Research Paper / 23-34

Nucleotide variation and secondary structures of nuclear ribosomal ITS2 in phylogeny relationships of some wheat species

Sedigheh Fabriki-Ourang[1*], Jafar Ahmadi[1], Hamid Karimi[1]

[1]Department of Genetics and Plant Breeding, Imam Khomeini International University, P. O. Box: 34148-96818, Qazvin, Iran.
[*]Corresponding author, Email: ourang@eng.ikiu.ac.ir. Tel: +98-28-33901227.

Received: 16 Mar 2021; Accepted: 13 Jun 2021.
DOI: 10.30479/ijgpb.2021.15244.1295

Abstract

This research aimed to estimate the nucleotide polymorphism and point mutation/substitutions for ITS2 region in 12 different Triticum and Aegilops species, and to demonstrate their phylogenetic relationships. The aligned ITS2 dataset showed a total of 67 SNPs and transitions (75%) predominant over transversions (25%). The C/T transitions were observed at highest level in the studied species. The highest nucleotide substitutions occurred in ITS2 sequences of *T. aestivum*, *T. turgidum*, and *A. speltoides* that showed similar substitutions on the base sites 76, 193, 215, 230 and 241 (TCCAG changed to ATTTA). *T. boeticum* and *T. urartu* showed similar substitutions on the base sites 85, 201, 231, 234 and 252 (GTGCC changed to AGATA). Remarkably, the nucleotide of base site 146 was C in all Triticum species, but T in all Aegilops species. The substitution rate in *A. neglecta*, and *A. umbellulata* was zero, it was concluded that their ITS2 is conserved. The secondary structures of ITS2 in the studied species was similar to that of the other flowering plants with four helices and the variation in helices length. The helix III had the longest length compared to the other helices. The length ranges of helices I, II, IV varied from 9 (*T. turgidum*, and *T. urartu*) to 10, 11 (*A. cylindrica*) 13 and 6 (*A. speltoides*) and 10 (*T. boeticum*) paired bases, respectively. Using pairwise ITS2 nucleotide comparisons, the closeness of *A. caudata* with *A. cylindrica*, *T. boeticum* with *T. urartu, A. neglecta* with *A. umbellulata* was observed. Phylogeny trees classified 12 species into four main clades. The highest point mutations occurred in ITS2 sequences related to the species grouped in the first clade.

***Key words*:** Aegilops, ITS, Point mutation, Phylogeny, Wheat.

INTRODUCTION

The wild relatives of crops constitute an increasingly genetic resource for improving agricultural production. Wild species of Aegilops and Triticum provide a useful source of new genetic variation for wheat improvement. Bread wheat is one of the most important cereal crops in the world consisting of A, B, and D genomes. About their original donors, the genome D is commonly accepted from *T. tauschii* (Jones *et al*., 1982) but the sources of A and B genomes are almost indefinite. The genome A was reported to be originated from *T. urartu* (Dvogak *et al*., 1993). Some reports supported *T. speltoides as* the donor of genome B, however, the results from molecular researches such as repeated DNA sequence analysis (Konarev, 1983) proved *T. searsii* or *T. sharonense* the probable donor of genome B. It was reported that the genome B may be originated from Aegilops species, such as *Ae. speltoides* (Kerby and Kuspira, 1987). The genus Aegilops has played a significant role in the evolution of wheat through natural hybridization (Sliai and Amer, 2011). It has been found that B and G genomes of allopolyploid wheats were closely similar to the S genome of Aegilops (Kilian *et al*., 2011; Haider, 2013), while *Ae. tauschii*

served as a donor of the D genome (Golovnina *et al*., 2009; Bordbar *et al*., 2011). The crossing of donor species with A and B genomes formed the tetraploid wheat *i.e*. *T. turgidum*, then hexaploid bread wheat was formed through the hybridization between *T. turgidum* and *T. tauschii* (Chen, 1997). Mehrabi and Safari (2019) results using internal transcribed spacer (ITS) regions confirmed that the genome A of bread wheat is more related to *T. urartu* than *T. boeticum*. Also their results revealed that *Ae. speltoides* was separated from Aegilops cluster and grouped with polyploid wheats. The close relationship between *Ae. speltoides* and polyploid wheats indicates that it is the most likely donor of the B genome to wheats (Mehrabi and Safari, 2019). The ITS1, ITS2 and 5.8S rDNA repeat segments from four species of Triticum and two subspecies of *Aegilops speltoides* were sequenced and the nucleotide sequences were phyletically analyzed. The results postulated that *T. boeoticum* is likely to be a genome donor of both, *T. dicoccoides* (AB) and *T. aestivum* (ABD) (Gulbitti-Onarici *et al*., 2009).

Genetic relationships among Aegilops and Triticum accessions have been investigated based on morphological, cytogenetic, and molecular markers (Kilian *et al*., 2011; Alnaddaf *et al*., 2012). DNA markers have provided new information to discover genetic relationships on different Triticum/Aegilops species (Petersen *et al*., 2006; Poczai and Hyvonen, 2010). The rDNA spacer regions could have a potential in the study of genetic relationships because of their simplicity, and small size (Nepolo *et al*., 2010; Poczai and Hyvonen, 2010). The nuclear ribosomal DNA-internal transcribed spacer (ITS) has been established as a suitable genomic region to clarify genomic and phylogenetic relationships of plants (Wang *et al.,* 2013). rDNAs is transcribed to rRNAs as a single precursor, which is subsequently cleaved by a series of nucleolar events leading to the mature small rRNA subunits *i.e.* the first internal transcribed spacer (ITS1), 5.8S rRNA, the second internal transcribed spacer (ITS2) and 26S/28S rRNA, respectively (Lenaers *et al*., 1989). In the maturation processing of rRNAs, ITS1 and ITS2 play important roles (Liu and Schard, 1994; Michot *et al*., 1999), which require secondary structures formation. The ITS2, which separates the nuclear ribosomal genes 5.8S rRNA and 28S rRNA, provides completely different associated characteristics for phylogenetic analyses (Alvarez and Wendel, 2003; Coleman, 2003; Feliner and Rossello, 2007). ITSs, one of the most important molecular markers in systematic and evolution studies, shows significant sequence diversity at the levels of species (Coleman, 2003; Thornhill *et al.,* 2007). Compared with rbcL, matK and psbA-trnH, ITS and ITS2 had a higher mutation rate and provided more information sites, and ITS2 had higher interspecific diversity and lower intraspecific variation in Rehmannia, but the interspecific genetic variation of rbcL and matK was lower (Duan *et al*., 2019). The diversity in ITSs structural information permits higher taxonomic level analysis (Coleman, 2003; Aguilar and Sanchez, 2007; Thornhill *et al.,* 2007; Schultz and Wolf, 2009; Coleman, 2009), which provides additional information for high accuracy and strength in the reconstruction of phylogenetic trees (Keller *et al.,* 2010).

Studies on the secondary structure of ribosomal RNAs have increased since the beginning of nucleic acid sequencing. General models of the secondary structure have been proposed for the different rRNAs that appear to be conserved among all eukaryotes. Many reports have emphasized that the secondary structure of the ITS regions is conserved at higher systematic levels (Mai and Coleman, 1997; Coleman *et al.,* 1998). Comparative studies in the secondary structure of ITS regions are little, whereas they could be useful for improving alignments at higher systematic levels, and the consisted secondary structure framework is considered as a tool for expanding the molecular phylogenies of plant species (Coleman and Mai, 1997; Joseph *et al*., 1999; Gottschling *et al*., 2001).

The ITS has already been proposed and widely used as universal barcode marker for plants, but a comprehensive, updated and accurate reference dataset of plant ITS sequences has not been available so far (Banchi *et al*., 2020). ITS2 is potentially useful as a standard DNA barcode to identify plant species (Chen *et al.,* 2010; Luo *et al.,* 2010). Also, it is one of the molecular markers which serves as a candidate DNA barcode because of its valuable characteristics, including the sufficient variability to distinguish closely related species and the availability of conserved regions for designing universal primers. Another advantage of the ITS2 is that they examine and confirm phylogenetic relationships. On the other hand, because of the high amount of nucleotide substitutions in ITS2, it could be used as a barcoding marker to distinguish species by a short fragment of the genomic DNA. The present study aimed to estimate nucleotide polymorphisms or SNPs of ITS2, variation of point mutation/substitutions in ITS2 region of different Triticum and Aegilops species, and to demonstrate their phylogenetic relationships.

MATERIALS AND METHODS

Plant materials and DNA extraction

Twelve species of Aegilops and Triticum were selected for this study, which were collected from different parts of Zagros Mountains in Northwes to Southwest of Iran (provided by the Gene Bank of Ilam University located in the west of Iran). A summary of geographical distributions and genome information of these species are shown in Table 1. For each species, total genomic DNA was extracted from young fresh leaves according to the CTAB method (Piccolo *et al.,* 2012) with minor modification. DNA quantity and quality was determined by NanoDrop-2000c spectrophotometer (BIO-RAD) as well as 0.8% agarose gel electrophoresis. Extracted DNA samples were stored at −20 °C prior to amplification.

PCR amplification and sequencing of ITS2 region

Nuclear rDNA-ITS2 region in the 12 species of Aegilops and Triticum were amplified by polymerase chain reaction with specific primer pairs ITS2F: 5'-ATGCGATACTTGGTGTGAAT-3' and ITS2R: 5'-GACGCTTCTCCAGACTACAAT-3 (Chen *et al.,* 2010). The PCR amplifications were performed using Taq-Amplicon PCR Master-Mix in total volumes of 50 µL reaction mixture according to the manufacturer's instructions. The PCR thermal cycling program consisted of an initial denaturation step at 95 °C for 5 min, followed by 40 cycles of denaturation (30 s at 95 °C), annealing (30 s at 56 °C), extension (45 s at 72 °C), and a final extension at 72 °C for 10 min. The PCR products for each species were sequenced by Bioneer company with both forward and reverse primer pairs.

Data analysis

The descriptive statistics for ITS2 sequences were calculated with EditSeq software (DNAStar, Inc., Madison, WI, USA) (Thompson *et al.,* 1994) and CLC Main Workbench 5.5. Genetic distance between studied species was estimated using MEGA 7.0 (Kumar *et al.,* 2016) based on the ITS2 sequences according to the Tamura-Nei model (Tamura and Nei, 1993). Internal transcribed spacer2 secondary structures predicted based on ITS2 online databases (http://its2.bioapps.biozentrum.uni-wuerzburg.de/) (Schultz *et al.,* 2006). Multiple sequences alignment was carried out using the ClustalW algorithm (Thompson *et al.,* 1994) followed by manual adjustment implemented by CLC Main Workbench 5.5. Phylogenetic analyses were conducted using Neighbor-Joining (Saitou and Nei, 1987), Maximum Likelihood, Minimum evolution (Rzhetsky and Nei, 1992), and Maximun Parsimony methods. Phylogenetic analyses were performed in CLC Main Workbench 5.5. Topological robustness was assessed by bootstrap analysis with 1000 replicates.

RESULTS AND DISCUSSION

PCR amplification and sequence analysis

After amplification of the ITS2 region, all studied species displayed a single band of PCR product (Figure 1). The length of ITS segments including the partial boundaries (5.8S rRNA and 26S rRNA) was 458-468 bp in all species. The ITS2 boundaries were determined using the conserved sequence of the two flanking regions (5.8S rRNA and 26S rRNA) by comparison with the available sequences of ITS2 databases (Ankenbrand *et al.,* 2015). After trimming and accurate analysis, all ITS2 sequences for examined Triticum and Aegilops species were submitted to GenBank and registered (listed in Table 1) with

Table 1. The list, genome information and registered ITS2 Gene-Bank accession no. for 12 Triticum/ Aegilops species used in this study.

Species	Genome	Collection site Province	Gene-Bank accession no.
Triticum aestivum	AABBDD	West Azerbaijan, Iran	MK611977
Aegilops speltoides	SS	Kermanshah, Iran	MK387299
Aegilops triuncialis	CCUU	Chaharmahal Bakhtiari, Iran	MK387301
Aegilops neglecta	UUMM	Ilam, Iran	MK387302
Aegilops tauschii	DD	Gilan, Iran	MK387303
Aegilops umbellulata	UU	Ilam, Iran	MK387304
Aegilops crassa	DDMM	Lorestan, Iran	MK612080
Aegilops cylindrica	DDCC	Zanjan, Iran	MK612093
Triticum urartu	$A^u A^u$	Kurdistan, Iran	MK612079
Triticum turgidum	AABB	Isfahan, Iran	MK612075
Triticum boeticum	$A^b A^b$	Kohgiluyeh and Boyer-Ahmad, Iran	MK612045
Aegilops caudata	CC	Lorestan, Iran	MK387300

the accession numbers of MK611977, MK387299, MK387301, MK387302, MK387303, MK387304, MK612080, MK612093, MK612079, MK612075, MK612045, MK387300. Base compositions of ITS2 sequences were highly similar (20–21% A, 14–17% T, 28–30% G, 28–35% C). The mean of GC content was 63.6% ranging from 62% in both *T. aestivum* and *A. speltoides* to 65% in both *A. caudata* and *A. sylindrica* (Table 2).

Sequence alignment and characteristics of ITS2 region

The multiple sequences alignment of the ITS2 sequences for the 12 species is presented in Figure 2. As expected, based on complete ITS2 and the adjacent partial 5.8S and 26S rRNA genes, the sequences alignment indicated the presence of conserved structural elements in these regions, further supporting homology of the observed sequence similarities in these regions. Length variation is a difficulty often encountered in non-coding regions of the rDNA and makes homology determination among sites a complex procedure (Wheeler, 1994). ITS sequences differ between species, but are generally conserved in length and composition within species (Hillis and Dixon, 1991). However, intra/inter-specific variations have also been reported in the form of single base point mutations or the insertion of a few base repeats. Analysis of variation in ITS2 sequences revealed transversion/transition substitutions. The aligned ITS2 dataset (Figure 2 and Table 3) showed a total of 67 variable positions or SNPs. The variable positions were found on ITS2 base sites of 71, 76, 80, 85, 89, 112, 115, 119, 132, 135, 142, 146, 169, 193, 200, 206, 215, 217, 221, 223, 229, 230, 231, 234, 241, 248, 252, 255 and 258 (Figure 2). In particular, the transition substitutions (75%) was predominant over transversion substitutions (25%) (Table 3). The C/T transitions play an important role in the ITS evolution of plants by achieving part of the GC balance recognized in the spacers (Torres *et al.,* 1990). In accordance, the C/T transitions were observed at the highest level (50%) in the studied Triticum and Aegilops species (Table 3). SNPs are excellent tools to obtain useful DNA markers for quickly identification of a single base variation within the genome. Also,

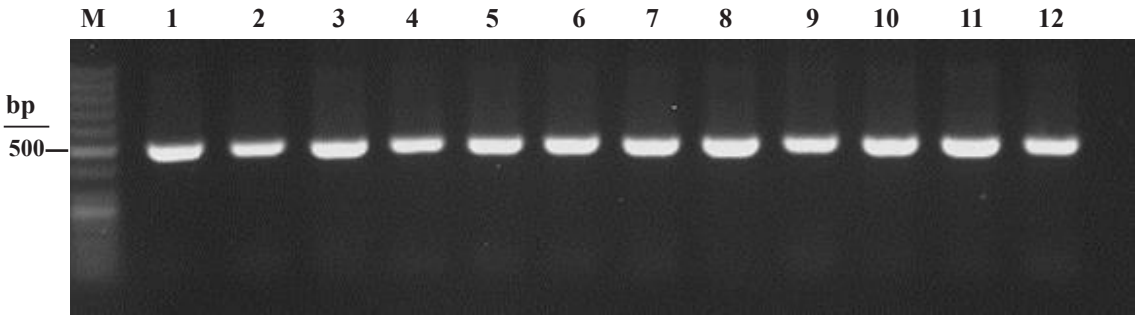

Figure 1. PCR products of fragments including the ITS2 region from Triticum and Aegilops species (1-12). M: DNA size marker.

Table 2. Analysis of ITS2 sequences and statistical values of the secondary structures in 12 Triticum/Aegilops species.

Species names	Product size (bp)	Size of ITS2 (bp)	Helices length (in paired bases)				CG (%)	ΔG
			I	II	III	IV		
T. aestivum	459	221	10	13	35	7	62	-85.9
A. speltoides	464	220	10	13	34	6	62	-74.6
A. triuncialis	466	218	10	13	33	7	63	-75.5
A. neglecta	460	219	10	13	34	7	64	-81.2
A. tauschii	460	219	10	13	34	7	64	-80.6
A. umbellulata	464	221	10	13	34	7	64	-81.2
A. crassa	464	221	10	13	32	7	64	-75.4
A. cylindrica	461	221	10	11	34	7	65	-81.2
T. urartu	464	220	9	13	34	8	64	-80.1
T. turgidum	463	221	9	13	35	7	62	-83.5
T. boeticum	468	220	10	13	35	10	64	-70.4
A. caudata	458	221	10	13	34	7	65	-81.2

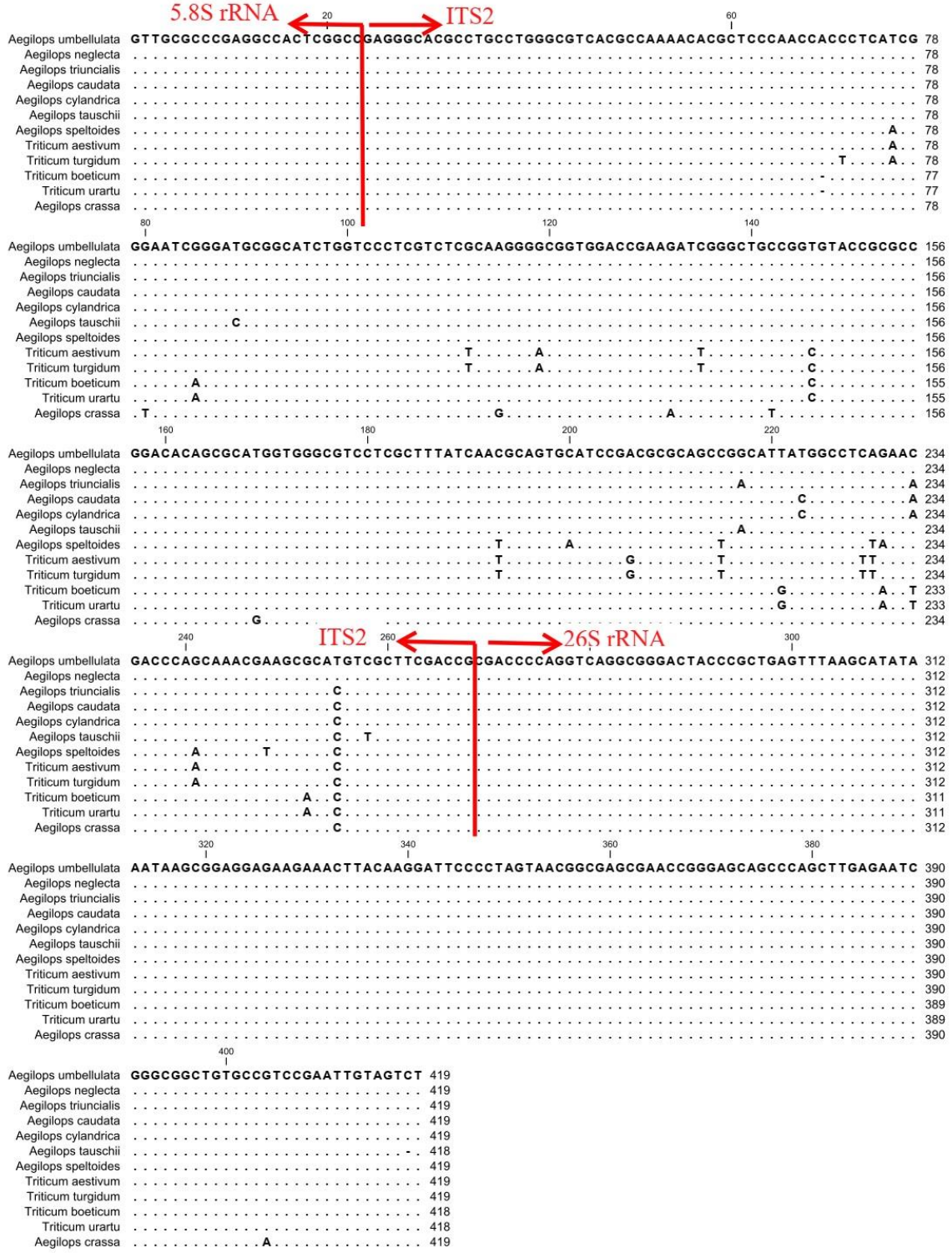

Figure 2. Multiple sequences alignment of 5.8S (partial), ITS2 (complete) and 26S (partial) ribosomal DNA for studied species of Triticum/Aegilops. Identities with the first sequence are indicated by dots.

they are the most used molecular markers because of their high density throughout the genome. On the other hand, ITSs show a relatively high number of SNPs and are easily isolated by PCR_ because they are located between highly conserved sequences (between 18S/5.8S or 5.8S/26S genes) that constitute the target site for the construction of primers. These two feature make ITSs efficient molecular markers for phylogenies and fingerprinting studies in plants (Balasubramani *et al.,* 2010).

Table 3. Point mutations and SNPs counted in ITS2 sequences of 12 Triticum/Aegilops species.

Species	Transition from				Transversion from							
	C	T	A	G	A	A	G	G	C	C	T	T
	to				to							
	T	C	G	A	C	T	C	T	A	G	A	G
A. neglecta	.	.	.	.	.	.	.	.	.	.	.	.
A. triuncialis	.	1	.	1	.	.	.	.	1	.	.	.
A. caudata	.	2	.	.	.	.	.	.	1	.	.	.
A. cylindrica	.	2	.	.	.	.	.	.	1	.	.	.
A. tauschii	1	2	.	1	.	.	.	.		.	.	.
A. speltoides	2	1	.	2	.	2	.	.	1	.	1	.
T. aestivum	5	2	1	2	.	1	.	1	.	.	1	.
T. turgidum	5	2	1	2	.	1	.	1	.	.	1	.
T. boeticum	1	2	.	3	.	.	.	.	.	.	.	1
T. urartu	1	2	.	3	.	.	.	.	.	.	.	1
A. crassa	1	1	1	1	.	.	.	1	.	.	.	1
Total	16	17	3	15		4		3	4		3	3

Secondary structures of ITS2

To identify the wheat species, we focused not only on the divergence of primary sequences of ITS2, but also on the use of variations in the secondary structures of ITS2. With prediction the secondary structure of ITS2 in the studied Triticum/Aegilops species (Figure 3) using the online ITS2 databases (Ankenbrand *et al.,* 2015) and counting the number of nucleotide base pairs of each helix and the number of unpaired loops (Table 2 and Figure 3), it was found that the secondary structures of ITS2 in the Triticum/Aegilops species were similar to the conventional structure of ITS2, a general secondary structure model consisting of a multi-branch loop with several paired regions, in the flowering plants that have four helixes (Schultz *et al.,* 2005). In line with the results of this research, Goertzen *et al.* (2003) reported that the secondary structure of the wheat ITS2 sequences mainly correspond to the common secondary structure (with four helices I–IV) in angiosperms.

The studied Triticum/Aegilops species showed a unique secondary structure that differed in this structure in two respects with each other, the length of helices and the number of loops on their helices (Figure 2). The variation in helices length in the secondary structure of ITS2 was observed at different Triticum/Aegilops species (Table 2 and Figure 3). The helix III had the longest length, varied from 32 (*A. crassa*) to 35 (*T. aestivum, T. turgidum,* and *T. boeticum*) paired bases, compared to other helices. The length ranges of helices I, II, IV varied from 9 (*T. turgidum*, and *T. urartu*) to 10 (Other species), 11 (*A. cylindrica*) 13 (Other species) 6 (*A. speltoides*) and 10 (*T. boeticum*) paired bases, respectively (Table 2 and Figure 3). All species, with the exception for *T. boeticum* possessing 4 loops, had 3 loops on their helix I. All species had 2 loops on their helix II. The number of loops in helix III were 7 for *T. aestivum*, *T. turgidum* and 8 in the rest of species. Also, the number of loops in helix IV was 2 in *T. urartu* and *A. speltoides*, 4 in *T. boeticum* and 1 in all other species.

One of the most widely used DNA markers in the fingerprinting and barcoding for plants, which in many cases has good efficiency for the correct diagnosis of species of the same genus, was used in this research. Considering the efficiency of ITS2 area, for distinguishing plant species, it was expected that this marker would be efficient in identifying the various species of Triticum/Aegilops. The secondary structure of ITS2 in the studied Triticum/Aegilops species was similar to that of other flowering plants with four helices, and also the length of the third helix as maximum length in flowering plants (Wilson, 2003; Schultz *et al.,* 2005; Coleman, 2007). In several studies on ITS2 in different genera (Coleman, 2003; Keller *et al.,* 2008; Chen *et al.,* 2010), the secondary structure of ITS2 has always been one of the main tools for species identification and differentiation within genus from each other. In accordance with the other reports, the secondary structure of ITS2 could support the identification of different Triticum/Aegilops species and also could be considered as an informative DNA marker. In fact, in the genus whose nucleotide substitutions in the ITS region occur more often, the

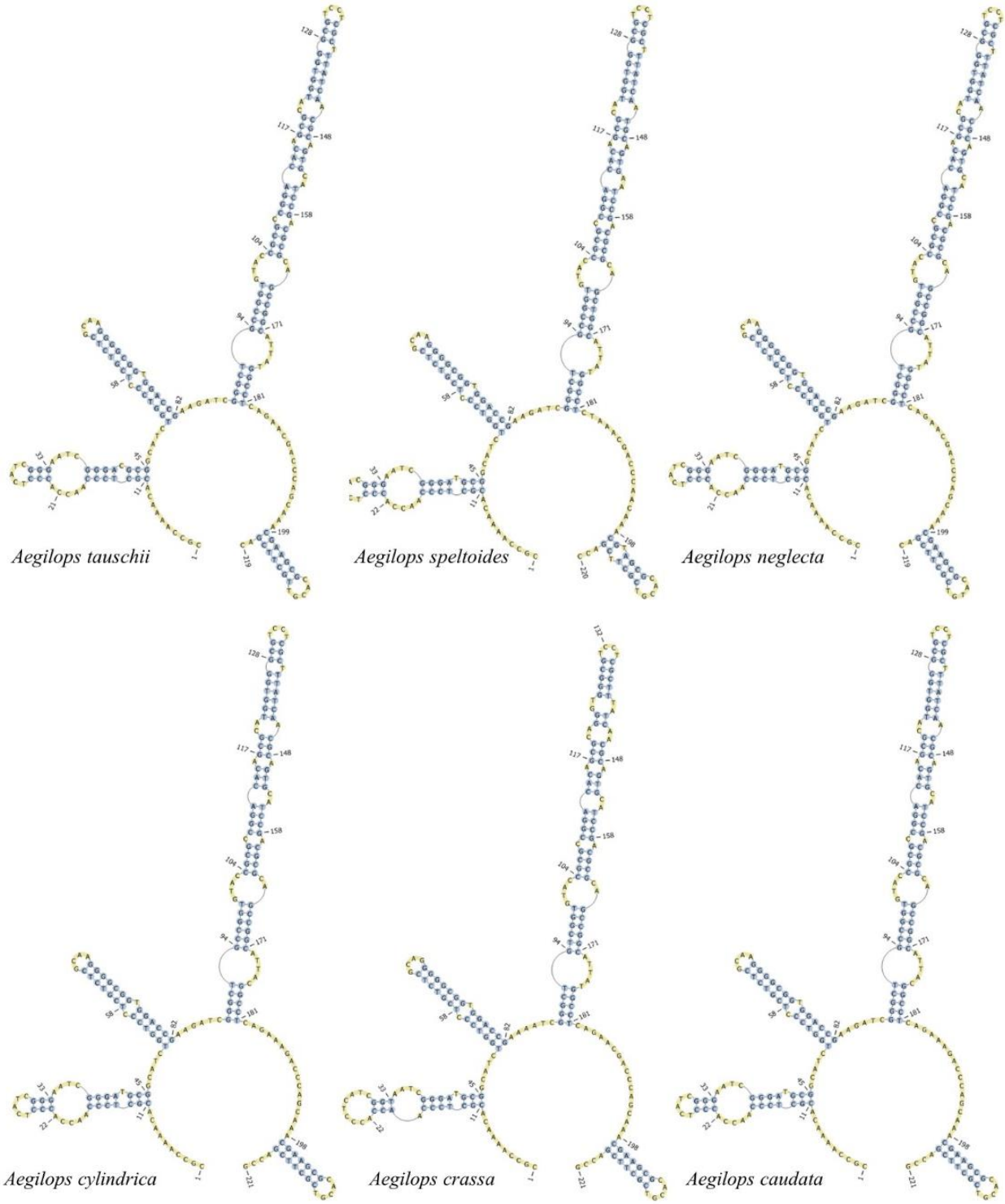

Figure 3. The predicted secondary structures of the internal transcribed spacer (ITS2) regions for 12 species of Triticum/Aegilops.

ITS2 secondary structure can succeed in differentiating different species (Meyer and Paulay, 2005).

Phylogenetic analyses

In distance matrix (Table 4), *T. aestivum* and *T. turgidum* revealed the highest different nucleotide residues with all other species. The lowest nucleotide differences were observed between *A. caudata* and *A. cylindrica*, *T. boeticum* and *T. urartu, A. neglecta* and *A. umbellulata*. This displays a higher nucleotide identity in ITS2 sequences in the mentioned species pairs. The phylogenetic relatedness of these species agrees with other reports (Gulbitti-Onarici *et al*., 2009; Mehrabi and safari, 2019).

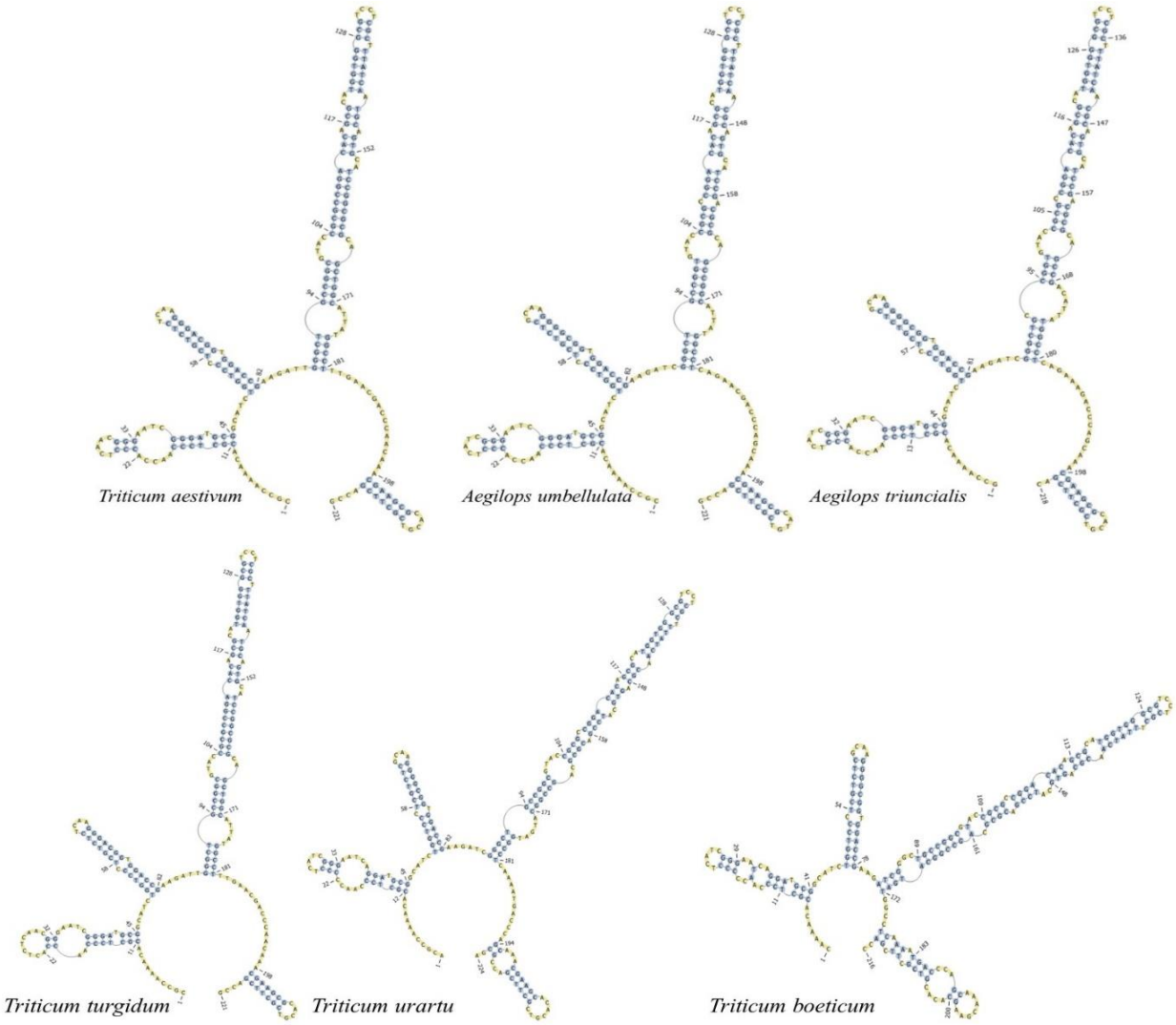

Figure 3 (Continued). The predicted secondary structures of the internal transcribed spacer (ITS2) regions for 12 species of Triticum/Aegilops.

Phylogeny trees of the studied Triticum/Aegilops species were constructed using ITS2 sequences with the methods of Neighbor-Joining (Figure 4A), Maximum Likelihood (Figure 4B), Minimum evolution (Figure 4C), and Maximun Parsimony (Figure 4D). The four trees had almost a consistent topological structure. The 12 species were classified into four main clades. The first clade included the species *T. aestivum*, *T. turgidum*. *A. speltoides T. boeticum*, and *T. urartu* were located in clade II. The third clade consisted of *A. crassa*, *A. neglecta*, and *A. umbellulata* species. The clade IV with two sub-groups included the species *A. triuncialis* and *A. tauschii* in sub-group IVa and *A. caudata* and *A. cylindrica* in sub-group IVb. The classified species in the same group had similar substitutions compared with the other group's species. The highest point mutations occurred in ITS2 sequences related to species grouped in the first clade. *T. aestivum*, *T. turgidum*, and *A. speltoides* showed similar substitutions on the base sites 76, 193, 215, 230 and 241 (TCCAG changed to ATTTA) (Figure 2). Also *T. aestivum* and *T. turgidum* contained the same substitutions on the base sites 112, 119, 135, 206, 229 and 230 (GGGACA changed to TATGTT) (Figure 2). The clade II species *T. boeticum* and *T. urartu* showed similar substitutions on the base sites 85, 201, 231, 234 and 252 (GTGCC changed AGATA) (Figure 2). The nucleotide substitutions on the base sites 223 and 234 (from TC to CA) were observed in the sub-group IVb species i.e. *A. caudata* and *A. cylindrica*. Remarkably, the nucleotide resided on the base site 146 was C in all Triticum species, but T in all Aegilops species (Figure 2). The substitution rate in *A. neglecta*, and *A. umbellulata* (clade III) ITS2 nucleotides was zero, it was concluded that they

have conserved ITS2 sequences compared to the other studied species.

CONCLUSION

The aligned ITS2 sequences in twelve Triticum and Aegilops species showed a total of 67 SNPs and transitions were predominant over transversions. The C/T transitions were observed at the highest level in the studied species. The highest nucleotide substitutions occurred in the ITS2 sequences of *T. aestivum*, *T. turgidum*, and *A. speltoides*, but the substitution rate in *A. neglecta*, and *A. umbellulata* was zero, it was concluded that they have conservedITS2. The secondary structures of ITS2 in the studied species was similar to that of the other flowering plants with four helices and the variation in helices length.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Ali-Ashraf Mehrabi for providing us with plant material. Also we gratefully acknowledge the cooperation of the Imam Khomeini International University of Iran for providing financial supports for this study.

Table 4. Pairwise nucleotide difference (upper diameter) and the percentage of identical residues (lower diameter) in alignment positions between two sequences of Triticum/Aegilops species.

Species	*A. cylindrica*	*A. caudata*	*A. triuncialis*	*A. neglecta*	*A. umbellulata*	*A. speltoides*	*A. tauschii*	*A. crassa*	*T. urartu*	*T. boeticum*	*T. aestivum*	*T. turgidum*
A. cylindrica	-	0	5	5	3	11	7	8	8	8	13	14
A. caudata	100	-	5	5	3	11	7	8	8	8	13	14
A. triuncialis	97.74	97.74	-	4	6	12	4	10	11	11	16	17
A. neglecta	97.74	97.74	98.17	-	2	10	4	8	10	10	14	15
A. umbellulata	98.64	98.64	97.29	99.1	-	10	6	6	8	8	12	13
A. speltoides	95.02	95.02	94.55	95.45	95.48	-	12	14	14	14	10	11
A. tauschii	96.83	96.83	98.17	98.17	97.29	94.55	-	10	12	12	16	17
A. crassa	96.83	96.83	95.48	96.38	97.29	93.67	95.47	-	12	12	16	17
T. urartu	96.38	96.38	95.02	95.48	96.38	93.67	94.57	94.57	-	0	16	17
T. boeticum	96.38	96.38	95.02	95.48	96.38	93.67	94.57	94.57	100	-	16	17
T. aestivum	94.12	94.12	92.76	93.67	94.57	95.48	92.76	92.76	92.76	92.76	-	1
T. turgidum	93.67	93.67	92.31	93.21	94.12	95.02	92.31	92.31	92.31	92.31	99.55	-

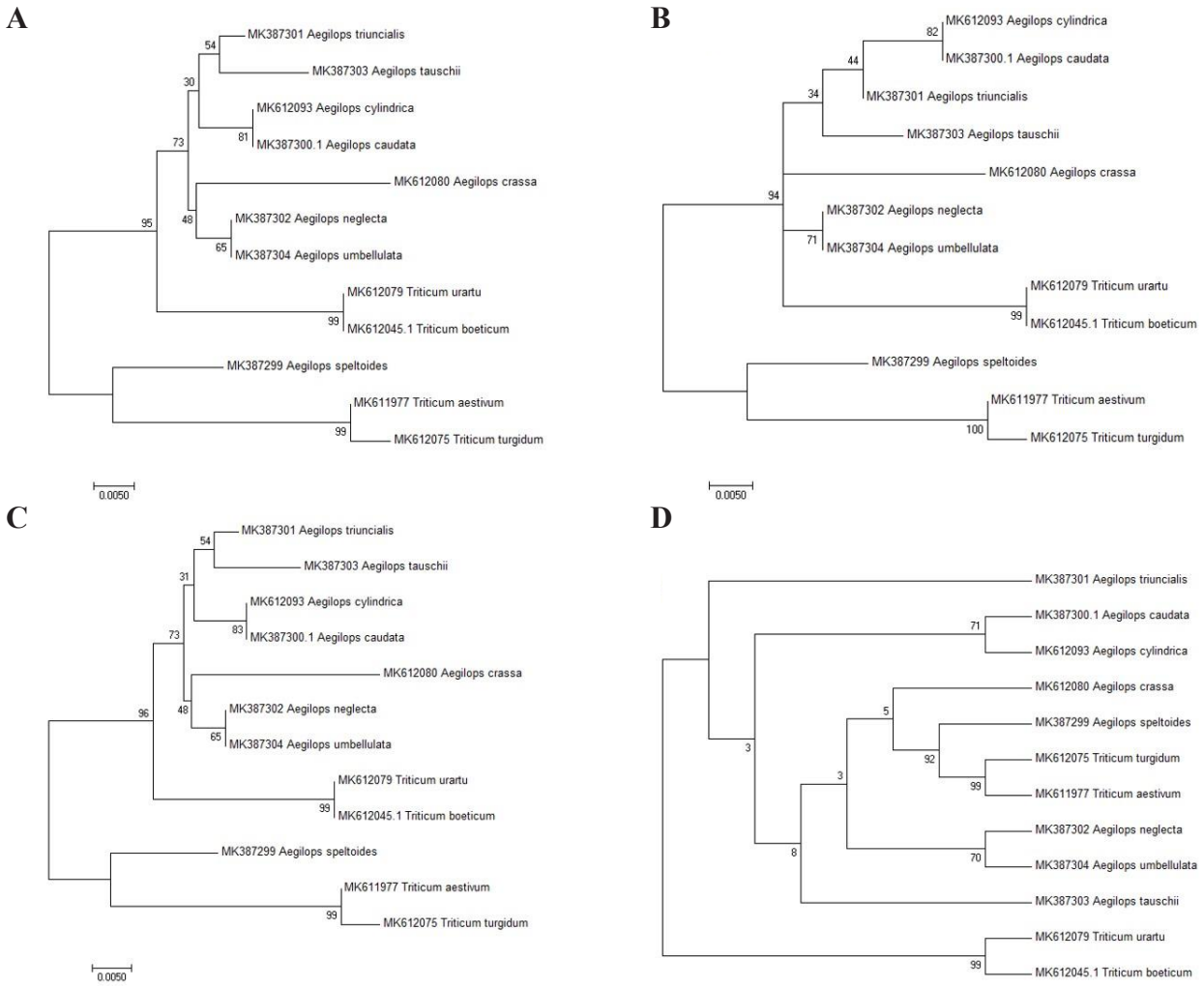

Figure 4. Phylogenetic tree of the species included in this study using **A:** Neighbor-Joining, **B:** Maximum Likelihood, **C:** Minimum Evolution, **D:** Maximum Parsimony methods. Numbers on branches are bootstrap values in percent after 1000 replicated tests.

REFERENCES

Aguilar C., and Sanchez J. A. (2007). Phylogenetic hypotheses of *Gorgonii doctocorals* according to ITS2 and their predicted RNA secondary structures. *Molecular Phylogenetics and Evolution,* 43: 774–786.

Alnaddaf L. M., Moualla M. Y., and Haider N. (2012). Resolving genetic relationships among *Aegilops* L. and *Triticum* L. species using analysis of chloroplast DNA by Cleaved Amplified Polymorphic Sequence (CAPS). *Asian Journal of Agricultural Science*, 4: 270–279.

Alvarez I., and Wendel J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29(3): 417–434.

Ankenbrand M. J., Keller A., Wolf M., Schultz J., and Förster F. (2015). ITS2 database V: Twice as much. *Molecular Biology and Evolution,* 32(11): 3030–32. DOI: https://doi:10.1093/molbev/msv174.

Balasubramani S. P., Murugan R., Ravikumar K., and Venkatasubramanian P. (2010). Development of ITS sequence based molecular marker to distinguish *Tribulu sterrestris* L. (Zygophyllaceae) from its adulterants. *Fitoterapia,* 81(6): 503–508. DOI: 10.1016/j.fitote.2010.01.002.

Banchi E., Ametrano C. G., Greco S., Stankovic D., Muggia L., and Pallavicini A. (2020). PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding, *Database*, 2020: 1–9. DOI: 10.1093/database/baz155.

Bordbar F., Rahiminejad M. R., Saeidi H., and Blattner F. R. (2011). Phylogeny and genetic diversity of D-Genome species of *Aegilops* and *Triticum* (Triticeae, Poaceae) from Iran based on microsatellites, ITS, and trnL-F. *Plant Systematics and* Evolution, 291: 117–131.

Chen Q. F. (1997). Inquisition about the origin and evolution of wheat genomes. *Guihaia*, 17: 276–282.

Chen S. L., Yao H., Han J. P., Liu C., and Song J. Y. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*, 5(1): e8613. DOI: https://doi.org/10.1371/journal.

pone.0008613.

Coleman A. W. (2003). ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends in Genetics*, 19: 370–375.

Coleman A. W. (2007). Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Research,* 35: 3322-3329. DOI: 10.1093/nar/gkm233.

Coleman A. W. (2009). Is there a molecular key to the level of ‘‘biological species’’ in eukaryotes? *Molecular Phylogenetics and Evolution,* 50: 197–203. DOI: 10.1016/j.ympev.2008.10.008.

Coleman A. W., and Mai J. C. (1997). Ribosomal DNA ITS-1 and ITS-2 sequence comparisons as a tool for predicting genetic relatedness. *Journal of Molecular Evolution*, 45: 168–177.

Coleman A. W., Preparata R. M., Mehrotra B., and Mai J. C. (1998). Derivation of the secondary structure of the ITS-1 transcript in Volvocales and its taxonomical correlations. *Protist*, 149: 135–146.

Duan H., Wang W., Zeng Y. Gue M., Zhou Y. (2019). The screening and identification of DNA barcode sequences for Rehmannia. *Scientific Reports*, 9: 17295. DOI: https://doi.org/10.1038/s41598-019-53752-8.

Dvogak J., Terlizzi P., Zheng H. B., and Resta P. (1993). The evolution of polyploidwheats: identification of the A genome donor species. *Genome*, 36: 21–30.

Feliner G. N., and Rossello J. A. (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution,* 44(2): 911–919. DOI: 10.1016/j.ympev.2007.01.013.

Goertzen L. R., Cannone J. J., Gutell R. R., and Jansen R. K. (2003). ITS secondary structure derived from comparative analysis: implications for sequence alignment and phylogeny of the Asteraceae. *Molecular Phylogenetics Evolution*, 29: 216–234.

Golovnina K. A., Kondratenko E. Y., Blinov A. G., and Goncharov N. P. (2009). Phylogeny of the a genomes of wild and cultivated wheat species. *Russian Journal of* Genetics, 45: 1360–1367.

Gottschling M., Hilger H. H., Wolf M., and Diane N. (2001). Secondary structure of the ITS1 transcript and its application in a reconstruction of the phylogeny of Boraginales. *Plant Biology*, 3: 629–636.

Gulbitti-Onarici S., Sancak C., Sumer S., and Ozcan S. (2009). Phylogenetic relationships of some wild wheat species based on the internal transcribed spacer sequences of nrDNA. *Current Science*, 96 (60): 794–800.

Haider N. (2013). The origin of the B genome of bread wheat (*Triticum aestivum* L.). *Russian Journal of* Genetics, 49: 263–274.

Hillis D. M., and Dixon M. T. (1991). Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Bio*logy, 66: 411–453.

Jones B. L., Lookhart G. L., Mak A., and Cooper D. B. (1982). Sequences of purothionins and their inheritance in diploid, tetraploid, and hexapliod wheats. *Journal of Heredity,* 73: 143–144.

Joseph N., Krauskopf E., Vera M. I., and Michot B. (1999). Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. *Nucleic Acids Research*, 27: 4533–4540.

Keller A., Forster F., Muller T., Dandekar T., and Schultz J. (2010). Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. *Biology Direct,* 5: 4. DOI: 10.1186/1745-6150-5-4. PMID: 20078867.

Keller A., Schleicher T., Forster F., Ruderisch B., Dandekar T., Muller T., and Wolf M. (2008). ITS2 data corroborate a monophyletic chlorophycean group (Sphaeropleales). *Biomed Central Evolutionary Biology*, 8: 218. DOI: http://www.biomedcentral.com/1471-2148/8/218.

Kerby K., and Kuspira J. (1987). The phylogeny of polyploidy wheats *Triticum aestivum* (bread wheat) and *Triticum turgidum* (macaroni wheat). *Genome*, 29: 722–737.

Kilian B., Mammen, K. Millet E., Sharma R., Graner A., Salamini F., Hammer K., and Ozkan H. (2011). Aegilops. In: Kole C. (Eds.) Wild Crop Relatives: Genomic and Breeding Resources. Springer, Berlin, Heidelberg, 1–76. DOI: https://doi.org/10.1007/978-3-642-14228-4_1.

Konarev V. G. (1983). The nature and origin of wheat genomes on the data of grain protein immunochemistry and electrophoresis. In proceeding of 6th International Wheat Genetics Symposium, Kyoto, Japan, 65–75.

Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33: 1870–1874.

Lenaers G., Maroteaux L., Michot B., and Herzog M. (1989). Dinoflagellates in evolution. A molecular phylogenetic analysis of large subunit ribosomal RNA. *Journal of Molecular Evolution*, 29: 40–51.

Liu J. S., and Schardl C., L. (1994). A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Molecular Biology*, 26: 775–778.

Luo K., Chen S. L., Chen K. L., Song J. Y., and Yao H. (2010). Assessment of candidate plant DNA barcodes using the Rutaceae family. *China Life Sciences*, 40: 342–351.

Mai J. C., and Coleman A. W. (1997). The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. *Journal* of *Molecular Evolution*, 44: 258–271.

Maxted N., and Kell S. (2009). CWR in crop improvement: To what extent are they used? *Crop Wild Relative Newsletter*, 7: 7–8.

Mehrabi A. A., and safari Z. (2019). Molecular phylogeny of *Aegilops* L. and *Triticum* L. species revealed by internal transcribed spacers of ribosomal genes. *Journal of Agricultural Science and Technology*, 21(3): 699–714.

Meyer C. P., and Paulay G. (2005). DNA Barcoding: error rates based on comprehensive sampling. *PLoS Biology*, 3: e422. DOI: https://doi.org/10.1371/journal.pbio.0030422.

Michot B., Joseph N., Sazan S., and Bachellerie J. P. (1999).

Evolutionary conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. *Nucleic Acids Research*, 27: 2271–2282.

Nepolo E., Chimwamurombe P. M., Cullisand C. A., and Kandawa-Schulz M. A. (2010). Determining genetic diversity based on ribosomal intergenic spacer length variation in Marama bean (*Tylosemae sculentum*) from the Omipanda Area, Eastern Namibia. *African Journal of. Plant Science*, 4: 368–373.

Petersen G., Seberg O., Yde M., and Berthelsen K. (2006). Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Molecular Phylogenetics and Evolution*, 39: 70–82.

Piccolo S. L., Alfonzo A., Conigliaro G., Moschetti G., Burruano S., and Barone A. (2012). A simple and rapid DNA extraction method from leaves of grapevine suitable for polymerase chain reaction analysis. *African Journal of Biotechnology*, 11(45): 10305–10309. DOI: https://doi: 10.5897/AJB11.3023.

Poczai P., and Hyvonen J. (2010). Nuclear ribosomal spacer regions in plant phylogenetics, problems and prospects. *Molecular Biology Reports*, 37: 1897–1912.

Rzhetsky A., and Nei M. (1992). A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution*, 9: 945–967.

Saitou N., and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* 4: 406–425.

Schultz J., and Wolf M. (2009). ITS2 sequence structure analysis in phylogenetics: a how to manual for molecular systematics. *Molecular Phylogenetics and Evolution*, 52: 520–523.

Schultz J., Muller T., Achtziger M., Seibel P. N., Dandekar T., and Wolf M. (2006). The internal transcribed spacer 2 database- a web server for (not only) low level phylogenetic analyses. *Nucleic Acids Research*, 34: 704–707.

Sliai, A. M., and Amer S. A. M. (2011). Contribution of chloroplast DNA in the biodiversity of some *Aegilops* species. *African Journal of. Biotechnology*, 10: 2212–2215.

Tamura K., and Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10:512–526.

Thompson J. D., Higgins D. G., and Gibson T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22: 4673–4680.

Thornhill D. J., Lajeunesse T. C., and Santos S. R. (2007). Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates. *Molecular Ecology*, 16: 5326–5340.

Torres R. A., Ganal M., and Hembleden V. (1990). GC balance in the internal transcribed spacers ITS1 and ITS2 of nuclear ribosomal RNA genes. *Journal of Molecular Evolution*, 30: 170–181.

Wang M., Zhao H. X., and Wang L. (2013). Potential use of DNA barcoding for the identification of Salvia based on cpDNA and nrDNAsequences. *Gene*, 528: 206–215.

Wheeler W. C. (1994). Sources of ambiguity in nucleic acid sequences alignment. In Molecular Ecologyand Evolution: *Approaches and Applications*, Schierwater B., Streit B., Wagner G. P., and DeSalle R. (Eds.), Birkhauser Verlag, Basel, 323–354.