Chloroplast DNA analysis of Tunisian cork oak populations (*Quercus suber* L.): sequence variations and molecular evolution of the *trn*L (UAA)-*trn*F (GAA) region

A. Abdessamad[1], G. Baraket[1], H. Sakka[1], Y. Ammari[2], M. Ksontini[2] and A. Salhi Hannachi[1]

[1]Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie LR99ES12, Campus Universitaire Farhat Hached El Manar, Tunis, Tunisie
[2]Laboratoire d'Écologie Forestière, Institut National de Recherche en Génie Rural Eaux et Forêts, Ariana, Tunisie

Corresponding author: A. Salhi Hannachi
E-mail: Amel.SalhiHannachi@fsb.rnu.tn

ABSTRACT. Sequences of the *trn*L-*trn*F spacer and combined *trn*L-*trn*F region in chloroplast DNA of cork oak (*Quercus suber* L.) were analyzed to detect polymorphisms and to elucidate molecular evolution and demographic history. The aligned sequences varied in length and nucleotide composition. The overall ratio of transition/transversion (ti/tv) of 0.724 for the intergenic spacer and 0.258 for the pooled sequences were estimated, and indicated that transversions are more frequent than transitions. The molecular evolution and demographic history of *Q. suber* were investigated. Neutrality tests (Tajima's *D* and Fu and Li)

ruled out the null hypothesis of a strictly neutral model, and Fu's *Fs* and Ramos-Onsins and Rozas' *R*2 confirmed the recent expansion of cork oak trees, validating its persistency in North Africa since the last glaciation during the Quaternary. The observed uni-modal mismatch distribution and the Harpending's raggedness index confirmed the demographic history model for cork oak. A phylogenetic dendrogram showed that the distribution of *Q. suber* trees occurs independently of geographical origin, the relief of the population site, and the bioclimatic stages. The molecular history and cytoplasmic diversity suggest that *in situ* and *ex situ* conservation strategies can be recommended for preserving landscape value and facing predictable future climatic changes.

Key words: Chloroplast DNA; Demographic history; *Quercus suber* L.; Molecular evolution; Polymorphism; *trn*L-*trn*F spacer.

INTRODUCTION

The genus *Quercus* comprises approximately 500 species of trees and shrub, which are distributed throughout much of the Northern hemisphere (Nixon, 1993). Oaks are conspicuous members of the temperate deciduous forests of Europe, North America, and Asia. In addition, *Quercus* is an important evergreen element of Mediterranean woodlands and subtropical forests. The largely temperate species of the subgenus *Quercus* are distinguished from the strictly South-East Asian members of the subgenus *Cyclobalanopsis* by several characters, notably the presence of expanded stigmatic surfaces on the pistillate flowers and small, inconspicuous bracts, which subtend single-staminate flowers (Nixon, 1993). *Quercus suber* L. is an emblematic Mediterranean evergreen sclerophyllous tree. Cork oak is a slow-growing, extremely long-lived evergreen tree (Zucca, 2012). It is a monoecious wind-pollinated species with a protandrous system to ensure cross-pollination (Zucca, 2012). The reproductive system (mostly sexual) and dissemination (gravity and zoochory) of cork oak are consistent with those of most oak species (Magri et al., 2007). These trees may reach about 20 m in height, with massive branches forming a round crown. The thickness of the soft bark is the source of cork, which is stripped every 10-12 years from the outer layer of the bark along the lower portion of the trunk. The modern distribution of *Q. suber* is rather discontinuous, and ranges from the Atlantic coasts of North Africa and the Iberian Peninsula to the Southern regions of Italy (especially Sardinia and Sicily), and includes the main West Mediterranean island as well as Morocco, coastal belts of Maghreb (Algeria and Tunisia), Provence (France), and Catalonia (Spain) (Zucca, 2012). *Q. suber* is widely cultivated within its natural range for cork production (Zucca, 2012). According to Carrion-Vazquez et al. (2000), without human activity, *Q. suber* would never develop pure stands in the Iberian Peninsula, and would form mixed forests with xerophyllous and deciduous oaks together with *Pinus pinaster*. In 2006, cork oak landscapes covered an area of nearly 2,277,700 ha, in Portugal, Spain, Italy, France, Morocco, Algeria, and Tunisia (ANF, 2006). Analysis of the world distribution of cork oak forests (ANF, 2006) reveals that Portugal has the largest area (736,700 ha), followed by Spain (506,000 ha), Algeria (414,000 ha), and Morocco (350,000 ha). France, Tunisia, and Italy have the smallest areas (92,000 ha). This reveals that Portugal contains nearly one-third of the world cork oak forest area; therefore, the environmental services and functions related to this species have

special importance in this country. Cork oak (*Q. suber*) is an endemic feature of the western Mediterranean basin. It occupies large areas and impacts sylvatic populations and diverse vegetation structures in the southern (Morocco, Algeria, Tunisia) and the north-western Mediterranean basin [Italy (Sardinia, Sicily), southwestern France, Spain, and Portugal].

In Tunisia, the cork oak accounts for 25% of hardwood species and occupies the northern part of the country, including two natural regions of Kroumirie and Mogods. *Q. suber* currently covers an area of approximately 90.423 ha or 5.5% of the total area of forest species. The cork oak is currently located mainly in Jendouba, Beja, Bizerte, and Nabeul localities of Northern Tunisia. Renovation of the cork oak forests in the Mediterranean represents a comprehensive approach to the development of space, including fire protection and regeneration of the species. These priorities are supported by the development of forestry policy in the countries concerned. In Tunisia, more than one attempt at cork oak regeneration has been made, including regeneration by seedlings with the development of breeding techniques for oak seedlings in modern forest nurseries. The economic importance of the cork and ecological cork forests in the Mediterranean and globally is widely reported by ecologists. Current yields indicate cork production of approximately 307.500 tons, with Tunisia currently producing only 3% of total world production.

In the *Quercus* genus, phylogenetic relationships based on morphology occur due to pronounced variation in vegetation, which is sharply contrasted with a stabilized, seemingly constant set of floral characters (Tucker, 1974). Most studies on the genetic variation of *Quercus* forest trees have been carried out using nuclear markers. The chloroplast genome typically shows low intraspecific variation, although in the last decade, the development of new marker systems has allowed the detection of sufficient variation to assess phylogeographical patterns.

In recent years, cork oak has been the subject of intensive genetic studies, which have utilized different markers (Olalde et al., 2002). Strong differentiation among populations has been detected around the Mediterranean basin by investigating both chloroplast and mitochondrial DNA, which are maternally inherited in oaks, as demonstrated by Dumolin et al. (1995), as well as allozyme variation. Moreover, evidence of cytoplasmic introgression of *Q. suber* by *Quercus ilex* genes has been reported (Lumaret et al., 2005). Although nuclear microsatellites have been available for oaks for many years (Kampfer et al., 1998), only two studies have reported multiplexing at no more than five loci (Lepais et al., 2006). Thus, analyzing large oak populations using multiple markers remains expensive and time consuming.

Chloroplast DNA (cpDNA) sequence variation is widely used in systematics and to make phylogenetic inferences at different taxonomic levels. Introns and intergenic spacers show high rates of mutation (Baraket et al., 2010). The *trn*T-*trn*L and *trn*L-*trn*F intergenic spacers and *trn*L intron are useful for evolutionary studies at low taxonomic levels, and these regions have been used extensively in phylogenetic studies for the analysis of cytoplasmic polymorphism, and to study the demographic history of several species (Baraket et al., 2010, 2015; Mustapha et al., 2015). Several hypotheses have been developed regarding the evolutionary history of cork oak as well as its center of origin, which has been proposed to lie in the eastern countries of the western Mediterranean Basin (Lumaret et al., 2005; Magri et al., 2007). Variation in cpDNA has been studied by PCR-RFLP and microsatellite markers to determine the route and pattern of postglacial recolonization of native oak throughout Europe and the Mediterranean Basin (Lumaret et al., 2005; Magri et al., 2007).

Here, we studied the genetic relationships and molecular evolution among six Tunisian populations of cork oak based on non-coding regions of cpDNA, i.e., the intergenic

spacer between *trn*L (UAA) and *trn*F (GAA) and the combined region [*trn*L (UAA) intron and the intergenic spacer *trn*L-*trn*F]. Here, we explored cpDNA to analyze population genetic differentiation using non-coding regions and to establish genetic relationships among cork oak trees. This report also describes a molecular evolution scenario and demographic history in order to understand how this species responded to past global and severe climatic changes with particular attention given to recommendations for conservation strategies of this forestry feature.

MATERIAL AND METHODS

Plant material

Plant material composed of young leaves was collected from Northern Tunisia (Kroumirie, Mogods), Cap Bon (Figure 1). We analyzed a set of 10 individuals from each population as follows: Population 1: Hammam Bourguiba (HB), Population 2: Béni Mtir (BM), Population 3: Bellif (B) (Nefza), Population 4: Jebel Khroufa (DK) (Tabarka), Population 5: Keff El Rand (KR) (El Haouaria), and Population 6: Hammam Jdidi (HJ) (Hammamet) (Table 1 and Figure 1).

Table 1. Geographical, bioclimatic, and relief characteristics of different populations of Cork oak (*Quercus suber* L.) in Tunisia.

Code	Site	Area	Geographical distribution	Relief	Altitude (m)
BM	Beni M'ttir	Ain Drahem	Western North (Kroumirie)	Mountain	800
HB	Hammam Bourguiba	Ain Drahem	Western North (Kroumirie)	Mountain	570
KR	Keff El Rand	Haouaria	North East (Cap Bon)	Mountain	642
HJ	Hammam Jdidi	Hammamet	North East (Cap Bon)	Mountain	174
DK	Djebel Khroufa	Tabarka	Western North (Mogods)	Tell Atlas	160
B	Bellif	Tabarka	Western North (Mogods)	Tell Atlas	88
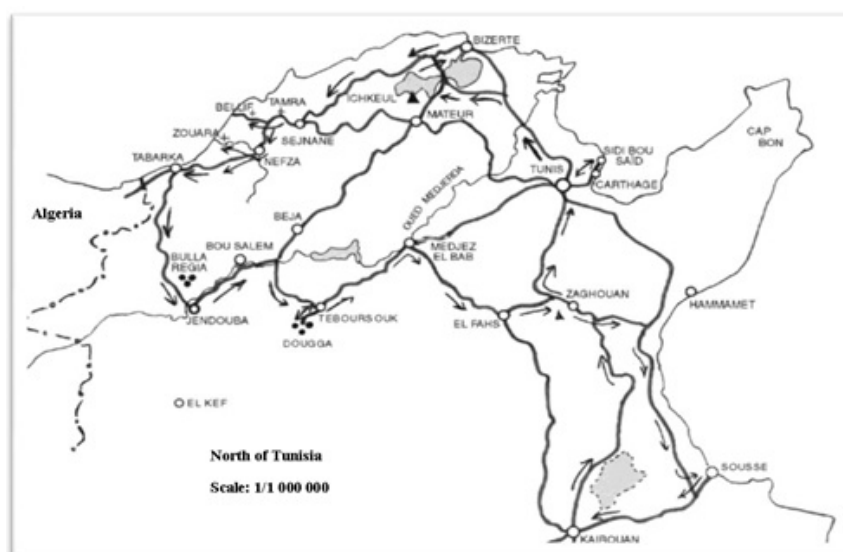
Figure 1. Tunisian map showing the geographic origin of *Quercus suber* L. in North Tunisia.

DNA preparation

Total cellular DNA was extracted from frozen leaves according the procedure (Qiacube Qiagen) described by Bernatzky and Tanksley (1986). DNA quality was examined by electrophoresis on 0.8% agarose gel. DNA concentration was estimated using a Gene Quant minispectrophotometry (Qubit) at 260 nm and DNA was stored at -4°C until use in PCR.

Amplification and sequencing

The cpDNA spacer region between the *trn*L (UAA) - *trn*F (GAA) and the *trn*L (UAA) introns was amplified from total genomic DNA as a template using PCR. Amplification of the *trn*L intron and the adjacent *trn*L-*trn*F spacer was achieved using the forward and reverse primers designed by Taberlet et al. (1991). The primer pairs, denoted 'c' (5'-CGAAATCGGTAGACGCTACG-3'), 'd' (5'-GGGGATAGAGGGACTTGAAC-3') 'e' (5'-GGTTCAAGTCCCTCTATCCC-3'), and 'f' (5'-ATTTGAACTGGTGACACGAG-3'), were used to amplify the intron of *trn*L (UAA) 5'-exon, and the *trn*L (UAA) 3'-exon (i.e., the intron *trn*L) and the spacer between the *trn*L (UAA) 3'-exon and the *trn*F (GAA) exon (i.e., the *trn*L-*trn*F spacer), respectively. The two non-coding regions of the cpDNA were amplified by PCR in a DNA thermocycler (Gene Amp® PCR System 9700: PE Applied Biosystem). Conditions for PCR amplification were as follows: 25 mM $MgCl_2$, 2 mM dNTP mix, 1 µM each primer, and 1 U DNA Taq polymerase (Fermentas) with 20 ng DNA in a final reaction volume of 25 µL. Cycling conditions were: 94°C for 4 min as an initial denaturation step before entering 35 cycles each composed of 1 min at 94°C, 1 min at 51°C, and 2 min at 72°C. A final extension step of 10 min at 72°C was usually programmed at the last cycle. Agarose gel electrophoresis (1.5%) and ethidium bromide staining were used to check the PCR products. The sequences obtained were published in GenBank with accession Nos. KR780486 to KR780545 for the intergenic spacer and KP693717 to KP693776 for the intron (http://www.ncbi.nih.gov/) (Table 2).

Sequence analysis

Sequences were aligned using the ClustalW package and analyzed with MEGA program version 6 (Tamura et al., 2011). Aligned sequences were analyzed with the DnaSP software version 5.10.01 (Librado and Rozas, 2009) to estimate polymorphism indices, molecular evolution, and past demographic history. Indices of haplotype diversity (Hd) (Nei and Tajima, 1983) and pairwise estimates of nucleotide diversity (Pi) (Jukes and Cantor, 1969) were used to evaluate genetic diversity among *Q. suber* populations. The average pairwise nucleotide differences (K), Theta (per site), and the minimum number of recombination events (Rm) were also estimated. Selective neutrality was tested by both Tajima's D (Tajima, 1989) and Fu and Li's D* and F* methods (Fu and Li, 1993). Demographic parameters were assessed using the distribution of pairwise sequence differences (Mismatch distribution) of Rogers and Harpending (1992) and site-frequency spectra (distribution of the allelic frequency at a site) of Tajima (1989). Fu's *Fs* (Fu, 1997) and Ramos-Onsins and Rozas' *R*2 (Ramos-Onsins and Rozas, 2002) tests for mutation/drift equilibrium were also performed in DnaSP version 5.10.01 with 10,000 simulations (Librado and Rozas, 2009). The overall validity of the estimated demographic model was evaluated by tests of raggedness index r (Harpending, 1994). The significance of r was assessed by parametric bootstraps (10,000 replicates), and the

significant value was taken as evidence for departure from the estimated demographic model of sudden population expansion. For each sequence, length, and contents, the proportion of AT and GC were estimated. The observed ratio of the mutational event was also calculated.

Phylogenetic reconstruction

The alignment was manually checked and pairwise sequence divergence between trees in *trn*L-*trn*F intergenic spacer and the combined *trn*L-F region was calculated according to the maximum composite likelihood (MCL) method (Tamura et al., 2004). The evolutionary history was inferred using the maximum likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). Initial trees for the heuristic search were obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using the MCL approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured by the number of substitutions per site. All positions containing gaps and missing data were eliminated using the MEGA Version 6 software (Tamura et al., 2011). The maximum parsimony (MP) tree was obtained using the Subtree-Pruning-Regrafting algorithm (Nei and Kumar, 2000) with search level 1, in which the initial trees were obtained by the random addition of sequences. To evaluate the relative robustness of the groups found in the relationship analysis, 1000 bootstrap replicates were calculated.

RESULTS

Sequence variation in the *trn*L-*trn*F spacer

The length of the trnL-trnF intergenic spacer ranged from 440 bp for 'BM8-4' to 452 bp for 'KR9-4', 'DK6-4', and 'BM9-4' cork oak with an average length of 448.7 bp. In addition, the AT content of the amplified sequences varied from 66.7 to 69.6% in the *trn*L-*trn*F spacer, and the GC content ranged from 30.4 to 33.3% (Table 2). The *trn*L-*trn*F spacer has undergone single mutation events such as indel (insertion/deletion) and substitution events (transversion and transition). The transition/transversion rate ratios were $K1 = 1.431$ (purines) and $K2 = 1.883$ (pyrimidines). An overall transition/transversion ratio R of 0.724 occurred in the cpDNA *trn*L-*trn*F spacer. This result shows that transversions are more frequent than transitions in the intergenic spacer (Figure 2).

In addition, the nucleotide frequencies were 38.84, 28.44, 15.72, and 16.99% for A, T, C, and G, respectively. The different substitutions detected are presented in Table 3, which shows that in Tunisian cork oak, $A \rightarrow G$ and $T \rightarrow C$ transitions are more frequent than $G \rightarrow A$ and $C \rightarrow T$ transitions. The cpDNA data set for the broad investigation of *Q. suber* individuals consisted of 412 aligned nucleotide positions with 309 conserved sites, 103 variable sites, 67 of which were parsimony informative and 36 of which were singletons.

Sequence variation in the combined sequence of the *trn*L intron and *trn*L-*trn*F spacer

The length of the combined region ranged from 932 bp for 'BM9' to 972 bp for 'B3', 'BM10', 'DK6', 'HJ5', and 'KR1' trees, among the analyzed cork oaks. In addition, the AT content of the amplified sequences varied from 67.0 to 68.5% in the combined region and the GC content ranged from 31.5 to 33.0% (Table 2).

Table 2. Localities, accession numbers, variations in length, GC and AT contents of the *trn*L-*trn*F spacer and combined region, *trn*L-*trn*F spacer and *trn*L intron in the studied *Quercus* trees.

Relief	Code	Locality	Accession No.	*trn*L-*trn*F spacer			Combined: *trn*L-*trn*F spacer and *trn*L intron		
				%AT	%GC	Length (bp)	%AT	%GC	Length (bp)
Tell Atlas	B2-4	Nefza	KR780487	67.6	32.4	450.0	67.3	32.7	970.0
Tell Atlas	B3-4	Nefza	KR780488	67.4	32.6	451.0	67.4	32.6	972.0
Tell Atlas	B4-4	Nefza	KR780489	68.6	31.4	449.0	67.8	32.2	966.0
Tell Atlas	B5-4	Nefza	KR780490	68.4	31.6	449.0	67.8	32.2	968.0
Tell Atlas	B6-4	Nefza	KR780491	67.6	32.4	448.0	67.3	32.7	967.0
Mountain	BM2-4	Ain Drahem	KR780497	67.7	32.3	449.0	67.4	32.6	969.0
Mountain	BM3-4	Ain Drahem	KR780498	67.8	32.2	447.0	67.5	32.5	968.0
Mountain	BM4-4	Ain Drahem	KR780499	68.0	32.0	447.0	67.4	32.6	959.0
Mountain	BM5-4	Ain Drahem	KR780500	68.0	32.0	447.0	67.5	32.5	967.0
Mountain	BM6-4	Ain Drahem	KR780501	67.6	32.4	448.0	67.2	32.8	964.0
Mountain	BM7-4	Ain Drahem	KR780502	67.6	32.4	451.0	67.4	32.6	970.0
Mountain	BM8-4	Ain Drahem	KR780503	67.5	32.5	440.0	67.3	32.7	956.0
Mountain	BM9-4	Ain Drahem	KR780504	68.1	31.9	452.0	68.5	31.5	932.0
Mountain	BM10-4	Ain Drahem	KR780505	67.8	32.2	451.0	67.5	32.5	972.0
Tell Atlas	DK1-4	Tabarka	KR780506	68.3	31.7	448.0	67.8	32.2	968.0
Tell Atlas	DK2-4	Tabarka	KR780507	68.1	31.9	448.0	67.8	32.2	971.0
Tell Atlas	DK3-4	Tabarka	KR780508	69.6	30.4	448.0	68.2	31.8	969.0
Tell Atlas	B1-4	Nefza	KR780486	67.4	32.6	448.0	67.1	32.9	968.0
Tell Atlas	B9-4	Nefza	KR780494	68.1	31.9	448.0	67.5	32.5	968.0
Tell Atlas	DK6-4	Tabarka	KR780511	68.6	31.4	452.0	67.7	32.3	972.0
Tell Atlas	DK7-4	Tabarka	KR780512	68.1	31.9	451.0	67.5	32.5	971.0
Tell Atlas	DK8-4	Tabarka	KR780513	68.5	31.5	448.0	67.8	32.2	968.0
Tell Atlas	DK5-4	Tabarka	KR780510	67.5	32.5	449.0	67.3	32.7	969.0
Tell Atlas	DK9-4	Tabarka	KR780514	67.8	32.2	451.0	67.5	32.5	971.0
Tell Atlas	B8-4	Nefza	KR780493	68.1	31.9	451.0	67.7	32.3	971.0
Tell Atlas	B10-4	Nefza	KR780495	68.1	31.9	445.0	67.7	32.3	965.0
Tell Atlas	DK10-4	Tabarka	KR780515	67.4	32.6	451.0	67.4	32.6	971.0
Mountain	HB1-4	Ain Drahem	KR780516	67.6	32.4	448.0	67.4	32.6	969.0
Mountain	HB2-4	Ain Drahem	KR780517	67.3	32.7	450.0	67.3	32.7	962.0
Mountain	HB3-4	Ain Drahem	KR780518	67.8	32.2	450.0	67.7	32.3	961.0
Mountain	HB4-4	Ain Drahem	KR780519	67.4	32.6	448.0	67.3	32.7	967.0
Mountain	HB5-4	Ain Drahem	KR780520	67.6	32.4	450.0	67.0	33.0	962.0
Mountain	HB6-4	Ain Drahem	KR780521	67.6	32.4	447.0	67.4	32.6	966.0
Mountain	HB7-4	Ain Drahem	KR780522	68.8	31.2	449.0	67.9	32.1	968.0
Mountain	HB8-4	Ain Drahem	KR780523	67.5	32.5	449.0	67.5	32.5	960.0
Mountain	HB9-4	Ain Drahem	KR780524	67.9	32.1	448.0	67.3	32.7	964.0
Tell Atlas	B7-4	Ain Drahem	KR780492	67.9	32.1	448.0	67.5	32.5	968.0
Tell Atlas	BM1-4	Ain Drahem	KR780496	68.2	31.8	443.0	67.6	32.4	963.0
Tell Atlas	DK4-4	Tabarka	KR780509	67.6	32.4	450.0	67.3	32.7	970.0
Mountain	HB10-4	Ain Drahem	KR780525	67.7	32.3	449.0	67.4	32.6	969.0
Mountain	HJ2-4	Hammamet	KR780527	67.9	32.1	448.0	67.6	32.4	968.0
Mountain	HJ3-4	Hammamet	KR780528	67.1	32.9	450.0	67.2	32.8	970.0
Mountain	HJ4-4	Hammamet	KR780529	68.1	31.9	448.0	67.7	32.3	968.0
Mountain	HJ5-4	Hammamet	KR780530	66.7	33.3	451.0	67.0	33.0	972.0
Mountain	HJ6-4	Hammamet	KR780531	68.0	32.0	450.0	67.7	32.3	969.0
Mountain	HJ7-4	Hammamet	KR780532	67.6	32.4	447.0	67.5	32.5	969.0
Mountain	HJ8-4	Hammamet	KR780533	67.9	32.1	449.0	67.7	32.3	966.0
Mountain	HJ9-4	Hammamet	KR780534	67.7	32.3	449.0	67.5	32.5	969.0
Mountain	HJ10-4	Hammamet	KR780535	67.8	32.2	447.0	67.7	32.3	968.0
Mountain	KR1-4	Haouaria	KR780536	68.1	31.9	451.0	67.6	32.4	972.0
Mountain	KR2-4	Haouaria	KR780537	68.3	31.7	445.0	67.7	32.3	957.0
Mountain	KR3-4	Haouaria	KR780538	68.4	31.6	446.0	67.9	32.1	965.0
Mountain	KR4-4	Haouaria	KR780539	67.6	32.4	450.0	67.5	32.5	969.0
Mountain	KR5-4	Haouaria	KR780540	68.6	31.4	446.0	67.8	32.2	965.0
Mountain	HJ1-4	Hammamet	KR780526	67.8	32.2	450.0	67.7	32.3	970.0
Mountain	KR7-4	Haouaria	KR780542	68.0	32.0	450.0	67.8	32.2	970.0
Mountain	KR6-4	Haouaria	KR780541	68.5	31.5	448.0	67.9	32.1	967.0
Mountain	KR8-4	Haouaria	KR780543	68.1	31.9	451.0	67.7	32.3	970.0
Mountain	KR9-4	Haouaria	KR780544	68.4	31.6	452.0	67.8	32.2	971.0
Mountain	KR10-4	Haouaria	KR780545	67.8	32.2	450.0	67.6	32.4	970.0
Average				67.9	32.1	448.7	67.6	32.4	966.9
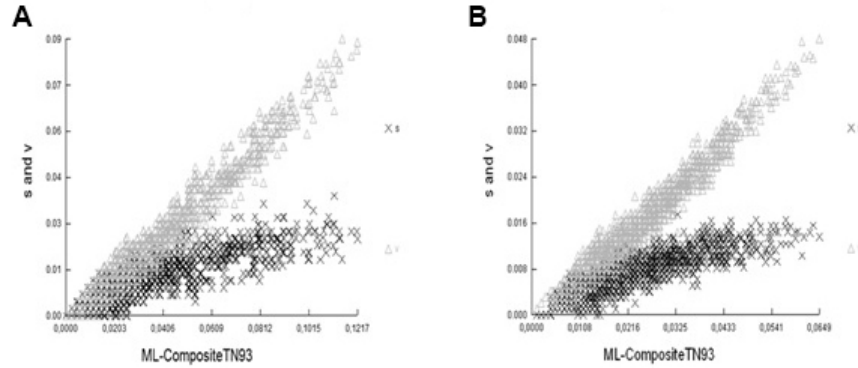
Figure 2. Maximum composite likelihood divergence plotted against the number of transitions (s: crosses) and transversions (v: triangles) for pairwise comparisons of the *trn*L-*trn*F spacer (**A**) and the combined region (*trn*L-*trn*F spacer and *trn*L intron) (**B**).

Table 3. Composite likelihood estimate of the pattern of nucleotide substitution in the *trn*L-*trn*F spacer and the combined sequences *trn*L-*trn*F spacer and *trn*L intron.

*trn*L-*trn*F spacer	A	T	C	G
A	-	7.83	4.33	6.7
T	10.7	-	8.15	4.68
C	10.7	14.75	-	4.68
G	15.31	7.83	4.33	-
Combined sequences				
A	-	13.45	6.09	6.26
T	12.79	-	1.18	6.21
C	12.79	2.6	-	6.21
G	12.89	13.45	6.09	-

The transition/transversion ratios were K1 = 1.008 (purines) and K2 = 0.194 (pyrimidines). The overall transition/transversion ratio, R, was of 0.258 in the combined region. The rate of the transitional substitutions (A/G: 12.89, T/C: 2.6, C/T: 1.18, and G/A: 6.26) were lower than the rates of transversions (A/T: 12.79, A/C: 12.79, T/A: 13.45, T/G: 13.45, C/A: 6.09, C/G: 6.09, G/T: 6.21, and G/C: 6.21) (Table 3), which confirms the high AT content. This may explain the high proportion of identified transversions (Figure 2). Of the 102 variable sites, 70 were parsimony-informative characters and 32 were singleton variable sites in *Quercus* sequences (Table 4). The nucleotide frequencies were 33.19, 34.90, 15.80, and 16.12% for A, T, C, and G, respectively.

Table 4. Comparison of data from the *trn*L-*trn*F spacer and the combined region of the *trn*L-*trn*F spacer and the *trn*L intron.

	Total informative	No. of tree	Length	CI	RI	HI	BS ≥ 50%
*trn*L-*trn*F	412/67 = 6.14	137	171	0.888	0.875	0.112	11
Combined	856/70 = 12.22	110	251	0.548	0.751	0.452	14

Total informative = total of characters (excluding gaps)/number of informative characters; CI = consistency index; RI = retention index; HI = homoplasy index; BS = number of branches that contained bootstrap support close to 50%.

Genetic diversity and differentiation

A high level of variation in cpDNA was detected in cork oak. Sequence alignment permitted the identification of 57 cpDNA haplotypes among 60 examined sequences for the intergenic spacer and 48 for the combined sequences. Haplotype diversity (Hd) values of 0.998 for the *trn*L-*trn*F spacer and 0.989 for the combined sequences were obtained in our data set (Table 5). Nucleotide diversity of cork oak was 0.03152 (3.152%) for the *trn*L-*trn*F spacer, and 0.01210 (1.210%) for the combined sequences (Table 5). Theta values (per site) and the average number of pairwise differences (K) were 0.05778 and 12.986, respectively, for the *trn*L-*trn*F spacer, and 0.02656 and 10.356, respectively, for the combined region (Table 5). The combined region *trn*L-*trn*F showed more polymorphic sites than the *trn*L-*trn*F spacer. Hence, it was more informative and proved to be very effective in differentiating the cork oak populations. Moreover, the *trn*L intron has been used successfully in Angiosperms (Nei and Kumar, 2000). The nucleotide diversity (Pi) and segregating sites (S), illustrated in Figure 3a and b, show the occurrence of substitutions along the analyzed sequences. A high level of nucleotide diversity was observed in the *trn*L-*trn*F spacer, and in the combined sequences. This illustrates the levels of polymorphism in Tunisian *Q. suber* and demonstrates the utility of the *trn*L-*trn*F spacer and the pooled sequences of the *trn*L intron and *trn*L-*trn*F spacer to reveal divergence at the intra-specific level. A 100-bp sliding window shows the distribution of nucleotide substitution diversity and segregating sites over the cpDNA region assayed.

Table 5. Polymorphism parameters and Tajima's neutrality tests for 60 sequences of the *trn*L-*trn*F spacer, and the combined region of the Tunisian cork oak accessions used in this study.
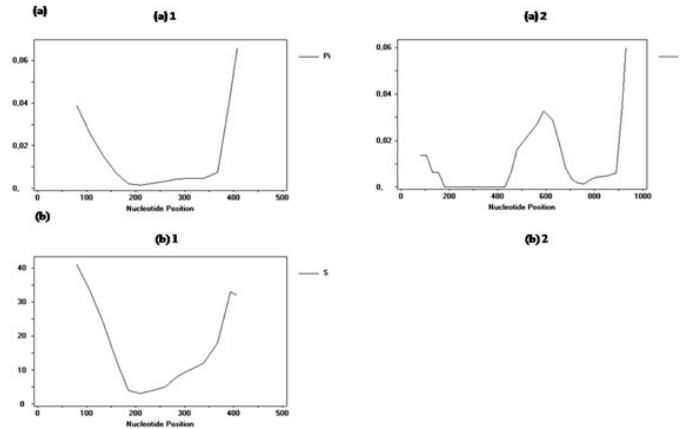
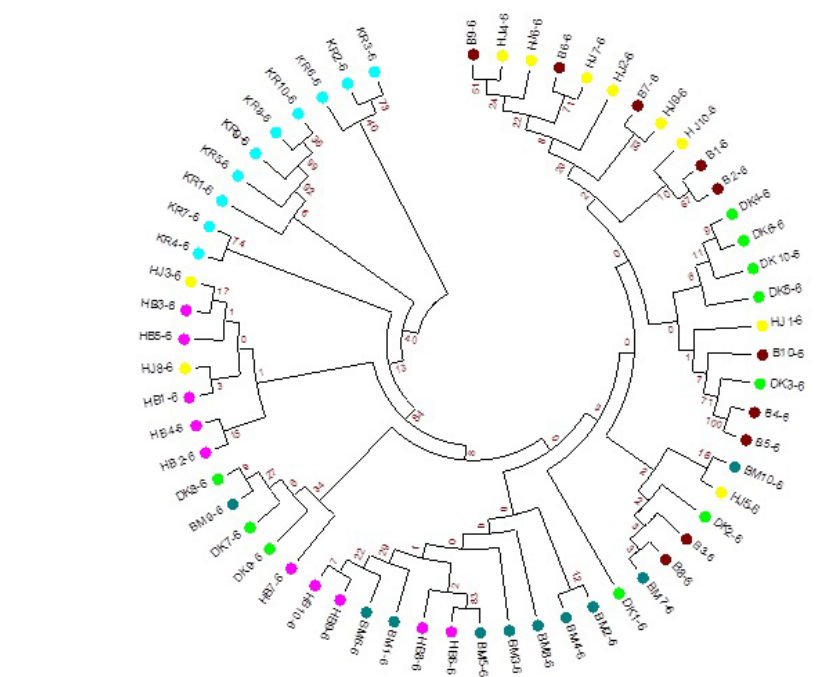|	*trn*L-*trn*F spacer	Combined: *trn*L-*trn*F spacer and *trn*L intron
Number of sequences	60	60
Alignment length (bp)	454	980
Monomorphic characters	309	754
Variable characters	103	102
Parsimony informative characters	67	70
Singleton variable sites	36	32
Total number of mutation	111	106
Number of polymorphic sites (S)	103	102
Number of haplotypes (H)	57	48
Haplotype diversity (Hd) ± SD	0.998 ± 0.003	0.989 ± 0.006
Variance of haplotype diversity	0.00001	0.00004
Nucleotide diversity (Pi)	0.03152	0.01210
Theta (per site)	0.05778	0.02656
Average of pairwise differences (K)	12.986	10.356
Minimum number of recombination events (Rm)	14	13
Tajima's D	-1.58503 NS (0.10 > P > 0.05)	-1.89622 *(P < 0.05)
Fu and Li's D*	-1.44212 NS (P > 0.10)	-1.22759 NS (P > 0.10)
Fu and Li's F*	-1.79379 NS (P > 0.10)	-1.78053 NS (P > 0.10)
Fu's *Fs* statistic	-52,249 (P = 0.000)	-33.722 (P = 0.000)

Genetic relationships

MP using the trnL-trnF spacer

Relationships between cork oak accessions were examined taking into account variation in the intergenic *trn*L-*trn*F spacer. To clarify the genetic relationships between haplotypes, the dendrogram was constructed using MP. Haplotype parsimony analysis

identified 240 parsimonious trees. Our analysis yielded one most parsimonious tree, 240 steps long, with a consistency index (CI) of 0.446, a retention index (RI) of 0.745, and a homoplasy index (HI) of 0.554 (Table 4 and Figure 4). Consequently, all data sets exhibited low homoplasy levels, demonstrating the utility of the *trn*L-*trn*F spacer region to elucidate genetic relationships among *Q. suber* populations (Figure 4).


Figure 3. Base-pair sliding window of the *trn*L-*trn*F spacer (a1, b1) and the combined region: *trn*L-*trn*F spacer and *trn*L intron (a2, b2) of cpDNA of *Quercus suber* L. **a.** Variability of the nucleotide diversity (Pi). **b.** Segregating sites (S).

Figure 4. Strict consensus tree of 140 most parsimonious trees (length = 240 steps, CI = 0. 446; RI = 0. 745) obtained from the *trn*L-*trn*F spacer data from Tunisian cork oak. Numbers below nodes represent bootstrap values.

MP using combined sequences

Parsimony analysis identified 167 parsimonious trees. Our analysis yielded one most parsimonious tree, 167 steps long, with CI of 0.548, RI of 0.751, and HI of 0.452 (Table 4 and Figure 5). Consequently, all data sets showed low homoplasy levels, demonstrating the utility of the combined sequences to elucidate genetic relationships among *Q. suber* populations (Table 4).
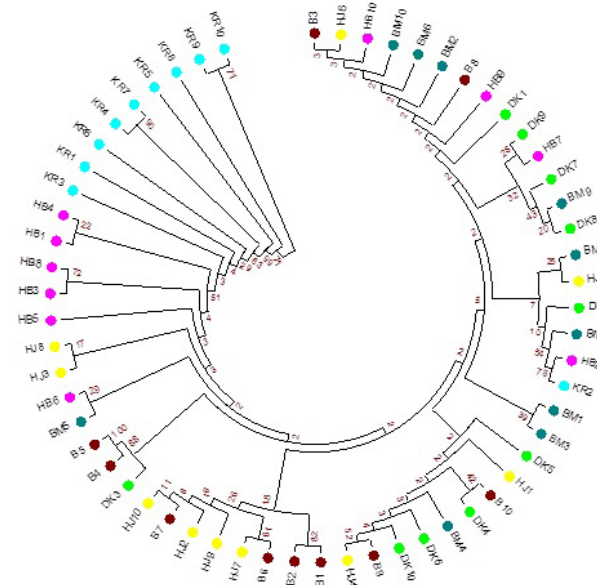

Figure 5. Strict consensus tree of 167 most parsimonious trees (length = 167 steps, CI = 0.548, RI = 0.751) obtained from the combined *trn*L-*trn*F region data from Tunisian cork oak. Numbers below nodes represent bootstrap values.

Genetic relationships using the trnL-trnF spacer

The pairwise sequence divergence ranged from 0.000 to 1.107, with an average of 0.025 indicating polymorphism among *Q. suber* populations. The low level of plastid DNA variation exhibited by *Q. suber* populations is probably related to its slow rate of evolution. The minimum distance (0.000) was observed between 'Bellif' (3-6), 'Béni Mtir' (2-6; 3-6; 6-6; 7-6; 9-6), 'Djebel Khroufa' (2-6; 4-6; 5-6; 7-6; 10-6), 'Hammam Bourguiba' (8-6; 9-6; 10-6), 'Keff El Rand' (1-6; 6-6), and 'Hammam Jdidi' (1-6) accessions, and the maximum distance (1.107) was observed between 'Bellif' (4-6) and 'Keff El Rand' (8-6; 10-6) population suggesting the important dissimilarity between sequences of the *trn*L-*trn*F spacer. To elucidate genetic relationships among Tunisian *Q. suber* populations, a dendrogram was constructed using the ML method. The dendrogram classified the 60 individuals studied into two main clusters (BS = 77) (Figure 6A). The first group designated I, included 'DK3-6', 'B4-6', and 'B5-6'. All the remaining accessions are clustered in the second group (II). These accessions showed a high level of similarity between each other. In fact, it is important to note that *Q. suber* individuals are distributed independently to the geographical origin, the bioclimatic stage, and the relief of populations.

Of note, the present study has shown that a sequencing approach is more efficient at examining the phylogenetic relationships in this species, and may be extended to other populations of Tunisian *Q. suber*. Mediterranean evergreen *Quercus* species are a group with overlapping habitats, which often leads to their consideration as a homogeneous entity in botanical, biogeographical, or paleo-historical studies. In the Western Mediterranean Basin, *Q. suber* (cork oak), *Q. ilex* (holm oak), and *Q. coccifera* (kermes oak) are the dominant broad-leaved species. The three species are sympatric in many areas, but have some different evolutionary histories. Previous studies have shown differences in their genetic variation patterns at both nuclear and cytoplasmic levels (Lumaret et al., 2005).
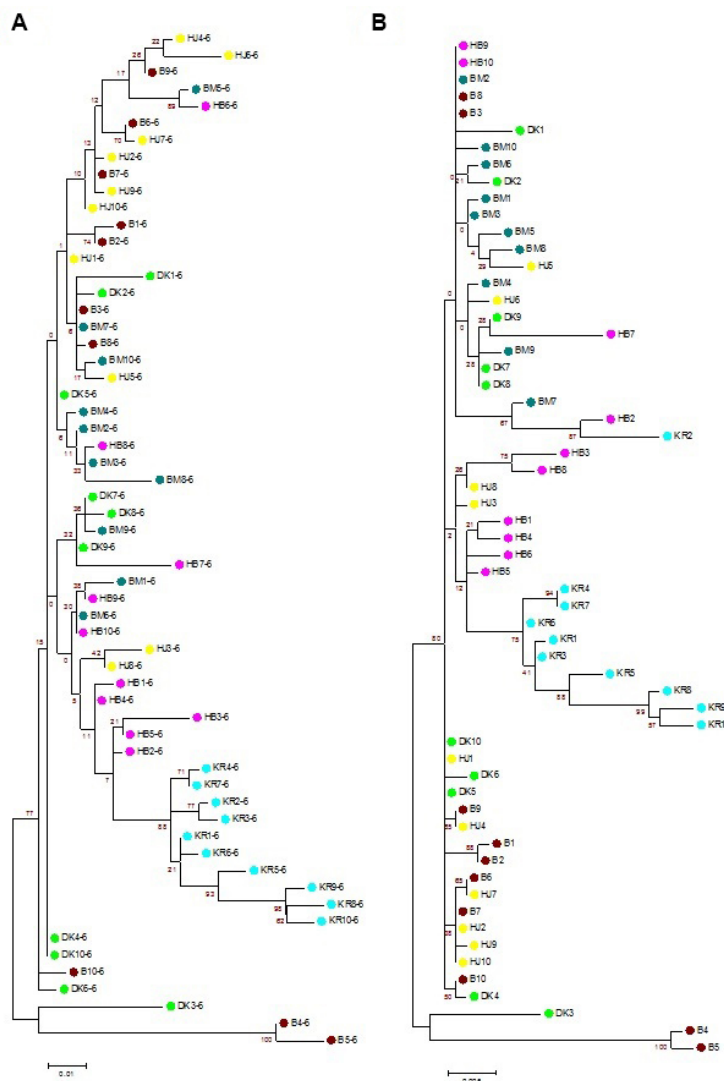
Figure 6. Maximum-likelihood tree obtained from the *trn*L-*trn*F spacer (**A**) and the combined region (*trn*L-*trn*F spacer and *trn*L intron) (**B**) represents genetic relationships between accessions of Tunisian cork oak.

Genetic relationships using the combined sequences

Pairwise sequence divergence ranged from 0.000 to 0.056 with an average of 0.012. The ML dendrogram showed two main groups (BS = 80) (Figure 6B). The first group consisted of 'DK3', 'B4', and 'B5' oaks, which differs from the other trees. All the remaining accessions clustered in the second group. However, the dendrogram topology of the *trn*L-*trn*F spacer is congruent with data provided by the combined sequences. Moreover, the combined sequences showed the same informative characters with the spacer. In our study, the combined data provided the same supported resolution with the intergenic spacer.

Molecular evolution

Tajima's and Fu and Li's tests

The selective neutrality of the detected mutations was tested using the methods of Tajima (1989) and Fu and Li (1993) to examine the null hypothesis. Selective neutrality tests show that both tests were negative and not significant in the total studied sample, with the exception that the Tajima test was negative and significant for the combined sequences (Table 5). These parameters reject neutrality in the regions studied. The deviation from selective neutrality, as evidenced by Tajima tests, is explained by an excess of rare mutations as identified in the studied sequences. The excess of low frequency polymorphism can be explained for *Quercus* populations in Tunisia.

Fu's Fs and R2 statistics

To describe the cause of the deviation from neutrality, the R2 and Fu's *Fs* statistic were estimated. For cork oak trees, Fu's *Fs* were negative (Fu's *Fs* = -52,249 for the *trn*L-*trn*F spacer, and Fu's *Fs* = -33.722, for the collective sequences). Conversely, R2 showed positive values (R2 = 0.0597 for the *trn*L-*trn*F spacer, and R2 = 0.0477, for the pooled sequences). The results reject a model of constant population size, and the observed patterns of variation provide evidence that cork oak trees have been undergoing rapid expansion. Fu's statistic is more powerful at detecting the deviation from neutrality and thereby attesting population expansion or a genetic hitchhiking in the data, which produces a similar pattern like that of recent population expansion. Our results demonstrate that the genetic effect of hitchhiking was ruled out by the Fu test, since the observed values were negative (Fu's *Fs* = -52,249; P = 0.000 for the intergenic spacer, and Fu's *Fs* = -33.722; P = 0.000 for the combined sequences, respectively) (Table 5 and Figure 7).

Tajima's and Fu and Li's parameters do not confirm the presence of background selection in the analyzed region, and Fu's *Fs* provides evidence for a recent population expansion of Tunisian cork oak trees. Observed variation in the nucleotide diversity (Pi) and the segregating sites (S), the footprint of selection, which reduced diversity, was limited to 170 bp for the *trn*L-*trn*F spacer and to 220 bp for the *trn*L intron in the combined data set (Figure 3). The recombination rate is also an important parameter affecting patterns of DNA polymorphism. The Rm values recorded for the *trn*L-*trn*F spacer and the *trn*L-*trn*F combined region in *Quercus* species were 14 and 13, respectively (Table 5). Recombination has been detected between sites: (5,6) (6,14) (20,21) (21,25) (48,62) (67,76) (122,175) (175,420)

(420,424) (425,426) (435,436) (436,437) (437,438), and (438,445) for the *trn*L-*trn*F spacer and between sites, and (59,110) (543,568) (574,588) (648,701) (701,933) (933,946) (946,950) (950,956) (956,960) (961,962) (962,963) (963,964), and (967,971) for the *trn*L-*trn*F combined region. The recombination has a greater role in generate variation and there seems to be a high positive correlation between the rate of recombination and genetic diversity. Conversely, in *Ficus carica*, only four recombination events were observed in the *trn*L intron, five in the segregated sequences and one in the *trn*L-*trn*F spacer (Baraket et al., 2010). In date palm (*Phoenix dactylefera* L.), two recombination events were detected in the spacer and three in the segregated sequences (Sakka et al., 2013).
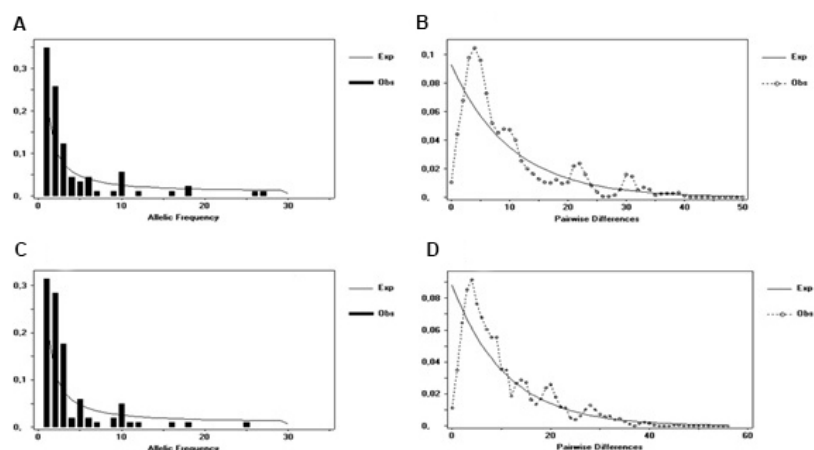
Figure 7. A and **C.** Site-frequency spectra of the cpDNA sequences in *Quercus suber* L. covered in this study for the *trn*L-*trn*F spacer and combined *trn*L-*trn*F spacer and *trn*L intron. Solid lines on the site-frequency spectra indicate the expected distributions under neutrality and at equilibrium. Fu's *Fs* statistics and corresponding P values are given. **B** and **D.** Mismatch distribution of the *trn*L-*trn*F spacer and combined *trn*L-*trn*F spacer and *trn*L intron of *Q. suber* L. individuals based on pairwise nucleotide differences in the *trn*L-*trn*F spacer and combined *trn*L-*trn*F spacer and *trn*L intron. Solid lines in the curves indicate the expected distribution under expansion and dotted lines the observed distribution under population expansion. The raggedness statistic, r, and corresponding P value are given.

Demographic history

Changes in population size leave particular footprints that may eventually be detected in DNA sequence data (Tajima, 1989; Rogers and Harpending, 1992). The empirical distribution of the pairwise number of nucleotide differences and segregating sites among pairs of individuals shows a skew in the mismatch distribution that appears uni-modal for the spacer, and the aggregated sequences (Figure 7). The observed uni-modal mismatch distribution and the low Harpending's raggedness index (r = 0.0019 for the *trn*L-*trn*F spacer and r = 0.0034 for the pooled sequences) (Figure 7) calculated for cork oak cpDNA haplotypes confirm the model of recent expansion. Usually, average values of Tajima's D found in most plant species studied so far are likely to represent the results of demographic processes (such as expanding populations, introgression, extinction, and recolonization events) and/or weak selection, which allows deleterious variants to give rise to a low population frequency (Wachowiak et al., 2009).

Linkage disequilibrium

To further examine whether the genetic hitchhiking acts in the *trn*L-*trn*F region, linkage disequilibrium was estimated among sites. The spacer has lower nucleotide diversity and lower widespread linkage disequilibrium compared to the *trn*L intron. Figure 8 shows the increase in linkage disequilibrium in the *trn*L intron. We noted a significant reduction in the level of linkage disequilibrium at the spacer-linked site (Figure 8), consistent with sweep selection and the lacking of a hitchhiking genetic effect.
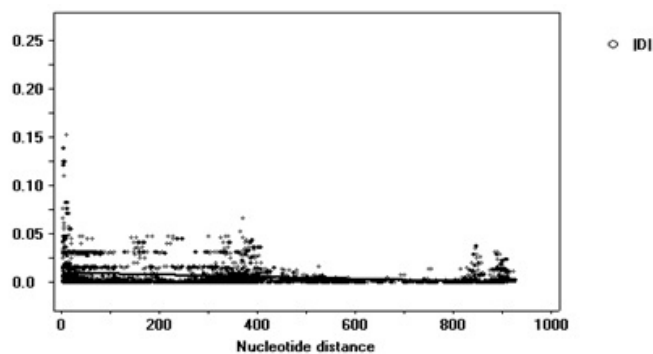
Figure 8. Linkage disequilibrium estimated from the pooled data for the *trn*L intron and *trn*L-*trn*F intergenic spacer of Tunisian *Quercus suber* L.

DISCUSSION

In this study, the *trn*L-*trn*F spacer and the combined region (*trn*L intron + *trn*L-*trn*F spacer) of cpDNA were used to study the genetic diversity and molecular evolution of cork oak populations, the establishment of relationships, and to elaborate on a scenario for the origin of the Tunisian *Q. suber* forest. The sequence analysis showed variations in both the lengths and nucleotide compositions for all cork individuals. The average size of the *trn*L-*trn*F spacer of the *Q. suber* tree was 448.7 bp. Similar results have been reported in *P. dactylefera* (364-397 bp) (Sakka et al., 2013) and in other plants such as the *Pinus* genus (467-471 bp) (Chen et al., 2002). The nucleotide composition of the *trn*L-*trn*F spacer in the Tunisian cork oak (AT = 67.9%; GC = 32.1%) is similar to that observed in the total *trn*L-*trn*F region of Angiosperm species (*Actaea*, *Digitalis*, *Drosera*) (Bakker et al., 2000), in *Ficus carica* (Baraket et al., 2010), in *P. dactylefera* (Sakka et al., 2013), and in 19 other cpDNA non-coding regions in Poaceae (Chen et al., 2002) and, in Rubiaceae, for the *atp*B-*rbc*L (Manen and Natali, 1995). In fact, the GC contents varied from 28 to 32% for the *psb*A-*trn*H intergenic spacer of date palm, from 39 to 41% for the *rpo*B spacer as reported by Al-Qurainy et al. (2011) and from 33 to 34.4% for the *trn*L-*trn*F spacer as reported by Sakka et al. (2013). The plastome genome was AT rich (63%) in *Q. suber*.

The transition/transversion (ti/tv) ratio was estimated, and was similar to values previously reported for the *trn*L-*trn*F spacer of cpDNA among *Fagus sylvatica*, *Q. macrolepis*, *Q. trojana*, *Q. ilex*, *Q. suber*, *Q. pubescens*, and *Q. robur*, horny oak, and *Q. ilex* (ti/tv = 0.5) (Paffetti et al., 2001), in *Ficus carica* L. (ti/tv = 0.9) (Baraket et al., 2010), and in other Angiosperms for the *trn*L-F and the *atp*B-*rbc*L intergenic spacers (Manen and Natali, 1995;

Bakker et al., 2000), and in cork oak (ti/tv = 7.60) (Cosimo et al., 2009). However, this differs from the values observed in *P. dactylefera* (ti/tv = 2.19) (Sakka et al., 2013).

Nucleotide frequencies were calculated and a similar composition has been reported in cork oak, in which the base frequencies were A = 18.33%, C = 33.66%, G = 30.71%, and T = 17.30% (Cosimo et al., 2009).

The numbers of conserved and variable sites determined in this study were found to be similar to values observed for the *trn*L-*trn*F intergenic region of six additional *Quercus* species: *Q. suber* L., *Q. trojana* Webb, *Q. macrolepis* Kotschy, *Q. ilex* L., *Q. coccijera* L., and *Q. calliprinos*. Different results have been reported in the *Quercus* genus, for which the data matrix listed 53 taxa (with four outgroups) and 595 characters (545 constant, 21 variable but parsimony-uninformative and 29 parsimony-informative). Six characters (four parsimony-informative) derived from indel coding were added to the MP analysis (heuristic search and bootstrap) (Cosimo et al., 2009). In the second part, the average size of the *trn*L-*trn*F region of the Tunisian *Q. suber* tree was estimated at 966.9 bp. Similar results have been reported for *P. dactylefera* (897-981 bp) (Sakka et al., 2013), *Ficus carica* (989-1022 bp) (Baraket et al., 2010), and *Q. rubra* (902 bp) (Gielly and Taberlet, 1994). The nucleotide composition in the pooled region of the Tunisian cork oak (AT = 67.6%; GC = 32.4%) was similar to that estimated for the combined sequences in *Q. rubra* by Gielly and Taberlet (1994) and in *F. carica* by Baraket et al. (2010). The ti/tv ratio was estimated (0.258) in the combined region. Conversely, in other studies, high levels of ti/tv have been found in *P. dactylefera* (ti/tv = 1.282) (Sakka et al., 2013), which show that transitions are slightly more common compared to transversions in the combined *trn*L-*trn*F spacer and *trn*L intron. We interpret these values as indicating different rates of transition and transversion. This result is consistent with findings in Angiosperm species, with a ti/tv ratio not exceeding 1 for any of the examined groups (Bakker et al., 2000). The haplotype diversity for the *trn*L-*trn*F spacer (Hd = 0.998) and for the combined sequences (Hd = 0.989) was estimated in our data set (Table 5). The *trn*L-*trn*F spacer and the *trn*L-*trn*F combined region of *P. dactylefera* showed the highest haplotype diversity of 0.848 and 0.955, respectively (Sakka et al., 2013). Nucleotide diversity of cork oak was 0.03152 (3.152%) for the *trn*L-*trn*F spacer, and 0.01210 (1.210%) for the combined sequences (Table 5). However, high nucleotide diversity has been found in other chloroplast non-coding regions of date palm such as *psb*A-*trn*H, *rpo*B genes (6.86 and 2.76%, respectively) (Manen and Natali, 1995), the *trn*L-*trn*F spacer (1.22%), and for the *trn*L-*tr*nF region (1.19%) (Sakka et al., 2013). This is higher than the nucleotide diversity obtained for the accessions of *Fagopyrum cymosum* in the three non-coding regions of the *rbc*L-*acc*D and spacer region (0.00260); *trn*C-*rpo*B (0.00311) and *trn*K-*mat*K (0.00304), and similar to the nucleotide diversity obtained for the accessions of *P. dactylefera* in the *trn*L-*trn*F spacer and the *trn*L-*trn*F combined region (Pi = 0.01220 and Pi = 0.01190). The Theta (per site) and average number of pairwise differences (K) were 0.05778 and 12.986, respectively, for the *trn*L-*trn*F spacer, and 0.02656 and 10.356, respectively, for the combined region (Table 5). However, high values for average pairwise differences (K) have been found in other chloroplast non-coding regions of date palm such as *psb*A-*trn*H and *rpo*B (45.893 and 11.321, respectively) (Al-Qurainy et al., 2011), of *P. dactylefera* for the *trn*L-*trn*F spacer (0.01870 and 3.697, respectively), and for the *trn*L-*trn*F region (0.01940 and 8.803, respectively) (Sakka et al., 2013).

Genetic relationships of the *trn*L-*trn*F spacer and combined region show that the distribution of *Q. suber* individuals is independent of the geographical origin, the bioclimatic stage, and the relief of populations. The MP tree, with an RI index of 0.745 for the *trn*L-*trn*F

spacer and of 0.751 for the pooled sequences, indicated minimal homoplasy within the data set. Furthermore, our results demonstrate that the *trn*L-trnF non-coding region is the location of numerous substitutions. Similar results were obtained by Cosimo et al. (2009) and show that the MP analysis produced 13,320 most parsimonious trees with 85 steps in length, CI of 0.682, and RI of 0.866.

Both Tajima's D and Fu's *Fs* tests revealed negative values, but only Fu's *Fs* was statistically significant from intron and spacer sequences (Fs = -52.249, P = 0.000 for the intron, and Fs = -33.722, P = 0.000 for the spacer). However, only Tajima's D tests revealed negative and significant values for the pooled sequences (D* = -1.89622; P < 0.05). This could be due to Fu's *Fs* test having more power than Tajima's D test to detect population expansion (Tamura and Nei, 1993). Neutrality tests reflect the degree of active population expansion or positive selection as evidenced by the Fu and Li test values for the *trn*L-trnF spacer and combined region. In fact, sites at 170 bp for the *trn*L-*trn*F spacer and to 220 bp for the *trn*L intron in the combined data set are indicated as the sites where selection occurs. This result is similar to those found in *Q. crispula* Blume by Quang et al. (2008), which shows that no significant Tajima's D and Fu and Li's D* and F* were observed, and similar to within-populations of *Q. crispula* (Quang et al., 2008). Furthermore, palynological studies have suggested that oak forests covered a broad area in southern Japan during the last glacial maximum and have just recently reached the Northern areas. In Southern Japan, the frequency of oak pollen rapidly increased 12,000 years ago and then began to decrease from 8500 years ago (Takahara, 1998). As suggested by Magri et al. (2007) and Lumaret et al. (2005), the distribution of cpDNA haplotypes of cork oak demonstrates its postglacial population expansion from potential glacial refuges in Italy, North Africa, and Iberia. Our results confirm this assumption, propose Tunisia as a refuge zone of cork oak, and validate its persistency in North Africa since the last glaciation during the Quaternary. Its long-term stability and persistence are demonstrated, and there appears to be preserved footprints of past expansion.

In summary, findings emerging from the present study reveal considerable intraspecific variation in the *trn*L intron and *trn*L-*trn*F spacer as non-coding regions of plastid DNA of Tunisian *Q. suber* populations and indicate a recent and rapid expansion of cork oaks. Results supported this scenario and rejected the neutral model of molecular evolution for the *trn*L-*trn*F spacer and combined region (*trn*L-*trn*F spacer and *trn*L intron) of *Q. suber* L. Cytoplasmic diversity seems to be unstructured with any biogeographic repartition, bioclimatic stage, or relief of populations. The results of the present study highlight that cytoplasmic DNA markers are reliable for use to elaborate molecular databases to implement breeding and rational management programs, in order to conserve and improve Tunisian cork oak forests. The results suggest that *Q. suber* populations possess considerable genetic diversity and should be given the highest priority for conservation. This is the first report concerning the detection of plastome mutational events for Tunisian Cork oaks and the markers identified are useful for studying cytoplasmic genetic diversity and differentiation of this species, and provide basic genetic information that should generate proposals for conservation and management of forest genetic resources. The present results provide managers with precise information that is essential for the evaluation of resources for gene conservation, and for prioritizing future actions. To successfully manage climatic changes and to address problems of deforestation, we recommend *in situ* and *ex situ* conservation strategies for all the populations studied herein in order to maintain their evolutionary potential, which has been present since quaternary glaciations, in Tunisia and the Mediterranean basin.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Al-Qurainy F, Khan S, Al-Hemaid FM, Ali MA, et al. (2011). Assessing molecular signature for some potential date (*Phoenix dactylifera* L.) cultivars from Saudi Arabia, based on chloroplast DNA sequences *rpo*B and *psb*A-*trn*H. *Int. J. Mol. Sci.* 12: 6871-6880. http://dx.doi.org/10.3390/ijms12106871

ANF (Inventário Florestal Nacional) (2006). Autoridade Nacional Florestal. Available at [http://www.afn.minagricultura. pt/portal/ifn]. Accessed May 27, 2011.

Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, et al. (2000). Patterns of nucleotide substitution in angiosperm cpDNA *trn*L (UAA)-*trn*F (GAA) regions. *Mol. Biol. Evol.* 17: 1146-1155. http://dx.doi.org/10.1093/oxfordjournals. molbev.a026397

Baraket G, Ben Abdelkrim A, Chatti K, Saddoud O, et al. (2010). Molecular evolution of chloroplast DNA in fig (*Ficus carica* L.): footprints of sweep selection and recent expansion. *Biochem. Syst. Ecol.* 38: 563-575. http://dx.doi. org/10.1016/j.bse.2010.06.011

Baraket G, Abdelkrim AB and Salhi-Hannachi A (2015). tRNALeu intron (UAA) of *Ficus carica* L.: genetic diversity and evolutionary patterns. *Genet. Mol. Res.* 14: 3817-3832. http://dx.doi.org/10.4238/2015.April.22.11

Bernatzky R and Tanksley SD (1986). Genetics of actin-related sequences in tomato. *Theor. Appl. Genet.* 72: 314-321. http://dx.doi.org/10.1007/BF00288567

Carrion-Vazquez M, Oberhauser AF, Fisher TE, Marszalek PE, et al. (2000). Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* 74: 63-91. http://dx.doi. org/10.1016/S0079-6107(00)00017-1

Chen J, Tauer G and Huang Y (2002). Paternal chloroplast inheritance patterns in pine hybrids detected with *trn*L-*trn*F intergenic region polymorphism. *Theor. Appl. Genet.* 104: 1307-1311. http://dx.doi.org/10.1007/s00122-002-0893-5

Cosimo SM, Papini A, Vessella F, Bellarosa R, et al. (2009). Multiple genome relationships and a complex biogeographic history in the eastern range of *Quercus suber* L. (Fagaceae) implied by nuclear and chloroplast DNA variation. *Caryologia* 62: 236-252.

Dumolin S, Demesure B and Petit RJ (1995). Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor. Appl. Genet.* 91: 1253-1256. http://dx.doi.org/10.1007/ BF00220937

Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.

Fu YX and Li WH (1993). Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.

Gielly L and Taberlet P (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences. *Mol. Biol. Evol.* 11: 769-777.

Harpending HC (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66: 591-600.

Jukes TH and Cantor CR (1969). Evolution of protein molecules. In: Mammalian Protein Metabolism (Munroled HN, ed.). Academy Press, New York, 31-132.

Kampfer S, Lexer C, Glössl J, Steinkellner H, et al. (1998). Characterization of (GA) n microsatellite loci from *Quercus robur. Hereditas* 129: 183-186. http://dx.doi.org/10.1111/j.1601-5223.1998.00183.x

Lepais O, Leger V and Gerber S (2006). Short note: high throughput microsatellite genotyping in oak species. *Silvae Genet.* 55: 238-240.

Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452. http://dx.doi.org/10.1093/bioinformatics/btp187

Lumaret R, Tryphon-Dionnet M, Michaud H, Sanuy A, et al. (2005). Phylogeographical variation of chloroplast DNA in cork oak (*Quercus suber*). *Ann. Bot.* 96: 853-861. http://dx.doi.org/10.1093/aob/mci237

Magri D, Fineschi S, Bellarosa R, Buonamici A, et al. (2007). The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Mol. Ecol.* 16: 5259-5266. http://dx.doi.org/10.1111/j.1365-294X.2007.03587.x

Manen JF and Natali A (1995). Comparison of the evolution of ribulose-1, 5-biphosphate carboxylase (rbcL) and atpB-rbcL noncoding spacer sequences in a recent plant group, the tribe Rubieae (Rubiaceae). *J. Mol. Evol.* 41: 920-927. http://dx.doi.org/10.1007/BF00173172

Mustapha SB, Ben Tamarzizt H, Baraket G, Abdallah D, et al. (2015). Cytoplasmic polymorphism and evolutionary history of plum cultivars: Insights from chloroplast DNA sequence variation of *trn*L-*trn*F spacer and aggregated *trn*L intron & *trn*L-*trn*F spacer. *Genet. Mol. Res.* 14: 3964-3979. http://dx.doi.org/10.4238/2015.April.27.11

Nei M and Tajima F (1983). Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics* 105: 207-217.

Nei M and Kumar S (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nixon KC (1993). Infrageneric classification of *Quercus* (Fagaceae) and typification of section names. *Ann. Sci. For.* 50: 25-34. http://dx.doi.org/10.1051/forest:19930701

Olalde M, Herrán A, Espinel S, Goicoechea PG, et al. (2002). White oaks phylogeography in the Iberian Peninsula. *For. Ecol. Manage.* 156: 89-102. http://dx.doi.org/10.1016/S0378-1127(01)00636-3

Paffetti D, Vettori C and Giannini R (2001). Relict populations of *Quercus culliprinos* Webb on Sardinia island identified by chloroplast DNA sequences. *For. Genet.* 8: 1-11.

Quang ND, Ikeda S and Harada K (2008). Nucleotide variation in *Quercus crispula* Blume. *Heredity (Edinb)* 101: 166-174. http://dx.doi.org/10.1038/hdy.2008.42

Ramos-Onsins SE and Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19: 2092-2100. http://dx.doi.org/10.1093/oxfordjournals.molbev.a004034

Rogers AR and Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9: 552-569.

Sakka H, Baraket G, Dakhlaoui Dkhil S, Zehdi Azzouzi S, et al. (2013). Chloroplast DNA analysis in Tunisian date-palm cultivars (*Phoenix dactylifera* L.): Sequence variations and molecular evolution of *trn*L (UAA) intron and *trn*L (UAA) *trn*F (GAA) intergenic spacer. *Sci. Hortic.* 164: 256-269. http://dx.doi.org/10.1016/j.scienta.2013.09.038

Taberlet P, Gielly L, Pautou G and Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17: 1105-1109. http://dx.doi.org/10.1007/BF00037152

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Takahara H (1998). Vegetation history of the Kinki district. In: Illustration: Vegetation History of the Japanese Archipelago (Yasuda Y and Miyoshi N, eds.). Asakura-Shoten, Tokyo, 114-134.

Tamura K and Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10: 512-526.

Tamura K, Nei M and Kumar S (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* 101: 11030-11035. http://dx.doi.org/10.1073/pnas.0404206101

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739. http://dx.doi.org/10.1093/molbev/msr121

Tucker JM (1974). Patterns of parallel evolution of leaf form in new world oaks. *Taxon* 23: 129-154. http://dx.doi.org/10.2307/1218095

Wachowiak W, Balk PA and Savolainen O (2009). Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvetsris* L.) *Tree Genet. Genomes* 5: 117-132. http://dx.doi.org/10.1007/s11295-008-0188-3

Zucca GM (2012). Molecular and phenotypic characterization of *Quercus suber* L. and *Pinus uncinata* R. populations in the Mediterranean basin. Doctoral thesis, Università degli Studi di Sassari, Sassari.