Journal of Medicinal Plants and By-products (2021) 1: 19-35

Original Article

Using Morphological and Phytochemical Traits and *ITS* (1, 4) and *rbcl* DNA Barcodes in the Assessment of Different *Malva sylvestris* L. Genotypes

Anahita Pahlavan[1], Mahmoud Solouki[2], Baratali Fakheri[2] and Bahman Fazeli-Nasab[3*]

[1]*Department of Horticultural, Faculty of Agriculture, university of Zabol, Zabol, Iran*
[2]*Department of Breeding and Biotechnology, Faculty of Agriculture, University of Zabol, Zabol, Iran*
[3]*Research Department of Agronomy and Plant Breeding, Agricultural Research Institute, University of Zabol, Zabol, Iran*

Article History: Received: 01 July 2020/Accepted in revised form: 23 September 2020

Abstract

Since *rbcl* gene and protein, as well as internal transcribed spacer (ITS1,4) regions, have not been used in the evaluation of *Malva sylvestris* L., this study was aimed to assess different Iranian *Malva sylvestris L.* genotypes by evaluation of morphological and phytochemical traits, along with *rbcl* and *ITS*1,4 gene sequences. Furthermore, three-dimensional and functional structures of *Malva sylvestris L. rbcl* protein were examined, as well. Nine *Malva sylvestris* L. genotype samples were collected from different regions of Iran. After species identification, completely randomized design with three replicates was used to evaluate different genotypes, based on morphological and phytochemical traits. DNA extraction was carried out using the SDS method and then final PCR products were sent to Macrogen Company, South Korea for sequencing. Sequence quality was assessed using Chromas 2.1.1 software. Then, the sequences were aligned using the ClustalW method in MegAlign 5 software and the dendrogram of the phylogenetic relationships and similarity matrices were plotted, as well. SWISS-MODEL and QMEAN servers were used for modeling and validation of *rbcl* protein. Ramachandran plot analysis and Pro-SA servers were used to evaluate the structure and chemical quality of the protein. Comparison of the mean physio-morphological traits between different genotypes showed the highest stem diameter (9.58 mm), root length (61.22 cm), root fresh weight (18.86 g), root dry weight (4.84 g) as well as proline content (0.614mg/gDW) in Mashhad genotype. Based on stepwise regression results in the presented models, root fresh weight and plant dry weight had the most positive effect on root length, but stem diameter and plant fresh weight had the most negative effect. Moreover, while chlorophyll b had the most negative and direct effect on proline function, chlorophyll a, carotenoids, carbohydrates and total protein contents had the most positive effects, respectively. Assessment of protein-protein interaction networks revealed that most proteins encoded by *matK*, *psb-tranH* genes interact with *rbcl* protein. The results of cluster analysis, similarity matrix as well as dN/dS ratio showed high similarity and conservation of *ITS* and *rbcl* sequences among different *Malva sylvestris* L. genotypes. *rbcl* and *ITS* sequences were not suitable markers to evaluate phylogenetic relationships intraspecies (at subspecies level) but are useful to evaluate interspecies relationships. Furthermore, the Mashhad genotype is suitable for dry and water deficit conditions, withstanding such conditions. Therefore, it is recommended that this genotype can be used both as a parent and or directly to breed, develop and modify *Malva sylvestris* L. species.

Keywords: *Malva sylvestris,* dN/dS ratio, *ITS,* Phylogeny, *rbcl*

Introduction

Malva sylvestris L. (*Malvaceae*) is an annual, biennial or rarely perennial plant [1]. Though Central Asia has been reported as it's the origin, it is growing almost everywhere in the world [2]. The medicinal Malvaceae family has 10 species [3] of which 7 species have been reported in Iran and three of them are endemic [4]. It is

[*]Corresponding author: Research Department of Agronomy and Plant Breeding, Agricultural Research Institute, University of Zabol, Zabol, Iran
Email Address: bfazeli@uoz.ac.ir

known as Panirak, Nun-kalaagh, Khabaazi and Tooleh in Iran [5], Common mellow in Italy [6], Marva in Europe, Ebegumeci and Gomecotu in Turkey [7] and Vilayatti-kangi or Gulkhaira in India and Pakistan [8].

Malva sylvestris L. is rich in vitamins A, B and C and is one of the herbs effective in reducing the complications of the common cold especially cough, as well as in the treatment of respiratory inflammation, urinary tract infection, digestive tract problems and skin rashes [9]. The chemical constituents of *Malva sylvestris* L. leaf include polyphenol, vitamin C and E, beta-carotene, glosid as well as flavonoid [10].

Iran's climatic diversity has created diverse habitats of plant species [11]. On the other hand, evaluation and specification of genetic diversity not only conserve genetic resources but also is essential to achieve initial diversity in order to promote the efficiency of the breeding programs [12]. Furthermore, genetic diversity is crucial for species adaptation to environmental changes and long-term survival. Therefore, awareness and knowledge of inter- and intra-population variation is of particular importance in natural resources conservation management [13].

As environmental factors influence the quality and quantity of crop essential oils by affecting plant vegetative growth and physiology, besides genetic factors, culture environmental conditions are also important to grow medicinal plants [14]. In this regard, though *Malva sylvestris* L. is one of the medicinal plants for the therapy of a variety of diseases, limited research on this herb has been conducted in Iran and even in the world. Therefore, it seems quite necessary to carefully investigate Iranian endemic populations and to prepare a seed certificate for future plant breeding researches.

In plant systematics, one of the concerns of specialists in identifying plant species is the decline in biodiversity, the occurrence of hybrids and chromosomal duplications and the acquisition of a quick, precise and inexpensive method [15-17]. The study of genetic diversity of medicinal plants is of particular importance[18]. Genetic diversity reflects inter and intraspecies differences which its study plays an important role in breeding programs in order to improve the quantitative and qualitative traits of plants [19]. Moreover, genetic identification and registration of different plant varieties are the major aspects of the utilization and conservation of genetic resources [20]. Since studying morphological traits is not very accurate due to the involvement of environmental impacts, DNA-based markers are considered as good tools to assess genetic diversity as they study variation at the DNA level and independently of the environment [21].

Some researchers have recognized DNA-based molecular markers as the best tool to identify plant species [19,22, 23]. *rbcl* and *matk* gene loci in the chloroplast genome have been accepted as the standard barcodes in plants [24]. Both genes play a significant role in the phylogenetic regeneration of terrestrial plants due to their strong phylogenetic signal [25-27]. Chloroplast DNA sequences are suitable barcodes for phylogenetic studies due to the involvement of a number of traits including presence in high copy number in the cell, occurrence of conserved genes and having general primers and high reproducibility [28]. Therefore, a number of studies have used chloroplast diversity to inspect intra and inter-species evolution and to specify kinship at the species level [25, 27]. Among the various chloroplast genes, the *rbcl* gene (encoding the large subunit of Rubisco enzyme, playing an important role in the photosynthetic carbon reversal cycle during photosynthesis) has been used more to inspect kinship relationships [26]. *rbcl*, *matK*, *psbA-trnH* and *ITS* barcodes have been investigated in the identification and documentation of plant diversity [28]. *ITS* molecular marker has been applied in phylogenetic studies for more than 10 years [29]. *rbcl* gene is also a valuable tool to assess phylogenetic relationships, being present in the chloroplasts of most photosynthetic organisms. *rbcl* is an abundant protein in leaf tissue and a common factor among photosynthetic organisms. It can be compared among different plants to determine genetic similarities and diversities.

Since *rbcl* gene and protein and *ITS*1,4 regions have not been used in *Malva* evaluation and this study was aimed Using morphological and phytochemical traits and *ITS* (1, 4) and *rbcl* DNA barcodes to assessment of different *Malva sylvestris* L. genotypes and also, three-dimensional and functional structures of *Malva sylvestris L. rbcl* protein were examined, as well.

Material and Methods

Nine *Malva sylvestris* L. genotypes were collected from different regions of Iran (Mashhad, Torbat Heydariyeh, Freeman, Zabol, Zarand, Jiroft, Rudbar, Bandar Abbas and Khorramdasht) in 2018 (Table 1) and subsequently identified in the botanical laboratory of Torbat-e-Heydariyeh University. Then, a completely randomized design with three replicates was conducted in the Agricultural Research Institute of Zabol University, Zabol, Iran in order to evaluate morphological and phytochemical traits. Seeds of each genotype were sown in five-liter pots containing a mixture of cropland soil, coco peat, perlite, and fully decayed animal manure.

Measurement of Morphological Traits

At the full flowering stage, the stem diameter and length of three plants from each pot were measured randomly.

Table 1 Characteristics and origin of *Malva sylvestris* L. genotypes

Name	Accession Number		Geographic coordinate			Address	
	rbcl	*ITS*	Latitude	Longitude	Elevation	City	Province
Malva sylvestris	MK521567	MK495371	6, 20′, 9.8′′	59, 32′, 40′′	1050	Mashahd	Razavi Khorasan
Malva sylvestris	MK521568	MK495430	27, 11′, 43′′	56, 17′, 1.4′′	25	Bandar Abbas	Hormozgan
Malva sylvestris	MK521565	MK495365	35, 16′, 56.6′′	59, 13′, 16.8′′	1365	Torbat Heydariyeh	Razavi Khorasan
Malva sylvestris	MK521571	MK497803	35, 16′, 56.6′′	59, 13′, 16.8′′	1404	Fariman	Razavi Khorasan
Malva sylvestris	MK521566	MK495364	30, 48′, 38.7′′	56, 33′, 57.09′′	1657	Zarand	Kerman
Malva sylvestris	MK521569	MK497792	28, 2′, 19.5′′	57, 59′, 13.8′′	488	Rudbar	Kerman
Malva sylvestris	MK521570	MK496052	28, 40′, 46.8′′	57, 44′, 43.11′′	680	Jiroft	Kerman
Malva sylvestris	MK521572	MK496176	35, 43′, 20.8′′	51, 15′, 49.4′′	1277	Khorramdasht	Tehran
Malva sylvestris	MK521573	MK497803	31, 1′, 51.07′′	61, 29′, 49.8′′	482	Zabol	Sistan and Baluchestan

Stem diameter was determined using a caliper and stem length from crown to the end of the tallest stem was measured using a standard ruler. The mean diameter and plant height were considered for each flower pot. The number of flowers with seeds and the number of leaves per plant were counted, as well. After separating 3 plants from each flower pot, fresh and dry root weights and plant weights were determined. Root and plant fresh and dry weight was measured using a digital scale with an accuracy of 0.01g. to scale dry weight Fresh samples were placed in an oven at 70 °C for 48 hours [30].

Measurement of Physiological Traits

The concentration of photosynthetic pigments was determined according to the Arnon method [31], using the following equations. The absorbance rate of the samples for chlorophyll a, b and carotenoid was determined using a spectrophotometer at the wavelengths of 663 nm, 645 nm and 470 nm, respectively.

Chlorophyll a = (12.7 (A663)) – ((2.69) (A645)) (V/1000(W))

Chlorophyll b = (22.9(A645)) – (4.88(A663)) (V/1000(W))

Carotenoids = A480 +(0/114(A663)- (0/638(A645)

, where A is the absorbance intensity at the corresponding wavelength in nanometers, V is the volume of the solution and W is the leaf sample weight.

Measurement of proline was carried out according to Bates *et al*. method [32] soluble carbohydrates according to Kells and Ancel method [33], and protein according to the Bradford method [34]

DNA Extraction and Gel Electrophoresis

The Leaves of different *Malva sylvestris* L. genotypes were harvested at the four-leaf stage. DNA extraction was carried out using the SDS method [35]. Quantification and qualification of the extracted DNA samples were performed using 260/280 OD ratio analysis and 1% agarose gel electrophoresis, respectively.

Polymerase Chain Reaction (PCR)

ITS and *rbcl* gene loci were amplified using primers as follows: *ITS*1 forward primer: TCCGTAGGTGAACCTGCGG, *ITS*2 reverse primer: TCCTCCGCTTATTGATATGC [36]; *rbcl*; forward primer: ATGTCACCACAAACAGAGACTAAAGC [37], *rbcl*; reverse primer: GTAAAATCAAGTCCACCRCG [38]. PCR reaction was carried out in a final volume of 45µl containing 24µl master mix (Amplicon Denmark), 6µl template DNA and 1.5µl of both forward and reverse primers. The reaction mixture volume reached to 45µl by adding 12µL of deionized sterile water. PCR reaction was performed in a thermocycler (Hamburg, Germany) with the following conditions: initial denaturation at 95 °C for 60s, followed by 35 cycles of denaturation at 95 °C for 60s, annealing at 72 °C for *ITS* and at 56.1 °C for *rbcl* for 60s, and final extension at 72 °C for 10 min. Finally, to ensure successful amplification, PCR products were loaded on 2% agarose gel, stained with Gel Red Staining.

Sequencing and Data Analysis

Final PCR products were purified from agarose gel using the AccuPrep DNA Gel Purification Kit. Nine samples were sent to Macrogen Company, South Korea for paired-end sequencing. Sequence quality was assessed using Chromas 2.1.1 software. Because each nucleotide is read twice during the paired-end sequencing process, probable errors, especially at the beginning and end of the sequence, as well as other errors including weak sequence, signal drop at the beginning or end of the sequence, signal drop after a homopolymer region, top-heavy signal, noisy read, delayed and unusual reads before correct sequence readings, and large raw signal spikes were corrected with the above-mentioned softwares. Each sequence was Blasted with the NCBI database sequences using the BLAST online tool [39], based on E Value, Query Cover, etc.

Sequence alignment, clustering and determination of genetic homology and distance were performed by the Clustal W method using MegAlign5 software from the Dnastar5 software package and Mega6 software. DnaSP5 software was used for further analysis including dn and ds calculation. Nucleotide substitution was determined based on the Tamura-Nei model [40] as transitional and transversal type substitutions.

Protein Modeling and Validation

Predicting the 3D structure of the *rbcl* protein through amino acid sequences is one of the major issues in structural computational biology. The accuracy of the predicted model was examined by the SWISS-MODEL server [41]. Ramachandran and Pro-SA server maps were used to assess the structure and quality of the stereochemistry of both wild and mutated models. The Pro-SA server determines the quality of the model by calculating the Z-score. Z-score is a measure of the overall quality of the model [42]. To inspect the proteins that interacted with *rbcl*, the amino acid sequence of *rbcl* was run in string software [42], then the protein network was predicted.

Statistical Analysis

Analysis of Variance (ANOVA) test was used to analyze morphological and phytochemical data. Mean comparisons were performed by Duncan's test at 1 and 5% probability level. All statistical analysis was conducted using SAS9 and Excel software.

Results

The results of variance analysis showed the difference in various *Malva sylvestris* L. genotypes based on the morphological and phytochemical traits (P <0.01). Duncan's post hoc test results showed a significant advantage of some genotypes over others in terms of different plant characteristics: Mashhad genotype for stem diameter, root length, root fresh weight and root dry weight; Rudbar genotype for plant weight; Bandar Abbas genotype for plant dry weight, plant weight, number of leaves per plant and number of flowers and seeds; Torbat Heydariyeh genotype for number of flowers and seeds (Table 2).

The highest (40.55 cm) and the lowest (1.81 cm) stem height belonged to the Jiroft genotype and Mashhad genotype, respectively. Increasing plant height promotes sunlight absorption by the formation of more leaves and improves plant yield by augmenting photosynthetic materials production [43]. The highest and the lowest stem diameter were shown in the Mashhad genotype and Khorramdasht genotype, respectively. The highest (61.22 cm) and the lowest (9.55 cm) root length were observed in the Mashhad genotype and Rudbar genotype, respectively (Table 2).

The highest and the lowest fresh and dry root weight belonged to the Mashhad genotype and Jiroft genotype, respectively. The highest and the lowest fresh plant weight were observed in the Rudbar genotype and Khorramdasht genotype, respectively. The highest and the lowest plant dry weight and leaf number per plant belonged to the Bandar Abbas genotype and Khorramdasht genotype, respectively. The highest number of flowers was observed in the Torbat-e-Heydariyeh genotype and the highest number of seeds was observed in Bandar Abbas and Torbat-e-Heydariyeh genotypes (Table 2).

Duncan post hoc test results showed the highest chlorophyll a, chlorophyll b and carotenoids content in the Freeman genotype, the lowest chlorophyll a and b content in Mashhad and Rudbar genotypes and the lowest carotenoids content in Bandar Abbas genotype. The highest proline and carbohydrate content were observed in Mashhad and Rudbar genotypes, respectively, the lowest proline content in the Zabol genotype, and the lowest carbohydrate content in the Torbat Heydariyeh genotype. While the highest total protein content was found in the Jiroft genotype, the lowest total protein content was observed in Freeman, Torbat Heydariyeh, Bandar Abbas and Khorramdasht genotypes (Table 3).

The results of stepwise regression coefficients showed that stem diameter, root weight, plant weight, plant dry weight had the greatest effect on root length and explained 0.92% of the changes in this trait. Root length and fresh root weight and the number of leaves per plant, the number of seeds per plant, and plant weight had the greatest effect on plant dry weight and explained 0.76% of the changes. Also, plant height from the crown, root length and root weight, weight Plant dryness, number of leaves per plant, and number of flowers had the greatest effect on plant fresh weight and 0.84% explained the changes in these traits (data not shown).

The most direct and positive effects related to root weight and plant dry weight had root effect on root length yield and stem diameter and plant weight had a negative effect on root length. The negative weight of plant weight and stem diameter on root length trait confirms that with increasing root length, stem diameter, and plant weight decrease. Root length and number of leaves per plant and number of seeds per plant and plant weight had a positive effect and root weight had a negative effect on plant dry weight. Also, the height of plant collar and root weight and plant dry weight and number of leaves per plant, and the number of flowers had an effect Positive and root length had a negative effect on the heavier yield of the plant. Negative root length on the heavier trait of the plant confirms that with increasing plant weight, the root length decreases (data not shown).

Regression analysis for proline function as a variable of the results showed that 5 traits of chlorophyll a, chlorophyll b, carotenoids, carbohydrates, and protein were entered into the model, respectively, and explained a total of 0.51% of the changes in proline trait. According to the obtained regression coefficients, it can be seen that the highest regression coefficients are related to the carotenoid trait (data not shown).

The simple correlation coefficient was calculated for the studied traits. A significant positive correlation was observed between carotenoid and chlorophyll a (0.612) and chlorophyll b contents (0.783). The same was observed between chlorophyll a and chlorophyll b (0.532) and between proline and carbohydrate (0.404), as well (Table 4). It should be noted that no significant correlation was observed between carbohydrate, total protein, carotenoid, chlorophyll b and chlorophyll content.

After sequencing, the homology of *rbcl* and *ITS* sequences were measured with KM360873 and EF419482 reference sequences, respectively, available at the NCBI database.

The homology percentage was recognized as 100% for *ITS* and 99.3% for *rbcl*. BLASTn analysis results of *ITS* and *rbcl* sequences in collected plant samples confirmed all samples as *Malva Sylvestris* L., a result that was in accordance with species identification based on plant morphological traits. All sequences with the relevant accession number were submitted to the NCBI database (Table 1). The sequences were aligned following verification and submission to the NCBI database. The results showed a 552 bp conserved region in *ITS* locus and two 440 bps and 127 bps conserved regions in *rbcl* locus.

Table 2 Mean comparison of the mean of the morphological traits among different *Malva sylvestris* L. genotypes

Genotypes	Plant height from crown	Stem diameter	Root Length	Fresh root weight	dry root weight	fresh plant weights	Dry plant weight	number of leaves per plant	Number of flowers	Number of seed
Zabol	33.66 b	3.54 d	13.44 ef	1.18 ef	0.48 f	18.18 c	4.08 b	30.66 ab	20.11 a	17.11 b
Bandar Abbas	36.77 ab	3.93 d	15.00 f	0.62 f	0.27 de	24.23 a	5.64 a	30.77 a	23.77 a	20.55 a
Jiroft	40.55 a	3.75 d	17.22 de	0.48 f	0.20 e	19.02 bc	3.72 bc	22.88 c	14.33 b	12.33 c
Rudbar	33.55 b	4.14 cd	9.55 f	0.65 f	0.27 de	25.49 a	3.14 bcd	24.00 c	14.22 b	12.11 c
Zarand	3.61 c	5.51 c	21.88 cd	5.62 bc	2.23 bc	13.63 d	2.25 d	12.00 d	20.11 a	17.11 b
Torbat Heydariyeh	3.33 c	3.88 d	23.44 c	3.16 de	0.82 d	10.41 de	2.84 cd	8.22 d	23.77 a	20.55 a
Fariman	2.27 c	7.74 b	29.22 b	6.85 b	2.70 b	18.44 c	3.55 bc	24.44 c	14.33 b	12.33 c
Mashhad	1.81 c	9.58 a	61.22 a	18.86 a	4.84 a	23.23 ab	3.37 bc	24.66 bc	14.22 b	12.11 c
Khorramdasht	3.55 c	3.71 d	32.66 b	4.69 cd	1.69 c	7.59 e	2.23 d	7.44 d	20.11 a	17.11 b

Meanings with similar letters were not statistically significant.

Table 3 Comparison of the mean of the physiological traits among different *Malva sylvestris* L. genotypes

Genotypes	chlorophyll A	chlorophyll b	carotenoid	protein	carbohydrate	proline
Zabol	0.877 Bcd	0.429 cdef	1.114 bc	0.237 b	1.350 ef	0.154 f
Bandar Abbas	0.937 bcd	0.401 def	0.975 c	0.143 c	1.748 ab	0.465 cd
Jiroft	0.907 bcd	0.565 ab	1.136 bc	0.369 a	1.625 abc	0.401 d
Rudbar	0.862 cd	0.492 bcd	1.222 b	0.271 b	1.850 a	0.552 ab
Zarand	0.754 d	0.390 ef	1.076 bc	0.243 b	1.570 bcd	0.501 bc
Torbat Heydariyeh	0.902 bcd	0.465 cde	1.129 bc	0.180 c	1.125 f	0.233 ef
Fariman	1.330 a	0.634 a	1.481 a	0.136 c	1.427 cde	0.554 ab
Mashhad	1.166 ab	0.345 f	1.067 bc	0.238 b	1.315 ef	0.614 a
Khorramdasht	1.056 cd	0.512 bc	1.135 bc	0.133 c	1.305 ef	0.254 e

Meanings with similar letters were not statistically significant.

Table 4 Correlation coefficients among traits in *Malva sylvestris* L. genotypes

S.O.V	chlorophyll A	chlorophyll b	carotenoid	protein	carbohydrate
chlorophyll b	0.532^{**}	-	-	-	-
Carotenoid	0.612^{**}	0.783^{**}	-	-	-
Protein	-0.355^{ns}	-0.006^{ns}	-0.979^{ns}	-	-
carbohydrate	-0.231^{ns}	0.037^{ns}	-0.061^{ns}	0.283^{ns}	-
Proline	0.285^{ns}	-0.021^{ns}	0.205^{ns}	0.113^{ns}	0.404^{*}

** and * were significant at the 1 and 5% probability levels, and ^{ns} non-significant, respectively

In the amplified *ITS* region, a total of 943 sites (872 excluding sites with gaps/missing data (or with insertions and deletions) (297 polymorphic sites, 575 monomorphic sites) and 71 without insertions and deletions) along with 174 singletons and 8 haplotypes with a haplotype diversity index of 1 were identified.

Moreover, in *rbcL* locus, a total of 610 sites (602 excluding sites with gaps/missing data (or with insertions and deletions) (14 polymorphic sites, 588 monomorphic sites), and 8 without deletions and insertions) along with 5 singletons and 9 haplotypes with a haplotype diversity index of 1 were identified. The number of haplotypes varies depending on the organism type and plant species so that it is related to 26 haplotypes in the chytrid fungus [44], 135 haplotypes in the Trichoderma [45] and 4 Haplotypes in grasshopper [46].

Estimating the nucleotide substitution pattern based on the Tamura-Nie Model [40], the transition/transversion ratio was calculated. In *ITS* locus, the substitution rate of pyrimidine bases was estimated as 13.61% for T to C substitution and 9.48% for C to T substitution, so that the same value for the purine bases was measured as 72.11% for A to G substitution and 24.8% for G to A substitution (Table 5). The highest (12.9%) and the lowest (5.63%) transition rate at *ITS* marker locus belonged to pyrimidine and purine bases, respectively. *ITS* transition/transversion ratio was calculated as R=0.76 and the average nucleotide ratio was obtained as T (19.9%), C (29.4%), A (21.1%) and G (29.6%). Furthermore, estimating *rbcl* nucleotide substitution pattern based on Nie pattern showed purine substitution rate as 11.25% for G to A substitution and 9.61% for A to G substitution and pyrimidine substitution rate as 9.19% for C to T substitution and 6.7% for T to C substitution (Table 6).

The highest and the lowest transition rate at *rbcl* marker locus were observed in purine (11.25%) and pyrimidine (6.57%) bases, respectively. *rbcl* transition/transversion ratio was calculated as R = 0.76 and the average nucleotide ratio was obtained as T (28.4%), C (20.9%), A (27.4%) and G (23.2%).dN/dS ratio for *ITS* and *rbcl* markers was measured as 0.554 and 0.033, respectively, a value which is less than one, indicating purifying selection with no key changes. Our results are in the same with previous research [47].

Tajima's D test was performed to identify any deviation from the null hypothesis of neutrality and the effects of natural selection on these genes in different *Malva sylvestris* L. genotypes.

Genotypes which show population size expansion or those that purifying selection has acted on them represent significant negative D values, whereas significant positive D values indicate genetic drift, genetic bottlenecks or a balancing selection over the evolutionary history of the genotype. Our results showed significant positive D values for both *rbcl* and *ITS* genes (Table 7), which may indicate that genetic drift or balancing selection has occurred throughout the evolutionary history of the genotypes.

Since genetic distance value indicates the extent of nucleotide substitution in the studied genotypes, it will be related to their kinship level, so that it can be used to figure out how closely the species are related.

Table 5 Transition and transversion substitutions based on Tamura model using *ITS* marker

	A	T	C	G
A	-	5.63	8.32	12.61
T	6.02	-	12.9	8.39
C	6.02	8.73	-	8.39
G	9.04	5.63	8.32	-

Table 6 Transition and transversion substitutions based on the Tamura model using rbcl marker

	A	T	C	G
A	-	8. 95	6.57	9.61
T	8.67	-	6.75	7.41
C	8.67	9.19	-	7.41
G	11.25	8.95	6.57	-

Table 7 Comparison of Tajima's D test results at *rbcl* and *ITS* loci

G	S	Ps	θ	π	D
ITS	297	0. 340596	0. 131359	0. 133437	0. 086745
rbcl	14	0. 023256	0. 008557	0. 008721	0. 092141

G: studied groups, S: segregated sites, Ps: S/m, θ: Ps/a1,π: nucleotide diversity, D: Neutrality test by Tajima's D test

Estimation of the genetic distance using the *ITS* marker showed a value range of 0.030 to 0.225%. The lowest genetic distance (0.030%) was between Kerman (Zarand) and Torbat Heydariyeh genotypes and the highest genetic distance (0.225%) was between Zabol and Jiroft genotypes (Table 8). Furthermore, the determination of the genetic distance using the *rbcl* marker showed a value range of 0.003 to 0.015. The highest genetic distance was observed between Kerman (Zarand) and Rudbar, Kerman (Zarand) and Mashhad, as well as Bandar Abbas and Torbat Heydariyeh and the lowest genetic distance, was observed between Jiroft and Mashhad as well as Rudbar and Zabol (Table 9). In other words, the more or less

genetic distance value indicates how similar or different these populations are.

The phylogenetic tree of the *ITS* and *rbcl* gene loci sequence was plotted using the Neighbor-Joining (NJ) method. The results of this study showed that *Malva sylvestris* L. genotypes are not geographically segregated using *ITS* and *rbcL* markers (Fig. 1 and 2). Another test was carried out in order to assess the capacity of *ITS* and *rbcl* markers in the evaluation of inter-species genetic diversity. So, NCBI sequences of other Malvaceae family species from other countries were used and compared with the present study sequences in order to plot the phylogenetic tree (Fig. 3 and 4). The results showed that similar species of each genotype from the *Malvaceae* family are placed together within the same cluster. Therefore, it can be concluded that *ITS* and *rbcl* were suitable tools for inter-species identifications and can specify genotypes at the species level, a fact which has been reported previously [47-49].

Evaluation of *rbcl* Protein Domains, Motifs, and 3D Structure

As the *rbcl* gene is recognized as one of the major genes in the plant's photosynthesis process, it is expected that it's sequence has domains and conserved regions. Consequently, the *rbcl* protein sequence was assessed in *Malva sylvestris* L. genotypes. Based on 6 to 50 amino acid motifs, the presence of one domain and three conserved motifs was confirmed (Fig. 5). As all sequences had identical and conserved motifs, one of the *rbcl* protein sequences was selected to predict 3D structure (Fig. 6A). Moreover, since *rbcl* is a secretory protein, it must have a membrane and cytoplasmic domains, as well. Therefore, to verify the validity of the assays, the domains were checked and it was indicated that the VacA gene has both target domains (Fig. 6 B).

Table 8 Genetic distance matrix of 9 studied *Malva sylvestris* L. genotypes using *ITS* marker

S.O.V	1	2	3	4	5	6	7	8
Torbat Heydariyeh	1	-	-	-	-	-	-	-
Zarand	2 0. 030	-	-	-	-	-	-	-
Mashhad	3 0. 079	0. 071	-	-	-	-	-	-
Bandar Abbas	4 0. 070	0. 058	0. 099	-	-	-	-	-
Rudbar	5 0. 136	0. 130	0. 161	0. 125	-	-	-	-
Jiroft	6 0. 127	0. 138	0. 171	0. 157	0. 206	-	-	-
Fariman	7 0. 216	0. 199	0. 214	0. 200	0. 196	0. 249	-	-
Khorramdasht	8 0. 031	0. 047	0. 081	0. 060	0. 137	0. 130	0. 213	-
Zabol	9 0. 208	0. 197	0. 228	0. 193	0. 184	0. 254	0. 155	0. 208

Table 9 Genetic distance matrix of 9 studied *Malva sylvestris* L. genotypes using *rbcl* marker

S.O.V	1	2	3	4	5	6	7	8
Zabol	1 -	-	-	-	-	-	-	-
Torbat Heydariyeh	2 0. 008	-	-	-	-	-	-	-
Khorramdasht	3 0. 007	0. 012	-	-	-	-	-	-
Rudbar	4 0. 003	0. 008	0. 010	-	-	-	-	-
Mashhad	5 0. 003	0. 012	0. 010	0. 005	-	-	-	-
Zarand	6 0. 012	0. 010	0. 008	0. 015	0. 015	-	-	-
Jiroft	7 0. 003	0. 008	0. 007	0. 007	0. 007	0. 008	-	-
Fariman	8 0. 005	0. 013	0. 008	0. 008	0. 008	0. 013	0. 005	-
Bandar Abbas	9 0. 007	0. 015	0. 010	0. 010	0. 010	0. 012	0. 007	0. 005

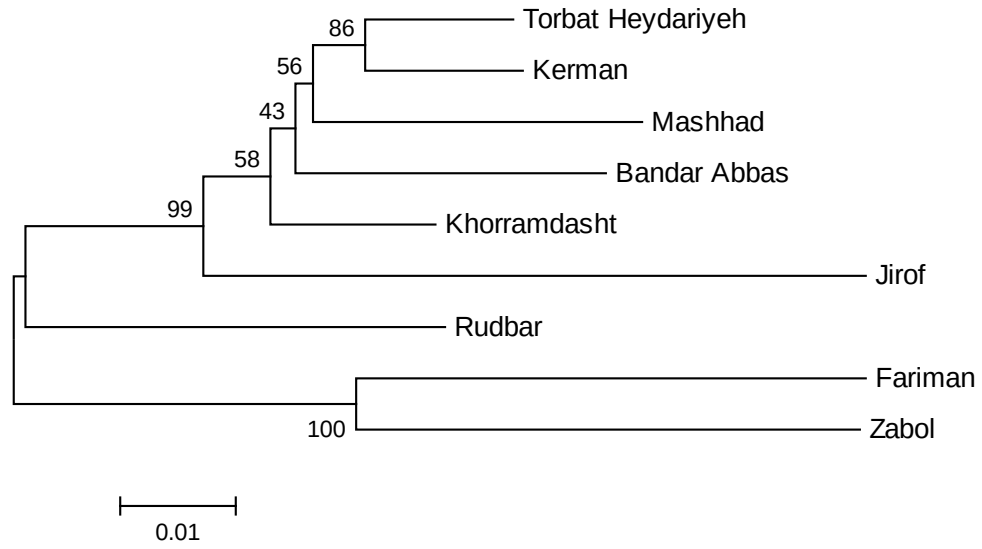

Fig. 1 *ITS*-based phylogenetic analysis of *Malva sylvestris* L. genotypes using NJ Method

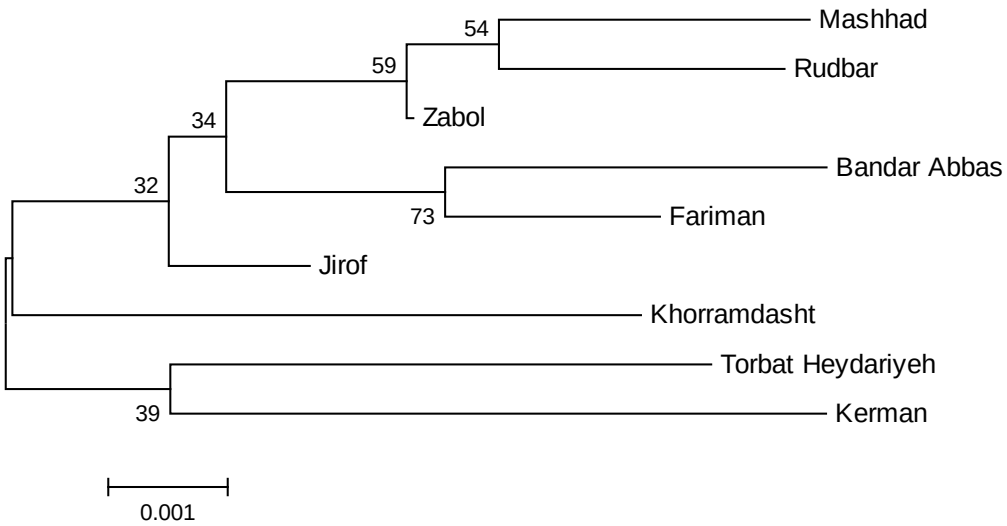

Fig. 2 *rbcl*-based phylogenetic analysis of *Malva sylvestris* L. genotypes using NJ Method

Rbcl Protein Model Validation

ProSA software was used to investigate the potential errors in the *rbcl* protein 3D model. The program indicates two features of the input structure: it's Z-score and a plot of its residue energies. A ProSA score of 0.8 indicates the overall quality of the *rbcl* protein model (Fig. 7). The Z score also measures the total energy deviation of the structure, respecting the energy distribution from random conformations.

The QMEAN and the Z scores of the model were 0.7 and 0.8, respectively, a value that was in the range of zero to one, indicating good model quality. The expected model reliability is expected to be 0 to 1 and this can be deduced from the density plot of the QMEAN scores of the reference set (Fig. 8). The scores characterize a highly reliable structure and are well within the range of scores normally found for proteins of a similar size. The energy plot demonstrates the quality of the local model by drawing a knowledge-based energy plot as a function of the position of the amino acid sequence.

The Ramachandran plot of *rbcl* secondary structure in *Malva sylvestris* L. genotypes indicates a good stereochemical quality chain, with 94.9% of the residues being placed in the red zone with the highest acceptance rate (Fig. 9).

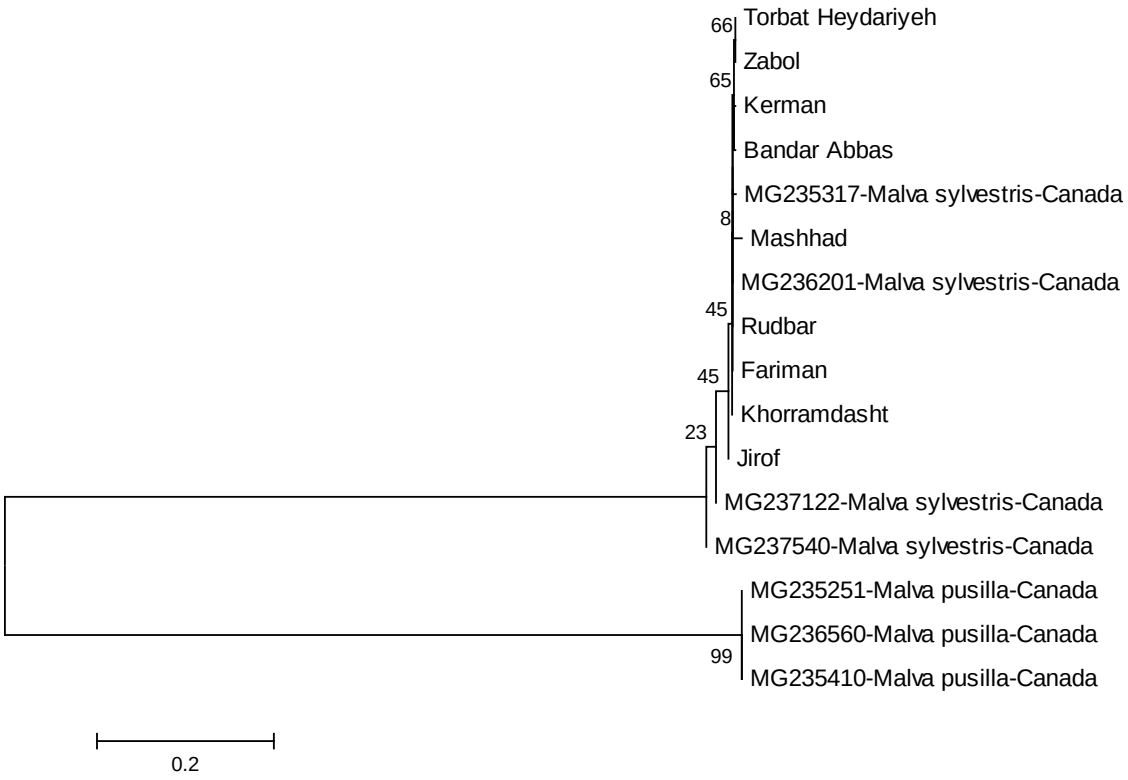

Fig. 3 *ITS*-based phylogenetic tree of different species of Malvaceae family using NJ Method

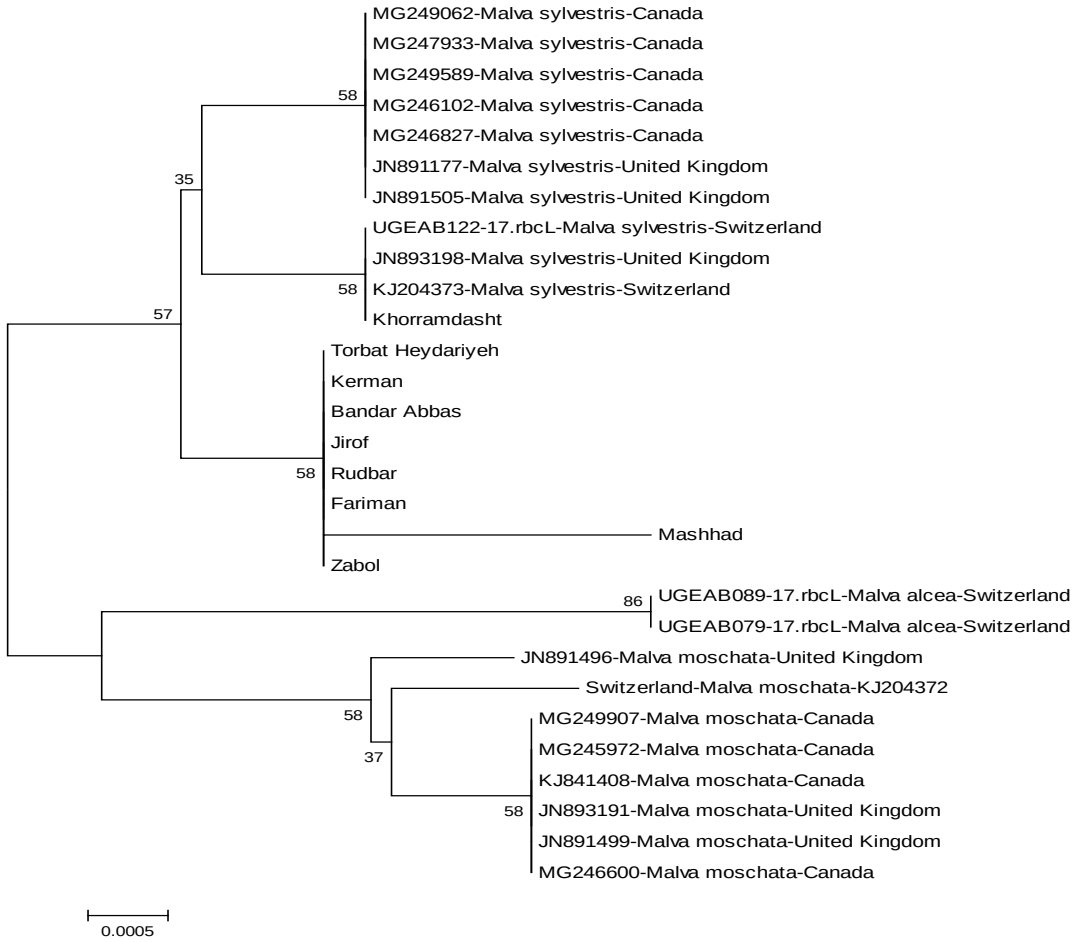

Fig. 4 *rbcl*-based phylogenetic tree of different species of Malvaceae family using NJ Method

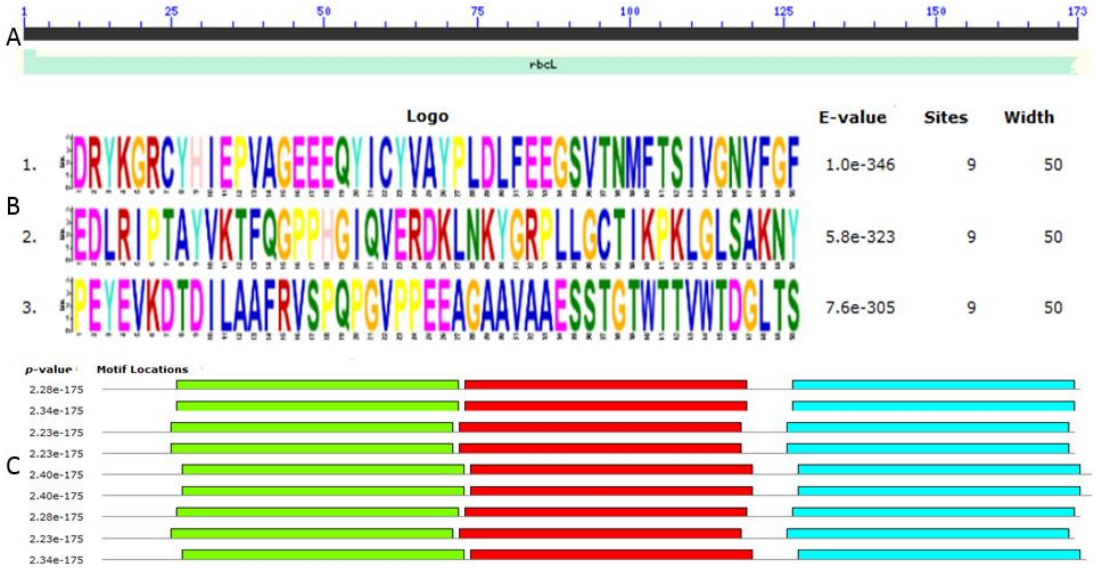

Fig. 5 (A) Conserved domain, (B and C) conserved motifs of the 3D structure of ribulose-1, 5-biphosphate carboxylase (*rbcl*) protein sequence in *Malva sylvestris* L. genotypes

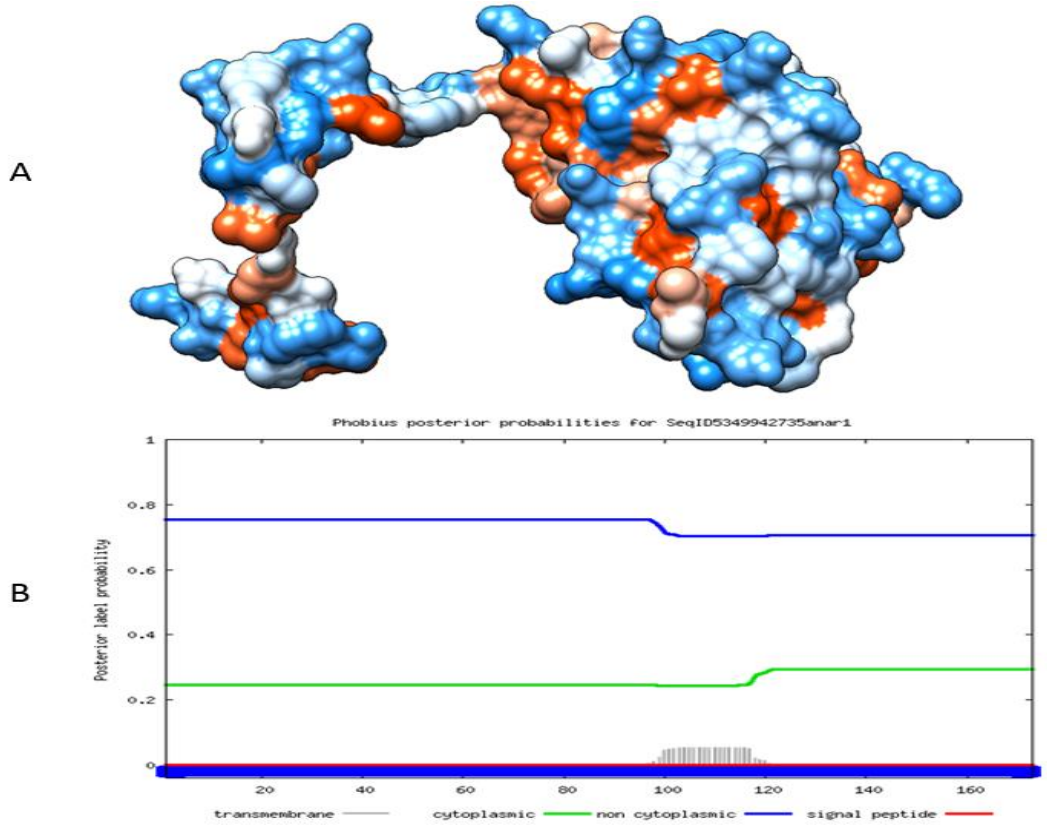

Fig. 6 (A) 3D structure, (B) membrane, cytoplasmic and kinase domains of ribulose-1, 5-biphosphate carboxylase (*rbcl*) protein sequence in *Malva sylvestris* L. genotypes

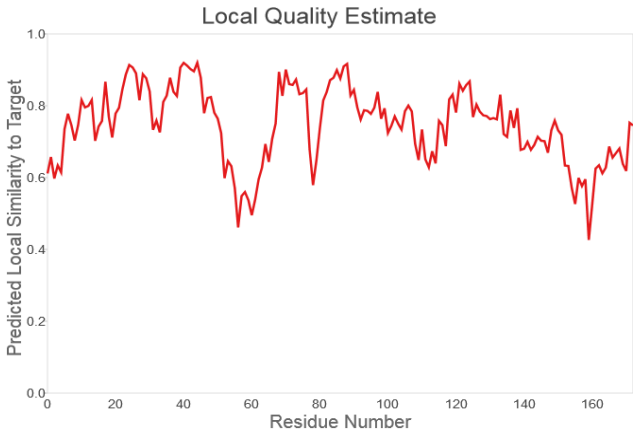

Fig. 7 Validation of the *rbcl* protein model with Pro-SA

Fig. 8 The quality of the generated model with QMEAN and Z-score value

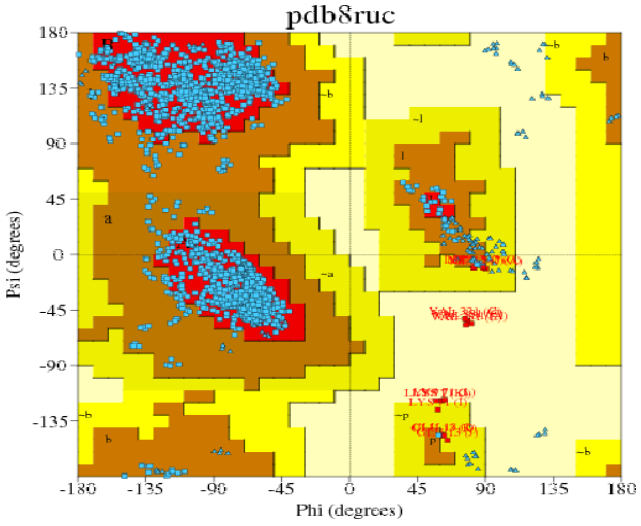

Fig. 9 The Ramachandran plot of *rbcl* secondary structure in *Malva sylvestris* L. genotypes

Evaluation of *rbcl* protein gene network with other genes: Assessment of protein-protein interactions is a fundamental approach to comprehend the function of protein networks [50].

In this study, it was predicted that *rbcl* interacts with *rbcs*, *prk*, matk, *psb* and *pb* genes (Fig. 10). The interaction of the predicted proteins with *rbcl* showed the participation of chloroplast-specific metabolic pathways in plants so that any changes in protein structure and function could affect these pathways.

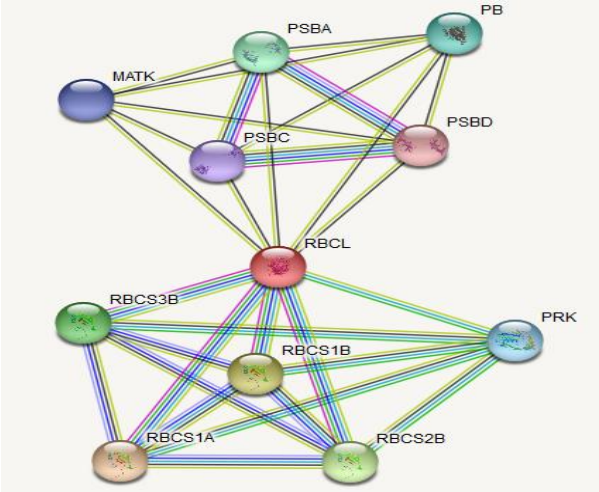

Fig. 10 Schematic representation of *rbcl* protein interaction with other proteins (*rbcs*, *PSB*, *matK*, *pb* and *prk*) in Arabidopsis thaliana.

Discussion

Plants growth and function in ecosystems is affected by a number of factors including species type, climate, soil environment, altitude as well as geographic location. The high diversity of traits in the studied genotypes demonstrates that the selection of individual genotypes will be effective in plant breeding in terms of studied traits [51]. In general, the significant difference between morphological and physiological traits can indicate the presence of genetic variation, which is the first step in the optimization and ideal utilization of the native genotype [52]. In this study, it was also indicated that different *Malva sylvestris* L. genotypes show lots of differences and variations in terms of morphological and phytochemical traits.

Investigating the morphological characteristics of two *Mentha pulegium* L. ecotypes showed a significant difference ($p< 0.05$) in plant height, leaf length and the number of lateral stems among different ecotypes [53]. Furthermore, in a study on *Urtica dioical* L. leaves in Mazandaran and Golestan provinces, it was indicated that plant height and ecotype significantly affect morphological and phytochemical characteristics. On the other hand, leaf length, leaf width and shoot height were decreased with increasing altitude [54]. In the present study, no correlation was found between more foliage and phytochemicals content. On the other hand, plants with

higher fresh and dry root weight and root length had higher phytochemicals content. This may be because plants with longer root lengths are more resistant to dehydration conditions and have no stress, and as a result, are more stable in chemicals production.

Proline is one of the active amino acids in osmotic phenomena with a major role in establishing and preserving osmotic pressure inside the plant. There have also been reports of a positive correlation between proline accumulation and adaptation to osmotic and drought stress conditions in plants [55-57]. In this study, the Mashhad *Malva sylvestris* L. genotype showed the highest proline content, so it can be said that this genotype will be drought-tolerant. On the other hand, a correlation was found between proline content and drought stress so that it was reported that while drought stress has a negative effect on flowering branches, leaves and most morphological traits, it increases root length, proline content and soluble sugars in shoot parts [58]. In another study, the correlation of root length and proline content was reported [55]. In the present study, the Mashhad *Malva sylvestris* L. genotype showed higher proline content as well as longer roots length as compared with other genotypes.

In some plants, in the early stages of drought stress condition several amino acids are increased, but with prolonged dehydration only proline amino acids are more accumulated and stored. The increased concentration of proline and soluble sugars under drought stress conditions has been observed in a number of studies on *Dracocephalum* [59], *Moldavian balm* [60] *Salvia officinalis* L. [61], *Symbopogon martini* [62] and *Petriwinkle medicina* L. [63]. The increased concentration of compatible osmolytes such as carbohydrates and proline under environmental stress conditions such as drought has been established [64]. It seems that the accumulation of compounds such as proline and carbohydrates in plant green tissues under drought stress condition may partially provide the conditions for the plant to continue to absorb water from the root environment, however relying osmotic adjustment on these organic compounds is costly for the plants, being compensated by reducing plant performance [65]. However, in the present study, though Mashhad *Malva sylvestris* L. genotype had the second highest fresh plant weight after Bandar Abbas genotype, but it showed the highest value for all measured phytochemical traits. This means that the accumulation of compounds such as proline and carbohydrates in plant green tissues could not diminish the plant yield, which could be due to its longer root length and higher stem diameter as compared with other genotypes.

Correlation analysis and stepwise regression analysis of *Malva sylvestris* L. genotypes showed the highest positive regression coefficients for the impact of proline content, fresh root weight and plant dry weight on plant root length, indicating the major role of these traits in enhancing plant performance and breeding ability. Overall, this study showed that root characteristic in Mashhad *Malva sylvestris* L. genotype has a significant effect on the final plant yield. Moreover, Mashhad genotype showed the most desirable yield in terms of traits evaluated, for which this genotype is recommended to users such as research centers, universities and private companies that cultivate and domesticate medicinal plants.

Different results on the efficacy of the *ITS* marker in intra- or inter-species identification have been reported. In a study on *Lactuca* species classification using the *ITS* marker sequence [66], it was shown that while *Lactuca sativa* L., *Lactuca serriola* L., *Lactuca dregeana* L. and *Lactuca altaica* L. species have similar sequences, the *aculeata* L. sequence is slightly different. It was also concluded that the *ITS* marker shows inter-species diversity better than intra-species diversity, a result that is in accordance with that of this study. Conversely, Investigating intraspecies diversity of *Erysiphe aquilegia* [67] as well as *Fasciolidae* [68] using *ITS* marker sequence, it was shown that *ITS* marker can identify plants at the sub-species level. Therefore, it can be concluded that the ability of *ITS* in genotype differentiation and identification varies depending on the plant species.

Assessment of plastid DNA sequences in many plants indicated that the plastid genome has modified very slightly during evolutionary events [24, 25, 69]. Chloroplast genes due to their single-parent inheritance nature, structural stability and lack of inter and intra-genic recombination have been carefully studied to identify their potential as DNA barcodes[70]. There are no deletion/insertion mutations in *rbcl* coding and conserved sequences. On the other hand, as some of the chloroplast-encoding genes have been interrupted by introns, such deletion/insertion mutations may have occurred within the introns, but this is not the case with the *rbcl* gene as it lacks introns. Therefore, it can be deduced that modifications arise very slowly in the *rbcl* gene and that is why this gene has been used to estimate and evaluate the evolution [71,72]. On the other hand, because the *rbcl* gene is inherited maternally, it is conserved among the varieties of the same species, having the same maternal chloroplast genome [73,74].

Determination of the ratio of non-synonymous to synonymous substitutions (dN/dS ratio) is an efficient way to figure out the process of natural selection during gene evolution [75]. If this ratio is greater than one it indicates positive selection; if it is less than one, it represents pure selection; and if it is equal to one, it shows neutral selection during the evolution of these genes[76]. Our results showed a dN/dS ratio of 0.555 for

the *ITS* marker and 0.033 for *rbcl* marker, values which were less than one indicating that pure selection has been occurred on the genes under study, with no key changes. The results of this study on the inability to isolate the genotypes based on the geographical location indicate a remarkable genetic similarity and low genetic diversity among different genotypes, which may be due to the high conservation of *ITS* and *rbcl* and low variability of this region. In addition, a comparison of *rbcl* sequences in two cyanobacteria classes and several plant genera showed high sequence similarity [77,78], which is in agreement with the results of the present study.

Though it is reported that one of the limitations of the chloroplast genome is its inability in studying close species, however, several studies have used chloroplast diversity to determine the kinship at the genotype level and to assess the evolution process within genotypes [25, 79,80].

Cluster analysis based on *ITS* and *rbcl* markers failed to differentiate different *Malva sylvestris* L. genotypes by geographical area, which is in agreement with a study on black cumin (*Bunium persicum* (Boiss.) B.Fedtsch.) ecotypes [81]. The placement of different geographical specimens within the same cluster can be due to genetic similarity or physical exchange of seeds between different geographical regions, though the second reason seems more justified [82].

As the present study was not able to differentiate *Malva sylvestris* L. genotypes geographically using *ITS* and *rbcl* markers (Fig. 1 and 2), another test was carried out in order to assess the capacity of *ITS* and *rbcl* markers in the evaluation of inter-species genetic diversity. So, NCBI sequences of other Malvaceae family species were used and compared with those of the present study (Fig. 3 and 4), indicating the placement of similar species within the same cluster. Therefore, it can be concluded that *ITS* and *rbcl* are suitable tools for inter-species, a point which has been mentioned in other studies [48-50]. On the other hand, intra- and inter-species diversity is an important source of speciation and one of the indicators of biodiversity reserves. As the lack of inter- and intra-species diversity among plants can cause a lot of problems, it is very essential to study biodiversity.

It has been reported that plant resistance to pests, diseases and environmental stresses depends on plant diversity centers [83]. Iran, due to high climate diversity and rich plant genetic resources, is considered as one of the richest countries in terms of natural resources and potentials. As a result, assessment of genetic diversity plays an important role in plant breeding programs, and it will help a lot in advancing research programs [84].

Protein-protein interaction plays a crucial role in all biological processes. Biological network analysis helps to realize the biological mechanism by revealing patterns of functional cellular networks [85]. Assessment of the *rbcl* protein-protein network indicates that the genes of the RBC family are most closely related to each other and that these genes are also related to the prk gene. The *rbc* family genes are involved in two major reactions: the carboxylation of D-ribulose 1,5-bisphosphate during carbon dioxide fixation as well as the oxidative fragmentation of the pentose suspension [86]. On the other hand, the psb genes interact mostly with each other and with the *pb*, *matK* and *rbcl* genes and are involved in the electron transport chain and ATP production. The psb genes split the water electrons using light energy and produce O2 and a proton gradient that is ultimately used to produce ATP [87]. Moreover, some RNA transcripts that require *matK* for intron excision involve *trnK*, *trnA*, *trnL*, *rps12*, *rpl2*, and *atpF* whose protein or tRNA products are required for a number of chloroplast functions including photosynthesis. These results suggest that *matK* which acts as a factor following transcription has a key function in chloroplasts [88-90]. These findings implicate the essential function of these proteins in the chloroplast [91].

The *rbcl*, *matK*, and *psb* genes are three important marker genes in DNA barcoding, and many researchers have used these DNA barcodes to identify and document plant diversity [28]. On the other hand, the results of this study indicated that the protein products of these genes are not independent in the chloroplast, so that increasing and decreasing the expression of each gene affects the others. In addition, some studies have demonstrated that chloroplast genes are not always used alone to evaluate genotypes due to the need for additional information [92-94]. Assessment of the gene network and the association of the *rbcs*, *prk*, *matK*, *psb* and *pb* genes in the present study (Fig. 7) verify the above statements, as well.

rbcl has been used as the best gene locus for phylogenetic studies at above-species taxonomic levels in many studies, whereas at the species level it has not been recommended [95, 96]. On the other hand, due to low discrimination power at the species level, most studies have emphasized that *rbcl* should be used in combination with other plant DNA barcodes [38,77]. The consortium for the barcode of life (CBOL) has introduced *rbcl* + *matK* chloroplast genes as the standard plant barcodes, which has a high level of species discrimination and an optimal sequence quality for plants[97,98].

Generally, a DNA region can be considered an ideal DNA barcode if it has sufficient diversity among the species of the same genus. However, the *ITS* and *rbcl* loci were not suitable for the assessment of the differences among *Malva sylvestris* L. genotypes because of showing low diversity, high conservation as well as maternal inheritance pattern. On the other hand, a dN/dS ratio of less than indicates no key changes in *ITS* and *rbcl* loci of *Malva sylvestris* L. genotypes, as well. Furthermore,

protein-protein interaction analysis also showed that the *rbcl* gene interacts with *rbc*, *psb*, *matK*, *prk* and *pb* genes and these genes are not independent of each other. Therefore, it can be suggested that other genes and their interactions should be considered in the modifications of each gene.

Conclusion

Overall, this study showed that root characteristic in Mashhad *Malva sylvestris* L. genotype has a significant effect on the final plant yield. Moreover, Mashhad genotype showed the most desirable yield in terms of traits evaluated, for which this genotype is recommended to users such as research centers, universities and private companies that cultivate and domesticate medicinal plants.

References

1. Razavi M, Zarrini G, Molavi G, Ghasemi G. Bioactivity of Malva Sylvestris L., a Medicinal Plant from Iran. Iran. J Basic Med Sci. 2011;14:574-579.
2. Shoshtariyan S, Ghahramanlo A, Izadidarbandi E, Alirezaienoghondar M. Effect of cropping pattern (cultivar distance) on yield of bacon bacon. National Conference on Med Plants and Recognition of its Economic Potential and its Occupation, Islamic Azad University, Birjand Branch. 2010.
3. Riedl I. Flora Iranica: no. 120. Malvaceae. Graz: Akademische Druck 86p.-illus., keys.. Ge, La Icones. Geog. 1976; 2.
4. Pakravan M. New findings of the genus Malva L. in Iran. Iranian J of Botany. 2006;11:247-249.
5. Henry AG, Piperno DR. Using plant microfossils from dental calculus to recover human diet: a case study from Tell al-Raqā'i, Syria. J of Archaeological Sci. 2008;35:1943-1950.
6. Tabaraki R, Yousefi Z, Asadi-GHarneh H. Chemical composition and antioxidant properties of Malva sylvestris. J Agric Res. 2012;8:59-68.
7. Scherrer AM, Motti R, Weckerle CS. Traditional plant use in the areas of Monte Vesole and Ascea, Cilento National Park (Campania, Southern Italy). J Ethnopharmacology. 2005;97:129-143.
8. Kültür Ş. Medicinal plants used in Kırklareli province (Turkey). J of Ethnopharmacology. 2007;111:341-364.
9. Mustafa A, Ali M. New steroidal lactones and homomonoterpenic glucoside from fruits of Malva sylvestris L. L. Acta Pol Pharm. 2011;68:393-401.
10. Shale T, Stirk W, Van Staden J. Variation in antibacterial and anti-inflammatory activity of different growth forms of Malva parviflora and evidence for synergism of the anti-inflammatory compounds. J Ethnopharmacology. 2005;96:325-330.
11. Omidbeigi R. Production and processing of medicinal plants (Vol. 8). Mashahd: Beh nashr, Astan Ghods Razavi. 2015;348.
12. Ahmand Z, Fahmideh L, Fazeli-Nasab B. Genetic Evaluation of Cumin and Caraway Using Eryngium planum Microsatellite Markers. Sci J Manag Sys. 2017;8:59-71.
13. Ansari S, Solouki M, Fakheri B, Fazeli-Nasab B, Mahdinezhad N. Assessment of molecular diversity of Internal transcribed spacer region in some Lines and Landrace of Persian Clover (Trifolium resupinatum L). Potravinarstvo Slovak J Food Sci. 2018;12:657-666.
14. Davari A, Solouki M, Fazeli-Nasab B. Effects of jasmonic acid and titanium dioxide nanoparticles on process of changes of phytochemical and antioxidant in genotypes of Satureja hortensis L. Eco-Phytochemical J Med Plants. 2018;5:1-20. (In Persian).
15. Blaxter M. Molecular systematics: counting angels with DNA. Nature. 2003;421:122-124.
16. Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B: Biological Sci. 2003;270:313-321.
17. Hebert PD, Ratnasingham S, de Waard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London B: Biological Sci. 2003;270(Suppl 1):S96-S99.
18. Keshvari T, Najaphy A, Kahrizi D, Zebarjadi A. Callus induction and somatic embryogenesis in Stevia rebaudiana Bertoni as a Medicinal Plant. Cell Mol Biology. 2018;64:46-49.
19. Fazeli-Nasab B, Mehrabi AA, Eizadi-Drabandi A. Genetic diversity of different wheat genotypes using seed storage proteins and SSRs markers. Modern Gen J. 2010;5:83-91.
20. Ghahramanzadeh R, Esselink G, Kodde L, Duistermaat H, Van Valkenburg J, Marashi S, Smulders M, Van de Wiel C. Efficient distinction of invasive aquatic plant species from non-invasive related species using DNA barcoding. Mol Ecol Reso. 2013;13:21-31.
21. Fazeli-Nasab B, Naghavi MR, Mehrabi AA. Allelic Variation of Microsatellite Markers from Linkage Group A Genome in Wild Populations of Einkorn and Hexaploid Wheat. Agric Biotech. 2013;4:53-62.
22. Balasaravanan T, Chezhian P, Kamalakannan R, Yasodha R, Varghese M, Gurumurthi K, Ghosh M. Identification of species-diagnostic ISSR markers for six Eucalyptus species. Silvae Gen. 2006;55:119-122.
23. Fazeli-Nasab B, Naghavi MR, Mardi M, Yazdi-Samadi B, Kazemi M. Genetic diversity and relationships of Iranian wheat cultivars using microsatellite markers. Iranian J of Agric Sci. 2006;37:93-99.
24. Group CPW, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL. A DNA barcode for land plants. Proceedings National Academy Sci. 2009;106:12794-12797.
25. Chang CC, Lin HC, Lin IP, Chow TY, Chen HH, Chen WH, Cheng CH, Lin CY, Liu SM, Chang CC. The chloroplast genome of Phalaenopsis aphrodite (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. Mol Biol Evolution. 2005;23:279-291.
26. Gamache J, Sun G. Phylogenetic analysis of the genus Pseudoroegneria and the Triticeae tribe using the rbcL gene. Biochem Sys Ecol. 2015;62:73-81.
27. Liu L-X, Li R, Worth J R, Li X, Li P, Cameron K M, Fu C-X. The complete chloroplast genome of Chinese bayberry (Morella rubra, Myricaceae): implications for understanding the evolution of Fagales. Front Plant Sci. 2017;8:968.
28. Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One. 2010;5:e8613.

29. Park JH, Kim YH, Ham CS, Shin SE, Lee HJ, Ko KS, Choi J, Son GH, Park SH. Molecular identification of forensically important calliphoridae and sarcophagidae species using ITS2 nucleotide sequences. Forensic Sci Int. 2018;284:1-4. 10.1016/j.forsciint.2017.12.017
30. Paykarestan B, Seify M, Anajafi M, Chavoshi S. Effect of mint essence and ultraviolet on thyme and hock seed germination. New findings in agriculture J. 2012;6:215-220.
31. Arnon DI. Ferredoxin and photosynthesis. Sci. 1965;149:1460-1470.
32. Bates LS, Waldren RP, Teare I. Rapid determination of free proline for water-stress studies. Plant & Soil. 1973;39:205-207.
33. Keleş Y, Öncel I. Response of antioxidative defence system to temperature and water stress combinations in wheat seedlings. Plant Sci. 2002;163:783-790.
34. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry. 1976;72:248-254.
35. Dellaporta SL, Wood J, Hicks JB. A plant DNA minipreparation: version II. Plant Mol Biol Rep. 1983;1:19-21.
36. White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and appli. 1990;18:315-322.
37. Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA, Sytsma KJ. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. American J Botany. 2003;90:107-115.
38. Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. Proceedings of the National Academy of Sci of the United States of America. 2009;106:18621-18626.
39. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T L. BLAST+: Architecture and Appli. BMC Bioinformatics. 2009;10:421.
40. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 2013;30:2725-2729.
41. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 2003;31:3381-3385.
42. Brohee S, Van Helden J. Evaluation of clustering algorithms for protein-protein interaction networks. BMC Bioinformatics. 2006;7:488.
43. Feizi M, Fahmideh L. Evaluation of agronomic traits and Yield Potential Diversity Inbreed Wheat Inbred Lines Triticum aestivum L. Derived from Roshan×Falat Cultivar. [Research]. J Crop Breeding. 2017;8:30-24.
44. Goka K, Yokoyama J, Une Y, Kuroki T, Suzuki K, Nakahara M, Kobayashi A, Inaba S, Mizutani T, Hyatt AD. Amphibian chytridiomycosis in Japan: distribution, haplotypes and possible route of entry into Japan. Mol Ecol. 2009;18:4757-4774.
45. Druzhinina IS, Kopchinskiy AG, Komoń M, Bissett J, Szakacs G, Kubicek CP. An oligonucleotide barcode for species identification in Trichoderma and Hypocrea. Fungal Gen Biol 2005;42:813-828.
46. Parkin EJ, Butlin RK. Within-and between-individual sequence variation among ITS1 copies in the meadow grasshopper Chorthippus parallelus indicates frequent intrachromosomal gene conversion. Mol Biol Evol. 2004;21:1595-1601.
47. Jomeh Ghasem Abadi Z, Fakheri B, Fazeli-nasab B. Study of the Molecular Diversity of Internal Transcribed Spacer Region (ITS1.4) in Some Lettuce Genotypes. [Research]. J of Crop Breeding. 2019;11:29-39.
48. Varela ES, Lima JP, Galdino AS, Pinto LdS, Bezerra WM, Nunes EP, Alves MA, Grangeiro TB. Relationships in subtribe Diocleinae (Leguminosae; Papilionoideae) inferred from internal transcribed spacer sequences from nuclear ribosomal DNA. Phytochemistry. 2004;65:59-69.
49. Wu CT, Hsieh CC, Lin WC, Tang CY, Yang CH, Huang YC, Ko YJ. Internal transcribed spacer sequence-based identification and phylogenic relationship of I-Tiao-Gung originating from Flemingia and Glycine (Leguminosae) in Taiwan. J. Food. Drug. Anal. 2013;21:356-362.
50. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou K P. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic. Acids. Res. 2014;43:D447-D452.
51. Omidbeigi R. Approaches to the production and processing of Medicinal plants (Vol. 1). Tehran: Fekreroz. 1996.
52. Shaykh F, Bazi H. Investigation on the variation of indigenous populations of squash in Golestan province. Paper presented at the National Symposium on Natural Products and Med Plants. 2013.
53. Abbaszadeh B, Rezaee M, Layegh haghighi M. Investigation Morphological Characteristics and Essential Oil Component of 2 Ecotypes of Watermint (Mentha aquatica L.). [Res]. J Med Plants. 2012;1:248-257.
54. Najarfiruzjaie M, Hemati K, Khorasaninejad S, khani AD g, fard AB. The effects of altitude on morphological characteristicsand biochemical plant leaf Urtica dioica L. In Mazandaran and Golestan. J Iranian Plant Ecophysiological Res. 2014;9:1-11.
55. Babaie k, Aminidehghi M, Modares sanavi A, M, Jabari R. Effect of drought stress on morphological traits, proline content and thymol percentage in Thyme. Iranian J Med Aromatic Plants Res. 2010;26:239-251.
56. Larkunthod P, Nounjan N, Siangliw JL, Toojinda T, Sanitchon J, Jongdee B, Theerakulpisut P. Physiological responses under drought stress of improved drought-tolerant rice lines and their parents. Notulae Botanicae Hortic Agrobotanici Cluj-Napoca. 2018;46:679-687.
57. Li J, Phan T-T, Li YR, Xing YX, Yang LT. Isolation, Transformation and Overexpression of Sugarcane SoP5CS Gene for Drought Tolerance Improvement. Sugar tech. 2018;20:464-473.
58. Lotfi M, Abbaszadeh B, Mirza M. The effect of drought stress on morphology, proline content and soluble carbohydrates of tarragon (Artemisia dracunculus L.). Iranian J Med Aromatic Plants Res. 2014;30:19-29.
59. Ardekani MR, Abbaszadeh B, Sharifi-Ashourabadi E, Lebaschi MH, Paknejad F. The effect of water deficit on quantitative and qualitative characters of balm (Melissa officinalis L.). Iranian J Med Aromatic Plants. 2007;23:251-261.
60. Safikhani F. Investigation on physiological resistance aspects under drought stress of Dracocephalum moldavica L. (ph.D), Chamran University. 2006

61. Bettaieb I, Zakhama N, Wannes WA, Kchouk M, Marzouk B. Water deficit effects on Salvia officinalis fatty acids and essential oils composition. Sci Hortic. 2009;120:271-275.
62. Fatima S, Abad Farooqi A, Ansari S, Sharma S. Effect of water stress on growth and essential oil metabolism in Cymbopogon martinii (palmarosa) cultivars. J Essen Oil Res. 1999;11:491-496.
63. Sreevalli Y, Baskaran K, Chandrashekara R, Kulkarni R. Preliminary observations on the effect of irrigation frequency and genotypes on yield and alkaloid concentration in periwinkle. Preliminary observations on the effect of irrigation frequency and genotypes on yield and alkaloid concentration in periwinkle. 2000;22:356-358.
64. Heidari M, Karami V. Effects of different mycorrhiza species on grain yield, nutrient uptake and oil content of sunflower under water stress. J Saudi Aociety Agric Sci. 2014;13:9-13.
65. Good AG, Zaplachinski ST. The effects of drought stress on free amino acid accumulation and protein synthesis in Brassica napus. Physiol Plantarum. 1994;90:9-14.
66. Koopman WJ, Guetta E, van de Wiel CC, Vosman B, van den Berg RG. Phylogenetic relationships among Lactuca (Asteraceae) species and related genera based on ITS-1 DNA sequences. American J Botany. 1998;85:1517-1530.
67. Cunnington JH, Lawrie AC, Pascoe IG. Unexpected ribosomal DNA internal transcribed spacer sequence variation within Erysiphe aquilegiae sensu lato. Fungal Diversity, 2004; 16(16): 1-10.
68. Adlard RD, Barker SC, Blair D, Cribb TH. Comparison of the second internal transcribed spacer (ribosomal DNA) from populations and species of Fasciolidae (Digenea). Int J Parasitology. 1993;23:423-425.
69. Besendahl A, Qiu YL, Lee J, Palmer JD, Bhattacharya D. The cyanobacterial origin and vertical transmission of the plastid tRNA Leu group-I intron. Curr Genet. 2000;37:12-23.
70. Ledford H. Botanical identities. Nature Publishing Group. 2008;451.
71. Guldhe A, Renuka N, Singh P, Bux F. Effect of phytohormones from different classes on gene expression of Chlorella sorokiniana under nitrogen limitation for enhanced biomass and lipid production. Algal Res. 2019;40:101518.
72. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings National Academy Sci. 1987;84:9054-9058.
73. Hapsari L, Trimanto T, Wahyudi D. Species diversity and phylogenetic analysis of Heliconia spp. collections of Purwodadi Botanic Garden (East Java, Indonesia) inferred by rbcL gene sequences. Biodiversitas J Biol Diversity. 2019;20:1266-1283.
74. Sarhan S, Hamed F, Al-Youssef W. The rbcL Gene Sequence Variations among and within Prunus Species. J of Agricultural Sci and Technol. 2016;18:1105-1115.
75. Nei M, Kumar S. Molecular evolution and phylogenetics: Oxford university press. 2000; 329 Pages, ISBN: 0-19-513584-9.
76. LI WH. Molecular evolution. Sinauer, Sunder-land, Massachusetts, ISBN : 0878934634 , 487 Pages. 1997;
77. Newmaster S, Fazekas A, Steeves R, Janovec J. Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Resour. 2008;8:480-490. https://doi.org/10.1111/j.1471-8286.2007.02002.x
78. Steven GN, Subramanyam R. Testing plant barcoding in a sister species complex of pantropical Acacia (Mimosoideae, Fabaceae). Molecular Ecology Resources. 2009;9 Suppl s1:172-180. 10.1111/j.1755-0998.2009.02642.x
79. Barker N, Howis S, Nordenstam B, Källersjö M, Eldenäs P, Griffioen C, Linder H. Nuclear and chloroplast DNA-based phylogenies of Chrysanthemoides Tourn. ex Medik.(Calenduleae; Asteraceae) reveal extensive incongruence and generic paraphyly, but support the recognition of infraspecific taxa in C. monilifera. South African J Botany. 2009;75:560-572.
80. Pfanzelt S, Albach DC, von Hagen KB. Extremely low levels of chloroplast genome sequence variability in Astelia pumila (Asteliaceae, Asparagales). PeerJ. 2019;7:e6244.
81. Azimzade M, Amir R, Osare MH, Bihamta MR, Frotan M. Evaluation of genetic diversity of Iranian Cumin (Bunium persicum Boiss) of nuclear ribosomal DNA using ITS. J Gen Res Plant Breeding Pasture Forest. 2014;22:1-10.
82. Bahari Z, Shojaeiyan A, Rashidi Monfared S, Mirshekari A, Nasiri K, Amiriyan M. Investigation of Genetic Diversity among Some Iranian Dill (Anethum graveolens L.) Landraces, Using ISSR Markers. Plant Gen Resour-C J. 2015;2:11-22.
83. Sindhu A, Tehlan SK, Chaudhury A. Analysis of genetic diversity among medicinal therapist Trigonella foenum-graecum L. genotypes through RAPD and SSR markers. Acta. Physiol. Plant. 2017;39:100. 10.1007/s11738-017-2395-8
84. Dehdashtian Z, Wahabi M r, Fazilati M, Ghaedi K, Salamian A. Analysis of the Genetic Diversity of Astragalus Gossypinus Population in Isfahan. [Research]. Gen in the 3rd Millennium. 2011;9:2474-2467.
85. Raman K. Construction and analysis of protein–protein interaction networks. Automated Experimentation. 2010;2:2.
86. Dietz KJ, Wesemann C, Wegener M, Seidel T. Toward an integrated understanding of retrograde control of photosynthesis. Antioxidants & Redox Signaling. 2019;30:1186-1205.
87. Nelson N, Ben-Shem A. The complex architecture of oxygenic photosynthesis. Nature Reviews Molecular Cell Biology, 2004; 5: 971.
88. Barthet MM. Expression and Function of the Chloroplast-encoded Gene matK. Virginia Tech. Retrieved from http://hdl.handle.net/10919/26287. 2006
89. Barthet MM, Hilu KW. Expression of matK: functional and evolutionary implications. American J Botany. 2007;94:1402-1412.
90. Barthet MM, Hilu KW. Evaluating evolutionary constraint on the rapidly evolving gene matK using protein composition. J Mol Evol. 2008;66:85-97.
91. Duan RY, Xiang GH, Lin YX, Luo YC, Peng RM. The complete chloroplast genome of the invasive plant Senecio vulgaris L.(Asteraceae). Mitochondrial DNA Part B. 2019;4:1794-1795.
92. Li Y, GAO LM, Poudel RC, Li DZ, Forrest A. High universality of matK primers for barcoding gymnosperms. J of Sys and Evolution, 2011; 49: 169-175.
93. Staff N. MatK Gene Is A "Barcode" DNA For Plants. 2008.
94. Williams BR, Mitchell TC, Wood JR, Harris DJ, Scotland RW, Carine MA. Integrating DNA barcode data in a monographic study of Convolvulus. Taxon. 2014;63:1287-1306.

95. Soltis DE, Soltis PS. Choosing an approach and an appropriate gene for phylogenetic analysis Molecular systematics of plants II (pp. 1-42): Springer. 1998.
96. Zhai W, Duan X, Zhang R, Guo C, Li L, Xu G, Shan H, Kong H, Ren Y. Chloroplast genomic data provide new and robust insights into the phylogeny and evolution of the Ranunculaceae. Mol Phylogenet Evol. 2019;135:12-21.
97. Federici S, Galimberti A, Bartolucci F, Bruni I, De Mattia F, Cortis P, Labra M. DNA barcoding to analyse taxonomically complex groups in plants: the case of Thymus (Lamiaceae). Botanical J Linnean Society. 2013;171:687-699.
98. Mahadani P, Ghosh S K. DNA Barcoding: A tool for species identification from herbal juices. DNA Barcodes. 2013;1:35-38.