

## Application of the ITS2 Region for Barcoding Plants of the Genus *Triticum* L. and *Aegilops* L.

I. GANOPOULOS<sup>2</sup>, A. KAPAZOGLOU<sup>1</sup>, I. BOSMALI<sup>1</sup>, A. XANTHOPOULOU<sup>1,2</sup>, I. NIANIOU-OBEIDAT<sup>2</sup>,  
A. TSAFTARIS<sup>1,2</sup> and P. MADEISIS<sup>1\*</sup>

<sup>1</sup>Institute of Applied Biosciences, CERTH, Thessaloniki, 570 01, Greece

<sup>2</sup>Department of Genetics and Plant Breeding, School of Agriculture, Aristotle University of Thessaloniki, Thessaloniki 54 124, Greece

(Received 4 September 2016; Accepted 31 January 2017;

Communicated by A. Börner)

Molecular taxonomic studies have been performed in the past in order to identify different wheat species and construct a molecular phylogeny. These were based on universal but sufficiently divergent sequences from both the nuclear and chloroplastic genomes of wheat. They included two short plastid sequences from the plastid genes *rbcL* and *matK* which have been proposed as the core “barcode” sequences by the “barcoding” guidelines for general plant identification. Historically, in molecular plant taxonomy, plastid sequences had been favored over nuclear sequences, due to their uniparental inheritance and consequently lower intra-molecular recombination. However recently, the short nuclear sequence from the internal transcribed spacer 2 (ITS2) has been used successfully for the accurate identification of many medicinal and other plant species. Herein, we have used the plastidic *matK*, *rbcL* *trnL*, and the nuclear ITS2 region for the identification of different wheat species of *Triticum* L. and goatgrass species of *Aegilops* L. We have successfully discriminated all species that were examined from both genera, thus, validating the ITS2 region as a ‘barcode tool’ for accurate distinction of plants in the genus *Triticum* L. and *Aegilops* L. The success rate of PCR amplification and sequencing of the ITS2 region was 100%. We report also that *matK*, *rbcL* and *trnL* regions could not discriminate all species in contrast to the ITS2 region which demonstrated full discriminatory capacity.

**Keywords:** ITS2, *Triticum*, *Aegilops*, DNA barcoding, identification

### Introduction

Modern wheat species rank first in global grain production, representing the most important caloric intake worldwide (>20% human food calories) and staple food for more than 40% of the world’s population (<http://www.fao.org/faostat>). Allohexaploid, common or bread wheat *T. aestivum* (BBAADD,  $2n = 6x = 42$ ) accounts for 95% of wheat production worldwide whereas allotetraploid hard or durum wheat *T. durum* (or *T. turgidum* ( $2n = 4x = 28$ )) represents the rest 5%. Species of the genus *Triticum* L. exist as a polyploid series of di-, tetra- and hexaploid wheat with a basic number  $n = 7$  (Sakamura

\*Corresponding author; E-mail: [pmadesis@certh.gr](mailto:pmadesis@certh.gr)

1918). Four basic genomes defined as A, B, D and G contribute to the genomic make up of all *Triticum* species (Kihara 1924; Lilienfeld and Kihara 1934). Intergenic hybridization and allopolyploidy involving the most closely related genus *Aegilops* L. have been crucial for the speciation and adaptation of wheat *Triticum* L. (Golovnina et al. 2007.) Species identification at the DNA level contributes to accurate species discrimination and can tackle the limitations of traditional taxonomy that is based purely on morphological characteristics. “DNA barcoding” a term first by (Hebert et al. 2003), entails the use of a short DNA sequence from a standard locus as an individual species identification signature (Kress et al. 2005; Hollingsworth 2011). The *Consortium for the Barcode of Life (CBOL) Plant Working Group*, proposed the use of two coding regions from the plastid (chloroplast) genome, *rbcl* and *matK*, as a “core barcode” for plant identification (Group et al. 2009).

However, since these sequences displayed some disadvantages regarding their primer universality and PCR efficiencies, a nuclear region, the ribosomal internal transcribed spacer (ITS), has been given considerable attention, the last few years (Hollingsworth 2011; Li et al. 2011a; Wang et al. 2015). ITS exhibits high molecular resolution for a variety of seed plant species (Hollingsworth 2011; Li et al. 2011a; Wang et al. 2015) and recently for commercial woody angiosperms (Bolson et al. 2015). Nevertheless, despite its high discriminatory power, ITS may not work well for all plant groups due to low universality mainly attributed to the quality of primers (Han et al. 2013).

To circumvent this problem, the ITS2 region which is part of the internal transcribed spacer has been widely used as a molecular taxonomic marker (Group et al. 2009; Chen et al. 2010; Gao et al. 2010; Hollingsworth 2011; Li et al. 2011b; Han et al. 2013). The ITS2 region has definite advantages in that it is shorter (average length in angiosperms ~ 412 bp), easily amplifiable using one pair of universal primers, and with high discriminatory capacity (Han et al. 2013). Chen et al. (2010) validated ITS2 as a powerful barcoding marker discriminating more than 6,600 plant samples belonging to 4,800 species from 753 distinct genera and demonstrated that the rate of successful identification using the ITS2 barcode was 92.7% and 99.8% at the species and genus level, respectively.

The *China Plant BOL Group (CPBG)* recommended that the ITS/ITS2 regions should be incorporated into the core barcode for seed plants (Hollingsworth 2011; Li et al. 2011a). In recent years several DNA barcodes including ITS2 (*matK*, *rbcl*, *psbA-trnH*, ITS, ITS2, *rpoCI*, etc.) have been tested and validated for the identification of species belonging to various medicinal plant families and other leguminous forage and pasture species (Chen et al. 2010; Shi et al. 2011; Ganopoulos et al. 2012). Phylogenetic relationships among different wheat species based on the ITS region had been demonstrated in the past (Zhang et al. 1998; Baum et al. 2001; Zhang et al. 2002; Goryunova et al. 2005; Boscato et al. 2008; Gulbitti-Onarici et al. 2009). However, in the present study we validated the potential of the ITS2 region as a suitable barcoding marker for the identification of closely related species of the Triticeae family. Our study indicates that the ITS2 region can be used as an effective barcode for the discrimination of plant species of the genus *Triticum* L. and *Aegilops* L.

## Materials and Methods

### *Plant material*

Plant material was obtained from the Greek Gene Bank (ELGO-DEMETER) (Table 1). The species selected were Greek *Triticum* and *Aegilops* species which are used as crops either for human nutrition or as animal feed, or species that are wild crop relatives.

Isolation of DNA was performed with 0.1 g of seeds as starting material of fine powder with Qiagen DNeasy plant mini kit according to the manufacturer's instructions. DNA concentration and purity was determined by standard spectrophotometric methods at 260 nm and 280 nm UV wavelengths with an Eppendorf BioPhotometer and DNA integrity was assessed by gel electrophoresis in a 0.8% agarose gel. Samples were then diluted to 20 ng/ $\mu$ L working concentration.

### *PCR amplification*

PCR amplification was performed in a total volume of 25  $\mu$ L in a MJ research thermocycler. The ITS2 region was amplified using the following pair of universal primers (Yao et al. 2010): ITS-S2F (forward), 5'-ATGCGATACTTGGTGTGAAT, and ITS-S3R (reverse), 5'-GACGCTTCTCCAGACTACAAT. The reaction mixture contained 20 ng genomic DNA, 1X PCR buffer, 2.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 300 nM forward and reverse primers and 0.5 U Kapa Taq DNA polymerase (Kapa Biosystems, USA). Initial denaturing step of 95 °C for 3 min followed by 30 cycles of 95 °C for 20 s, 54 °C for 40 s and 72 °C for 20–40 s, then a final extension step of 72 °C for 2 min.

PCR analyses for *matK*, *rbcL* and *trnL* barcoding regions were performed according to Madesis et al. (2012).

### *Sequence analysis*

PCR products were directly sequenced in two directions of each fragment with Big Dye terminator v3.1 Cycle sequencing kit (PE Applied Biosystems, Foster City, CA, USA) in an automated ABI 3730 sequencer (PE Applied Biosystems). The sequences were aligned with the CLUSTAL W program.

### *Data analysis*

Species identification based on the sequence similarity approach was performed using GenBank databases and Blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The genetic distances were then calculated using MEGA 5.2.2 (Tamura et al. 2011), according to the Kimura 2-Parameter (K2P) model. The average intra-specific distance, coalescent depth and theta were calculated to evaluate the intra-specific variation using the K2P model. The average inter-specific distance, the minimum inter-specific distance and Theta prime were used to represent inter-specific divergences (Meyer and Paulay 2005; Chen et al. 2010). Indels were coded with the simple indel coding method of (Simmons and Ochoter-

ena 2000). Neighbor-joining (NJ) dendrogram was conducted using MEGA 5.2.2 with 1,000 bootstrap replicates. ITS2 sequences with different sequence divergence were subjected to secondary structure prediction in genus *Triticum* L. using tools from the ITS2 database (Koetschan et al. 2010). Additionally, the sequence character-based method (Rach et al. 2008) was used with DnaSPv5 (Librado and Rozas 2009), and the information from each site was treated as a character to distinguish the species from each other.

## Results

*MatK*, *trnL*, *rbcL* and ITS2 regions were assessed as potential *Triticum* barcoding regions. All sequences have been deposited in NCBI database (Table S1\*). *MatK* amplification was attempted with two sets of standard primers with no results; thus this marker region was subsequently eliminated from the study (data not shown). Each of the two remaining barcoding loci was successfully amplified using standard primer pairs and PCR protocols. We obtained *trnL* and *rbcL* DNA sequences from every species analyzed. The results from the sequence analysis showed that *rbcL* and *trnL* do not have sufficient variation to be suitable for barcoding the genus *Triticum* (species identification rate 4/12 for both barcoding markers).

On the contrary the PCR amplification rate of the ITS2 sequences from plants of the genus *Triticum* L. was 100%, and the sequencing success rate was 100% for all 12 taxa examined. The amplified sequence length ranged from 433 to 460 bp. The GenBank accession numbers are listed in Table S1. The lengths of the ITS2 regions ranged from 145 to 189 bp, with an average length of 162 bp; the mean GC content was 56%, with a range of 46% to 67%.

The mean GC content was 62.7% and ranged from 46% to 67%. Together these data show that the length and GC content of *Triticum* ITS2 sequences are relatively variable. The ITS2 matrix aligned sequence length was 409 bp; the distribution of 24 parsimony informative sites and 38 variable sites was intensive and dense across the matrix and included 1 Indel.

We used six metrics to characterize inter- vs. intra-specific variations (Meyer and Paulay 2005; Chen et al. 2010). As depicted in (Fig. S1), the dataset showed significant levels of inter-specific divergence within ITS2 sequences. Relatively lower levels of intra-specific divergence were found with calculations for all three metrics. The average difference considering the entire dataset was 0.041 and the smallest interspecific divergence was 0.007 (Fig. 1). When calculated according to the K2P model, the intraspecific genetic distance (0.004) is less than the interspecific genetic distance (0.041). The interspecific percentage differences among the *Triticum* species were greater than the intra-specific variations. Therefore, the ITS2 region of the *Triticum* species, with lower levels of genetic divergence within species than between species, may be used as a genomic marker for the identification of these species.

The NJ tree intuitively displayed the relationship among the species in the genus *Triticum* L. and *Aegilops* L. (Fig. 1). ITS2 was able to discriminate 19 out of 29 taxa analyzed

\*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

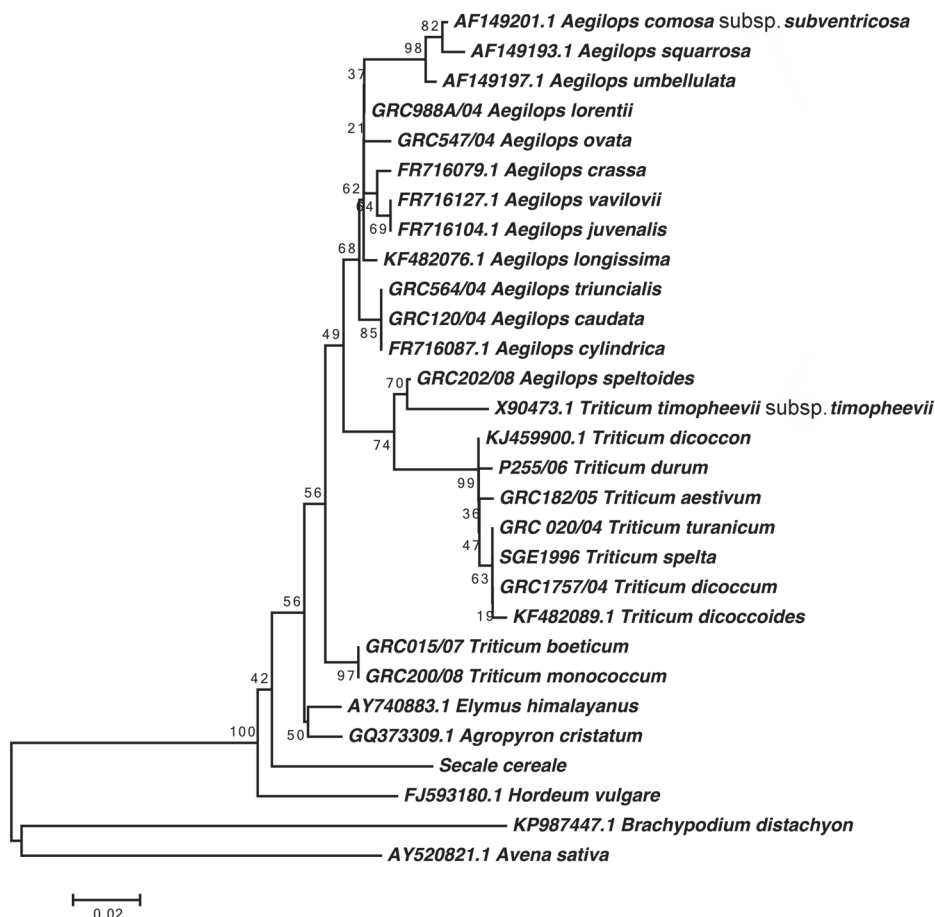


Figure 1. Neighbor-joining (NJ) tree based on the nuclear internal transcribed spacer 2 (ITS2) region

(65.51%). The evolutionary history was inferred using the Neighbor-joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 0.62970968 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004) and are in the units of the number of base substitutions per site (Table S2). The analysis involved 29 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 258 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016). The twenty specific informative

and discriminatory sites are presented in Table S3. *T. durum* has two different diagnostic sites 66 (-), 79 (T) while *T. aestivum* has also two 68 (C) and 116 (A). *A. speltoides* has two informative sites 66 (-) and 228 (A) whereas *A. ovata* has one informative site at 252 (T). Moreover, *T. monococcum* which is very closely related to *T. boeoticum* has one informative site at 395 (C). In support to this no obvious differences in elementary genomes of *T. boeoticum* and *T. monococcum* were found based on nuclear gene variability (Goncharov et al. 2009)

To identify the effect of the primary sequence divergences, secondary structures were constructed (Fig. S2). All of the secondary structures of ITS2 in these species contained a central ring (primary ring) and four similar helices (I, II, III and IV).

### Discussion

More than one haplotype from the same species clustered into one branch, and there was a clear boundary between species. Interestingly, *T. monococcum* and *T. boeoticum* are grouped together. The results were basically consistent with traditional plant morphotaxonomy (Golovnina et al. 2007).

ITS2 secondary structures among the different plants of the genus *Triticum* L. differed significantly in the four helical regions in stem loop number, size, position, and screw angle. The secondary structures of the ITS2 sequences for some species could not be displayed because they had no reference models. On the basis of the ITS2 secondary structure, the plants of the genus *Triticum* L. could be discriminated well.

The advantages of using ITS2 as the second tier as proposed by Chen et al. (2010) for identifying *Triticum* species include the fact that i) ITS2 provides high species resolution, ii) it is a sequence originating from the nuclear genome which evolves faster than the plastid genome with higher substitution rates and therefore provides increased sequence divergence and iii) it is a much shorter sequence allowing higher recovery from processed plant materials such as herbal products (Newmaster et al. 2013) which could also benefit analysis of processed wheat products.

In summary, the present study investigated the potential of the *matK*, *rbcL*, *trnL* and ITS2 region to be used as a universal barcode for the phylogenetic analysis of wheat species. We have demonstrated for the first time that the ITS2 region can be used effectively for the identification and discrimination of different economically important species of *Triticum* L. The successful rate of PCR amplification and sequencing of the ITS2 region was 100%. Significant divergence was evident between the interspecific and intraspecific genetic distances of the ITS2 regions, while the barcoding gap was more obvious. Noticeably, apart from primary structure divergence the secondary structure of ITS2 was significantly different among different *Triticum* L. species in the helical region, providing additional discriminatory power. Overall, cluster analysis using the ITS2 barcode was consistent with traditional morphology-based plant taxonomy. Hence, we have shown that the ITS2 region can serve as a powerful barcode marker for efficient identification of wheat species.

In conclusion, the *matK*, *rbcl*, *trnL* and ITS2 regions were examined for their efficacy in identifying different species of wheat. Our results provide solid evidence that the ITS2 region can be used as a barcoding tool for rapid and accurate taxonomic analysis of the genus *Triticum* L and *Aegilops* L. In addition, the *matK* region could not amplify the corresponding barcoding region and thus could not serve as a barcoding marker, whereas *rbcl* and *trnL* regions could only discriminate a fraction of the species studied. Taxonomic analysis through the ITS2 region could complement morphology-based taxonomy and may have potential applications in the agro-food sector related to diagnostic food technologies and consumer protection.

### Acknowledgement

We would like to acknowledge the Greek Gene Bank (ELGO-DEMETER) for their kind help in supplying the plant material used in this study.

### References

- Baum, B.R., Johnson, D.A., Bailey, L.G. 2001. Defining orthologous groups among multicopy genes prior to inferring phylogeny, with special emphasis on the *Triticeae* (*Poaceae*). *Hereditas* **135**:123–138.
- Bolson, M., de Camargo Smidt, E., Brotto, M.L., Silva-Pereira, V. 2015. ITS and *trnH-psbA* as efficient DNA barcodes to identify threatened commercial woody angiosperms from southern Brazilian Atlantic rainforests. *PLoS one* **10**:e0143049.
- Boscato, P., Carioni, C., Brandolini, A., Sadori, L., Rottoli, M. 2008. Molecular markers for the discrimination of *Triticum turgidum* L. subsp. *dicoccum* (Schrank ex Schübl.) Thell. and *Triticum timopheevii* (Zhuk.) Zhuk. subsp. *timopheevii*. *J. Archaeological Sci.* **35**:239–246.
- Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS one* **5**:e8613.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Ganopoulos, I., Madesis, P., Tsaftaris, A. 2012. Universal ITS2 barcoding DNA region coupled with high-resolution melting (HRM) analysis for seed authentication and adulteration testing in leguminous forage and pasture species. *Plant Mol. Biol. Reporter* **30**:1322–1328.
- Gao, T., Yao, H., Song, J., Liu, C., Zhu, Y., Ma, X., Pang, X., Xu, H., Chen, S. 2010. Identification of medicinal plants in the family *Fabaceae* using a potential DNA barcode ITS2. *J. of Ethnopharmacology* **130**:116–121.
- Golovnina, K.A., Glushkov, S.A., Blinov, A.G., Mayorov, V.I., Adkison, L.R., Goncharov, N.P. 2007. Molecular phylogeny of the genus *Triticum* L. *Plant Systematics and Evolution* **264**:195–216.
- Goncharov, N.P., Golovnina, K.A., Kondratenko, E.Y. 2009. Taxonomy and molecular phylogeny of natural and artificial wheat species. *Breeding Sci.* **59**:492–498.
- Goryunova, S.V., Chikida, N.N., Gori, M., Kochieva, E.Z. 2005. Analysis of nucleotide sequence polymorphism of internal transcribed spacers of ribosomal genes in diploid *Aegilops* (L.) species. *Mol. Biol.* **39**:173–176.
- Group, C.P.W., Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J. 2009. A DNA barcode for land plants. *Proc. Am. Natl Acad. of Sci.* **106**:12794–12797.
- Gulbitti-Onarici, S., Sancakz, C., Sumer, S., Ozcan, S. 2009. Phylogenetic relationships of some wild wheat species based on the internal transcribed. *Current Science* **96**:6.
- Han, J., Zhu, Y., Chen, X., Liao, B., Yao, H., Song, J., Chen, S., Meng, F. 2013. The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *BioMed Research International* 2013. Article ID 741476, <http://dx.doi.org/10.1155/2013/741476>.

- Hebert, P.D.N., Cywinska, A., Ball, S.L. 2003. Biological identifications through DNA barcodes. Proc. of the Royal Society of London B: Biological Sciences **270**:313–321.
- Hollingsworth, P.M. 2011. Refining the DNA barcode for land plants. Proc. of the National Academy of Sciences **108**:19451–19452.
- Kihara, H. 1924. Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rücksicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden (Cytological and genetic studies in major cereals with particular reference to the behavior of chromosomes and sterility in hybrids). Memoirs of the College of Science. Kyoto Imperial University, Series B, vol. 1. Kyoto, Japan. 200 p. (in German)
- Koetschan, C., Hackl, T., Müller, T., Wolf, M., Frank, F., Schultz, J. 2010. ITS2 Database IV: Interactive taxon sampling for internal transcribed spacer 2 based phylogenies. Molecular Phylogenetics and Evolution **63**:585–588.
- Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Janzen, D.H. 2005. Use of DNA barcodes to identify flowering plants. Proc. Am. Natl Acad. of Sci. **102**:8369–8374.
- Kumar, S., Stecher, G., Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. and Evol. **33**:1870–1874.
- Li, D.-Z., Gao, L.M., Li, H.T., Wang, H., Ge, X.J., Liu, J.Q., Chen, Z.D., Zhou, S.L., Chen, S.L. Yang, J.B. 2011a. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Am. Natl Acad. of Sci. **108**:19641–19646.
- Li, D.Z., Liu, J.Q., Chen, Z.D., Wang, H., GE, X.J., Zhou, S.L., Gao, L.M., Fu, C.X., Chen, S.L. 2011b. Plant DNA barcoding in China. J. Systematics and Evol. **49**:165–168.
- Librado, P., Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**:1451–1452.
- Lilienfeld, F., Kihara, H. 1934. Genomanalyse bei *Triticum* und *Aegilops*. Cytologia **6**:87–122.
- Madesis, P., Ganopoulos, I., Ralli, P., Tsaftaris, A. 2012. Barcoding the major Mediterranean leguminous crops by combining universal chloroplast and nuclear DNA sequence targets. Genet. and Mol. Res. **11**:2548–2558.
- Meyer, C.P., Paulay, G. 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biology **3**:e422.
- Newmaster, S., Grguric, M., Shanmughanandhan, D., Ramalingam, S., Ragupathy, S. 2013. DNA barcoding detects contamination and substitution in North American herbal products. BMC Medicine **11**:222.
- Rach, J., DeSalle, R., Sarkar, I.N., Schierwater, B., Hadrys, H. 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proceedings of the Royal Society B: Biological Sciences **275**:237–247.
- Saitou, N., Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. and Evol. **4**:406–425.
- Sakamura, T. 1918: Kurze Mitteilung über die Chromosomenzahlen und die Verwandtschaftsverhältnisse der *Triticum*-Arten (A brief communication on the chromosome numbers and the relationships of the *Triticum* species). Bot. Mag. Tokyo **32**:151–154. (in German)
- Shi, L.C., Zhang, J., Han, J.P., Song, J.Y., Yao, H., Zhu, Y.J., Li, J.C., Wang, Z.Z., Xiao, W., Lin, Y.L. Xie, C.X. 2011. Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. J. of Systematics and Evol. **49**:261–266.
- Simmons, M.P., Ochoterena, H. 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biol. **49**:369–381.
- Tamura, K., Nei, M., Kumar, S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Am. Natl Acad. of Sci. **101**:11030–11035.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. and Evol. **28**:2731–2739.
- Wang, X.C., Liu, C., Huang, L., Bengtsson Palme, J., Chen, H., Zhang, J.H., Cai, D., Li, J.Q. 2015. ITS1: a DNA barcode better than ITS2 in eukaryotes? Molecular Ecol. Resour. **15**:573–586.
- Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P., Chen, S. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PloS one **5**:e13102.



- Zhang, W., Qu, L., Gu, H., Gao, W., Liu, M., Chen, J., Chen, Z. 2002. Studies on the origin and evolution of tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. *Theor. Appl. Genet.* **104**:1099–1106.
- Zhang, W.-J., L.-J. Qu, W. Gao, H.-Y. Gu, J. Chen, Chen, Z. L. 1998. ITS1 and ITS2 sequences of four possible donors to bread wheat genome and their phylogenetic relationships. *Acta Botanica Sinica* **40**:994–1000.

### Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at <http://www.akademai.com/content/120427/>

Electronic Supplementary *Table S1*. List of taxa, collection information, and Genbank accession numbers of ITS2 sequences analyzed

Electronic Supplementary *Table S2*. Maximum Likelihood Estimate of Substitution Matrix

Electronic Supplementary *Table S3*. Character based DNA database for *Triticum* L. and *Aegilops* L. species from the ITS2 region. Character states (nucleotides) at 20 selected positions (ranging from 66 to 252) are shown. Grey cells show important diagnostic character sites; ‘-’ denotes the indel site

Electronic Supplementary *Figure S1*. Inter- and intra-specific genetic divergence and identification efficiency in ITS2 sequences of genus *Triticum* L. and *Aegilops* L.

Electronic Supplementary *Figure S2*. The secondary structures of the ITS2 regions in five species of the genus *Triticum* L. and *Aegilops* L.