J Plant Biotechnol (2020) 47:15–25
DOI:https://doi.org/10.5010/JPB.2020.47.1.015

ISSN 1229-2818 (Print)
ISSN 2384-1397 (Online)

Research Article

Phylogenetic relationships of Iranian *Allium* species using the *matK* (cpDNA gene) region

Hemadollah Zarei · Barat Ali Fakheri · Mohammad Reza Naghavi · Nafiseh Mahdinezhad

Received: 6 December 2019 / Revised: 25 December 2019 / Accepted: 27 January 2020

Abstract *Allium* L. is one of the largest genera of the *Amaryllidaceae* family, with more than 920 species including many economically important species used as vegetables, spices, medicines, or ornamental plants. Currently, DNA barcoding tools are being successfully used for the molecular taxonomy of Allium. A total of 46 Allium species were collected from their native areas, and DNA was extracted using the IBRC DNA extraction kit. We used specific primers to PCR amplify *matK*. DNA sequences were edited and aligned for homology, and a phylogenetic tree was constructed using the neighbor-joining method. The results show thymine (38.5%) was the most frequent and guanine (13.9%) the least frequent nucleotide. The *matK* regions of the populations were quite highly conserved, and the amount of C and CT was calculated at 0.162 and 0.26, respectively. Analysis of the nucleotide substitution showed C-T (26.22%) and A-G (8.08%) to have the highest and lowest percent, respectively. The natural selection process dN/dS was 1.16, and the naturality test results were -1.5 for Tajima's D and -1.19 for Fu's Fs. The NJ dendrogram generated three distinct clades: the first contained *Allium austroiranicum* and *A. ampeloprasum*; the second contained *A. iranshahrii, A. bisotunense,* and *A. cf assadi*; and the third contained *A. rubellum* and other species. In this study, we tested the utility of the *matK* region as a DNA barcode for discriminating *Allium.* species.

Keywords cpDNA, marker, Molecular phylogeny, *matK*, *Allium*, Taxa

H. Zarei (✉) · B. A. Fakheri · N. Mahdinezhad
Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Zabol, Zabol, Iran
e-mail: h.zarei1989@gmail.com

M. R. Naghavi
Department of Agronomy and Plant Breeding, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

H. Zarei · M. R. Naghavi
Plant Bank, Iranian Biological Resource Center (IBRC), ACECR, Tehran, Iran

Introduction

Allium L. falls under one of the widespread and main genera in the *Amaryllidaceae* family (Friesen et al. 2006; Li et al. 2010; Fritsch et al. 2010). As of Linnaeus, the number of species have risen from 30 to over 920 species at present (Govaerts et al. 2005-2014). There are some species of economic importance in the genus, namely onion, garlic, leek, shallot, bunching onion, and chives planted as vegetables or spices, as well as species utilized as seasoning crops, traditionally used medications, and ornamental plants (Fritsch and Friesen, 2002).

The main focus of investigators has been on existing natural plants that are recently cultivated in investigations, and most of *Allium* newly introduced species have been found in Iran. A number of recent species and subspecies achieved scientific qualification from some regions in Iran (Akhani 1999; Fritsch et al. 2001; Fritsch et al. 2002; Fritsch et al. 2006; Fritsch and Abbasi 2008; Fritsch and Maroofi 2010; Kamelin and Seisums 1996; Khassanov and Memariani 2006; Khassanov et al. 2006; Mashayekhi et al. 2005; Memariani et al. 2007; Neshati et al. 2009; Razyfard et al. 2011). A comparison was made among 170 species and subspecies, and their diverse types in Southwest Asia, mainly in Iran and Turkey, with over 120 species identified in Iran classified into seven subgenera and 30 sections (Friesen et al. 2006; Fritsch and Maroofi 2010; Fritsch and Abbasi 2013; Memariani et al. 2012). *Allium* genus is generic for the Irano-Turanian Phyto-geographic zone and represents a highly endemic rate (Matin 1992).

To increase the number of qualitative traits for tightly

related species, molecular markers have greater capability than morphologic characters, most of which present only quantitatively diverse attributes (Harpke et al. 2013). Linne von Berg pioneered in a publication for the organization of the *Allium* genus by molecular markers (von Berg et al. 1996). Friesen then innovated the classification of *Allium* in publication according to molecular techniques (Friesen et al. 2006). Thereafter, *Allium* was affirmed monophyletically by the entire scientific reports (Choi et al. 2012; Li et al. 2010; Nguyen et al. 2008). Despite a huge body of research, perfect studies have not been conducted on the whole subgenera regarding the phylogenetic condition of their species (Friesen et al. 2006).

Currently, the DNA barcoding technique has been proven as an instructive and efficacious procedure for assessing plant phylogenies and presented successful applications in the molecular classification of *Allium* (Abdulina, 1999; von Berg et al. 1996). The basis of barcoding technique is on the alignment of short sequences of DNA markers of the nuclear and plastid genomes (Kress, 2017; Li et al. 2015). The analysis of phylogenetic associations and plant identification surveys have initially used variabilities of the nucleotide sequence in plastid DNA (cpDNA) at inter-specific (among family or genus) and intra-specific levels (into the species or varieties) (Tamura et al. 2004).

The *matK* is a chloroplast gene that encodes a locus in the intron of the *trnk* gene region. The gene encodes a maturase on the large single-copy part at the adjacency of the inverted repeat of plant species. The *matk* has vast substitution rates in comparison to other chloroplast genes, and its gene sequence lies among the lowest conserved plastid genes; hence, it has received effective uses in plant evolution and solving phylogenetic equations in a variety of taxonomic levels (Fuse and Tamura, 2000; Ito et al. 1999).

matK gene is very beneficial over other genes that contain the organelle genome genes. Firstly, this gene undergoes evolution nearly thrice as fast as the vastly applied plastid genes such as *rbcL* and *atpB. matk* gene is in the chloroplast genome and, in general, it has maternal inheritance. The gene is sensibly sizeable, with an extensive substitution rate, a large ratio of alterations at the first and the second codon positions, a low transition/transversion ratio, and the occurrence of mutationally conserved segments. It is also powerful and efficacious in species discrimination, and has excellent sequence recovery rate, a simple technique experimentally, a facile sequence alignment, and nonexistence of allelic polymorphisms or manifold paralogous copies against the nuclear DNA genome. It was demonstrated that the conversion at nucleic acid (DNA) and amino acid levels have even distribution throughout the whole gene. Apparently, the 5' region of the *matK* gene contains a greater variation than the 3' region in some monocotyledons and dicotyledons. To address family and even species-level associations, *matK* gene sequences (at both nucleic acids and amino acid sequence levels) has had successful applications for these specific attributes (Brochmann et al. 1998; Burgess et al. 2011; Hollingsworth et al. 2009; Koch et al. 2001; Lahaye et al. 2008; Steele and Vigalys, 1994; Tamura et al. 2004). To distinguish numerous plant species, the *matK* DNA gene has been utilized as a barcode gene alone or combined with other plant barcode sequences in current investigations (Bandara et al. 2013; De Mattia et al. 2011; Guo et al. 2011; Jing et al. 2011; Lahaye et al. 2008; Li et al. 2011; Pang et al. 2011; Seberg et al. 2012; Zhi-Yuan et al. 2011).

Material and Method

Plant materials

Totally, 46 species in 11 sections representative of five subgenera of *Allium* were gathered from their natural habitat and relocated to the Iranian Biological Resource Center (IBRC), located in Karaj province of Iran during 2017-2018 (Table 1). The identification of the specimens was according to morphologic traits and diagnostic explanations of the species in related publications (Fritsch and Abbasi, 2013; Wendelbo, 1971). The genus was classified according to up-to-date infrageneric concepts presented previously (Friesen et al. 2006, Fritsch et al. 2010, and Fritsch and Abbasi 2013).

DNA extraction, PCR Amplification and DNA Sequencing

The entire DNA was extracted from fresh leaf tissues of *Allium* species by the "IBRC plant extraction kit" (IBRC IND.) as recommended by the manufacturer's protocol. The extracted DNA was assessed spectrophotometrically in terms of concentration and quality by determination of the absorbance at 260 nm and 280 nm. Different concentrations of template DNA and Mg and different annealing temperatures were used to optimize the PCR condition. Next, the PCR reaction materials were set in a final volume of 50 μL using *Taq* DNA Polymerase Master Mix RED (Amplicon), 1 ng of the primers, and 0.5 μL of genomic DNA.

The *matK* region was amplified using the primers 5'-CGA TCT ATT CAT TCA ATA TTT C- 3' and 5'- TCT AGC ACA CGA AAG TCG AAG T -3', *matK* 390F and

Table 1 *Allium* species analyzed using nucleotide sequences of *matK* DNA regions

Code	Species	Subgenus	Section	Accession number
1	*A. akaka sub akaka*	*Melanocrommyum*	*Acanthoprason*	P1009858
2	*A. akaka sub bozghushense*	*Melanocrommyum*	*Acanthoprason*	P1011312
3	*A. alamutense*	*Melanocrommyum*	*Acanthoprason*	P1010457
4	*A. ampeloprasum sub ampeloprasum*	*Allium*	*Allium*	P1010913
5	*A. ampeloprasum sub porrum*	*Allium*	*Allium*	P1010908
6	*A. asarence*	*Cepa*	*Cepa*	P1004474
7	*A. atroviolaceum*	*Allium*	*Allium*	P1006775
8	*A. austroiranicum*	*Melanocrommyum*	*Acanthoprason*	P1009666
9	*A. bisotunense*	*Melanocrommyum*	*Melanocrommyum*	P1011153
10	*A. breviscapum*	*Melanocrommyum*	*Acanthoprason*	P1010598
11	*A. breviscapum*	*Melanocrommyum*	*Acanthoprason*	P1010598
12	*A. cepa*	*Cepa*	*Cepa*	P1010909
13	*A. cf. assadi*	*Melanocrommyum*	*Megaloprason*	P1011015
14	*A. cf. cardiostemon*	*Melanocrommyum*	*Melanocrommyum*	P1009998
15	*A. cf. latifolium*	*Melanocrommyum*	*Acanthoprason*	P1010022
16	*A. chrysantherum*	*Melanocrommyum*	*Melanocrommyum*	P1010382
17	*A. derderianum*	*Melanocrommyum*	*Acanthoprason*	P1009475
18	*A. egorovae*	*Melanocrommyum*	*Acanthoprason*	P1011306
19	*A. fistulosum*	*Cepa*	*Cepa*	P1010910
20	*A. haemanthoides*	*Melanocrommyum*	*Acanthoprason*	P1011119
21	*A. iranicum*	*Allium*	*Allium*	P1009733
22	*A. iranshahrii*	*Melanocrommyum*	*Acanthoprason*	P1011271
23	*A. kazerouni*	*Melanocrommyum*	*Procerallium*	P1010341
24	*A. keusgenii*	*Melanocrommyum*	*Melanocrommyum*	P1010423
25	*A. koelzii*	*Melanocrommyum*	*Pseudoprason*	P1010355
26	*A. latifolium*	*Melanocrommyum*	*Acanthoprason*	P1011289
27	*A. mahneshanense*	*Melanocrommyum*	*Acanthoprason*	P1010524
28	*A. materculae*	*Melanocrommyum*	*Acanthoprason*	P1009955
29	*A. cf. minutiflorum*	*Melanocrommyum*	*Acanthoprason*	P1010986
30	*A. moderense*	*Melanocrommyum*	*Melanocrommyum*	P1011021
31	*A. mozaffarianii*	*Melanocrommyum*	*Melanocrommyum*	P1011232
32	*A. pesodoholandicum*	*Melanocrommyum*	*Procerallium*	P1010545
33	*A. psedoampeloprasum*	*Allium*	*Allium*	P1009877
34	*A. rubellum*	*Allium*	*Avulsea*	P1009972
35	*A. sabalanense*	*Melanocrommyum*	*Acanthoprason*	P1009837
36	*A. sativum*	*Allium*	*Allium*	P1010653
37	*A. scabriscapum*	*Reticulatobulbosa*	*Scabriscapa*	P1009668
38	*A. scabriscapum*	*Reticulatobulbosa*	*Scabriscapa*	P1009668
39	*A. stamineum*	*Allium*	*Codonoprasum*	P1009964
40	*A. stipitatum*	*Melanocrommyum*	*Procerallium*	P1010429
41	*A. sub akaka*	*Melanocrommyum*	*Acanthoprason*	P1009965
42	*A. tripedale*	*Nectaroscordum*	*Nectaroscordum*	P1009676
43	*A. ubipetrense*	*Melanocrommyum*	*Acanthoprason*	P1010544
44	*A. umbilicatum*	*Allium*	*Avulsea*	P1009439
45	*A. zagricum*	*Melanocrommyum*	*Acanthoprason*	P1011079
46	*Allium Sp.*	*Melanocrommyum*	*Acanthoprason*	P1009928

matK 1326R primers for *matK* gene. DNA was amplified on an Eppendorf Master cycler gradient (Eppendorf Scientific, Germany) with the setting below: initial denaturation for 2 min at 92°C, followed by 35 cycles of denaturation (94°C, 1 min), annealing (55°C, 1 min), extension (72°C, 55 sec); and a final extension for 5 min at 72°C. After completion of the PCR reaction, 5 µl of PCR solution with 5 µL loading buffer was decanted into 1.5% agarose gel well holding TBE buffer. The PCR products were subjected to electrophoresis at 90 V for 85 min. Thereafter, the agarose gel was stained for 20 min in 0.50 mg/l of ethidium bromide and replicated fragments were visualized under UV light

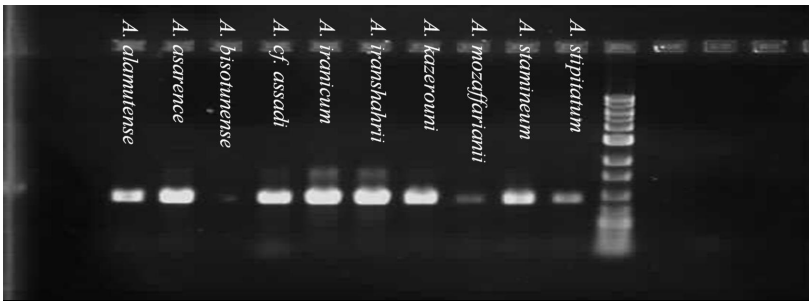

Fig. 1 Electrophoresis of purified PCR products of *matK* gene from *Allium* species

followed by gel document imaging. These PCR products were purified by the IBRC DNA purification kit (IBRC, Iran) based on the manufacturer's instructions. A 10 kbp DNA ladder (Thermo Scientific, USA) was utilized as a molecular size standard (Fig. 1). PCR amplification was redone two times or sometimes more for each primer to ensure that the results were reproducible. The sequencing products were produced with the Sanger method by MWG Co. (Germany) that performed sequencing in both directions by the PCR primers.

Editing, Sequence Alignment and Phylogenetic Reconstruction

Extraction of DNA sequences was performed from the chromatograms of the company by the use of Chromas v2, which were edited by the BioEdit program (Hall et al. 1999). The sequences were then put together with the CAP3 tool (Huang et al. 1999) and alignment of the homologous sequences was carried out using the EMBL-EBI CLUSTAL W tool (Edgar, 2004). The MEGA v6.0 software was used to analyze the multiple sequence alignment (MSA) file (Tamura et al. 2011). The genetic distance per loci among the accessions was estimated based on the number of base-pair replacement among the sequences. All the positions with missing data were eliminated by Kimura's 2-parameter model. Moreover, the distance between matrices from the three loci by DNAsp was analyzed with Pearson's correlation. The phylogenetic tree was developed by the Maximum Parsimony (MP) and Neighbor Joining (NJ) approach with a 1000 replicate bootstrap by the MEGA 6.0 software.

For estimating the resolution of DNA barcode, the percentage of produced monophyletic groups was determined using a bootstrap greater than 50% as a factor for defining the nodes, as recommended by Tripathi in DNAsp (Tripathi et al. 2013). The Tajima's D and Fu's Fs naturality tests were calculated by the DNAsp software. The amount of the dN/dS ratio was obtained numerically by the use of HIV databases. Haplotype network was scrutinized by the popART software (Leigh and Beryant, 2015).

Results

The scientific progressive innovations in molecular science and sequencing approaches has empowered the recognition of organismal genomes. Besides, important data are provided by continuing variety of ongoing genome projects for multiple species concerning their classification, gene structure, and application scientifically. Here, nucleotide polymorphisms of the *matK* gene quality are applied for 46 species of *Allium* to recognize the levels and patterns of interspecific and distinctions.

According to our findings on the analyzed *matK* gene sequence in *Allium* species, thymine (38.5%) and guanine (13.9%) bases had the most remarkable and the lowest nucleotide rates (Table 2). In the nucleotide substitution recorded in the *Allium* species, substitution rates were detected significantly in pyrimidine, namely 26.22 for C-T and 10.68% for T-C conversions (Table 3). The above levels were less for purines for the G-A and A-G conversions (19.22% and 8.08%, respectively). These findings correspond to those of other investigators reporting consecutive cases of pyrimidine substitutions, which most probably result from cytosine methylation (Picoult et al. 1999).

A total of 595 mutations was identified for genetic indicators of the *matK* gene in *Allium* species, which had different distributions all over the genome. Polymorphisms were detected in 243 sites, demonstrating the process of

Table 2 Nucleotide abundances derived from the *matK* gene of *Allium* species

G	C	T/U	A	Nucleotide
13.9	15.2	38.5	32.4	Frequency

Table 3 Nucleotide substitution pattern estimation matrix of the *matK* gene in *Allium* species

G	C	T	A	From/To
8.08	2.75	6.74	-	A
2.49	**10.68**	-	5.92	T
8.49	-	**26.22**	5.92	C
-	2.75	6.74	**19.22**	G

Each value represents the frequency of substitutions from one base (row) to another base (column). In this table, the percentage of transition mutations (purine-purine substitution, pyrimidine-pyrimidine substitution) and transversion mutations (purine-pyrimidine substitution and vice versa) are shown in bold and italics, respectively.

Table 4 Gene polymorphism of *matK* location in *Allium* species

K	Eta	Pi	Hd	H	S	Population
80.112	595	0.160	0.9874	39	423	*Allium*

S: Number of polymorphic positions, H: number of haplotypes, Pi: nucleotide diversity, Eta: total number of mutations, K: number of nucleotide differences between populations or species (nucleotide divergence)

Table 5 Conserved DNA regions of the *matK* gene in *Allium* species

CT	MWL	C	Population
0.26	50	0.162	*Allium*

C: Sequence conservation, MWL: Minimum conservation Length, CT: Conservation threshold

Table 6 Identification of the *matK* gene natural selection process in *Allium* species

Numerical value	Parameter
2.3254	dN
1.9970	dS
1.1644	dN/dS[†]

[†]The numerical value of dn / ds represents the natural selection process

Table 7 Results obtained from an evaluation of the natural evolution of the *matk* gene in *Allium* species

Fu's Fs	Tajima's D	Gene
-1.19	-1.50	*matK*

a positive selection of the *matK* gene sequence (Table 4). Searching for conserved DNA sections of the *matK* gene in the *Allium* species revealed a 0.26 CT region, a MWL of 50 bases, and a sequence conservation of 0.162 (Table 5). These conserved sections comprised a minor portion of the *matK* gene sequence, indicating that this site has discriminated differently and is susceptible to nucleotide alterations and mutations among various species, leading to variability among species.

An estimated value of 1.16 was obtained for the dN/dS proportion in *Allium* species (Table 6), suggesting the positive selection of the *matK* gene in *Allium* species in the course of evolution. This kind of evolution has led to new species and stabilized better refinement of their efficacy during the evolution, which arises from the conversion of non-coding gene sites to the gene coding sites. The same as Tajima's D and Fu's Fs, neutrality tests were obtained to determine deviations from the null theory on the neutral evolution and identify the impacts of natural selection on these genes in *Allium* species. Significantly negative estimates of D (-1.50) and Fu's Fs (-1.19) are made by groups of people under the influence of recently advanced or highly developed effective population size or the directional selection. Positively estimated values of D and Fs indicate the influences of the genetic drift, genetic dilemma, or a balancing effect over the evolutionary history of the population. The findings of the present research proved negative estimates of D and Fs (Table 7). The negative and positive results represent a significantly slight dissimilarity between polymorphisms with regard to their frequency, respectively. According to the outcomes of both neutrality tests, the continuing development of the *Allium* species has affected the planet or directional selection has influenced this gene throughout evolution. Altogether, Fu's Fs and Tajima's D tests have been proven to have higher effectiveness for small size and larger estimated populations.

It is of paramount importance to determine haplotype groups (by the popART software) for determination of the geographic regions of the examined breeds in comparison to other breeds. The entire 46 samples grouped in haplotype

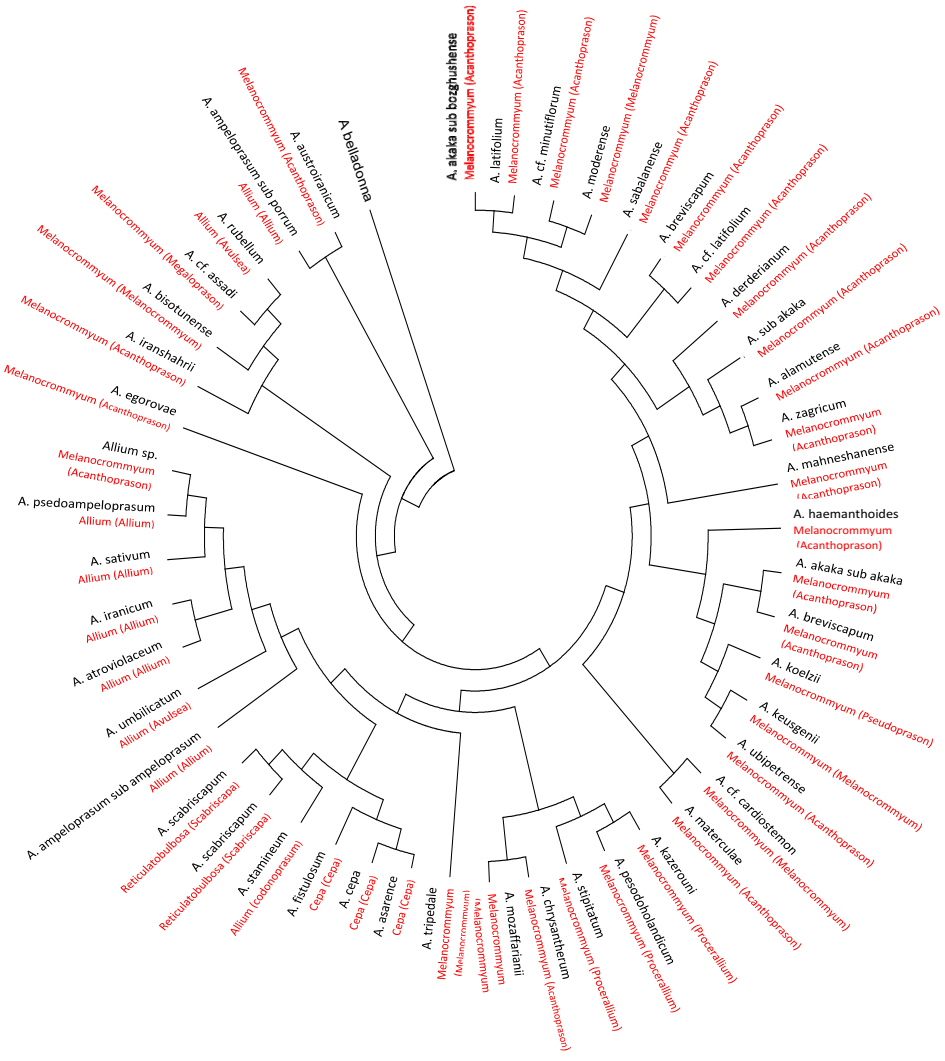

Fig. 2 Neighbor-joining phylogenetic tree analysis of 46 *Allium* species based on the nucleotide sequences of a *matK* region

bunch B is the largest in various species around the world. Based on our observations, the haplogroup A is usually present in all continents, and the haplogroup B may similarly have originated from Asia (Fig. 3) (Ghanbari et al., 2018).

The NJ tree was drawn by the Kimura distance (determination of distances aimed at building NJ tree using the MEGA 6.0 with the program defaults). The nucleotide sequences of the *matK* site presented no significantly different topology from the MP tree (MP tree was produced using the MEGA 6.0 by the program defaults), which differed in their branch-length (Kimura, 1980; Nei and Kumar, 2000).

In the current investigation, the *matK* gene sequence of *Amaryllis belladonna* was utilized for the outside group. Phylogenetic assays that depended on the nucleotide sequences of *matK* could mostly differentiate subgenera and sections in the genus *Allium*; though, few accessions of species were located outside of the section. Three discrete clades were produced by the NJ dendrogram. The first clade consisted of species *A. austroiranicum* and *A. ampeloprasum*. The second clade comprised the species *A. iranshahrii*, *A. bisotunense*, *A. cf assadi*, and *A. rubellum*, and the rest of species lied in the third clade.

A. egorovae presented middle sites between groups of the first and the second evolutionary lines in the tree topology. The produced *matK* phylogenetic tree endorsed the theory of *Allium* as a monophyletically originated genus. Some species analyzed phylogenetically for the *matK* site revealed that accessions were placed in distant clades. *A. ampeloprasum* lied in a clade near *A. umbilicatum* and another accession of *A. ampeloprasum* had a place in a distant clade with *A. austroiranicum*. The majority of the species were also distinguished by *matK*, and few species with close relations were not distinguishable. *A. zagricum* and *A. alamutense* lied in a similar clade without any distinct dissimilarities.

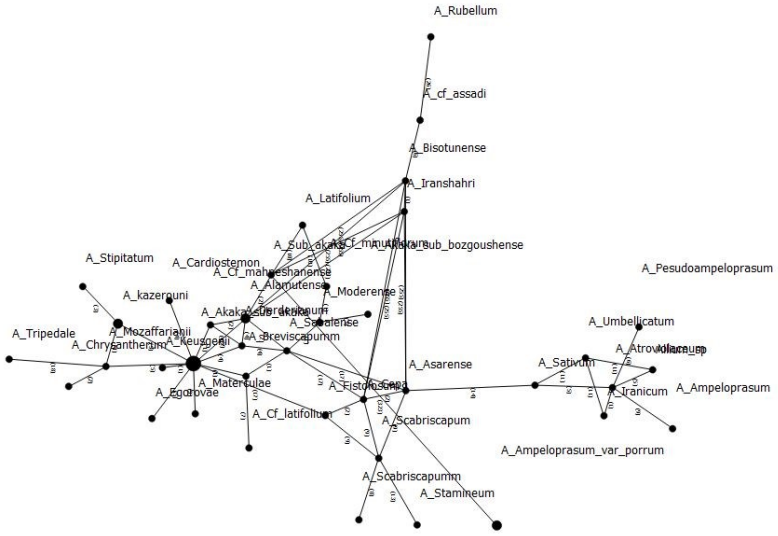

Fig. 3 Interconnectivity grid of the varieties of *Allium* species studied using the *matK* gene (popART software)

Discussion

In a previous study, the *matk* gene was proposed for resolving phylogenetic associations in plant species (Abugalieva et al. 2017; İpek et al. 2014 Ito et al. 1999; Kim et al. 2018; Son et al. 2010). The genes utilized for DNA barcoding in plant populations include *matK, rpoC1, rpoB, trnH-PsbA, rbcL, atpF-atpH, psbK-psbI*, and their combinations, of which, *matK* and *rbcL* were adopted as 2-locus DNA barcode by the CBOL Group (Hollingsworth et al. 2009). Polymorphic barcode nucleotide sequences are desirable at interspecific or upper classification levels, but not at the intraspecific level. Accordingly, comparison of the barcode nucleotide sequence can identify an unidentified plant accession as a species (Hebert et al. 2003; Kress and Erikson, 2008; Pang et al. 2011; Stoeckle, 2003). Additionally, the DNA fragment used as barcoding should have high recoverability, high proportion in species identification, and proper affordability (Burgess et al. 2011).

Here, the efficacy of the *matK* gene was examined as a DNA barcode for discriminating *Allium* species. Our findings indicated that *matK* sites could be easily amplified by PCR. DNA sections were mostly discrete parts in *Allium*. Nevertheless, some species were not distinguishable from one another by the *matK* sites (Abugalieva et al. 2017; Ince et al. 2005; İpek et al. 2014; Li et al. 2011; Son et al. 2010).

The length of *matK* nucleotide sequences varied between 540 bp in *A. porrum* and *A. austroiranicum* and 1023 bp in *Allium* sp. Genetic associations among *Allium* species were examined through the polymorphisms within the nucleotide sequences of *trnH-psbA*, ITS and *matK* in a recent report (İpek et al. 2014; Son et al. 2010). Over one haplotype in both organellar and nuclear genomes were present in a single plant of *Allium* species, making phylogeny analysis and barcoding implausible applying the nucleotide arrays of these DNA sites. At present, 39 haplotypes were noticed in *Allium* species for *matK* in the plastid genome.

Phylogenetic associations among *Allium* species on the basis of *matK* analyses are in line with earlier research by the use of the *matK* site (Abugalieva et al. 2017; Gurushidze et al. 2007; İpek et al. 2014; Li et al. 2010). *A. asarence* and *A. cepa* lied jointly in a group with no polymorphisms in some investigations (Li et al. 2010; Nguyen et al. 2008). Likewise, *A. asarence* and *A. cepa* put together in a cluster by analyzing *matK* in the present research. Moreover, an accession of *A. ampeloprasum* was tightly grouped with *A. umbilicatum* and *A. atroviolaceum*, which is almost similar to those of Nguyen et al. and Hirschegger et al. On the other side, an accession of *A. ampeloprasum (var. porrum)* lied in close grouping with *A. austroiranicum. A. Akaka* was placed tightly in a group with *A. breviscapum* and *A. sub akaka* was assigned to a group close to *A. alamutense* and A. *zagricum*, which have a closer relation to a recently published report (Akhavan et al. 2015; Li et al. 2010; Nguyen et al. 2008). Similar to a latest investigation, *A. pseudoampeloprasum* was clustered tightly to *A. atroviolaceum*, *A. iranicum,* and *A. sativum* in our research ((Hirschegger et al. 2010; Veiskarami et al. 2019). *A. umbilicatum* lied in a group with *A. ampeloprasum* and *A. atroviolaceum*, which corresponds to a latest research (Friesen et al. 2006; Li et al. 2010). *A. fistulosum* was in group close to *A. cepa* and *A. asarence* as reported previously (Veiskarami et al. 2019;

Li et al. 2010; Friesen et al. 2006). A similar finding was reported in an earlier study (Akhavan et al. 2015; Friesen et al. 2006; Fritsch and Abbasi, 2013; Fritsch et al. 2010; Gurushidze et al. 2010; Li et al. 2010; Sýkorová et al. 2006). Two accessions of *A. scabriscapum* were positioned in a similar clade and clustered closely with *A. asarence*, *A. cepa*, and *A. fistulosum* by analyzing *matK*, which is in agreement with a recently studied case (Friesen et al. 2006; Veiskarami et al. 2019). As detected in a recently published study, *A. sub Akaka*, *A. zagricum,* and *A. alamutense* were put together with no distance (Akhavan et al. 2015; Fritsch et al. 2010; Gurushidze et al. 2008; Li et al. 2010).

These observations indicated a close relation between *Allium* species in subgenera *Allium* and *Reticulatobulbosa*. Likewise, it was recently demonstrated that species in these subgenera were closely related phylogenetically (Friesen et al. 2006; Li et al. 2010). All accessions of *Allium* species in subgenus *Allium* were grouped in a similar clade. Yet, *A. rubellum* and *A. stamineum* from subgenus *Allium* were placed in the clade of subgenus *Melanocrommyum* and *Reticolatobulbosa* with *matK* examinations. This finding suggests that mixtures are present in these accessions. Accessions relating to the subgenus *Melanocrommyum* were put together in a clade and endorsed by a 100% bootstrap value. Similarly, the present outcomes according to the *mat*K site revealed that these species had close phylogenetic relations. To conclude on the basis of *matK* examination, phylogenetic associations among *Allium* species scrutinized here corresponded to those reported previously (Li et al. 2010; Veiskarami et al. 2019).

Here, *A. mahneshanense* was grouped with *A. zagricum* whereas it was clustered with *A. cf. minutiflorum* elsewhere. *A. cf. minutiflorum* was gathered with *A. moderense* while it was assigned to another *Allium* spp. elsewhere. In the current research, *A. iranshahrii*, *A. bisotunense*, and *A. cf. assadi* were put together tightly whereas *A. bisotunense* was clustered with *A. keusgenii* and *A. iranshahrii* was assigned closely to *A. haemanthoides* in other investigations. (Akhavan et al. 2015; Gurushidze et al. 2008; Li et al. 2010). It is proposed to utilize *matK* as an additional means for analyzing *Allium* phylogenetically as using the *matK* site had more simplicity for recouping and more affordability than other sites. The present information indicated that the sampled *Allium* species were related to the subgenera in the second and third evolutionary lines (Friesen et al. 2006; Fritsch and Abbasi 2013). The dendrogram in the *matK* examination recommends that species in the subgenus *Melanocrommyum* developed earlier than in the subgenus *Allium*, and these two subgenera possess a shared genetic node (Fig. 2).

As in the phylogenetic tree drawn in Friesen et al. (2006) in which subclades *Polyprason*, *Reticulatobulbosa*, and *Cepa* formed three sister subclades, three subclades were detected herein. The first subclade was species from the subgenera *Melanocrommyum*, whereas the succeeding one contained species of *Allium* and *cepa*. Our presented phylogeny of *matK* suggests that five subgroups can be distinguished within subgenus *Melanocrommyum*. *A. stipitatum*, *A. kazerouni*, and *A. pseudoholandicum* are the representatives of the first group (section *procelarium*). *A. chrysantherum*, *A. mozaffarianii*, *A. cf. cardiostemon, A. keusgenii, A. moderense*, and *A. bisotunense* (section *Melanocrommyum*) comprise the second group, *A. koelzii* (section *pseudoprason*) belongs to the third group, *A. cf. assadi* (section *megaloprason*) is related to the fourth group, and the rest of species from section *Acanthoprason* form the fifth group.

A research confirmed the *Allium* genus originated monophyletically, which was proved in some other literature (Friesen et al. 2006; Li et al. 2010; Nguyen et al. 2008). Additionally, a thorough research compared *Melanocrommyum* species and mainly focused on the morphological and molecular genetic explanation of species cultivated in Iran. In spite of the fact that Fritsch and Abbasi (Fritsch and Abbasi, 2013) studied fundamentally the phylogenetic taxonomy of the *Allium* genus, lots of ill-known *Allium* taxa remain accessible in all around the country. To evaluate the phylogeny of 46 endemics, scarce *Allium* species with economic importance from Iran, the *matK* DNA barcoding marker was employed in this research.

Reference

Abdulina SA. (1999) Checklist of vascular plants of Kazakhstan. *Institute of Botany and Plant Introduction, Almaty*, 1-187

Abugalieva S, Volkova L, Genievskaya Y, Ivaschenko A, Kotukhov Y, Sakauova G, and Turuspekov Y. (2017) Taxonomic assessment of *Allium* species from Kazakhstan based on ITS and *matK* markers. *BMC plant biology* 17, 258

Akhani H. (1999) Studies on the flora and vegetation of the Golestan National Park, NE Iran*. III. Three new species, one new subspecies and fifteen new records for Iran. *Edinburgh Journal of Botany* 56, 1-31

Akhavan A, Saeidi H, Rahiminejad MR, Zarre S, and Blattner FR. (2015) Interspecific relationships in *Allium* subgenus *Melanocrommyum* sections *Acanthoprason* and *Asteroprason* (*Amaryllidaceae*) revealed using ISSR markers. *Systematic botany* 40(3), 706-715

Bandara NL, Papini A, Mosti S, Brown T, and Smith LMJ. (2013) A phylogenetic analysis of genus *Onobrychis* and its relationships

within the tribe *Hedysareae* (*Fabaceae*). *Turkish Journal of Botany* 37(6), 891-992

Brochmann C, Xiang QY, Brunsfeld SJ, Soltis DE, and Soltis PS, (1998) Molecular evidence for polyploid origins in *Saxifraga* (*Saxifragaceae*): the narrow arctic endemic *S. svalbardensis* and its widespread allies. *American Journal of Botany* 85, 135-143

Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, and Barrett SC. (2011). Discriminating plant species in a local temperate flora using the rbcL+ *matK* DNA barcode. *Methods in Ecology and Evolution* 2(4), 333-340

Choi HJ, Giussani LM, Jang CG, Oh BU, and Cota-Sánchez JH. (2012) Systematics of disjunct northeastern Asian and northern north American *Allium* (*Amaryllidaceae*). *Botany* 90(6), 491-508

De Mattia F, Bruni I, Galimberti A, Cattaneo F, Casiraghi M, and Labra M. (2011) A comparative study of different DNA barcoding markers for the identification of some members of *Lamiacaea*. *Food Research International* 44(3), 693-702

Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32(5), 1792-1797

Friesen N, Fritsch RM, and Blattner FR. (2006) Phylogeny and new intrageneric classification of *Allium* (*Alliaceae*) based on nuclear ribosomal DNA ITS sequences. *Aliso: A Journal of Systematic and Evolutionary Botany* 22, 372-395

Fritsch RM, and Friesen N. (2002) Evolution, domestication and taxonomy. *Allium crop science: recent advances*, 5-30

Fritsch RM, Khassanov FO, and Matin F. (2002). New *Allium* taxa from Middle Asia and Iran. na

Fritsch RM, Salmaki Y, Zarre S, and Joharchi M. (2006) The genus *Allium* (*Alliaceae*) in Iran: current state, new taxa and new records. *Rostaniha* 7(2), 255-281

Fritsch RM, and Abbasi M. (2008) New taxa and other contributions to the taxonomy of *Allium* L. (*Alliaceae*) in Iran

Fritsch RM, and Abbasi M. (2013) A taxonomic review of *Allium* subg. *Melanocrommyum* in Iran. *Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung*

Fritsch RM, Blattner FR, and Gurushidze M. (2010) New classification of *Allium* L. subg. *Melanocrommyum* (Webb & Berthel.) Rouy (*Alliaceae*) based on molecular and morphological characters. *Phyton (Horn)* 49(2), 145-220

Fritsch RM, and Maroofi H. (2010) New species and new records of *Allium* L. (*Alliaceae*) from Iran. *Phyton (Horn)* 50, 1-26

Fritsch RM, Matin F, and Klaas M. (2001) *Allium vavilovii* M. Popov et Vved. and a new Iranian species are the closest among the known relatives of the common onion *A. cepa* L. (*Alliaceae*). *Genetic Resources and Crop Evolution* 48(4), 401-408

Fuse S, and Tamura MN. (2000) A phylogenetic analysis of the plastid *matK* gene with emphasis on *Melanthiaceae* sensu lato. *Plant Biology* 2(4), 415-427

Ghanbari S, Fakheri BA, Naghavi MR, and Mahdinezhad N. (2018) Evaluating phylogenetic relationships in the *Lilium* family using the ITS marker. *Journal of Plant Biotechnology* 45(3), 236-241

Govaerts R, Kingto, S, Friesen N, Fritsch RM, Snijman DA, Marcucci R, Silverstone Sopkin PA. and Brullo S. (2005-2014). World checklist of Amaryllidaceae (WWW Document)

Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, and Fazekas AJ. (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences* 106(31), 12794-12797

Guo J, SU JX, LIN RZ, LI RQ, and XIAO PG. (2011) Testing four proposed barcoding markers for the identification of species within *Ligustrum* L. (*Oleaceae*). *Journal of Systematics and Evolution* 49(3), 213-224

Guo X, Simmons MP, BUT PPH, SHAW PC, and WANG RJ. (2011) Application of DNA barcodes in *Hedyotis* L. (*Spermacoceae*, *Rubiaceae*). *Journal of Systematics and Evolution* 49(3), 203-212

Gurushidze M, Fritsch RM, and Blattner FR. (2010) Species-level phylogeny of *Allium* subgenus *Melanocrommyum*: Incomplete lineage sorting, hybridization and trnF gene duplication. *Taxon* 59(3), 829-840

Gurushidze M, Mashayekhi S, Blattner FR, Friesen N, and Fritsch RM. (2007) Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. *Plant Systematics and Evolution* 269(3-4), 59-269.

Hall TA. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* 41, 95-98

Harpke D, Meng S, Rutten T, Kerndorff H, and Blattner FR. (2013) Phylogeny of *Crocus* (*Iridaceae*) based on one chloroplast and two nuclear loci: ancient hybridization and chromosome number evolution. *Molecular phylogenetics and evolution* 66(3), 617-627

Hebert PD, Cywinska A, Ball SL, and Dewaard JR. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313-321

Hirschegger P, Jakše J, Trontelj P, and Bohanec B. (2010) Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section *Allium* (*Allium*: *Alliaceae*). *Molecular Phylogenetics and Evolution* 54(2), 488-497

Huang X, and Madan A. (1999) CAP3: A DNA sequence assembly program. *Genome research*, 9(9), 868-877

Ipek M, İpek A, and Simon PW. (2014) Testing the utility of *matK* and ITS DNA regions for discrimination of *Allium* species. *Turkish journal of botany*, *38*(2), 203-212

Ito M, Kawamoto A, Kita Y, Yukawa T, and Kurita S. (1999) Phylogenetic relationships of *Amaryllidaceae* based on *matK* sequence data. *Journal of Plant Research* 112(2), 207-216

Jing YU, Jian-Hua XUE, and Shi-Liang ZHOU. 2011. New universal *matK* primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution*, 49(3), 176-181

Ince AG, Karaca M, Onus AN, and Bilgen M. (2005) Chloroplast matK gene phylogeny of some important species of plants.

Akdeniz Universites Iziraatfakultesi Dergisi 18, 157-162

Kamelin R, and Seisums A. (1996) "Tri novykh vida roda *Allium* L." *Alliaceae*) iz yugo-zapadnoi Azii. (Tres species novae generis *Allium* L. (*Alliaceae*) ex Asia austro-occidentali.) Novosti Sistematiki Vysshikh Rastenii 30, 29-33

Khassanov FO, and Memariani F. (2006) *Allium joharchii*, a new species from Khorasan Province (Iran). *Rostaniha* 7(2), 63-69

Khassanov FO, Noroozi J, and Akhani H. (2006) Two new species of the genus *Allium* (*Alliaceae*) from Iran. *Rostaniha* 7(2), 119-129

Kim YB, Ramekar RV, Choi SJ, Choi BG, Kim SW, Moon YK, and Choi IY. (2018) Molecular identification of *Allium ochotense* and *Allium microdictyon* using multiplex-PCR based on single nucleotide polymorphisms. *Horticulture, Environment, and Biotechnology* 59(6), 865-873

Kimura M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* 16(2), 111-120

Koch M, Haubold B, and Mitchell-Olds T. (2001) Molecular systematics of the *Brassicaceae*: evidence from coding plastidic *matK* and nuclear Chs sequences. *American Journal of Botany* 88(3), 534-544

Kress WJ. (2017) Plant DNA barcodes: Applications today and in the future. *Journal of systematics and evolution* 55(4), 291-307

Kress WJ, and Erickson DL. (2008) DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences* 105(8), 2761-2762

Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, and Savolainen V. (2008) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences* 105(8), 2923-2928

Leigh JW, and Bryant D. (2015) popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6(9), 1110-1116

Li FW, Kuo LY, Rothfels CJ, Ebihara A, Chiou WL, Windham MD, and Pryer KM. (2011) *rbcL* and *matK* earn two thumbs up as the core DNA barcode for ferns. *PLoS One* 6(10), e26597

Li QQ, Zhou SD, He XJ, Yu Y, Zhang YC, and Wei XQ. (2010) Phylogeny and biogeography of *Allium* (*Amaryllidaceae*: *Allieae*) based on nuclear ribosomal internal transcribed spacer and chloroplast *rps16* sequences, focusing on the inclusion of species endemic to China. *Annals of botany* 106(5), 709-733

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, and Chen S. (2015) Plant DNA barcoding: from gene to genome. *Biological Reviews* 90(1), 157-166

Mashayekhi S, Zarre S, Fritsch RM, and Attar F. (2005) A new species of *Allium* subgen. *Melanocrommyum* sect. *Compactoprason* (*Alliaceae*) from Iran. *Feddes Repertorium: Zeitschrift für botanische Taxonomie und Geobotanik* 116(3-4), 191-194

Matin F. (1992) The genus *Allium* in Iran, diversity, distribution and endemism. In The genus *Allium*-taxonomic problems and genetic resources. *Proceedings of an international symposium held at Gatersleben* 193-194

Memariani F, Joharchi MR, and Arjmandi AA. (2012) *Allium aladaghense* (*Amaryllidaceae*, *Allieae*), a new species of section *Asteroprason* from northeast of Iran. *Phytotaxa* 56(1), 28-34

Memariani F, Jouharchi M, and Khassanov FO. (2007) *Allium* L. subgen. *Rhizirideum* sensu lato in Iran, two new records and a synopsis of taxonomy and phytogeography

Nei M, and Kumar S. (2000) Molecular evolution and phylogenetics. Oxford university press

Neshati F, Zarre S, Fritsch RM, and Joharchi MR. (2009) *Allium oriento-iranicum* (*Alliaceae*), a new species from Iran. In *Annales Botanici Fennici* 46(6), 599-602

Nguyen NH, Driscoll HE, and Specht CD. (2008) A molecular phylogeny of the wild onions (*Allium*; *Alliaceae*) with a focus on the western North American center of diversity. *Molecular Phylogenetics and Evolution* 47(3), 1157-1172

Pang X, Song J, Zhu Y, Xu H, Huang L, and Chen S. (2011) Applying plant DNA barcodes for *Rosaceae* species identification. *Cladistics* 27(2), 165-170

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, and Boyce-Jacino M. (1999) Mining SNPs from EST databases. Genome Research (9), 167-174

Razyfard H, Zarre S, Fritsch RM, and Maroofi H. (2011) Four new species of *Allium* (*Alliaceae*) from Iran. In *Annales Botanici Fennici* 48(4), 352-361

Seberg O, Petersen G, Davis JI, Pires JC, Stevenson DW, Chase MW, and Pillon Y. (2012) Phylogeny of the *Asparagales* based on three plastid and two mitochondrial genes. *American Journal of Botany* 99(5), 875-889

Son JH, Park KC, Kim TW, Park YJ, Kang JH, and Kim NS. (2010) Sequence diversification of 45S rRNA ITS, *trnH-psbA* spacer, and *matK* genic regions in several *Allium* species. *Genes & Genomics* 32(2), 165-172

Steele KP, and Vilgalys R. (1994) Phylogenetic analyses of *Polemoniaceae* using nucleotide sequences of the plastid gene *matK*. *Systematic Botany* 126-142

Stoeckle M. (2003) Taxonomy, DNA, and the bar code of life. *BioScience* 53(9), 796-797

Sýkorová E, Fajkus J, Mezníková M, Lim KY, Neplechová K, Blattner FR, and Leitch AR. (2006) Minisatellite telomeres occur in the family *Alliaceae* but are lost in *Allium*. *American journal of botany* 93(6), 814-823

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* 28(10), 2731-2739

Tamura MN, Yamashita J, Fuse S, and Haraguchi M. (2004) Molecular phylogeny of monocotyledons inferred from combined analysis of plastid *matK* and *rbcL* gene sequences. *Journal of Plant Research* 117(2), 109-120

Tripathi AM, Tyagi A, Kumar A, Singh A, Singh S, Chaudhary LB, and Roy S. (2013) The internal transcribed spacer (ITS) region and *trnhH-psbA* are suitable candidate loci for DNA

barcoding of tropical tree species of India. *PloS one* 8(2), e57934

Veiskarami GH, Khodayari H, Heubl G, Weigend M, and Zarre S. (2019) Phylogenetic relationships in *Allium* sect. *Allium* (*Amaryllidaceae*, *Allioideae*) in Iran as inferred from nrDNA ITS, cpDNA rps16 and *trnL-F* sequences. *Nordic Journal of Botany*

von Berg GL, Samoylov A, Klaas M, and Hanelt P. (1996) Chloroplast DNA restriction analysis and the infrageneric grouping of *Allium* (*Alliaceae*). *Plant systematics and evolution*, 200(3-4), 253-261

Wendelbo P. (1971) Some distributional patterns within the Flora Iranica area. *Plant Life of South West Asia*

Zhi-Yuan DU, Qimike A, Chun-Feng Y, Jin-Ming C, and Qing-Feng W. (2011) Testing four barcoding markers for species identification of Potamogetonaceae. *Journal of Systematics and Evolution* 49(3), 246-251