rDNA-ITS2 characterization of *Trichogramma* species (Hymenoptera: Trichogrammatidae) in Turkey

Fahriye Sümer Ercan1*, Mevlüde Alev Ateş2 and Sevcan Öztemiz3

Abstract

**Background:** ITS2 sequences can be used in systematic studies and proved to give reliable results in the distinguishing of *Trichogramma* species (Hymenoptera: Trichogrammatidae). Correct identification of natural enemies forms the basis of a biological control program. The present study aimed to compare sequences of rDNA-ITS2 of *Trichogramma* samples with those deposited in GenBank by using ITS2, as a barcode for reliable species identification and assessment of genetic diversity.

**Results:** Molecular identification methods were used to differentiate 2 *Trichogramma* species collected from Adana province of Turkey; *Trichogramma brassicae* (Bezdenko) and *Trichogramma turkestanica* Meyer (Hymenoptera: Trichogrammatidae). ITS2 sequences of samples ranged in size from 378 to 406 bp. The ITS2 sequences were aligned using Clustal W, genetic distances and phylogenetic tree were calculated using MEGA V7.0. rDNA-ITS2 sequences of 37 specimens of *Trichogramma* confirmed in GenBank in the study. Also, secondary structures of ITS2 sequences were predicted with the help of Mfold web server. All secondary structure constructions were performed at 37 °C using RNA version 2.3 default parameters.

**Conclusions:** A molecular marker can be used successfully to distinguish closely related groups if it is a rapidly evolving and highly conserved gene region. In the study, it was shown that ITS2 was a reliable molecular marker in distinguishing species. Therefore, with rDNA-ITS2 sequence analysis, *Trichogramma* spp., which is a very important natural enemy in biological control, has been identified.

**Keywords:** Internal transcribed spacer 2, *Trichogramma*, Molecular systematic, GenBank, Biological control

Background

The internal transcribed spacer 2 (ITS2) region of nuclear ribosomal DNA is one of the important DNA barcodes because of the availability of conserved regions for designing universal primers. This region has been used for many years for species identification and phylogenetic examination of *Trichogramma*. Since this region is preserved, a wide variety of taxa can be studied with specific primers (Sumer et al. 2009). Stouthamer (2006) emphasized that the identification of protein and DNA regions would be much more useful in the detection and diagnosis of parasitoids.

*Trichogramma* species are one of the most commonly used groups of natural enemies of major crop pests. Although the diagnosis at the species level is based on the male genitalia, the presence of only female individuals under natural conditions is an important problem for many species. In addition to being quite small and morphologically indistinguishable, its morphology and physiology are significantly influenced by environmental factors. Therefore, identification of these wasps is problematic and systematic clarification is needed. A wide
variety of methods have been used to elucidate the classification of Trichogramma species. These include morphological features such as antennae structure and wing veining, followed by morphological comparison of the male reproductive system, and then a series of methods including enzymatic analysis (Owen and Pinto 2004). Esterase electrophoresis provided favorable results for the differentiation of some Trichogramma species (Sümer et al. 2008). However, limited variation in the esterase locus allows for differentiation between a limited number of species (Silva et al. 1999). At the same time, the samples must be fresh in order to be used in alloenzyme analysis or they must be stored at -70°C to prevent enzyme degradation. Therefore, researchers tried alternative methods and used various molecular markers, including RAPD (Ercan et al. 2012), RFLP (Vanlerbergh-Masutti 1994), COI (Ercan et al. 2013), microsatellite markers (Pizzol et al. 2005) and ITS regions of rDNA, for diagnostic purposes (Ercan et al. 2011).

With studies based on DNA sequence analysis, identification of cryptic species of Trichogramma, use of correct species in production and release studies can be ensured. The success of a biological control program is closely related to the correct identification of the natural enemy species (Silva 1999). Use of the most successful species in terms of host search, parasitism capacity and tolerance to environmental conditions will undoubtedly ensure success. In the present study, secondary structures of ITS2 sequences of 2 Trichogramma species were predicted and compared. Generally, secondary structure form of ITS2 was seen in RNA activities of cells. Even if it has different combinations of nucleotide sequences, eukaryotic ITS2 region is common with 4 helices and motifs (Coleman 2007). With all these conserved motifs and helices ITS2 secondary structure supports more reliable perspective of relationships at higher taxonomic levels (Zhang et al. 2015). Evaluation of base pair interactions in the secondary structure helps us to explain the phylogenetic information to advance phylogenetic estimations (Telford et al. 2005).

Generally, the aim of molecular systematic studies is to elucidate the structure of the target population, to determine its phylogenetic boundaries, and to elucidate intraspecies and inter-species relationships. For this purpose, in the present study, rDNA ITS2 gene region with nucleotide sequences and secondary structure form of these sequences was used to identify Trichogramma species collected from Adana province of Turkey.

**Methods**

**Trichogramma samples**

All samples were collected from Adana province of Turkey in August 2018 & 2021 (A1-A25 coded samples belong to the field study carried out in 2018, while A26-A37 coded samples that carried out in 2021). Host insect was Ostrinia nubilalis Hbn. (Lepidoptera: Crambidae) and host plant was corn. A total of 37 Trichogramma samples, hatched from the collected parasitized host eggs, was transferred into 70% alcohol for DNA isolation.

**DNA isolation**

To extract DNA from Trichogramma samples, one wasp from each sample was used. They were ground in 60 μl 5% Chelex-100 and 2 μl Proteinase K (20 mg/ml) and incubated at 1 h at 55 °C, followed by 10 min at 96 °C (Stouthamer et al. 1999).

**rDNA-ITS2 amplification**

The following primers were used for ITS2 amplification: ITS2 forward, 5’-TGTGAACTGAGGACACA TG-3’, and ITS2 reverse, 5’-GTCCTTGCCCTGCTGAGG-3’ (Stouthamer et al. 1999). The PCR was performed in a total volume of 25 μl. It contained 2 μl DNA template, 2.5 μl PCR buffer (10X buffer with (NH4)2SO4), 5 μl dNTPs (10 mM stock solution), 0.5 μl forward and reverse primers, 0.2 μl Taq Polymerase (5 u/μl) and 14.3 μl of sterile distilled water. The cycling program was also the same as used by Stouthamer et al. (1999). The size of PCR product was determined with 1% agarose gel electrophoresis with a size standard.

After electrophoresis, PCR products were purified by the Wizard® PCR Preps DNA Purification System. Following the purification, the PCR products were ligated into a Pgem-T® Vector (Promega). 2 μl of the ligation mix was transformed in the heat-shock cells of DH5-α Escherichia coli and plated in a LB agar medium containing Ampicillin, X-GAL and IPTG. The plates were stored overnight at 37 °C. The next day, white colonies on each plate were removed with sterile toothpicks, and the bacteria attached to the toothpicks were dispersed in Eppendorf tubes containing 50 μl sterile distilled water. Two μl of this solution was used for PCR reaction using the ITS2 primers in the PCR reaction described above to determine the correct size of insert in the Pgem plasmid in each sample (Ercan et al. 2013). The PCR products were loaded on a standard TAE buffered agarose gel. Then, the gel was stained with ethidium bromide and viewed and photographed under UV light. PCR products of 40 samples were then sent for automatic sequencing in a sequencing facility (Atlas Biotechnology).

**Secondary structure of ITS2 sequences**

Secondary structures of ITS2 sequences and ΔG (Gibbs) free energy calculations were predicted and calculated with the help of Mfold web server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form2.3) which
were predicted at 37 °C using RNA version 2.3 default parameters by the program parameters (Zuker 2003).

**Phylogenetic analysis**

Aligned sequences were analyzed using the Kimura 2-Parameter (K2P) distance model (Kimura 1980) on the Mega 7 program to determine the genetic diversity index and inter- and intra-specific nucleotide differences (Kumar et al. 2016). The probability of the best phylogenetic tree according to the polymorphism in the nucleotide sequences was determined by Maximum Likelihood (ML) analysis. The model with the lowest AIC value (Akaike Information Criteria) was used to determine the phylogenetic tree through the jModelTest v.0.1.1 program to determine the most suitable model for sequence change in Maximum Likelihood analyses (Posada 2008). In addition, the reliability of the model phylogenetic trees was checked with the 1000 replication Bootstrap test.

**Results**

*Trichogramma* samples were collected from *O. nubilalis* eggs in Adana province of Turkey. Twenty-five adults emerged from the collected host eggs in 2018, 12 adults were obtained from 2021 field collection. rDNA-ITS2 sequences of 37 specimens were blasted in GenBank database of National Center for Biotechnology Information to control the amplified region. The specimens were compared them with all obtained homologous sequences of other *Trichogramma* species in GenBank. Five of the 37 specimens were identified as *Trichogramma turkestanica* Meyer and 32 as *Trichogramma brassicae* (Bezdenko) (Fig. 1). The ITS2 sequences of *T. turkestanica* samples varied between 378 and 380 bp in length, while the ITS2 sequence length of all samples from *T. brassicae* was determined as 406 bp. The phylogenetic tree was constructed using the obtained sequence results. A dendrogram was also constructed with sequences of different *Trichogramma* species obtained from GenBank (Fig. 2).

In addition, secondary structures of both *T. turkestanica* and *T. brassicae* were predicted by Mfold web server (Fig. 3). Branched structures in *T. brassicae* were remarkable compared to *T. turkestanica*. Moreover, the similarities and differences in the constructed phylogenetic tree were also reflected in the secondary structure form of ITS2 with helices and angles among species. Therefore, calculation of ΔG (Gibbs) free energy values with Mfold program parameters for *T. turkestanica* and *T. brassicae* based on the helices and angles in the secondary structure with thermodynamic calculations (Santa Lucia 1998) were different based on species, as — 136.50 kcal/mol and — 146.60 kcal/mol, respectively. So, secondary structure form of ITS2 region was also used like a morphological characteristic of species due to its clear visually of nucleotide sequences.

**Discussion**

*Trichogramma* spp. are known to be distributed worldwide, represented by approximately 210 species (Pinto 2006). Up till now, 11 different *Trichogramma* species have been identified in different cultivated plants and forest areas in Turkey. Considering their synonyms, 8 species are known to exist in Turkey (Öztemiz et al. 2013). However, the existence of only *T. brassicae* (Koca et al. 2018), *T. euproctidis* Girault (Ercan et al. 2011) and *T. evanescens* (unpublished data) has been demonstrated in Turkey using molecular methods.

Among the genetic markers, rDNA is found in all organisms and contains various regions that evolve at different rates. Therefore, its use as a molecular marker is preferred in studies comparing closely related species and populations. It also plays an important role in the molecular diagnosis of *Trichogramma* spp. In addition, it is known that DNA-based methods are not affected by the life stage or sex of *Trichogramma* species (Vanlerberghe-Masutti 1994).

In the present study, 37 *Trichogramma* samples were identified based on ITS2 sequences and *T. brassicae* and *T. turkestanica* species were determined. *T. evanescens*, *T. brassicae* and *T. euproctidis* were included in the *Trichogramma* species *evanescens* group in Europe and *T. turkestanica* and *T. euproctidis* were used synonymously (Roji and Pintureau 2003). *T. brassicae* was first identified by Bezdenko (Bezdenko 1968). Later, it was determined as a valid species by morphometric (Pintureau and Keita 1989), biochemical (Pintureau and Keita 1989) and DNA-based studies (Laurent et al. 1996).

Silva et al. (1999) used ITS2 sequence analysis, as in this study, to identify the 5 *Trichogramma* species in Portugal. Honda et al. (2006) re-evaluated *Trichogramma* species from Japan and identified 3 new species based on the ITS2 sequence. In another study, new *Trichogramma* species, *T. itsybitis*, was found based on ITS2 sequences analysis (Pinto et al. 2002). It has been reported that classical methods are not sufficient to diagnose micro-hymenoptera species at the racial level (Landry et al. 1993). Thomson et al. (2003) used ITS2 sequence analysis to identify *Trichogramma* species from South-eastern Australia and found that the ITS2 size was different for each species. Similarly, the size of ITS2 differed between species in the present study.

In the study, while 25 of 37 *Trichogramma* specimens were collected in 2018, the remaining 12 specimens were collected in 2021. These *Trichogramma* samples were divided into 2 main groups. Five specimens were included in a group named: *T. turkestanica* according to
Fig. 1 Phylogenetic tree based on ITS2 gene region of *Trichogramma* specimens of 37 different studied samples (AD1-AD37) was constructed by using the Maximum Likelihood method based on the Kimura 2-parameter model with a uniform distribution to model evolutionary rate differences among sites. Numbers on the branches indicate the bootstrap values of MP.
Fig. 2 Phylogenetic tree based on ITS2 gene region of *Trichogramma* specimens was constructed by using the Maximum Likelihood method based on the Kimura 2-parameter model with a discrete Gamma distribution to model evolutionary rate differences among sites. Numbers on the branches indicate the bootstrap values of MP.
ITS2 sequence analysis. The other 32 samples were in the 2nd main group. These samples were named *T. brassicae* according to ITS2 sequence analysis. According to the genetic distance relationship, the samples in both groups were completely similar in themselves.

**Conclusions**

The ITS2 region is one of the most important molecular markers that can be used to compare closely related species, subspecies and populations. The correct identification of the natural enemy in a biological control program plays a key role in the success of this control. The absence of adequate and appropriate techniques for diagnosis makes the biological control program unsuccessful. Therefore, there is a need to develop molecular diagnostic tools to identify natural enemies and understand population dynamics.

The use of molecular techniques in *Trichogramma* systematics has been made by many researchers for many years (Karimi 2012). The molecular technique used in this study contributed to the correct identification of *Trichogramma* species distributed in Turkey. It is obvious that there are species spread in the country and waiting to be identified molecularly.

**Abbreviations**

ITS2: Internal transcribed spacer 2; rDNA-ITS2: Internal transcribed spacer regions of ribosomal DNA.

**Acknowledgements**

Not applicable.

**Author contributions**

FSE was responsible for molecular data and prepared the manuscript. SÖ collected insect samples. MAA prepared the data. All authors contributed to writing and editing the manuscript. All authors read and approved the final manuscript.

**Funding**

Not applicable.

**Availability of data and materials**

All data generated or analyzed during this study are included in this published article.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Fig. 3** Secondary structures predictions and ΔG (Gibbs) free energy values (at the bottom) of *T. turkestanica* (a) *T. brassicae* (b) were calculated and analyzed at mFOLD web server (http://www.unafold.org/mfold/applications/rna-folding-form-v2.php)
