

# Assessment of genetic diversity in Tunisian populations of *Medicago polymorpha* based on SSR markers

Loua Haddoudi<sup>1,2</sup>, Sabrine Hdira<sup>1,2</sup>, Najah Ben Cheikh<sup>1</sup>, Asma Mahjoub<sup>1</sup>, Chedly Abdelly<sup>1</sup>, Ndiko Ludidi<sup>3,4</sup>, and Mounawer Badri<sup>1\*</sup>

<sup>1</sup>Centre of Biotechnology of Borj Cedria, Laboratory of Extremophile Plants, B.P. 901, Hammam-Lif 2050, Tunisia.

\*Corresponding author (mounawer.badri@cbbc.rnrt.tn).

<sup>2</sup>University of Tunis El Manar, Faculty of Mathematical, Physical and Natural Sciences of Tunis, Campus Universitaire El-Manar, 2092 El Manar Tunis, Tunisia.

<sup>3</sup>University of the Western Cape, Department of Biotechnology, Robert Sobukwe Road, Bellville, 7530, South Africa.

<sup>4</sup>University of the Western Cape, DSI-NRF Centre of Excellence in Food Security, Robert Sobukwe Road, Bellville, 7530, South Africa.

Received: 29 June 2020; Accepted: 27 September 2020; doi:10.4067/S0718-58392021000100053

## ABSTRACT

*Medicago polymorpha* L. is a herbaceous legume that can be a useful pasture crop, especially in Mediterranean climates. This study aimed to analyze the genetic variation in five populations of *M. polymorpha* collected from different eco-geographic regions in Tunisia using eight SSR markers. The transferability of 112 SSR markers distributed on the eight chromosomes of *M. truncatula* Gaertn. showed that 50 SSR markers could be amplified in *M. polymorpha*. Among these 50 SSR markers, eight (8) markers were polymorphous. A high level of polymorphism (126 polymorphic alleles with an average of 5.3 alleles per locus) and a moderate level of genetic diversity were found in all the studied populations, with observed and expected heterozygosities averaging between 0 and 0.69, respectively. Results from analysis of molecular variance (AMOVA) revealed that the most variation was found within populations (76%). Moderate levels of population differentiation ( $F_{ST} = 0.12$  to  $0.19$ ) accompanied by a high rate of gene flow between populations ( $N_m = 1.08$  to  $1.83$ ) were recorded. This molecular differentiation ( $F_{ST}$ ) was not dependent on geographical distances ( $r = 0.395$ ,  $p = 0.524$ ), suggesting that studied populations are not geographically isolated. Our results showed that studied populations were clustered into three groups. A first group is formed by the populations of El Kef (TNP7) and Bulla Regia (TNP9), a second group of Enfidha (TNP1) and a third group of Soliman (TNP8) and Mateur (TNP11). Results obtained in our study could be helpful for breeders considering introduction of some lines of this species into *M. polymorpha* breeding programs.

**Key words:** Gene flow,  $F_{ST}$ , *Medicago polymorpha*, SSR markers, Tunisian populations.

## INTRODUCTION

*Medicago* is a genus of flowering plants in the Leguminosae family, commonly known as medick or burclover. It contains about 87 species and is distributed mainly around the Mediterranean basin (Gholami et al., 2014). The genus *Medicago*, well represented in Tunisia, has an extremely rich and diverse gene pool (Badri et al., 2016). *Medicago polymorpha* L. is one of the major annual *Medicago* species and it belongs to the same genus as the model forage legume *M. truncatula* Gaertn. It is native to the Mediterranean basin and successfully naturalized in Mediterranean-like climates. It is diploid ( $2n = 14$ ) and autogamous and spreads from humid to upper arid regions in Tunisia. Analysis of the diversity of populations of *M. polymorpha* is a prerequisite for any breeding program of lines of this species. In this context, Badri et al. (2016) studied the morpho-phenological diversity in Tunisian populations of this species while the current study aimed to analyze

the genetic variation in these populations using simple-sequence repeats (SSR) markers. An SSR marker consists of a tract of tandemly repeated (i.e., adjacent) DNA motifs that range in length from two to six nucleotides (Remya et al., 2010). One of the major advantages of SSR markers is their transferability between genetically similar species. Cross amplifications were obtained from different families including *Triticeae* (Zeid et al., 2010), *Fabaceae* (Badri et al., 2007; 2008a; 2008b; Liu et al., 2019), *Rosaceae* (Zhang et al., 2014), *Solanaceae* (Grover et al., 2009), *Brassicaceae* (Shirasawa et al., 2011), *Pinaceae* (Uchyama et al., 2013) and *Myrtaceae* (Nogueira et al., 2016).

The effect of natural selection on quantitative traits can be evaluated by an approach involving the comparison of two population differentiation indices:  $Q_{ST}$  and  $F_{ST}$  (Edelaar et al., 2011).  $Q_{ST}$  measures the divergence of quantitative traits between populations measured in controlled conditions and  $F_{ST}$  measures differentiation based on neutral molecular markers (Whitlock and Guillaume, 2009). When  $Q_{ST} > F_{ST}$ , the selection effect is greater than that of genetic drift, suggesting divergent selection between different habitats. If  $Q_{ST} = F_{ST}$ , then the effects of neutral processes and selection on population differentiation cannot be distinguished. Finally, a value where  $Q_{ST} < F_{ST}$  suggests a homogenizing or stabilizing selection effect that led to the establishment of similar phenotypes in different environments.

Several studies have investigated genetic diversity in *Medicago* species, but DNA-based genetic diversity studies in *M. polymorpha* are limited. This limitation may be due to this *M. polymorpha* being a spontaneous species that is not cultivated as a major crop. The current study aimed (i) to analyze the transferability of SSR markers developed in the model legume *M. truncatula* to *M. polymorpha*, (ii) to investigate the genetic structure within and among natural populations of *M. polymorpha* by using SSR markers, and (iii) to identify the evolutionary forces responsible for this variation based on the comparison between population differentiation through SSR markers ( $F_{ST}$ ) and phenotypic traits ( $Q_{ST}$ ).

## MATERIALS AND METHODS

### Plant material and molecular characterization

Five populations of *Medicago polymorpha* L. collected from different eco-geographic regions in Tunisia were used. The geographical location, sampling method of studied populations and creation of the lines have been described in the work of Badri et al. (2016).

DNA extraction of each line was performed from three of the youngest trifoliate leaves from each plant. The DNA was extracted from 200 mg fresh material according to the protocol of Geuna et al. (2003). The transferability of SSR markers developed on the genome of *M. truncatula* Gaertn. to *M. polymorpha* was performed using 112 SSR loci mapped on the genome of the recombinant inbred lines populations LR1 and LR5. Populations LR1 (233 RILs) and LR5 (173 RILs) were, respectively, obtained from the crosses DZA315.26 × DZA45.6 and Jemalong6 × F83005.5. DZA315.26 and DZA45.6 from Algeria, F83005.5 from France, and Jemalong6 from the Australian collection. Furthermore, the analysis of polymorphism within and among studied populations of *M. polymorpha* was performed based on eight polymorphic SSR markers (Table 1). The PCR was performed in a volume of 20 µL containing 30 ng genomic DNA, 9.36 µL distilled water, 2 ng each primer, 2.5 mM dNTP, 0.8 U Taq DNA polymerase and 1.28 mM MgCl<sub>2</sub>. Amplifications were performed in a thermal cycler (2720 Thermal Cycler; Applied Biosystems, Foster City, California, USA) according to the following program: 4 min at 94 °C, 40 cycles of 30 s at 55 °C and 30 s at 72 °C; accompanied by a final extension step of 6 min at 72 °C.

**Table 1. Name, linkage group, position on composite map of *Medicago truncatula*, primers and origin of eight used SSR loci.**

Name	Linkage group	Composite map	BAC	Primer left sequence (5'-3')	Primer right sequence (5'-3')	Origin
MTIC1044	7	MTE85	AC146554	GTATTGGACGACCGAGTATT	CTGTCCTCTCGAACTCCAT	Current study
MTIC976	6	MTE67	AC137553	CTAGAGAGAGAGCTTCCGATT	ATTCATTCACCGTCCAAG	Current study
MTIC531	1	MTE8	AC131240	GATGGATGAGAAATGGAGAA	GAATCAAATCCACCTATCCA	Current study
MTIC983	2	MTE72	AC132565	TGGGAGGAACCAAGTATAAAA	GGTTGAATTGTGTTGGATCT	Current study
MTIC333	7	MTE49	AC128638	GCAACCGCTTTATTACATTGTC	GCGGTTGAGTAACGATCTGG	Current study
MTIC068	1	MTE2	AC137824	AATTGCAGCAGCAACAATCA	TCCAAACCTCCCATGGTATC	Exbrayat et al., 2014
MTIC784	8	MTE71	AC137821	GTGGAAGTTTGATCATTTTCG	CAAATCTCACACCCACTACA	Current study
MTIC1050	3	MTE89	AC138014	GATTAAGCATTATTGGATGTG	ATCATGATCAAAACTATTGC	Exbrayat et al., 2014

Bacterial artificial chromosome (BAC).

The amplified fragments were separated by electrophoresis on an agarose gel of 3% for analysis of polymorphism using Tris Acetate EDTA (TAE 0.5x) as running buffer. The electric field used for migration was 110 V, which corresponds to an approximately 90 mA per gel. The migration time was 2 h for amplification tests and 4 h for the polymorphism analysis. The DNA samples amplified after deposition and migration, were photographed using a Gel-Doc (Vilber Lourmat, Marne-la-Vallée, France) equipped with a charge-coupled device (CCD) camera (M&S Instruments Inc., Osaka, Japan). The amplified fragments were scored against a 100 bp DNA ladder (Promega Corporation, Madison, Wisconsin, USA).

### Data analyses

Genetic diversity among populations of *M. polymorpha* was partitioned into within and among populations' levels. To describe the genetic diversity within population, the observed number of alleles per locus ( $N_a$ ), effective number of alleles ( $N_e$ ), Shannon's information index (I), polymorphism percentage (P), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), fixation index ( $F_{IS}$ ), number of private alleles, percentage of frequent (appearing in > 50% of lines) and rare (appeared in < 5% of genotypes) alleles were estimated using GenAIEx v. 6.5 (Peakall and Smouse, 2006).

The studied populations were split into three bioclimatic classes (from sub-humid to lower-semi-arid), according to Emberger's pluviothermic quotient (rain temperature coefficient; Emberger, 1955). The Shannon's index of diversity (I) and the expected unbiased heterozygosity ( $uH_e$ ) were estimated as described by Saoudi et al. (2019). In order to examine the distribution of variation among and within populations, and among regions, a hierarchical AMOVA was performed in GenAIEx v. 6.5 using populations nested within ecoregions. The same program was used to evaluate the pairwise genetic differentiation ( $F_{ST}$ ) and Nei's genetic distances (Nei, 1978) between populations, as well as the estimate of the effective number of migrants per generation ( $N_m$ ).

The potential structure of the studied populations was analyzed using STRUCTURE software v. 2.3.3 (Pritchard et al., 2000) as described by Targonska et al. (2016).

The principal coordinates analysis (PCoA) was performed on the basis of genetic distances using GenAIEx v. 6.5 while the genetic relationship between populations and lines was carried out using the scores of PCoA in XLSTAT software v. 7.5 (Addinsoft, Long Island City, New York, USA).

The correlation between matrices of log-transformed geographic distances and odds-transformed genetic distances  $F_{ST}/(1-F_{ST})$  was estimated by a nonparametric permutation Mantel test (Mantel, 1967) implemented in XLSTAT software v. 7.5.

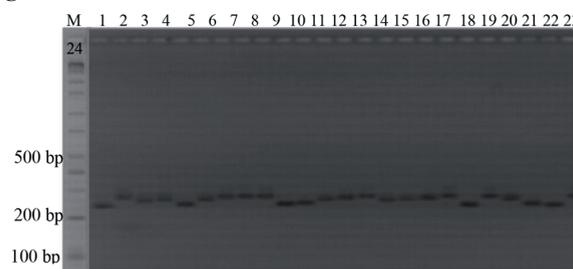
## RESULTS

### Cross-amplification of SSR markers and genetic variation in *M. polymorpha*

The cross-amplification of SSR markers developed on the genome of the model legume *M. truncatula* to *M. polymorpha* showed that among the 112 markers tested, 50 markers were amplified in *M. polymorpha*. Furthermore, only eight (8) out of the 50 SSR markers were polymorphous.

Using 3% agarose gels, we were able to detect a minimal size difference of 2 bp between alleles of each SSR marker (Figure 1). A total of 126 alleles were detected for the five studied populations with an average of 5.3 alleles per locus (Table 1). A high number (25) of alleles was found for the MTIC333 marker while the lowest value (7) was observed for marker MTIC976. Of the 126 alleles, 74 (59%) were private alleles. A high number (14) of private alleles was found

**Figure 1. Polymorphism observed for the population of Mateur (TNP11) of *Medicago polymorpha* for MTIC1044 locus. M: 100 bp DNA ladder (Promega).**



for MTIC333 and MTIC068 while the lowest number (3) was observed for MTIC976 with an average of 9.25 per locus. Nevertheless, only 18 (14%) out of the 126 scored alleles were rare alleles (< 5%). The highest numbers (7 and 5) of rare alleles were, respectively, recorded for MTIC333 and MTIC531 while the lowest value (0) was registered for MTIC1044, MTIC976, MTIC983 and MTIC1050.

The rates of observed and expected heterozygosities were calculated for each locus and population under the assumption of equilibrium Hardy-Weinberg. There is a lack of observed heterozygosity ( $H_o = 0$ ) for all analyzed SSR markers while the values of expected heterozygosity ( $H_e$ ) varied from 0.55 to 0.79, with an average of 0.69.

The highest values of alleles were found in the populations of Enfidha (42), El Kef (43), Bulla Regia (45) and Mateur (45) while a low value was observed in Soliman (37) (Table 2). This is also true for the remaining genetic diversity parameters, while the Soliman population was found to be the less variable. Moreover, the highest number of private alleles were found in Enfidha (19), El Kef (16), Bulla Regia (16) and Mateur (15); while less values was observed for Soliman (8). The mean number of local common alleles ( $\leq 50\%$ ) varied between 0.75 for Enfidha and Mateur and 1.13 for El Kef. The mean inbreeding coefficient estimated as an average across loci and populations was relatively high ( $F_{is} = 1.00$ ), which indicates an excess of homozygotes. Population differentiation ( $F_{ST}$ ) ranged from 0.13 for MTIC983 to 0.33 for MTIC1044 with an average of 0.21. The number of migrants ( $N_m$ ) per generation varied between 0.52 for MTIC1044 and 1.61 for MTIC983 (Table 3).

### Population structure and cluster analysis

AMOVA showed that most of the variation (76%) was observed within populations while only 24% was due to among-populations (Table 4). Therefore, there was nonsignificant variation among ecoregions. In addition, the pairwise matrix of  $F_{ST}$  among populations showed that all the populations are moderately differentiated (Table 5). The high value of  $F_{ST}$  (0.188) was found between populations of El Kef and Soliman while the low value (0.120) was registered between El Kef and Bulla Regia. Moreover, the low genetic variation among populations was accompanied by a relatively high gene flow ( $N_m = 1.47$ ). Furthermore, there was nonsignificant association ( $r = 0.395$ ,  $p = 0.524$ ) between pairwise population differentiation ( $F_{ST}$ ) and geographic distances, suggesting that studied populations are not geographically isolated.

**Table 2. Mean number of different ( $N_a$ ) and effective ( $N_e$ ) alleles, mean number of locally common alleles (frequency  $\leq 25\%$  and frequency  $\leq 50\%$ ), Shannon's information index (I), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), unbiased expected heterozygosity (uHe), fixation index ( $F_{is}$ ), and percentage of polymorphism (% P), for each population.**

Population	$N_a$	$N_a$ freq $\geq 5\%$	$N_e$	Nr private alleles	Nr locally common alleles ( $\leq 25\%$ )	Nr locally common alleles ( $\leq 50\%$ )	I	$H_o$	$H_e$	uHe	$F_{is}$	% P
Enfidha	5.25 (0.70)	4.00 (0.38)	3.50 (0.32)	2.38 (0.38)	0.00 (0.00)	0.75 (0.31)	1.37 (0.10)	0.00 (0.00)	0.70 (0.02)	0.71 (0.02)	1.00 (0.00)	100
El Kef	5.38 (0.84)	4.25 (0.49)	3.68 (0.54)	2.00 (0.65)	0.00 (0.00)	1.13 (0.30)	1.38 (0.15)	0.00 (0.00)	0.68 (0.05)	0.70 (0.05)	1.00 (0.00)	100
Soliman	4.63 (0.50)	3.88 (0.30)	3.14 (0.37)	1.00 (0.27)	0.00 (0.00)	1.00 (0.42)	1.23 (0.13)	0.00 (0.00)	0.64 (0.06)	0.65 (0.06)	1.00 (0.00)	100
Bulla Regia	5.63 (1.13)	4.25 (0.31)	4.18 (0.74)	2.00 (0.71)	0.00 (0.00)	1.00 (0.38)	1.46 (0.15)	0.00 (0.00)	0.72 (0.03)	0.74 (0.03)	1.00 (0.00)	100
Mateur	5.63 (0.84)	3.88 (0.48)	3.69 (0.45)	1.88 (0.58)	0.00 (0.00)	0.75 (0.25)	1.40 (0.15)	0.00 (0.00)	0.68 (0.06)	0.70 (0.06)	1.00 (0.00)	100
Mean	5.30 (0.36)	-	3.64 (0.22)	-	-	-	1.37 (0.06)	0.00 (0.00)	0.69 (0.02)	0.70 (0.02)	1.00 (0.00)	100 (0.00)

**Table 3. Mean number of different ( $N_a$ ) and effective ( $N_e$ ) alleles, number of private and rare (< 5%) alleles, Shannon's information index (I), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), unbiased expected heterozygosity (uHe), fixation index ( $F_{is}$ ), population differentiation ( $F_{ST}$ ), and gene flow ( $N_m$ ), for each locus.**

Locus	$N_a$	$N_e$	Nr Private alleles	Nr Rare alleles ( $< 5\%$ )	I	$H_o$	$H_e$	uHe	$F_{is}$	$F_{ST}$	$N_m$
MTIC1044	3.20 (0.20)	2.51 (0.30)	7	0	0.98 (0.11)	0.00 (0.00)	0.57 (0.07)	0.58 (0.07)	1.00 (0.00)	0.33	0.52
MTIC976	3.20 (0.20)	2.39 (0.26)	3	0	0.95 (0.11)	0.00 (0.00)	0.55 (0.07)	0.56 (0.07)	1.00 (0.00)	0.30	0.57
MTIC531	5.60 (1.08)	4.20 (0.67)	13	5	1.49 (0.17)	0.00 (0.00)	0.74 (0.04)	0.75 (0.04)	1.00 (0.00)	0.19	1.07
MTIC983	4.00 (0.00)	3.01 (0.09)	4	0	1.21 (0.01)	0.00 (0.00)	0.67 (0.01)	0.68 (0.01)	1.00 (0.00)	0.13	1.61
MTIC333	8.20 (1.53)	4.43 (1.23)	14	7	1.62 (0.25)	0.00 (0.00)	0.70 (0.08)	0.72 (0.08)	1.00 (0.00)	0.24	0.80
MTIC068	7.40 (0.51)	4.42 (0.33)	14	4	1.69 (0.06)	0.00 (0.00)	0.77 (0.02)	0.78 (0.02)	1.00 (0.00)	0.16	1.29
MTIC784	5.00 (0.00)	3.39 (0.20)	8	2	1.37 (0.04)	0.00 (0.00)	0.70 (0.02)	0.72 (0.02)	1.00 (0.00)	0.20	0.98
MTIC1050	5.80 (0.37)	4.73 (0.16)	11	0	1.63 (0.05)	0.00 (0.00)	0.79 (0.01)	0.80 (0.01)	1.00 (0.00)	0.16	1.34
Mean	5.30 (0.36)	3.64 (0.22)	9.25 (4.14)	2.25 (3.22)	1.37 (0.06)	0.00 (0.00)	0.69 (0.02)	0.70 (0.02)	1.00 (0.00)	0.21 (0.02)	1.02 (0.14)

**Table 4. Analysis of molecular variance at three hierarchical levels.**

Source of variation	df	SS	Expected variance	Percentage of variation
Among regions	2	80.800	0.000	0
Among populations	2	97.750	0.899	24
Among lines	115	657.833	2.860	76
Within lines	120	0.000	0.000	0
Total	239	836.383	3.759	100

SS: Sum of squared deviation.

**Table 5. Pairwise population differentiation ( $F_{ST}$ ) (above the diagonal) and gene flow (Nm) (below the diagonal) for studied populations of *Medicago polymorpha*.**

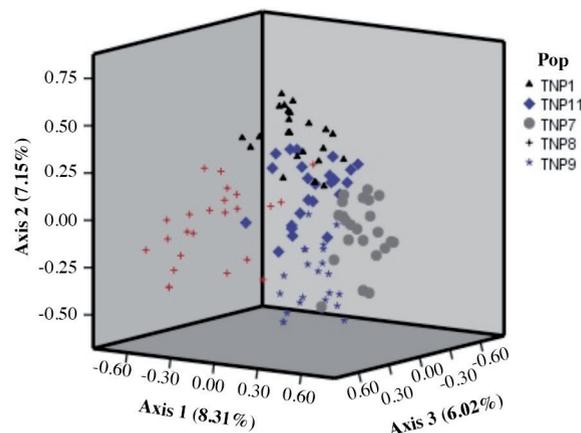
	Enfidha (TNP1)	El Kef (TNP7)	Soliman (TNP8)	Bulla Regia (TNP9)	Mateur (TNP11)
Enfidha (TNP1)	0.00	0.14	0.16	0.15	0.15
El Kef (TNP7)	1.49	0.00	0.19	0.12	0.16
Soliman (TNP8)	1.34	1.08	0.00	0.14	0.14
Bulla Regia (TNP9)	1.40	1.83	1.56	0.00	0.14
Mateur (TNP11)	1.47	1.34	1.60	1.60	0.00

On the other hand, to analyze similarities between studied populations, a principal coordinates analysis (PCoA) was performed using genetic distances between the lines. The population of Soliman formed a separate group while the populations of El Kef and Bulla Regia had the maximum of similarity (Figure 2). Furthermore, to examine genetic relationship of studied populations, hierarchical cluster analysis was performed on the basis of all scores of PCoA. The UPGMA dendrogram showed that the five populations formed three groups (Figure 3). A first group was constituted of Bulla Regia and El Kef, a second group was formed by Enfidha and a third group is composed of Soliman and Mateur.

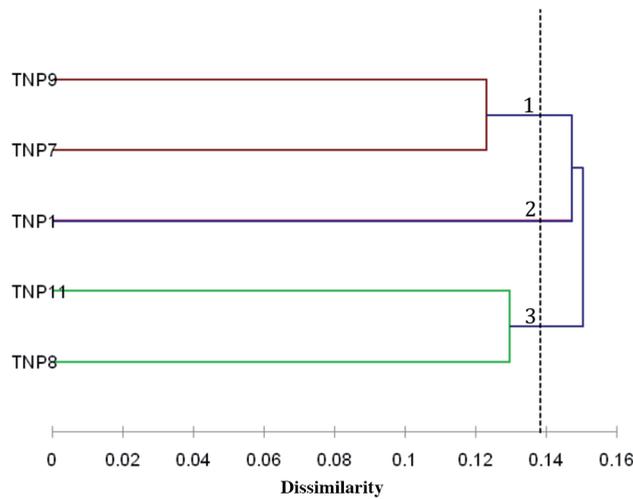
The genetic structure of the populations was also studied using the Bayesian method contained in the STRUCTURE program. Our results showed that the best inferred number of clusters was  $K = 3$  (Figure 4), obtained by the method of Evanno et al. (2005). The first group (green color) is constituted of most lines of Enfidha. The second group (blue color) is formed by most lines from populations of Soliman and Mateur. The third group (green color) was composed of most lines of El Kef and Bulla Regia. This topology was similar to that obtained by the UPGMA dendrogram.

### Molecular vs. quantitative traits

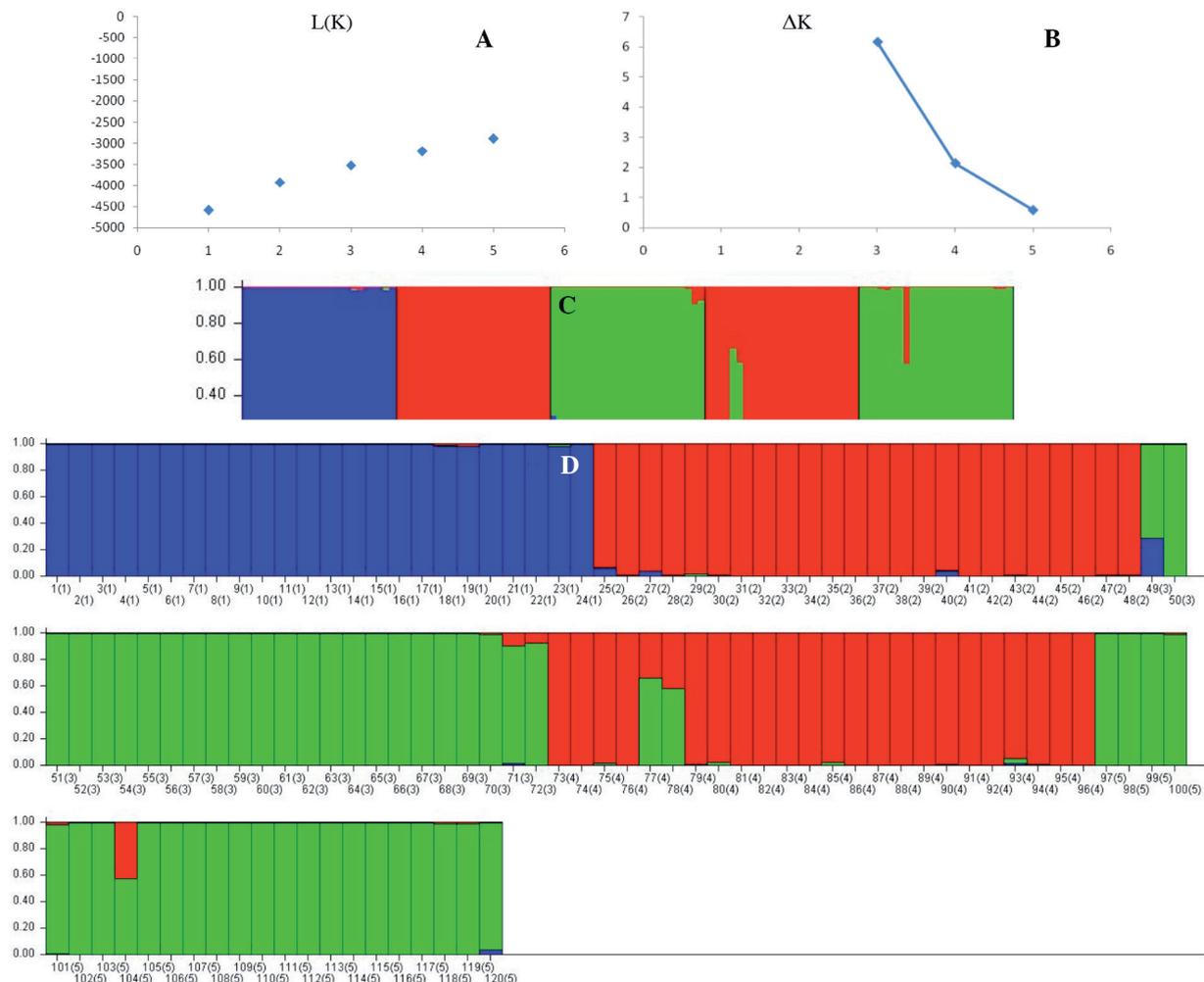
We found that population differentiation among populations is higher at all SSR markers ( $F_{ST} = 0.147$ ) than at quantitative traits ( $Q_{ST} = 0.034$ ), which is consistent with the action of uniform selection.

**Figure 2. Projection of 120 lines from populations of *Medicago polymorpha* in the three first axes of principal coordinate analysis (PCoA) based on Nei's genetic distances. Populations: Enfidha (TNP1), El Kef (TNP7), Soliman (TNP8), Bulla Regia (TNP9), and Mateur (TNP11).**

**Figure 3.** UPGMA derived dendrogram of the five populations of *Medicago polymorpha* based on the pairwise population differentiation ( $F_{ST}$ ) values. Populations: Enfidha (TNP1), El Kef (TNP7), Soliman (TNP8), Bulla Regia (TNP9), and Mateur (TNP11).



**Figure 4.** (A) Mean  $L(K)$  (SD) over 10 runs for each  $K$  value. (B) The ad hoc quantity  $\Delta K$  calculated as  $\Delta K = mL''(K)/s[L(K)]$ . (C) Estimated genetic structure of 120 lines of *Medicago polymorpha* assessed by STRUCTURE (Pritchard et al., 2000) for  $K = 3$ . Each individual is represented by a thin vertical bar, partitioned into up to  $k$  colored segments. Populations: Enfidha (TN1), El Kef (TN7), Soliman (TN8), Bulla Regia (TN9), and Mateur (TN11).



## DISCUSSION

The cross amplification of SSR markers in different studies showed contrasted results, according to the evolutionary distance and complexity of the genome (Kalia et al., 2011). In this study, our results showed that among the 112 SSR markers developed in *M. truncatula*, 50 SSRs could be amplified in *M. polymorpha*. This amplification percentage was relatively low compared to amplification rates reported for *M. laciniata* (Badri et al., 2007) and *M. ciliaris* (Badri et al., 2008a) where 39 and 17 out of 52 and 30 were amplified, respectively, in the two species. However, Gutierrez et al. (2005) reported amplification percentages of 39%-43% and 21%-24%, respectively, for the SSRs from ESTs (expressed sequence tags) banks and those of genomic origins in pea (*Pisum sativum*), beans (*Vicia faba*) and chickpea (*Cicer arietinum*). The moderate percentage of cross amplification of SSR markers in *M. polymorpha* suggests interruptions that may occur in the process of repeating sequences and mutations that can occur in sequences of repeated SSR motifs. Moreover, despite the moderate percentage of polymorphic markers (16%), our results showed a sufficient level of variation to study the diversity and genetic structure in studied populations.

Overall, the values of the observed heterozygosity ( $H_o$ ) are lower than those of the expected heterozygosity ( $H_e$ ) for all populations. These results are expected given the nature of the reproductive system of *M. polymorpha* which is preferentially autogamous (Badri et al., 2016). In the same sense, the values of the fixation index ( $F_{is}$ ) indicate an overall deficit in heterozygosity in this species. In addition, a high level of genetic diversity between populations was detected ( $I = 1.37$ ;  $uHe = 0.70$ ). The values vary considerably among populations, where the lowest values were recorded for the Soliman population ( $I = 1.23$ ;  $uHe = 0.65$ ). These values are higher than those observed in *Hordeum marinum* (on average  $I = 0.36$  and  $uHe = 0.25$ ) (Saoudi et al., 2019).

Results from AMOVA showed that most of the genetic variation (76%) was registered within populations. Accordingly, similar results were also reported for autogamous wild species (Badri et al., 2007; 2008a; 2008b; Neji et al., 2015; Shakhathreh et al., 2016; Saoudi et al., 2019). In addition, a moderate population differentiation ( $F_{ST} = 0.147$ ) was observed for SSR markers. This value is consistent with those detected for *M. ciliaris* (0.17) (Badri et al., 2008a) and *Hordeum vulgare* (0.13) (Guo et al., 2012). Nevertheless, higher values of differentiations were registered for *M. laciniata* ( $F_{ST} = 0.47$ ) (Badri et al., 2008b) and *M. truncatula* (Lazrek et al., 2009). Our results showed that there is not a significant association of population differentiation ( $F_{ST}$ ) with geographic distances. Similar observations have been reported in natural populations of *M. ciliaris* using SSR markers (Badri et al., 2008a; Saoudi et al., 2019) and in *M. polymorpha* using allozyme loci (Hannachi et al., 1998), and in autogamous poaceae using SSR markers (Neji et al., 2015). However, numerous studies have reported coherence between genetic distance and eco-geographic factors (Badri et al., 2008b; Guo et al., 2012; Shakhathreh et al., 2016; Pico-Mendoza et al., 2020).

PCoA and the Bayesian clustering results showed that the 120 lines did not cluster into distinct groups according to their sites of origin. Indeed, the populations of Enfidha and Soliman which grow naturally in saline sites did not show the highest level of similarities. Cluster analysis of molecular and morphological markers is useful to establish a core collection of *M. polymorpha* lines that will be analyzed for tolerance to salt and drought stress constraints in greenhouse. The most tolerant lines will be cultivated in the field on different sites in Tunisia and they will be analyzed for their forage quality. The lines with the best agronomic performance will be recorded and provided to farmers. In addition, the core collection of this species will be conserved at the Tunisian National Gene Bank. On the other hand, our results showed that population differentiation at molecular markers ( $F_{ST} = 0.147$ ) is higher than that at quantitative traits ( $Q_{ST} = 0.034$ ), which is consistent with the action of uniform selection. Furthermore, similar results were detected for *H. marinum* showing higher population differentiation at RAPD markers (0.315) (Saoudi et al., 2019) than at phenotypic characters ( $Q_{ST} = 0.149$ ) (Saoudi et al., 2017).

## CONCLUSIONS

This study aimed to analyze the molecular polymorphism in five populations of *Medicago polymorpha* sampled in different eco-geographic regions in Tunisia using SSR markers. A moderate percentage (44.64%) of microsatellite markers of *M. truncatula* amplified in *M. polymorpha*.

Our results showed that most of the polymorphism resides within populations (76%). Moderate levels of populations differentiation ( $F_{ST} = 0.12$  to  $0.19$ ) were recorded. Population differentiation is independent of geographic distances, which suggests that studied populations are not geographically isolated. The five populations were clustered into three groups. Furthermore, the comparison of the levels of population differentiation at molecular markers ( $F_{ST}$ ) and quantitative traits ( $Q_{ST}$ ) argues in favor of the action of stabilizing selection by promoting the same phenotype in different populations.

## ACKNOWLEDGEMENTS

We thank Dr. Marshall Keyster for a critical reading of the manuscript and useful comments. We are also grateful to Warda Saoudi, Imene Haddoudi and Asma Bourougaaoui for technical help. This study was funded by the Tunisian-Algerian bilateral project (2013-2015) and by the Tunisian Ministry of Higher Education and Scientific Research (CBBC02 LR15).

## REFERENCES

- Badri, M., Ben Cheikh, N., Mahjoub, A., and Abdelly, C. 2016. Morpho-phenological diversity among natural populations of *Medicago polymorpha* of different Tunisian ecological areas. *African Journal of Biotechnology* 15:1330-1338.
- Badri, M., Ilahi, H., Huguet, T., and Aouani, M.E. 2007. Quantitative and molecular genetic variation in sympatric populations of *Medicago laciniata* and *M. truncatula* (Fabaceae): relationships with eco-geographical factors. *Genetics Research* 89:107-122.
- Badri, M., Zitoun, A., Ilahi, H., Huguet, T., and Aouani, M.E. 2008a. Morphological and microsatellite diversity associated with ecological factors in natural populations of *Medicago laciniata* Mill. (Fabaceae). *Journal of Genetics* 87:241-255.
- Badri, M., Zitoun, A., Soula, S., Ilahi, H., Huguet, T., and Aouani, M.E. 2008b. Low levels of quantitative and molecular genetic differentiation among natural populations of *Medicago ciliaris* Kroch. (Fabaceae) of different Tunisian eco-geographical origin. *Conservation Genetics* 9:1509-1520.
- Edelaar, P., Burraco, P., and Gomez-Mestre, I. 2011. Comparisons between  $Q_{ST}$  and  $F_{ST}$ - how wrong have we been? *Molecular Ecology* 20:4830-4839.
- Emberger, L. 1955. Une classification biogéographique des climats. *Recueil des Travaux des Laboratoires de Botanique, Géologie, et Zoologie. Série Botanique* 7. 43 p. Faculté des Sciences, Université de Montpellier, Montpellier, France.
- Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14:2611-2620.
- Exbrayat, S., Bertoni, G., Naghavi, M.R., Peyghambari, A., Badri, M., and Debelle, F. 2014. Genetic variability and identification of quantitative trait loci affecting plant growth and chlorophyll fluorescence parameters in the model legume *Medicago truncatula* under control and salt stress conditions. *Functional Plant Biology* 41(9):983-1001.
- Geuna, F., Toschi, M., and Bassi, D. 2003. The use of AFLP markers for cultivar identification in apricot. *Plant Breeding* 122:526-531.
- Gholami, A., De Geyter, N., Pollier, J., Goormachtig, S., and Goossens, A. 2014. Natural product biosynthesis in *Medicago* species. *Natural Product Reports* 31(3):356-380.
- Grover, A., Ramesh, B., and Sharma, P.C. 2009. Development of microsatellite markers in potato and their transferability in some members of Solanaceae. *Physiology and Molecular Biology of Plants* 15(4):343-358.
- Guo, Y., Li, Y., Huang, Y., Jarvis, D., Sato, K., Kato, K., et al. 2012. Genetic diversity analysis of hullless barley from Shangri-la region revealed by SSR and AFLP markers. *Genetic Resources and Crop Evolution* 59:1543-1552.
- Gutierrez, M.V., Vaz Patto, M.C., Huguet, T., Cubero, J.I., Moreno, M.T., and Torres, A.M. 2005. Cross species amplification of *Medicago truncatula* microsatellites across three major pulse crops. *Theoretical and Applied Genetics* 110:1210-1217.
- Hannachi, A.S., Boussaid, M., and Marrakchi, M. 1998. Genetic variability organization and gene flow in natural populations of *Medicago polymorpha* L. prospected in Tunisia. *Genetic Selection Evolution* 30:121-135.
- Kalia, R.K., Rai, M.K., Kalia, S., Singh, R., and Dhawan, A.K. 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177:309-334.
- Lazrek, F., Roussel, V., Ronfort, J., Cardinet, G., Chardon, F., Aouani, M.E., et al. 2009. The use of neutral and non-neutral SSRs to analyze the genetic structure of a Tunisian collection of *Medicago truncatula* lines and to reveal associations with eco-environmental variables. *Genetica* 135:391-402.

- Liu, F.-M., Hong, Z., Yang, Z.-J., Zhang, N.-N., Liu, X.-J., and Xu, D.-P. 2019. De Novo transcriptome analysis of *Dalbergia odorifera* T. Chen (Fabaceae) and transferability of SSR markers developed from the transcriptome. *Forests* 10(2):98.
- Mantel, N. 1967. The detection of disease clustering as a generalized regression approach. *Cancer Research* 27:209-220.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583-590.
- Neji, M., Geuna, F., Taamalli, W., Ibrahim, Y., Chiozzotto, R., Abdelly, C., et al. 2015. Assessment of genetic diversity and population structure of Tunisian populations of *Brachypodium hybridum* by SSR markers. *Flora* 216:42-49.
- Nogueira, A.M., Ferreira, A., and da Silva Ferreira, M.F. 2016. Transferability of microsatellites from *Psidium guajava* to *Eugenia*, *Myrciaria*, *Campomanesia*, and *Syzygium* species (Myrtaceae). *Plant Molecular Biology Reporter* 34:249-256.
- Peakall, R., and Smouse, P.E. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-295.
- Pico-Mendoza, J., Garcia-González, R., Quiroz, K., Pinoargote, M., Rodríguez-Alvarez, Y., Chong, B., et al. 2020. Genetic diversity and population structure of Chilean blueberry *Gaultheria pumila* (L.f.) D.J. Middleton (Ericaceae). *Chilean Journal of Agricultural Research* 80:413-421.
- Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- Remya, K.S., Sigimol, J., Lakshmi, P.K., and Akhila, S. 2010. Microsatellites in varied arenas of research. *Journal of Pharmacy and Bioallied Sciences* 2(2):141-143.
- Saoudi, W., Badri, M., Gandour, M., Smaoui, A., Abdelly, C., and Taamalli, W. 2019. Analysis of genetic diversity and spatial structure in Tunisian populations of *Hordeum marinum* ssp. *marinum* based on molecular markers. *The Journal of Agricultural Science* 157:399-412.
- Saoudi, W., Badri, M., Gandour, M., Smaoui, A., Abdelly, C., and Taamalli, W. 2017. Assessment of genetic variability among Tunisian populations of *Hordeum marinum* using agronomic traits. *Crop Science* 57:302-309.
- Shakhatreh, Y., Baum, M., Haddad, N., Alrababah, M., and Ceccarelli, S. 2016. Assessment of genetic diversity among Jordanian wild barley (*Hordeum spontaneum*) genotypes revealed by SSR markers. *Genetic Resources and Crop Evolution* 63:813-822.
- Shirasawa, K., Oyama, M., Hirakawa, H., Sato, S., Tabata, S., Fujioka, T., et al. 2011. An EST-SSR linkage map of *Raphanus sativus* and comparative genomics of the Brassicaceae. *DNA Research* 18:221-232.
- Targonska, M., Bolibok-Bragoszewska, H., and Rakoczy-Trojanowska, M. 2016. Assessment of genetic diversity in *Secale cereale* based on SSR markers. *Plant Molecular Biology Reporter* 34:37-51.
- Uchiyama, K., Fujii, S., Ishizuka, W., Goto, S., and Tsumura, Y. 2013. Development of 32 EST-SSR Markers for *Abies firma* (Pinaceae) and their transferability to related species. *Applications in Plant Sciences* 1(2):1200464.
- Whitlock, M.C., and Guillaume, F. 2009. Testing for spatially divergent selection: Comparing  $Q_{ST}$  to  $F_{ST}$ . *Genetics* 183(3):1055-1063.
- Zeid, M., Yu, J.K., Goldowitz, I., Dentond, M.E., Costich, D.E., Jayasuriyaa, C.T., et al. 2010. Cross-amplification of EST-derived markers among 16 grass species. *Field Crops Research* 118:28-35.
- Zhang, M., Fan, L., Liu, Q., Song, Y., Wei, S., Zhang, S., et al. 2014. A novel set of EST-derived SSR markers for pear and cross-species transferability in Rosaceae. *Plant Molecular Biology Reporter* 32:290-302.