MOLECULAR PHYLOGENY AND SYSTEMATIC STATUS OF SOME TANACETUM L. (ASTERACEAE) TAXA FROM TURKEY

GULDEN KOCAK¹, ASLI YILMAZ² AND ALPASLAN KOCAK³

¹Department of Molecular Biology and Genetics, Faculty of Arts and Sciences, Bingol University, 12000 Bingol, Turkey
²Department of Biosciences, Faculty of Arts and Sciences, Ataturk University, 25240 Erzurum, Turkey
³Department of Biology, Faculty of Arts and Sciences, Bingol University, 12000 Bingol, Turkey

Corresponding author’s email: gulden.kocak23@gmail.com

Abstract

Tanacetum is an important member of the Asteraceae family and considered the most problematic genus and the phylogenetic position of some taxa is of great interest because of the high morphologic diversity and taxonomical complexity. Molecular phylogeny and systematic status of some Tanacetum taxa from Turkey has been carried out. T. heterotomum (endemic), T. cadmeum subsp. orientale (endemic), T. cappadocicum, and T. eginense have not been included in any molecular phylogenetic analysis yet. To determine the phylogenetic relationships and taxonomic status of Tanacetum L. taxa we analysed internal transcribed spacer (ITS) from nuclear ribosomal DNA (nrDNA) and trnL (UAA) intron, intergenic spacer between the trnL (UAA) 3′ exon and trnF (GAA) from chloroplast DNA (cpDNA). According to the phylogenetic trees two main clades were formed. First clade included T. eginense, T. cadmeum subsp. orientale and T. cappadocicum, the other clade included T. heterotomum according to the data based upon the nrDNA. On the other hand, according to the cpDNA data, all four taxa were located in the same branch.

Key words: Tanacetum, nrDNA, cpDNA, Phylogeny.

Introduction

The genus Tanacetum L. is an important member of the Asteraceae family and widespread mainly in Europe and western Asia through the northern temperate regions. Anthemideae is a medium-sized tribe in the family (Valles et al., 2005) comprising of 111 genera and about of 1,800 species worldwide (Oberprieler et al., 2007). Tanacetum consists of about 150-200 species and it is the third largest genus of the Anthemideae tribe after Artemisia (522 ssp.) and Anthemis (175 spp.) (Heywood 1976; Soreng & Cope, 1991). The Asteraceae family is the second largest family according to the Flora of Turkey with regard to number of genera (Ozhatay et al., 2011). In Flora of Turkey 143 genera and 1484 species are recorded that belong to Asteraceae family, among them 474 species are endemic and endemism ratio is approximately 38% (Davis 1982; Ozhatay et al., 2011). Tanacetum contains several annual and perennial taxa and it is represented in Turkey by 46 species, 18 of which are endemic and endemism rate is 40% (Güner et al., 2012).

Tanacetum has not been divided in sections but placed in three groups in Flora of Turkey (A, B and C) based largely on capitula and flower characters (Davis 1982). According to Flora of Turkey discrimination of this genus is done as follows:

1. Capitula heterogamous; marginal female flowers present, ligulate but sometimes inconspicuous and scarcely longer than disc flowers;
2. Female flowers white, pale sulphur yellow (but not bright yellow) or pinkish red, always with conspicuous ligules (Group A)
3. Female flowers bright or deep yellow, ligules sometimes inconspicuous (Group B)
4. Capitula homogamous, discoid; female flowers completely absent (Group C)

Until today many molecular-phylogenetic studies have been done to resolve generic delimitation and infrageneric classifications of many groups of Anthemideae (Watson et al., 2000; Masuda et al., 2009; Zhao et al., 2010; Sonboli et al., 2011; Sonboli et al., 2012). Tanacetum is considered to be one of the most problematic genera and the phylogenetic position of some taxa is of great interest because of the high morphologic diversity and taxonomical complexity. Tanacetum which is a polymorphic genus is described to have important variation in flowers, inflorescence morphology, and achenes (Sonboli et al., 2012). Among the published studies, Sonboli et al., (2012) was studied the phylogenetic position of 80 Tanacetum taxa within the tribe.

The use of the internal transcribed spacer of the nuclear ribosomal repeat (nrDNA ITS) region in plant molecular systematics has been reviewed by Baldwin et al., (1995). The entire ITS region is now a widely used data source in molecular systematic studies of plants at lower taxonomic levels for three principal reasons. First, the high copy number allows easy amplification of the region from total DNA. Second, the spacer sequences evolve rapidly and can therefore resolve lower level relationships better than slowly evolving genes, such as 18S and rbcl (Baldwin, 1992; Baldwin et al., 1995; Baker et al., 1999). Third, the availability of several sets of universal (or near so) PCR primers working with a large diversity of taxonomic groups is easy (White et al., 1990; Gardes & Bruns, 1993). In addition to nrDNA sequences, noncoding chloroplast sequences as the trnL (UAA) intron and the intergenic spacer between the trnL (UAA) 3′ exon and the trnF (GAA) gene also have phylogenetic potential (Taberlet et al., 1991).
The aim of this study is to provide first report on the systematic position of *T. heterotomum*, *T. cadmeum* subsp. *orientale*, *T. cappadocicum*, and *T. eginense* which have been not included in any molecular phylogenetic analysis yet. Among these *T. heterotomum* and *T. cadmeum* subsp. *orientale* are endemic taxa to Turkey (Davis, 1982). In this study, we used molecular data from entire nrDNA ITS region and we further included sequence information from the cpDNA non-coding regions trnL (UAA) intron, intergenic spacer between the trnL (UAA) 3’ exon and trnF (GAA) to provide a more comprehensive taxonomic and phylogenetic results and a more stable classification.

**Material and Methods**

**Plant material:** Plant material was obtained from silica-gel dried leaved of collected specimens in the wild. *T. abrotanifolium* was collected from natural habitats in Sancak (Bingol), 2008, *T. argenteum* subsp. *argenteum* was collected from (Harput) Elazig, 2008, *T. balsamita* subsp. *balsamita* was collected from Metan village (Bingol), 2014, *T. cappadocicum* was collected from Munzur valley (Tunceli), 2008, *T. cadmeum* subsp. *orientale* was collected from Kayalik Village-Palu (Elazig), 2013, *T. ciliicum* was collected from Saban village (Bingol), 2014, *T. chilophyllum* var. *chilophyllum* was collected from Sancak (Bingol), 2013, *T. chilophyllum* var. *chilophyllum* was collected from Sivrice (Elazig), 2008, *T. densum* subsp. *amani* was collected from Tecer mountain (Sivas), 2008, *T. eginense* was collected from Darende (Malatya), 2008, *T. heterotomum* was collected from Kangal (Sivas), 2008, *T. kotschyi* was collected from Saban village (Bingol), 2014, *T. mucroniferum* was collected from Munzur valley (Tunceli), 2008, *T. nitens* was collected from Baskil (Elazig), 2008, *T. parthenium* was collected from Pnaralı village (Bingol), 2013, *T. parthenifolium* was collected from Saban village (Bingol), 2014, *T. vulgare* was collected from Munzur mountain (Tunceli), 2008, *T. zahlbruckneri* was collected from Saban village (Bingol), 2014. The plant materials were identified by Dr. A. Kocak. Voucher specimens were deposited at the Molecular Biology and Genetics Laboratory of Bingol University and Plant Products and Biotechnology Research Laboratory of Firat University.

DNA extraction, amplification, and sequencing: Total genomic DNA was extracted by modified protocol of the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). Polymerase chain reaction (PCR) of the whole region of nrDNA ITS were performed using the ITS AB101 and ITS AB102 primers (Douzery et al., 1999) and ITS4 and ITS5 primers (White et al., 1990) in some cases. PCR amplifications were conducted according to the protocols described in Somboi et al., (2010). Amplification of intergenic spacer trnL (UAA) intron (B49317 and A49855 primers) and intergenic spacer between the trnL (UAA) 3’ exon and trnF (GAA) (B49873 and A50272 primers) were performed according to the protocols of Taberlet et al., (1991). Sequencing reactions were performed using ABI 3730 XL (Applied Biosystems).

Alignment and phylogenetic analyses: Phylogenetic analysis were undertaken using three data sets of samples and each included the sequences from the GenBank database of the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) were aligned using ClustalW (Thompson et al., 1994) software and subsequently checked visually. Indels were not treated in final datasets. Ultimately, evaluation carried out by grouping the data into three sets as nrDNA and two regions from cpDNA. The first dataset was comprised of both studied taxa ITS (ITS 1, 5.8S and ITS 2) sequences and the ITS sequences from closely related taxa retrieved from the NCBI database (Table 1). The second dataset was composed of the sequences of the species of the genera of *trnL* from current study, the third and the last dataset included both *trnL*-F and the GenBank sequences.

Variable sites, number of parsimony-informative sites, transition, transversion, genetic distance, nucleotide diversity, and divergence within species were computed as molecular diversity statistics for each dataset using Molecular Evolutionary Genetics Analysis software (MEGA 6.0; Tamura et al., 2013). Ultimately, pylonetic tree was constructed by Maximum Likelihood Method with 1000 bootstrap replicates.

**Table 1. Accession number from the NCBI database (out group).**

<table>
<thead>
<tr>
<th>Specimens</th>
<th>Internal transcribed spacer (ITS)</th>
<th>trnL</th>
<th>trnL-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. millefolium</em></td>
<td></td>
<td>AY603263</td>
<td>-</td>
</tr>
<tr>
<td><em>T. vulgar</em></td>
<td>EF577323</td>
<td>-</td>
<td>EF577378</td>
</tr>
<tr>
<td><em>T. vulgare</em></td>
<td>AY603264</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>T. cinerarifolium</em></td>
<td>EF577319</td>
<td>-</td>
<td>EF577374</td>
</tr>
<tr>
<td><em>T. macrophyllum</em></td>
<td>AY603262</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>T. ptarmiciflorum</em></td>
<td>EF577322</td>
<td>-</td>
<td>EF577377</td>
</tr>
<tr>
<td><em>T. parthenium</em></td>
<td>EF577320</td>
<td>-</td>
<td>EF577375</td>
</tr>
<tr>
<td>Achillea biebersteinii</td>
<td>AY603218</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Achillea millefolium</em> subsp. sudetica</td>
<td>AY603187</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Achillea millefolium</td>
<td>AY603186</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anthemis arvensis</td>
<td>EU179214</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anthemis melampodia</td>
<td>KJ004380</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anthemis cotula</td>
<td>EU179216</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2. Numeric information of ITS, trnL and trnL-F.

<table>
<thead>
<tr>
<th></th>
<th>ITS</th>
<th>trnL</th>
<th>trnL-F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the aligned sequence (including all taxa with out group)</td>
<td>826</td>
<td>555</td>
<td>514</td>
</tr>
<tr>
<td>GC% content (including all taxa with out group)</td>
<td>50.4</td>
<td>34.1</td>
<td>36.8</td>
</tr>
<tr>
<td>Parsimony informative sites (including all taxa with out group)</td>
<td>118</td>
<td>17</td>
<td>20</td>
</tr>
</tbody>
</table>

Results

The characteristics of sequences: The aligned data set of entire ITS, trnL and trnL-trnF included a total of 37 (10), 26 and 29 (4) taxa. The number in the parenthesis indicates the taxa taken from GenBank. ITS, trnL and trnL-trnF sequences length varied from 826, 555, and 514 respectively. The aligned data comprised of 593bp from nrDNA, 418 and 337bp from cpDNA without gaps and missing data including the taxa taken from GenBank and all data sets comprised different numbers of taxa from five different genera (Tanacetum, Anthemis, Achillea, Bellis and Paeonia) including outgroup taxa. The parsimony informative sites were 118 for nuclear DNA gene region, 17 for trnL and 20 for trnL-trnF of cpDNA (Table 2). Considering Tanacetum taxa without other taxa parsimony informative sites were 43, 8 and 5 respectively.

The evolutionary characteristics: According to the three phylogenetic trees, constructed by using both chloroplast and nuclear DNA showed that all taxa of Tanacetum supported as a monophyletic group, but basically two main clades were observed in the Tanacetum taxa (Figs. 1, 2 and 3). One clade, including most specimens of Tanacetum especially the taxa T. eginense, T. cadmeum subsp. orientale and T. cappadocicum, the other clade comprised of three subclades with T. heterotomum according to the data based upon the nuclear DNA (Fig. 1). On the other hand, based upon the chloroplast DNA data, all four taxa were located in the same clade (Fig. 2). In Flora of Turkey, members of Tanacetum fall into three main clades (Group A, B and C) according to the gender status of the capitula and colour of the flowers. It is clearly seen on phylogenetic tree obtained from ITS data that Clade 1 composed of mainly Group B members, on the other hand, Clade 2 encompassed Group A Tanacetum members.

Discussion

In this study identification of molecular systematic position of T. heterotomum, T. cappadocicum, T. eginense and T. cadmeum subsp. orientale tax was done by using nrDNA and cpDNA sequence analysis. Whereas these taxa were simply classified by their capitulum sexual status, ligular and tubular flower structure and leaf fragmentation characters. According to Flora of Turkey T. heterotomum is in the group A, T. cappadocicum, T. eginense and T. cadmeum subsp. orientale are in the group B (Davis, 1982).

The results of nuclear DNA analysis correspond to the group separation in Flora of Turkey. Phylogenetic trees which are based upon the nuclear DNA show that basically two main clades are formed among taxa of Tanacetum and they were collected together and separated from outgroup taxa completely. While first clade includes three species; T. eginense, T. cadmeum subsp. orientale and T. cappadocicum, the other clade includes T. heterotomum (Fig. 1). These results show conformity to the groups separation based on morphological charts in Flora of Turkey (Davis, 1982). However, according to the one type of chloroplast marker (cpDNA trnL), all four species were located in the same branch and this result was not compatible with the group recognized in Flora of Turkey (Fig. 2). These analysis results on the other hand don’t comply with the groups recognized in Flora of Turkey based on morphological charts (Davis, 1982).

On the other hand with the other chloroplast marker (cpDNA trnL-F) analysis shows that T. cappadocicum, T. eginense and T. cadmeum subsp. orientale are in the same clade but T. heterotomum could not be evaluated because PCR studies with primers for trnL-F region didn’t work (Fig. 3).

T. balsamita subsps. balsamita which is defined in both group A and group C according to the Flora of Turkey, located among the taxa with heterogamous capitulum (group A). Also it is obvious that the genus Tanacetum is closely related to the genera Achillea, and Anthemis.

Sonboli et al., (2012) studied molecular phylogeny of Tanacetum genus and based on the nrDNA (ITS), cpDNA (trnH-psbA) sequence variation they introduced the infrageneric taxonomy and Bayesian tree of Tanacetum. In addition to Sonboli et al., (2012), with this study the position of T. heterotomum (endemic), T. cadmeum subsp. orientale (endemic), T. cappadocicum, and T. eginense in the phylogenetic tree of Tanacetum was determined for the first time.

In conclusion, dendrograms drawn by the results of both nrDNA (ITS) and cpDNA (trnL and trnL-F) sequence analysis put forth immense diversity. In future detail evaluation with other variable markers between all members of the Tanacetum genus and with closely related genera will make this study more meaningful.

Acknowledgement

This study was supported by Bingol University (BAP Project No.712-159-2013).
Fig. 1. Maximum Likelihood tree based upon the Tamura-Nei model of nrDNA ITS region with 1000 bootstrap replicates. Bootstrap values are represented next to each node of the branches. Important clades and flower information are indicated in different colours on the tree.

Fig. 2. One of the cpDNA based Maximum Likelihood method tree specifically the sequence data of trnL region.
Fig. 3. Maximum Likelihood cpDNA tree based on the data obtained from the sequences of the \textit{trnL-trnF}.

**References**


National Centre for Biotechnology Information. Website: \url{http://www.ncbi.nlm.nih.gov}


(Received for publication 25 September 2017)