RESEARCH ARTICLE

Diversity and spatial genetic structure of natural Moroccan *Quercus suber* L. assessed by ISSR markers for conservation

Amal Laakili¹ · Bouchra Belkadi¹ · Leila Medraoui¹ · Mohammed Alami¹ ·
Chaimaa Yatrib¹ · Ouafae Pakhrou¹ · Mohamed Makhouloufi² · Salwa El Antry² ·
Ahmed Laamarti³ · Abdelkarim Filali-Maltouf¹

Received: 6 June 2017 / Revised: 27 November 2017 / Accepted: 13 April 2018 / Published online: 30 April 2018
© Prof. H.S. Srivastava Foundation for Science and Society 2018

Abstract Morocco is one of the most important regions of the world in terms of *Quercus suber* L. number and variation. This species is in decline due to several factors, which can lead to permanent loss of this resource. It would be essential to evaluate the genetic diversity in order to conserve maximum genetic variability of this species. The genetic diversity and differentiation of 16 sites from five regions representing the entire range of Moroccan Cork Oak were assessed. Twenty-three ISSR primers used generated 985 polymorphic fragments with an average of 42.8 bands per primer and showed 100% of polymorphism. The 173 individuals revealed a moderate level of genetic diversity at species level ($I = 0.27$, $H_e = 0.161$). The AMOVA showed that the highest level of diversity occurred within populations (64%), this was also confirmed by the coefficient of differentiation ($G_{ST} = 0.47$). The estimated gene flow ($N_m = 0.56$) and the Mantel test revealed a significant correlation between geographic and genetic diversity ($r = 0.266$; $p = 0.001$). This genetic structure was further shown by the topology of the UPGMA, sPCA and STRUCTURE analysis. In addition, a core collection of 34 genotypes was established representing the essential diversity detected. This research

advocates populations and individuals to preserve in order to improve and conserve this resource in the future.

Keywords *Quercus suber* L. · Genetic diversity · ISSR · Population structure · Core collection · Morocco

Introduction

The cork oak (*Quercus suber* L.) is a native species in Mediterranean forests with economic and social importance. It only grows on siliceous soils, avoids the limestone substrates and usually grows in hot regions of the humid, sub-humid zones and in semi-arid areas, with an average minimum annual rainfall of 450 mm and $> 4-5$ °C average temperature for the coldest month (Lumaret et al. 2005). In Morocco, cork oak has discontinuous distribution to exclusively northwest of the country and occupy 350.000 hectares (ha) (Hammoudi 2002). This ecosystem is weakening due to the biotic and abiotic attacks resulting in a gradual decrease in tree vigor and vulnerability of natural regeneration (Coelho et al. 2006). The cork-oak woodlands are considered as a reservoir of biodiversity and home to a variety of threatened and endangered species (Peinado 1987). They represent a valuable genetic resource that has to be preserved mainly through restoration and rehabilitation programs. Infact, good management of this species necessitates adequate genetic identification. Therefore, a description of the genetic structure of populations of this species provides a robust framework to evaluate the relationship among and within these populations and to define the priority populations for conservation.

Many authors have showed that *Q. suber* L. is highly polymorphous species using different markers (Elena-Rossello and Cabrera 1996; Cottrell et al. 2003; Belahbib

✉ Abdelkarim Filali-Maltouf
a.filalimaltouf@gmail.com

¹ Laboratory of Microbiology and Molecular Biology, Faculty of Sciences, Mohammed V University, Rabat, Morocco

² Department of Forestry and Forest Health, High Commission for Water, Forests, and Desertification Control (HCEFLCD), Forestry Research Centre (FRC), Agdal, Rabat, Morocco

³ Plant Biotechnology Team, Faculty of Sciences, Abdelmalek Essaadi University, Tétouan, Morocco

et al. 2004; Alfonso-Corrado et al. 2005; Lumaret et al. 2005; Coelho et al. 2006; López-Aljorna et al. 2007; Gandour et al. 2007; Simeone et al. 2009) and have suggested that multi-factorial mutations and hybridization played a determinant role in the phylogenetic development of this species (Natividade 1950).

To date in Morocco, information on levels of genetic diversity within and among populations of *Q. suber* L. remains low for every breeding programs. The study of genetic variability has been previously assessed in all distribution areas of Moroccan *Q. suber* L. using dendrometric traits and showed high levels of variability, depending on the geographical distribution (Laakili et al. 2016). Phenotypic characters are unable to provide a stable structure to implement adequate conservation strategy because of the influence of environmental factors. The use of molecular markers would give more information content that is not prone to variation due to environmental effects.

In the present study, ISSR markers (Inter-Simple Sequences Repeats) have been chosen to study genetic diversity of Moroccan Cork oak because they are ubiquitous, simple, affordable and polymorphic. These molecular markers have been used to assess genetic diversity and differentiation among cork oak populations of the Mediterranean Basin (Rubio de Casas et al. 2007; López-Aljorna et al. 2007). They have proved to be a reliable molecular tool for the study of hybrid complexes and of inter-populations relationships of several forest trees (Labra et al. 2006; Ansari et al. 2012; Patel et al. 2016; Pakhrou et al. 2017).

This study aims to assess the genetic diversity and genetic structuration within and among the Moroccan regions of *Q. suber* L. using ISSR markers and to construct a core collection in order to develop the genetic component for their conservation. In addition, genetic relationship and differentiation of *Q. suber* L. between Northern, Central and Atlantic regions in the country will be discussed.

Materials and methods

Plant materials

The present study incorporates pure populations as diverse as possible from the entire range of Moroccan natural cork oak forests. One hundred seventy-three tree were sampled from 5 biogeographic different regions in Morocco (Central Plateau, East Mamora, West Mamora, Oriental Middle Atlas and North) (16 selected sites) (Fig. 1, Table 1).

In this study, the cork oak sites were selected following the register of a Moroccan-German Cooperation (GTZ 1997) made in 1997. Other forested areas which are not

mentioned in this register were also selected (ex: Jbal Karn). Three hierarchical levels were considered; region, sites and individual for the choice of collected material. Mature leaves were collected between 2010 and 2011 and stored at -80°C . The samples were freeze-dried and ground by CryoMill or directly ground in liquid nitrogen.

DNA extraction

The total genomic DNA was extracted from frozen leaves as described by slightly modified method of Doyle and Doyle (1990). Modifications added to the protocol were basically the amount of leaf material used and the grinding technique applied. Therefore, fresh leaves were lyophilized and ground for 3–5 min using an electric grinder. 20 mg of ground material was used for DNA extraction. Purified DNA was then suspended in TE buffer. DNA concentration and quality were determined by spectrophotometry (ND-2000, NanoDrop, USA) and evaluated by electrophoresis on 1% agarose gel. The DNA concentrations were adjusted to 50 ng/ μL for all samples and then stored at -20°C for subsequent use.

ISSR-PCR

Twenty-three ISSR primers were selected out of 50 primers previously tested (Table 2), these primers showed consistent good amplification. Based on the protocol given by Zietkiewicz et al. (1994), the effects on the amplification of different concentrations, Mg^{2+} , dNTPs, DNA templates, primers and annealing temperature were tested and the final conditions of amplification were fixed. ISSR amplifications were performed in a total volume of 25 μL containing 2 μL total genomic DNA, 1 $\times$ PCR buffer, 2.5 mM MgCl_2 , 0.2 mM dNTPs, 0.2 μM primer and 1U Taq DNA polymerase (Promega, Madison WI, USA). PCR reaction was carried out in a TECHNE, TC-512 thermal cycler following the program: initial denaturation for 7 min at 94°C , followed by 45 cycles of 45 s at 94°C , 45 s at 45°C and 2 min at 72°C and 7 min final extension at 72°C . Amplified products were separated by standard horizontal electrophoresis in 2.8% agarose gels in 1X Tris–acetate EDTA buffer (TAE) for 2 h 30 min at 120 V. The agarose gel was then stained with ethidium bromide (BET) and visualized under ultraviolet light (UV) (Fig. 2).

Data analysis

For each ISSR reaction, only the clear and visible fragments were scored as present (1) and a binary matrix was generated (0–1 matrix). To analyze the performance of the ISSR markers, three parameters were measured for each primer: Polymorphic Information Content (PIC): $\text{PIC}_i = 2$

Fig. 1 Locations of sampled sites of *Quercus suber* L. in Morocco

Table 1 Environmental characteristics of 16 natural populations sampled of *Quercus suber* L. and sample size analyzed

Regions	Code	Populations (sites)	Code	Sample size	Size (ha)	Altitude (m)	Latitude (north)	Longitude (west)
CENTRAL PLATEAU	PC III2	ZITCHOUINE	ZT	5	65	1113	33°27'	6°05'
		TIMKSAOUINE	TS	5	65	410	33°36'	6°05'
		TILIOUINE	TL	8	145	760	33°33'	6°04'
		BENI ABID	BA	13	309	415	33°42'	6°51'
EAST MAMORA	MA I	CANTON A	CA	15	23.868	161	34°03'	6°38'
		CANTON B	CB	25	28.460	180	34°17'	6°30'
		CANTON C	CC	10	17.220	140	34°23'	6°10'
WEST MAMORA	MA II	CANTON D	CD	29	32.464	190	34°14'	6°12'
		CANTON E	CE	13	30.059	261	34°12'	6°05'
ORIENTAL MIDDLE ATLAS	MAO IV2	BAB AZHAR I	BZI	10	76	1250	34°05'	4°14'
		BAB AZHAR II	BZII	10	111	1290	34°03'	4°12'
NORTH	NO	LARACHE	LR	5	—	40	35°07'	6°09'
		BOUSSAFI	BS	5	175	80	35°13'	6°03'
		BOUHACHEM	BH	5	—	1220	35°16'	5°29'
		AIN RAMI	AR	5	86	300	35°07'	5°16'
		JBAL KARN	JK	10	1.200	1070	35°07'	3°42'

Table 2 ISSR primer used, sequences, number fragments scored, approximate size range (in base pairs), number of polymorphic and monomorphic bands and Markers performance indexes (PIC, MI and Rp) calculated for each primer

Primers	Sequences 5' → 3'	Size bands (bp)	TF ^a	PP ^b (%)	PIC ^c	MI ^d	Rp ^e
CH120	(GA) ₈ C	200–1500	37	100	0.31	11.48	16.50
CH121	(AG) ₈ TC	200–1500	38	100	0.34	13.00	18.71
CH122	(AG) ₈ CT	210–1500	44	100	0.32	14.20	20.32
CH123	(AG) ₈ TT	170–1500	46	100	0.28	13.09	17.83
CH124	(GA) ₈ TC	190–1175	38	100	0.33	12.58	18.11
CH125	(CA) ₈ G	160–1350	48	100	0.33	15.79	23.95
CH126	(AC) ₈ TG	220–1780	41	100	0.31	12.55	17.47
CH127	(GGAT) ₄	270–1240	39	100	0.28	10.76	14.77
CH128	(GGAGA) ₃	150–1200	40	100	0.33	13.10	19.24
F1	(CA) ₆ AT	160–1800	39	100	0.27	10.51	14.20
F2	(CA) ₆ GC	160–1675	44	100	0.26	11.50	15.15
F3	(CA) ₆ AG	140–1720	44	100	0.21	9.38	12.61
F4	(AGC) ₄ T	220–1700	49	100	0.26	12.51	17.62
FL1	(AC) ₈ CA	130–1730	50	100	0.26	13.19	18.37
FL2	(AC) ₈ CG	200–1760	51	100	0.28	14.39	19.26
FL3	(AC) ₈ CT	230–1760	41	100	0.31	12.60	18.02
FL4	(AG) ₈ CC	170–1770	45	100	0.25	11.26	15.78
FL6	(CA) ₈ AG	140–1830	44	100	0.25	10.79	14.93
FL7	(CA) ₈ AC	140–1670	40	100	0.26	10.40	13.82
FL8	(GA) ₈ CC	120–1730	45	100	0.30	13.45	18.28
FL9	(GA) ₈ CG	140–1700	41	100	0.25	10.12	14.33
FL10	(GA) ₈ CT	160–1700	45	100	0.24	10.99	15.25
FL11	(GT) ₈ CC	170–1780	36	100	0.30	10.70	15.26
Total			985		6.53	278.3	389.8
Max		120	51		0.34	12.61	23.95
Min		1830	36		0.21	9.38	15.79
Average		–	42.83		0.28	12.1	16.95

^aTotal number of fragments^bPercent polymorphism^cPolymorphic information content^dMarker index^eResolving power

$fi(1 - fi)$ (Roldán-Ruiz et al. 2000), where PIC_i is the polymorphic information content of marker *i*, *fi* is the frequency of the marker bands present and $1 - fi$ is the frequency of marker bands absent. Marker index (MI): $MI = PIC \times EMR$ where EMR is defined as the product of the fraction of polymorphic loci (β) and the number of polymorphic loci (*n*) ($EMR = \beta \times n$) (Powell et al. 1996). Resolving power (Rp): $Rp = \sum Ib$, where *Ib* (band informativeness) takes the values of: $1 - [2x|0.5 - p|]$ and *p* is the proportion of the genotypes containing the band (Prevost and Wilkinson 1999).

The generated data were analyzed using GenAlEx (v 6.5) program (Peakall and Smouse 2006, 2012) to estimate the genetic diversity. Different parameters were assessed including: observed number of alleles (*N_a*), effective number of alleles (*N_e*), Shannon's information index (*I*),

Nei's (1973) gene diversity index (*H_e*) and percentage of polymorphic loci (%P). The coefficient of genetic differentiation (*G_{st}*) (Nei 1973) was estimated Using POPGENE software v1.32 (Yeh et al. 1999). Gene flow (*N_m*) was computed by the formula: $N_m = (1 - G_{st})/2G_{st}$ (McDermott and McDonald 1993). The program GenAlEx (v 6.5) was used to describe genetic variation among regions, within and among the populations based on the analysis of molecular variance (AMOVA) (Excoffier et al. 1992; Peakall and Smouse 2006).

In addition to obtaining a clear representative distribution of five Moroccan *Q. suber* L. regions, biodiversity software DIVA GIS (Hijmans et al. 2001) was also used in order to map the diversity indexes, allelic richness, locally common alleles, frequent alleles and private alleles. The relationship between geographic distance and genetic

Fig. 2 Electrophoresis of DNA from some individuals of different Moroccan Cork oak populations using CH122 primers. M: 50 pb DNA Ladder

distance was performed with the Mantel test. Moreover, to visualize the distribution of individuals, the obtained genetic distance matrix obtained was employed for two analyses. First, to construct a dendrogram based on the unweighted pair-group method with arithmetic average (UPGMA), using Nei's unbiased genetic distances with POPGENE v1.32 software (Yeh et al. 1999). Second, to visualize the global genetic structure of *Q. suber* L. using spatial components analysis (sPCA) with biodiversity software DIVA GIS (Hijmans et al. 2001). This method allows illustrating the results of the genetic variability and the spatial structure of the samples.

Finally, to evaluate the population structure and to estimate possible genetic exchange between Moroccan populations of *Q. suber* L., the Bayesian model was performed using STRUCTURE software v2.1 (Pritchard et al. 2000). This analysis completes the result obtained with UPGMA and sPCA analysis. Bayesian analyses will identify genetically homogenous groups and the number of differentiated clusters without prior group designation. This analysis was run for genetic clusters range from $K = 1$ –15 with 20 independent simulations for each K , used a burn-in and MCMC (Markov Chain Monte Carlo) length of 10,000 based the work of Evanno et al. (2005) and a run length of 30,000. We determined the optimal value of K using Structure Harvester (Earl and vonHoldt 2011) based on the second order likelihood (ΔK) following the procedure described by Evanno et al. (2005).

Construction of genetic core collection

Genetic core collection was constructed allowing having a limited number of individuals and obtained a wide allelic diversity. Therefore, this method will allow the selection of finite individuals' number with the highest diversity. To

reach this objective, the parameters of genetic diversity of core collection were calculated. Thereafter, these parameters will be subsequently compared with the results of the initial collection. An approach of a heuristic algorithm implemented in the Power Core 1.0 software was used (Kim et al. 2007).

Results

Markers informativeness and performance

In the present study, twenty-three primers were chosen to amplify a total of 173 individuals from 16 populations of *Q. suber* L. and produced a total of 985 polymorphic loci with a polymorphism of a 100% (Table 2). In addition, the band size ranged from 120 to 1830 bp. Fragments of the same molecular weight were considered identical locus (Huang and Sun 2000). The primer FL11 produced the lowest number of bands whereas the primer FL2 presented the highest number (51) with an average of 42.83 fragments per primer. Polymorphic index content (PIC) values ranged from 0.21 (F3) to 0.34 (CH121) with an average of 0.28 (Table 2) and where all CH-primers specific to cork oak are above average. We also noticed that the highest number of polymorphic bands were associated with lowest values of polymorphic index content (PIC). Progressively the MI (marker index) and R_p (resolution power) parameters were also calculated in order to evaluate the efficiency and the usefulness of each ISSR markers primer (Table 2). The highest MI was observed for CH125 (15.79) and lowest for F3 (9.38) with an average of 12.1. Likewise, the highest R_p value calculated was observed for CH125 (23.95) and the lowest for F3 (12.61) with an average of 16.95 per ISSR primer (Table 2).

Genetic diversity

Concerning the diversity parameters, the observed number of alleles (N_a) per population ranged from 0.396 to 1.801 while the effective number of alleles (N_e) varied from 1.105 to 1.264 for the 985 loci detected. According to Hardy–Weinberg equilibrium assumption, Shannon's information index (I) ranged from 0.084 to 0.313 with an average of 0.168, Nei's genetic diversity index (H_e) (1973) varied from 0.058 to 0.186 with an average of 0.109 and percentage of polymorphic bands (%P) ranged from 13.71 to 90.05% with an average of 36.99% (Table 3). At the regions level, the percentage of polymorphic bands ranged from 28.73% (MAO IV2) to 92.59% (MA II).

Genetic differentiation and relationships

The analysis of molecular variance (AMOVA) showed a high proportion of genetic variation within populations (64%) and significant inter-population (25%) and inter-region (11%) differentiation in this species ($p < 0.001$) (Table 4). The proportion of genetic variation contributed

Table 3 Genetic diversity within populations of Moroccan *Quercus suber* L.

Populations	N^a	N_a^b	N_e^c	I^d	H_e^e	%P ^f
ZITCHOUIE	5	0.881	1.252	0.223	0.149	41.93
TIMKSAOUIE	5	0.889	1.231	0.215	0.141	43.15
TILIOUIE	8	1.111	1.252	0.249	0.160	53.91
BENI ABID	13	0.640	1.153	0.140	0.092	28.83
CANTON A	15	0.888	1.166	0.168	0.105	43.45
CANTON B	25	1.723	1.248	0.292	0.174	85.99
CANTON C	10	0.780	1.185	0.177	0.115	38.38
CANTON D	29	1.801	1.264	0.313	0.186	90.05
CANTON E	13	0.830	1.166	0.171	0.107	39.90
BAB AZHAR I	10	0.436	1.122	0.097	0.067	15.84
BAB AZHAR II	10	0.520	1.155	0.125	0.086	20.61
AIN RAMI	5	0.457	1.105	0.084	0.058	13.71
BOUSSAFI	5	0.396	1.111	0.087	0.061	13.81
LARACHE	5	0.422	1.118	0.094	0.065	15.74
JBAL KARN	10	0.718	1.184	0.167	0.111	31.78
BOUHACHEM	5	0.398	1.118	0.092	0.064	14.82
Mean	10.813	0.806	1.177	0.168	0.109	36.99

^aNumber of samples

^bObserved number of alleles

^cEffective number of alleles

^dShannon's information index

^eNei's (1973) gene diversity index

^fPercentage of polymorphic loci

Table 4 Analysis of molecular variance (AMOVA) for Moroccan *Quercus suber* L. by ISSR

Source of variation	df ^a	SS ^b	PV (%) ^c	p value ^d
Among regions	4	2506.184	11	<0.001
Among populations	12	2708.858	25	<0.001
Within populations	156	7985.143	64	<0.001
Total	172	13,200.185	100	<0.001

^aDegree of freedom

^bSum of squares

^cPercentage of variance

^dSignificance tests after 999 permutation

by the differences among populations (G_{st}) is 0.47. Also, the level of gene flow (N_m) was estimated to be 0.56. The Mantel test revealed that there was moderate significant correlation between geographic and genetic distance ($r = 0.266$, $p = 0.001$ Permutation: 999).

To map the allelic richness, locally common alleles, frequent and private alleles, the software DIVA GIS was used. The Allelic richness ranged from 243 in MAO IV2 to 471 in MAII. Also, the number of frequent alleles ranged from 243 to 422 in MAO IV2 and MAII respectively. Locally, common alleles ranged from 9 in MAO IV2 to 55 in Mamora region (Fig. 3).

Based on Nei's unbiased genetic distance, the pool of ISSR data was used to construct a UPGMA dendrogram (Fig. 4) to estimate the genetic relationships among the five Cork oak regions. All 16 populations were separated into two geographic groups. The two regions of Mamora (MAI, MAII) and populations of Central Plateau formed Group I. Group II consisted of the populations from North region (NO) and Oriental Middle Atlas.

The report of Nei's genetic distances between pairs of regions and geographic distance was calculated and ranged from 0.013 (MAI and MAII) to 0.115 (PC III2 and MAO IV2). The geographic distance values ranged from 0.299 to 2.291 (Table 5). This again confirms the correlation between geographic and genetic distances previously estimated by Mantel test.

The projection of special analysis of principal components (sPCA) on map regions was performed to provide spatial representation of the relative genetic distances among sampled trees (Fig. 5) and to determine the consistency of differentiation defined by the cluster analysis. We observed that the individuals of PC III2 and MA were separated from those of NO and MAO IV2.

The genetic structure obtained was further justified by the Bayesian analyses. In the ISSR admixture analysis using STRUCTURE (Fig. 6), the highest likelihood was $K = 2$. The second order statistics (ΔK) determined the

Fig. 3 Spatial variation of different genetic parameters, represented at a resolution of 30 s grid cells and circular areas of 10 km

Fig. 4 UPGMA dendrogram based on Nei's (1973) genetic distances among Moroccan regions of *Quercus suber* L.

Table 5 Nei's (1973) unbiased measures of genetic distance (below diagonal) and geographic distance (above diagonal) between regions of *Quercus suber* L.

	PC III2	MAI	MAII	MAO IV2	NO
PC III2	****	0.578	0.630	2.231	2.101
MAI	0.031	****	0.299	2.291	1.790
MAII	0.031	0.013	****	1.997	1.537
MAO IV2	0.115	0.100	0.091	****	1.421
NO	0.072	0.057	0.052	0.044	****

*** not significant

optimal value for $K = 2$. The first cluster (red) consisted mainly the population from Mamora (CA, CB, CC, CD CE CE) and those from Central plateau regions (ZT, TS, TL, BA), the second group (green) contains the populations from Oriental Middle Atlas (BZI, BZII) and North region (LR, BS, BH, AR, JK). These divisions were entirely reliable with those of the UPGMA clustering and sPCA

Fig. 5 Scores of sampled trees as projected on the first ordination axis of the biplot of a spatial analysis of principal components

analysis. The structure analysis assigned genotypes into respective groups on the basis of their allele frequencies.

Construction of genetic core collection

Based on the heuristic algorithm in Power Core software, we obtained a core collection with 34 genotypes from different regions (20% of the total collection) representing the genetic diversity existing with 99.80% of total generated alleles and an allelic richness varied from 157 to 266 (Table 6). The effective number of alleles, Shannon's index, Nei's (1973) gene diversity calculated for the core collection ($N_e = 1.236$, $I = 0.244$ and $H_e = 0.153$) were close to those values for the overall genotypes ($N_e = 1.237$, $I = 0.269$ and $H_e = 0.161$).

Fig. 6 Genetic relationships among the 16 populations of *Q. Suber* L. estimated using STRUCTURE program based on ISSR data. The model with $K = 2$ showed the highest ΔK value. Populations numbers see Table 1

Table 6 List of 34 trees out of 173 trees of entire collection included in core set formed by PowerCore

Region	Individuals	Latitude (north)	Longitude (west)	Allelic richness
PC III	ZT2	33°26'54"	6°05'01"	216
	TS2	33°35'47"	6°05'16"	188
	TS3	33°35'47"	6°05'17"	204
	TS5	33°35'48"	6°05'23"	188
	TL4	33°33'15"	6°03'46"	236
	TL5	33°33'14"	6°03'47"	233
	TL6	33°33'17"	6°03'51"	229
	BA3	33°39'41"	6°58'25"	171
MAI	BA12	33°59'47"	6°62'56"	189
	CA14	34°11'30"	6°53'20"	165
	CB2	34°16'56"	6°29'37"	205
	CB10	34°03'43"	6°33'10"	187
	CB11	34°03'26"	6°33'36"	231
	CB12	34°03'18"	6°33'44"	253
	CB15	34°03'16"	6°33'46"	251
	CB21	34°13'57"	6°47'05"	248
MAII	CB22	34°16'00"	6°41'36"	246
	CC10	34°08'56"	6°38'21"	157
	CD3	34°24'10"	6°25'30"	224
	CD8	34°18'07"	6°24'31"	260
	CD12	34°11'17"	6°11'25"	226
	CD15	34°13'16"	6°13'23"	210
	CD16	34°13'15"	6°13'14"	229
	CD21	34°16'55"	6°22'26"	235
MAO IV2	CD23	34°11'37"	6°15'50"	239
	CD26	34°12'56"	6°23'40"	165
	CE11	34°08'12"	6°26'38"	188
	BZI10	34°13'01"	4°17'25"	221
NO	BZII8	34°11'34"	4°21'23"	227
	BZII10	34°12'08"	4°21'20"	229
	AR1	35°25'53"	5°31'48"	266
	AR3	35°19'48"	5°43'07"	250
	LR1	35°07'49"	6°14'46"	210

Discussion

Markers informativeness and performance

In Morocco, focused on mixed areas of *Quercus*, the Oak has been studied using RFLP and SSR markers in order to quantify the level of cpDNA variation within *Q. Suber* L. and the level of introgression of this species with *Q. Ilex* (Belahbib et al. 2001, 2004). The sampling concerned all the distribution area of the national Cork oak in Morocco. The number of samples differs from one population to another, taking into account the density of each stand. In fact, several genetic diversity studies on forest species used molecular markers on unbalanced sampling without influencing the genetic results (Fofana et al. 2009).

Our latest study, using dendrometric characters (Laakili et al. 2016) was based on a large sample that allowed us to assess a morphological genetic structure of Moroccan cork oak forests. The evaluation of phenotypic diversity of Moroccan cork oak populations demonstrated a large diversity. Cork diversity has been justified by the observed differences between individuals of cork oak trees from different provenances. Several allometric relationships were highlighted between traits, but most of the variability of dendrometric characteristics variation of trees is due to changes in environmental conditions. Furthermore, research using molecular markers seems to be necessary to confirm the phenotypic variability that is independent of environmental factors. Therefore, our choice focused on ISSR markers because they are reproducible, present the entire genome variation and they are capable of detecting high level of genetic diversity.

Many studies based on ISSR analyses have shown that endemic and endangered species tend to hold low genetic diversity level (Xiao et al. 2004; Li and Xia 2005; Xu-Mei et al. 2012) and others showed opposite results (Ge et al. 2003; Li and Jin 2007). In fact, research work of Hamrick and Godt (1989) demonstrated that species with long life span tend to have an important genetic diversity level. In addition, the primers CH121, CH124 and CH128 were also successfully used for the amplification of ISSR markers in *Q. suber* L. individuals from different provenances (López-Aljorna et al. 2007). The oligonucleotide CH124 produced the highest number of fragments in López-Aljorna et al. study (2007) for *Q. suber* L. and for *Q. coccifera*. Considering that dominant marker may present a high PIC value of 0.5 (Riek et al. 2001), it reveals that the primers (CH121, 124, 125 and 128) used in this study were discriminative for the assessment of intra specific genetic diversity. Therefore, the PIC values can be used to compare the information content generated by ISSR primers while

MI and Rp parameters can be used as the most informative data to differentiate among different regions and sites.

Genetic diversity and population structure analysis

Twenty-three ISSR primers produced variable values regarding Nei's genetic diversity and Shannon's information index. ISSR primers reflected the differential abundance of their complementary sequences on genomic DNA across Moroccan cork oak populations. The highest value of Nei's gene diversity was detected in the two Mamora regions, while the lowest value was found in oriental middle atlas region and the North region. Whereas, Central Plateau populations represent an important genetic diversity.

The result of AMOVA analysis corroborates that long-lived, out-crossing tree with wide and continuous range retains most of their genetic variation within their populations (Hamrick and Godt 1989, 1996; Nybom and Bartish 2000; Nybom 2004). The existence of an important genetic diversity within populations compared to inter-population genetic diversity can be explained by environmental diversity effect within populations to generate micro local adaptation (Audigeos et al. 2013). Also, this differentiation may be the result of a dynamic relationship between erosive forces (selection, random drift) and generative forces of diversity (recombination, genes flow).

The double contingency (contingency environmental conditions and contingency to other characters) is also another reason for maintaining a wide variety within populations. Moderate observed gene diversity across various cork oak populations may be due to a influence that is exerted by anthropological activities and also due to a common gene pool. In addition, the genetic differentiation reflects the footprints of its tertiary history despite relatively intense, albeit recent, human exploitation (Magri et al. 2007). It is notable that the highest values of different genetic parameters are observed in Mamora region. We notice the presence of sixteen private alleles scattered in five regions (4 at PC III2, 4 at MAI, 3 at MAII, 3 at MAO IV2 and 2 at NO) (Fig. 3). The private bands detected in this study can be used for the authentication of this tree forest.

Moderate gene diversity among various populations adequately distinguishes Atlantic (Mamora) and Central plateau populations from North and Oriental Middle Atlas populations. This separation reflects geographic distribution of these populations. The genetic diversity of *Q. Suber* L. of Mamora (Atlantic Cork Oak) and central plateau region was very high compared to the Oriental Middle Atlas and the North suberaie's. It is well known that these two areas differ in topography and altitude, which may be another reason for genetic differentiation found in this

work. Also, during the major quaternary glaciations, the cork oak has migrated from north to south, which has allowed this species to have a high ecological plasticity and high adaptability to environment. Gene flow between cork oak populations was estimated to $N_m = 0.56$, which is sufficient to prevent local genetic differentiation by genetic drift (Wright 1931).

Furthermore, the result of UPGMA analysis was further confirmed by the finding of the sPCA. In agreement with the cluster analysis, regions and sites from each region formed a separate plot and could be clearly distinguished, the individuals of PC III2 and MA were separated from those of NO and MAO IV2, which indicated an important genetic diversity level among regions. Therefore, taking into account geographical proximities, the genetic clustering of the populations shows the presence of an important relation among neighboring regions. In addition, Bayesian analysis provided evidence of a great association between MA and PCIII regions and between MAO IV2 and NO region. This analysis allowed the presence of the admixture individuals from Ain Rami (North region). The deduction is in agreement with the recent genetic structure finding using phenotypic traits (Laakili et al. 2016), which assembled the Atlantic Rif (Boussafi and Ain Rami) with Mamora regions in one group.

Hence, based on the results of the Mantel test, the isolation could influence the present genetic structure. The low genetic diversity among cork oak populations is justified by the important gene flow level obtained. Also, the high genetic diversity within Moroccan *Q. Suber* L. could be explained by the emergence of genes, during the long evolutionary period, to adapt to environmental changes. Therefore, genetic structuration of *Q. suber* L. could be affected by multiple evolutionary forces including the gene flow, natural selection and reproduction mode.

Construction of genetic core collection

Genetic variability maintenance constitutes the first objective in conservation strategies of threatened species. In fact, genetic diversity conservation allows the preservation of the evolutionary potential of species. The core collection constitutes a bank of genotypes representative of the Cork oak national genetic material which must be conserved in situ and ex situ in order to preserve the genetic diversity of Moroccan cork oak populations.

All the results obtained in this study provide important information for conservation plan of Moroccan cork oak. Therefore, the genetic structure obtained must be considered when harvesting seedlings and during the reforestation. The Cork oak stands should be protected and the major factor of deterioration must be banned. Thus, the stands of small size and Mamora region that present an

important genetic diversity should receive more attention and thus a special importance in breeding programs.

In fact, the cork oak has a broad genetic base as evident from its distribution in different eco-climatic zones of Morocco. The climate models predict changes in temperature and rainfall, but they cannot take into account the adaptive capacity of forest ecosystems. The remarkable plasticity of the cork oak, especially in relation to a tree whose bark is taken at regular intervals ensures that the future of the cork oak forests is intact. Noticing that, cork oak remains one of the main bulwarks against desertification in many Mediterranean regions, as well as a habitat of exceptional diversity.

Conclusion

In summary, these molecular relationships of *Q. suber* L. populations showed a clear genetic structuration. Genetic structure of Moroccan Cork oak can be a foundation for further research on the evolution, breeding programs including marker-assisted selection and also will serve as an initial point to study genetic relationships and evolution. Our results indicated that high genetic differentiation was found mainly within populations and significant correlation was found between geographic and genetic distance. Cluster analysis using the UPGMA method, sPCA Analysis and Bayesian analysis grouped all populations into two geographic groups. That is, latter analysis demonstrated that ISSR technique was very effective in determining genetic diversity of the *Q. suber* L. populations. Furthermore, our study advocates the necessity to give importance in first place to Mamora region. Based on these results, an appropriate strategy and breeding programs can be anticipated for the conservation and amelioration of Moroccan *Q. suber* L.

Acknowledgements This work was supported by Hassan II Academy of Sciences and Technology (Morocco). We express our deep gratitude to all forest technicians who contributed to the collection of plant material throughout Morocco.

References

- Alfonso-Corrado C, Esteban-Jiménez R, Clark-Tapia R et al (2005) Clonal and genetic structure of two Mexican oaks: *Quercus eduardii* and *Quercus potosina* (Fagaceae). *Evol Ecol* 18:585–599
- Ansari SA, Narayanan C, Wali SA et al (2012) ISSR markers for analysis of molecular diversity and genetic structure of Indian teak (*Tectona grandis* Lf) populations. *Ann For Res* 55:3
- Audigeos D, Brousseau L, Traissac S et al (2013) Molecular divergence in tropical tree populations occupying environmental mosaics. *J Evol Biol* 26:529–544

- Belahbib N, Pemonge M-H, Ouassou A et al (2001) Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Mol Ecol* 10:2003–2012
- Belahbib N, Ouassou A, Dahmani J, Douira A (2004) Contribution à l'étude de l'introgession génétique entre *Quercus suber* L et *Q. rotundifolia* (Lamk.) Trabut au Maroc par l'utilisation des marqueurs microsatellites. *Bull Inst Sci Rabat Sect Sci Vie* 26–27:31–34
- Coelho AC, Lima MB, Neves D, Cravador A (2006) Genetic diversity of two evergreen oaks (*Quercus suber* L. and *Q. ilex rotundifolia* Lam.) in Portugal using AFLP markers. *Silvae Genet* 55:105–118
- Cottrell JE, Munro RC, Tabbener HE et al (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oak woods differing in population history. *For Ecol Manag* 176:287–303
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- Elena-Rossello JA, Cabrera E (1996) Isozyme variation in natural populations of cork-oak (*Quercus suber* L.) Population structure, diversity, differentiation and gene flow. *Silvae Genetica* 45:229–239
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Fofana JJ, Ofori D, Poitel M, Verhaegen D (2009) Diversity and genetic structure of teak (*Tectona grandis* L.f) in its natural range using DNA microsatellite markers. *New For* 37:175–195
- Gandour M, Khouja ML, Toumi L, Triki S (2007) Morphological evaluation of cork oak (*Quercus suber*): Mediterranean provenance variability in Tunisia. *Ann For Sci* 64:549–555. <https://doi.org/10.1051/forest:2007032>
- Ge X-J, Yu Y, Zhao N-X et al (2003) Genetic variation in the endangered Inner Mongolia endemic shrub *Tetraena mongolica* Maxim. (Zygophyllaceae). *Biol Conserv* 111:427–434
- GTZ (1997) Les ressources génétique forestières au Maroc peuplement à graines classés et arbres plus. Chêne liège (*Quercus suber* L.), vol. 6
- Hammoudi A (2002) La subéraie: biodiversité et paysage. Haut-commissariat aux eaux et forêts et à la lutte contre la desertification, Maroc. Haut-commissariat aux eaux et forêts et à la lutte contre la desertification, Maroc., VIVEXPO liege
- Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) *Plant population genetics, breeding and genetic resources*. Sinauer Associates, Sunderland, pp 43–63
- Hamrick JL, Godt MJW (1996) Conservation genetics of endemic plant species. In: Avise JC, Hamrick JL (eds) *Conservation genetics: case studies from nature*. Chapman & Hall, New York, pp 281–304
- Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genet Resour Newsl* 127:15–19
- Huang JC, Sun M (2000) Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. *Theor Appl Genet* 100:1050–1060
- Kim K-W, Chung H-K, Cho G-T et al (2007) PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinforma Oxf Engl* 23:2155–2162
- Laakili A, Belkadi B, Gaboun F et al (2016) Analysis of dendrometric diversity among natural populations of cork oak (*Quercus suber* L.) from Morocco. *Turk J Agric For* 40:127–135
- Labra M, Grassi F, Sgorbati S, Ferrari C (2006) Distribution of genetic variability in southern populations of Scots pine (*Pinus sylvestris* L.) from the Alps to the Apennines. *Flora Morphol Distrib Funct Ecol Plants* 201:468–476
- Li J, Jin Z (2007) Genetic variation and differentiation in *Torreya jackii* Chun, an endangered plant endemic to China. *Plant Sci* 172:1048–1053
- Li F, Xia N (2005) Population structure and genetic diversity of an endangered species, *Glyptostrobus pensilis* (Cupressaceae). *Bot Bull Acad Sin* 46:155–162
- López-Aljorna A, Bueno MA, Aguinalde I, Martín JP (2007) Fingerprinting and genetic variability in cork oak (*Quercus suber* L.) elite trees using ISSR and SSR markers. *Ann For Sci* 64:773–779
- Lumaret R, Tryphon-Dionnet M, Michaud H et al (2005) Phylogeographical variation of chloroplast DNA in Cork Oak (*Quercus suber*). *Ann Bot* 96:853–861
- Magri D, Fineschi S, Bellarosa R et al (2007) The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Mol Ecol* 16:5259–5266
- McDermott J, McDonald BA (1993) Gene flow in plant pathosystems. *Annu Rev Phytopathol* 31:353–373
- Natividade VJ (1950) Subericultura. Ministério da Agricultura, Pescas e Alimentação – Direcção Geral das Florestas, Lisbon
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321–3323
- Nybom H (2004) Nybom H comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol Ecol* 13:1143–1155
- Nybom H, Bartish IV (2000) Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Evol Syst* 3:93–114
- Pakhrou O, Medraoui L, Yatrib C et al (2017) Assessment of genetic diversity and population structure of an endemic Moroccan tree (*Argania spinosa* L.) based in IRAP and ISSR markers and implications for conservation. *Physiol Mol Biol*. <https://doi.org/10.1007/s12298-017-0446-7>
- Patel DM, Fougat RS, Sakure AA et al (2016) Detection of genetic variation in sandalwood using various DNA markers. *3 Biotech* 6:1–11
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Resour* 6:288–295
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Peinado M (1987) La vegetación de España. Serv Publ Univ Alcalá, Madrid
- Powell W, Morgante M, Andre C et al (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225–238
- Prevost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor Appl Genet* 98:107–112
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959

- Riek JD, Calsyn E, Everaert I et al (2001) AFLP based alternatives for the assessment of distinctness, uniformity and stability of sugar beet varieties. *Theor Appl Genet* 103:1254–1265
- Roldán-Ruiz I, Dendauw J, Bockstaele EV et al (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). *Mol Breed* 6:125–134
- Rubio de Casas R, Cano E, Balaguer L et al (2007) Taxonomic identity of *Quercus coccifera* L. in the Iberian Peninsula is maintained in spite of widespread hybridisation, as revealed by morphological, ISSR and ITS sequence data. *Flora Morphol Distrib Funct Ecol Plants* 202:488–499
- Simeone MC, Papini A, Vessella F 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
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159
- Xiao L, Ge X, Gong X et al (2004) ISSR variation in the endemic and endangered plant *Cycas guizhouensis* (Cycadaceae). *Ann Bot* 94:133–138
- Xu-Mei W, Hou X-Q, Zhang Y-Q et al (2012) Genetic diversity of the endemic and medicinally important plant *rheum officinale* as revealed by inter-simple sequence repeat (ISSR) markers. *Int J Mol Sci* 13:3900–3915
- Yeh FC, Yang RC, Boyle TBJ et al (1999) POPGENE, version 1.31. The user-friendly shareware for population genetic analysis
- Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20:176–183