive orchards. American Society for Horticultural Science 44: 499–502. htpps://doi.org/10.21273/HORTSCI.44.2.499

Rubio de Casas, R.; Besnard, G.; Schönswetter, P.; Balaguer, L.; Vargas, P. 2006. Extensive gene flow blurs phylogeographic but not phylogenetic signal in Olea europaea L. Theoretical and Applied Genetics 113: 575-583

Sabetta, W.; Miazzi, M.M.; Di Rienzo, V.; Fanelli, V.; Pasqualone, A.; Montemurro, C. 2017. Development and application of protocols to certify the authenticity and the traceability of Apulian typical products in olive sector. Rivista Italiana Delle Sostanze Grasse 94: 37–43.

Saddoud, O.M.M.D; Miazzi, O.; Elloumi, M.; Fendri, F.; Ben Amar, M.; Savoia, S.; Sion, H.; Souabni, S.; Rahmani, M.S. 2020. Recovery, assessment, and molecular characterization of minor olive genotypes in Tunisia. Plants 9: 382.

Saponari, M.; Giampetruzzi, A.; Loconsole, G.; Boscia, D.; Saldarelli, P. 2019. *Xylella fastidiosa* in olive in Apulia: where we stand. Phytopathology 109: 175-186. https://doi.org/10.1094/PHYTO-08-18-0319.FI

Sefc, K.M.; Lopes, M.S.; Mendonça, D.; Santos, M.R.; Machado, M.L.C.; Machado, A.C. 2000. Identification of microsatellite loci in olive (*Olea europaea* L.) and their characterization in Italian and Iberian olive trees. Molecular Ecology 9: 1171–1193. https://doi.org/10.1046/j.1365-294x.2000.00954.x

Sion, S.; Taranto, F.; Montemurro, C.; Mangini, G.; Camposeo, S.; Falco, V.; Gallo, A.; Mita, G.; Saddoud Debbabi, O.; Ben Amar, F.; Pavan, S.; Roseti, V.; Miazzi, M.M. 2019. Genetic characterization of apulian olive germplasm as potential source in new breeding programs. Plants 8: 268. https://doi.org/10.3390/plants8080268

Spadoni, A.; Sion, S.; Gadaleta, S.; Savoia, M.A.; Piarulli, L.; Fanelli, V.; Di Rienzo, V.; Taranto, F.; Miazzi, M.M.; Montemurro, C.; Sabetta, W. 2019. A simple and rapid method for genomic DNA extraction and microsatellite analysis in tree plants. Journal of Agricultural Science and Technology 21: 1215–1226. https://journals.modares.ac.ir/article-23-13213-en.html

Vossen, P. 2007. Olive oil: history, production, and characteristics of the world's classic oils. Horticultural Science 42: 1093–1100. https://doi.org/10.21273/HORTSCI.42.5.1093

Zohary, D.; Hopf, M. 1994. Domestication of Plants in the Old World. 2ed. Clarendon Press, Oxford, UK.