Acta Bot. Croat. 77 (1), 88–92, 2018
DOI: 110.1515/botcro-2017-0019

CODEN: ABCRA 25
ISSN 0365-0588
eISSN 1847-8476

Short communication

Assessment of the genetic relationship of Turkish olives (*Olea europaea* subsp. *europaea*) cultivars based on cpDNA *trnL-F* regions

Ergun Kaya[1], Recep Vatansever[2], Ertugrul Filiz[3]*

[1] Department of Molecular Biology and Genetics, Mugla Sitki Kocman University, 48000, Kotekli, Mugla, Turkey
[2] Department of Biology, Faculty of Science and Arts, Marmara University, 34722, Istanbul, Turkey
[3] Department of Crop and Animal Production, Cilimli Vocational School, Duzce University, 81750, Duzce, Turkey

Abstract – The olive tree (*Olea europaea* L.) is one of the major cultivated species in the world, and Mediterranean countries produce about 90% of world cultivated olives. In this study, the genetic relationship of seven Turkish olive cultivars was investigated using non-coding *trnL-F* regions in chloroplastic genome. Cultivars demonstrated a similar sequence length of 330-340 bp with an average 35.26% G+C content. Variable (polymorphic/segregating), parsimony informative and total numbers of the insertion or the deletion of bases in the DNA (indel sites) were 4, 3, and 28, respectively. Nucleotide diversities π and θ were found as 0.00631 and 0.00644 respectively, while Tajima's D was –0.786. cpDNA *trnL-F* regions of sequenced Turkish olive cultivars had a low level of genetic variations, and these non-coding regions were strictly conserved in all analyzed cultivars. Geographically distant shared more sequence similarities than relatively close cultivars. The phylogenetic analyses indicated that the biogeographic distribution of cultivars does not demonstrate any association inferring cultivar source. These results indicate the possibility of germplasm exchanges among countries or that some indel mutations contribute to variations of the Turkish olive gene pool. Thus, the authorities should develop the necessary programs to preserve the purity of native germplasms.

Keywords: germplasm, intergenic, lineage, non-coding *trnL-F*, pairwise, plastome

Introduction

The olive tree (*Olea europaea* L.) is one of the major cultivated species in the Mediterranean Basin. *O. europaea*, which belongs to section *Olea*, demonstrates a wide spectrum of distribution with six natural subspecies, including *O. europaea* subsp. *Europaea* (Mediterranean Basin), *O. europaea* subsp. *cuspidata* (from South Africa throughout East Africa and Arabia to South West China), *O. europaea* subsp. *guanchica* (Canaries), *O. europaea* subsp. *cerasiformis* (Madeira), *O. europaea* subsp. *maroccana* (Morocco), and *O. europaea* subsp. *laperrinei* (Algeria, Sudan and Niger). The Mediterranean form, *O. europaea* subsp. *europaea* contains two varieties, cultivated (var. *europaea*) and wild (var. *sylvestris*) (Green 2002). In *O. europaea* members, seven main cpDNA lineages were reported such as E1 (Mediterranean area and Saharan Mountains), E2 and E3 (Western Mediterranean area), M (Macaronesia), C1 and C2 (from Southern Asia to Eastern Africa), and A (Tropical African olives) (Green 2002).

In Turkey, the olive grows in the Aegean, Marmara, Mediterranean and Southeastern Anatolia regions. Turkey hosts a large number of cultivated and wild cultivars/germplasms, including very old cultivars such as "Gemlik", "Ayvalık" amd "Uslu"(Mendilcioglu 1999, Ercisli 2004). In addition, Turkey is one of the most important olive producers in the world, along with Spain, Italy and Greece. *Olea* species have been used in various molecular studies, including plastome sequencing (Besnard et al. 2011), restriction endonuclease studies (Amane et al. 2000), SSR (Hannachi et al. 2010), RAPD (Hess et al. 2000) and ISSR marker analyses (Kaya 2015). Besides, chloroplast DNA (cpDNA) sequences have been used as genomic resources in plant phylogenetic

* Corresponding author, e-mail: ertugrulfiliz@gmail.com

studies. The cpDNA *trnL–trnF* intergenic spacer region has been used for intra- and interspecific levels in plants (Mes et al. 2000). For example, four plastid sequences such as *trnL-trnF, trnT-trnL, trnS-trnG*, and *matK* were studied in 71 *Olea* samples, resulting in 261 variable sites and 121 potentially parsimony-informative characters (Besnard et al. 2009). In a different study, two non-coding chloroplast loci, *rps16* intron and *trnL-F* were used in the phylogenetic analysis of 76 species of the Oleaceae family; *rps16* and *trnL-F* datasets contained 265 and 240 informative characters, respectively (Wallander and Albert 2000). In addition, Besnard et al. (2011) reported that nucleotide divergence between olive cpDNA lineages was low (<0.07%) in eight complete cpDNA genomes of *Olea*. Moreover, chloroplast genome organization and gene order of *O. europaea*, subsp. *europaea* var. *europaea* was reported to be conserved among numerous Angiosperm species, indicating the lack of gene inversions, duplications, insertions, inverted repeat expansions and intron losses (Mariotti et al. 2010).

Recent studies have reported genetic analyses of some Turkish olive cultivars using DNA-based molecular markers (Ipek et al. 2012, Çelikkol et al. 2014, Kaya 2015). However, no studies are available regarding the usage of *trnL-F* intergenic spacer regions for genetic assessment in Turkish olive cultivars. Therefore, in this study, we have analyzed the seven Turkish olive cultivars using *trnL-F* intergenic spacer regions to investigate the genetic relationships among Turkish olive cultivars.

Materials and methods

Plant materials and genomic DNA isolation

Seven cultivated Turkish olive cultivars (*O. europaea* subsp. *europaea* var. *europaea*) were obtained from six different locations in Turkey (Tab. 1 and On-line Suppl. Fig. 1). Total genomic DNA was isolated from 0.5 g of powdered fresh leaves with the use of the CTAB method (Doyle and Doyle 1987). The DNA concentration of each sample was measured with the use of BioSpec-nano (Shimadzu, Japan) and then elutions were diluted with distilled water for a final concentration of 50 ng μL^{-1}.

trnL-F amplification and sequencing

The cpDNA *trnL-F* regions in seven Turkish olive cultivars were amplifiedwith the use of PCR with *trnL-F* primers. For PCR amplification, reaction mixture was prepared for 25 μL total volume, containing 2.0 mM $MgCl_2$, nucleotides dATP, dTTP, dCTP, and dGTP (200 μM each), 0.2 μM primers, 50 ng template DNA, and 0.5 units (U) of Taq DNA polymerase (Thermo Sci, USA). Thermal cycling conditions were chosen as: 3 min at 94 °C; 36 cycles of 45 s at 94 °C, 1 min at annealing temperature of each primer pair, and 1 min at 72 °C, and a final extension step of 5 min at 72 °C. PCR products were analyzed on 1% agarose gel in 1×TBE buffer, stained with Safeview DNA stain (NBS scientific, UK), and visualized with Quantum ST5 imaging system. Then, PCR products were purified using a GeneJET Gel Extraction Kit (Thermo Scientific, USA), and sequenced by Iontek Sequencing Service.

Data analysis

To supplement the seven Turkish olive cultivars, 25 additional *trnL-F* sequences from various *Olea* species were obtained from the NCBI database (www.ncbi.nlm.nih.gov). Thus, a total of 32 *trnL-F* sequences were used in analyses. Additional species include *O. borneensis*, *O. capensis* subsp. *capensis*, *O. capensis* subsp. *enervis*, *O. capensis* subsp. *macrocarpa*, *O. chimanimani*, *O. dioica*, *O. europaea* subsp. *cerasiformis*, *O. europaea* subsp. *cuspidate*, *O. europaea* subsp. *guanchica*, *O. europaea* subsp. *laperrinei*, *O. javanica*, *O. lancea*, *O. paniculata*, *O. neriifolia*, *O. rosea*, *O. salicifolia*, *O. schliebenii*, *O. tsoongii*, *O. welwitschii* and *O. europaea* subsp. *europaea*. The NCBI accession numbers of additional sequences were indicated on phylogenetic trees. The sequence alignment was done by CLUSTALW (Thompson et al. 2002) and a phylogenetic tree was constructed with MEGA v. 6.0 using the maximum likelihood (ML) method for 1000 bootstrap values (Tamura et al. 2013). Bootstrap values lower than 70% were not shown on the branch. The DNAsp v 5.10 was used to calculate the nucleotide diversity with π (Nei 1987) and θ (Watterson 1975), segregating/polymorphic sites (S), and Tajima's D (Tajima 1989, Librado and Rozas 2009). Genotype/cultivar comparisons were performed by Arlequin 3.5.2 software with the use of the population comparison tool with the following settings; pairwise differences (π) with Tajima and Nei's method for 100 permutations and 0.05 significance level (Excoffier and Lischer 2010).

Tab. 1. Geographic location and some gene features of seven Turkish olive cultivars (*Olea europaea* subsp. *europaea*).

NCBI access.	Genotype name	Locality	Geographical regions	GPS coordinates	*trnL-F* gene (bp)	G+C content (%)
KJ670690	Yaglık	Izmir	Aegean	38°27'53.7"N, 27°14'22.9"E	337	35.01
KJ670691	Mugla	Mugla	Aegean	37°07'43.1"N, 28°27'05.0"E	330	35.45
KJ670688	Gemlik	Bursa	Marmara	40°31'08.7"N, 29°06'05.0"E	339	35.69
KJ670692	Hatay	Hatay	Mediterranean	36°46'59.4"N, 36°15'47.7"E	339	35.10
KJ670693	Samsun	Samsun	Black sea	41°38'42.3"N, 35°26'07.9"E	337	35.01
KJ670694	Tekir	Izmir	Aegean	38°27'53.7"N, 27°14'22.9"E	340	35.59
KJ670689	Burhaniye	Izmir	Aegean	38°30'28.5"N, 27°01'58.9"E	340	35.00

Results and discussion

Sequence analysis of *trnL-F* genes

The cpDNA *trnL-F* regions of the seven sequenced Turkish olive cultivars demonstrated similar sequence lengths, ranging from 330 bp (Mugla cultivar) to 340 bp (Tekir and Burhaniye cultivars), with an average of 35.26% G+C content (Table 1). Similar results were also indicated in a previous study, in which the length of *trnL-F* sequences in *Olea* species was reported to range between 327-343 bp, and the number of variable/potentially informative characters and mean G+C content were calculated as 23/6, and 35%, respectively (Besnard et al. 2009). To analyze the conservancy or divergency degrees in Turkish olive cultivars, sequences were multiply aligned (Fig. 1). The alignment analysis demonstrated that non-coding *trnL-F* sequences are strictly conserved in all Turkish olive cultivars. Despite the high degree of similarities between sequences, we also searched for the presence of any informative variable site/s. The variable (polymorphic/segregating), parsimony informative sites and total numbers of indel sites were found to be 4, 3, and 28, respectively. In addition, nucleotide diversity π and θ values were 0.00631 and 0.00644 respectively, while Tajima's D was -0.786. The nucleotide diversity and parsimony informative sites in Turkish olive cultivars were found to be low. In the *Olea* subgenus, the cpDNA substitution rate was predicted to be between 1.2×10^{-10} and 2×10^{-10}. These values were ten times lower than plastid mutation rates previously reported in other plant species. This slower molecular evolution might be related with the long generation time of olive trees (Besnard et al. 2009). The plastome sequence comparisons also demonstrated that cpDNA of olive cultivars represent low level genetic variations (Mariotti et al. 2009). In a study from Moroccan olives, low cpDNA variation was reported using various restriction endonuclease enzymes (Amane et al. 2000). Therefore, the low level of genetic variations in the cpDNA *trnL-F* regions of the seven Turkish olive cultivars sequenced in this study complies with the reports of previous studies. Furthermore, negative and positive Tajimas's D values were reported to show an excess of low-frequency and intermediate polymorphisms (Luo et al. 2012). In this study, Tajimas's D was found to be negative (–0.786), suggesting a possibility of selection at *trnL-F* loci in analyzed *Olea* members.

Pairwise-comparison of Turkish olive cultivars

In order to have insights about the genetic relationships of Turkish olives, pairwise similarity and difference comparisons of *trnL-F* sequences were conducted. The similarity matrix (On-line Suppl. Tab. 1) ranged between lowest 0.927 (Yaglık-Mugla cultivars) and highest 0.988 (Samsun-Hatay cultivars) values. However, it was interesting that different regional cultivars such as those from the Samsun, Hatay, Black Sea and Mediterranean regions had the highest similarity value, while the same regional cultivars such as those from the Mugla, Yaglık and Aegean region had the lowest value. Besides, the difference matrix (Fig. 2) demonstrated that Hatay and Samsun have lowest (0.00) pairwise difference, therefore the highest similarity, while Mugla and Gemlik have highest (4.052) pairwise difference, and therefore the lowest similarity in analyzed samples. Accordingly, it seems that geographically distant cultivars share more sequence similarities than relatively close varieties. Previous studies have also demonstrated similar results. For example, in SSR analysis, the Turkish olive cultivars analyzed were not grouped according to their their locations of cultivation; Southeastern Anatolia and Aegean Region cultivars were grouped together (Ipek et al. 2012). In a different study, Sarri et al. (2006) reported that 118 olive cultivars from different regions were genetically identical in the Mediterranean Basin. In another study, all Gemlik (Turkish cultivar) individuals collected from the same region indicated 100% identity (Çelikkol et al. 2014). The low level of genetic variations in cpDNA *trnL-F* regions in Turkish olive cultivars analyzed could have arisen from indel events. In this study, a total of 28 indel events were found in cultivars.

Fig. 1. The alignment of *trnL-F* regions in seven Turkish olive cultivars. The sequences were aligned by CLUSTALW, and identical and similar nucleotide residues were shaded black and grey, respectively, with 100% threshold value. The sequences were annotated with NCBI accession numbers and cultivar names.

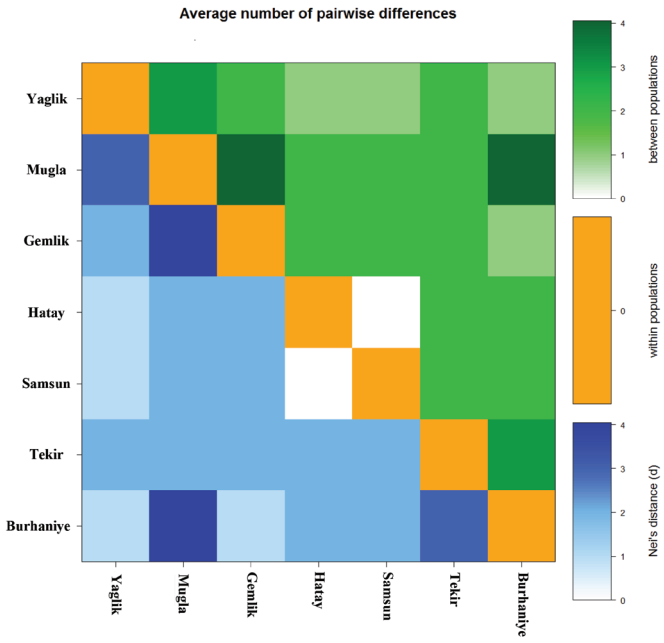

Fig. 2. The pairwise difference comparisons of seven Turkish olive cultivars. The blue, orange and green colors represent Nei's distance, within population and between population variations, respectively.

Phylogenetic analysis of *trnL-F* genes

In addition to seven Turkish cultivars, 25 additional *trnL-F* sequences from various *Olea* species were obtained from NCBI. Thus, a total of 32 *trnL-F* sequences were comparatively analyzed in three separate phylogenetic trees. The phylogenetic tree of Turkish olive cultivars (On-line Suppl. Fig. 2) demonstrated two main groups, A and B. Group A was then subdivided into two subgroups based on the tree topology, A1 and A2. The subgroup A1 included Mugla and Tekir cultivars, subgroup A2 contained Hatay and Samsun cultivars, and group B had Yaglık, Gemlik and Burhaniye cultivars. Thus, phylogenetic distribution of cultivars did not demonstrate any association inferring the regional localization in terms of cpDNA *trnL-F* regions.

To further understand the phylogenetic distribution of Turkish olive cultivars along with various *O. europaea* subsp. *europaea* species, a phylogenetic tree was constructed using 13 *trnL-F* sequences (Fig. 3). Phylogeny showed two major groups, A and B. The group A was then subdivided into two subgroups based on tree topology, named A1 and A2. All *O. europaea* subsp. *europaea* individuals from various countries such as Australia, Morocco, Algeria, Egypt, Italy and Syria were clustered in subgroup A1, while five Turkish cultivars, including Gemlik, Yaglık, Burhaniye, Hatay and Samsun were in subgroup A2, with two of the Turkish cultivars, Mugla and Tekir, being separated from others and clustered in group B. The close phylogenetic relationship of five Turkish cultivars (Gemlik, Yaglık, Burhaniye, Hatay and Samsun) with various *O. europaea* subsp. *europaea* members from different countries indicated their common ancestral origin. However, two Turkish Aegean cultivars, Mugla and Tekir demonstrated a clear divergence from all other *Olea* individuals, indicating that either these two cultivars are native to Turkey or they share a totally different origin. Overall, germplasm exchanges between countries or some mutations may have been a major contributing factor in variations of Turkish olive gene pool. Thus, the relevant authorities should develop the necessary breeding programs in order to preserve the purity of of germplasms.

Moreover, we also constructed an *Olea* genus-level phylogenetic tree using a total of 32 *trnL-F* sequences (On-line Suppl. Fig. 3). The phylogeny included two major groups, A and B. The group A was then subdivided into three subgroups A1, A2 and A3, based on the tree topology. Subgroup A1 included cultivars from various geographic locations such as Algeria, Italy, Egypt, Pakistan, Morocco, Portugal, Spain, Australia and Syria, subgroup A2 had cultivars from neighboring countries such as Tanzania, Kenya, Zimbabwe, South Africa and Madagascar, subgroup A3 only contained Turkish varieties, and group B had cultivars from countries relatively close to each other, such as Laos, Philippines, Thailand, Indonesia and China. The phylogenetic analysis indicated that germplasms could be relatively more conserved in group A2, A3 and B countries. However, clustering of different cultivars in the same group could indicate the possibility of germplasm exchanges be-

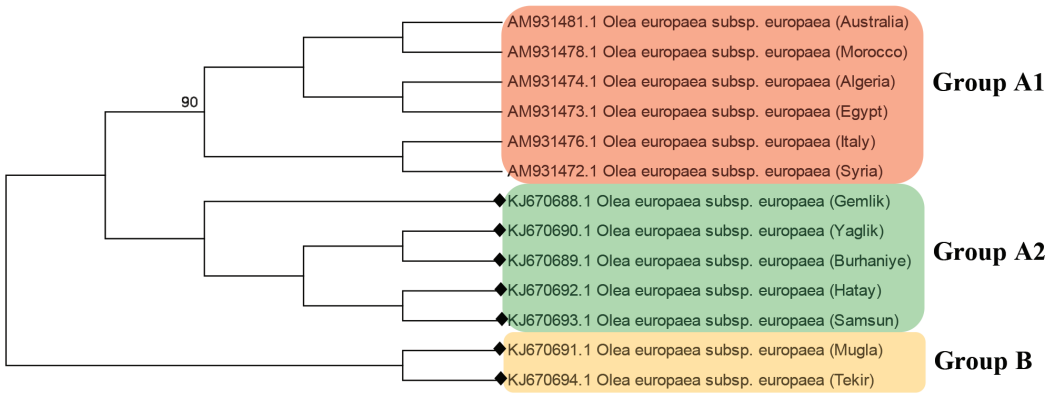

Fig. 3. Phylogenetic tree of various *O. europaea* subsp. *europaea* taxa, including Turkish olive cultivars. The tree was constructed using *trnL-F* sequences with maximum likelihood method for 1000 bootstraps. Turkish olive cultivars are indicated with a diamond symbol, "◆".

tween countries. Previous studies have also reported similar results. Besnard et al. (2001) showed that olive genotypes from different countries grouped together without respect to geographical origins. Slovenian and Croatian olive cultivars were clustered together based on SSR markers (Poljuha et al. 2008). A phylogenetic analysis of approximately 40 *Olea* taxa revealed that different country cultivars were combined in a phylogenetic tree based on four plastid regions (Besnard et al. 2009). Hess et al. (2000) reported that different country cultivars were grouped in a phylogenetic tree in terms of *ITS1* sequences. All these studies indicate that geographical origin was not so effective for olive cultivars as to allow any inference for phylogenetic relationships to be made.

Overall, the analyzed Turkish olive cultivars demonstrated a low level of genetic variation in terms of cpDNA *trnL-F* regions. This could indicate the genetic stability of plastid genomes. In addition, no particular relationship was observed between biogeographic distribution and cultivar localization. What is more, since Turkey has 88 different local olive cultivars (Iooc 2011), further studies with more cultivars are required for a better understanding of the phylogenetic relationships of Turkish olive cultivars with the use of chloroplast genomes.

References

Amane, M., Ouazzani, N., Lumaret, L., Debain, C., 2000: Chloroplast-DNA variation in the wild and cultivated olives (*Olea europaea* L.) of Morocco. Euphytica 116, 59–64.

Besnard, G., Breton, C., Baradat, P., Khadari, B., Berville, A., 2001: Cultivar identification in olive based on RAPD markers. Journal of the American Society for Horticultural Science 126, 668–675.

Besnard, G., Casas, R. R., Christin, P. A., Vargas, P., 2009: Phylogenetics of *Olea* (Oleaceae) based on plastid and nuclear ribosomal DNA sequences: Tertiary climatic shifts and lineage differentiation times. Annals of Botany 104, 143–160.

Besnard, G., Hernández, P., Khadari, B., Dorado, G., Savolainen, V., 2011: Genomic profiling of plastid DNA variation in the Mediterranean olive tree. BMC Plant Biology 11, 80.

Çelikkol, A., Özkan, U., Şan, G., Dolgun, B., Dağdelen, O., Bozdoğan A., Konuşkan, D., 2014: Genetic stability in a predominating Turkish olive cultivar, Gemlik, assessed by RAPD, microsatellite, and AFLP marker systems. Turkish Journal of Botany 38, 430–438.

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

Ercisli, S., 2004: A short review of the fruit germplasm resources of Turkey. Genetic Resources and Crop Evolution 51, 419–435.

Excoffier, L., Lischer, H. E., 2010: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10, 564–567.

Green, P. S., 2002: A revision of *Olea* L. Kew Bulletin 57, 91–140.

Hannachi, H., Breton, C., Msallem, M., Hadj, S. B., Gazzah, M., Bervillé, A., 2010: Genetic Relationships between cultivated and wild olive trees (*Olea europaea* L. var. *europaea* and var. *sylvestris*) based on nuclear and chloroplast SSR markers. Natural Resources 1, 95–103.

Hess, J., Kadereit, J. W., Vargas, P., 2000: The colonization history of *Olea europaea* L. in Macaronesia based on internal transcribed spacer 1 (*ITS-1*) sequences, randomly amplified polymorphic DNAs (RAPD), and intersimple sequence repeats (ISSR). Molecular Ecology 9, 857–868.

Iooc, 2011: International Olive Council. Retrieved from http://www.internationaloliveoil.org/.

Ipek, A., Barut, E., Gulen, H., Ipek, M., 2012: Assessment of inter- and intra-cultivar variations in olive using SSR markers. Scientia Agricola 69, 327–335.

Kaya, E., 2015: ISSR analysis for determination of genetic diversity and relationship in eight Turkish olive (*Olea europaea* L.) cultivars. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 43, 96–99.

Librado, P., Rozas, J., 2009: DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Luo, N., Yu, X., Liu, J., Jiang, Y., 2012: Nucleotide diversity and linkage disequilibrium in antioxidant genes of *Brachypodium distachyon*. Plant Science 197, 122–129.

Mariotti, R., Cultrera, N.G.M., Muñoz Díez, C., Baldoni, L., Rubini, A., 2010: Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea* L.) through plastome sequence comparison. BMC Plant Biology 10, 211.

Mendilcioglu, K., 1999: Subtropical fruit species (Olive). Ege University Agricultural Faculty 12, 43.

Mes, T. H., Kuperus, P., Kirschner, J., Stepanek, J., Oosterveld, P., Storchova, H., Den Nijs, J. C., 2000: Hairpins involving both inverted and direct repeats are associated with homoplasious indels in non-coding chloroplast DNA of *Taraxacum* (Lactuceae: Asteraceae). Genome 43, 634–641.

Nei, M., 1987: Molecular evolutionary genetics. Columbia University, New York.

Poljuha, D., Sladonja, B., Setic, E., Milotic, A., Bandelj, D., Jakse, J., Javornik, B., 2008: DNA fingerprinting of olive varieties in Istria (Croatia) by microsatellite markers. Scientific Horticulture 115, 223–230.

Sarri, V., Baldon, L., Porceddu, A., Cultrera, N. G. M., Contento, A., Frediani, M., Belaj, A., Trujillo, I., Cionini, P. G., 2006: Microsatellite markers are powerful tools for discriminating among olive cultivars and assigning them to geographically defined populations. Genome 49, 1606–1615.

Tajima, F., 1989: Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013: MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30, 2725–2729.

Thompson, J. D., Gibson, T., Higgins, D. G., 2002: Multiple sequence alignment using ClustalW and ClustalX. Current Protocols in Bioinformatics, 2–3.

Wallander, E., Albert, V. A., 2000: Phylogeny and classification of Oleaceae based on *rps16* and *trnL-F* sequence data. American Journal of Botany 87, 1827–1841.

Watterson, G. A., 1975: On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7, 188–193.