DNA barcoding to analyse taxonomically complex groups in plants: the case of *Thymus* (Lamiaceae)

SILVIA FEDERICI1, ANDREA GALIMBERTI1, FABRIZIO BARTOLUCCI2, ILARIA BRUNI1, FABRIZIO DE MATTIA1, PIERLUIGI CORTIS3 and MASSIMO LABRA1*

1ZooPlantLab, Dipartimento di Biotecnologie e Bioscienze,Università degli Studi di Milano Bicocca, Piazza della Scienza 2, 20126 Milano, Italy
2Centro Ricerche Floristiche dell’Appennino (Parco Nazionale del Gran Sasso e Monti della Laga – Scuola di Scienze Ambientali dell’Università di Camerino), San Colombo, Via Provinciale Km 4.2, 67021 Barisciano (AQ), Italy
3Dipartimento di Scienze della Vita e dell’Ambiente Macrosezione Botanica ed Orto Botanico,Università degli Studi di Cagliari, Viale S. Ignazio 13, 09123 Cagliari, Italy

Received 30 December 2012; revised 10 January 2013; accepted for publication 18 January 2013

We evaluated the utility of the core barcode regions (*matK* and *rbcL*) and the plastid intergenic spacer *trnH-psbA* to distinguish between *Thymus* spp. This is a taxonomically complex group that has been investigated so far mainly using morphological approaches. Thirty-six samples representing nine different morphospecies were collected and used for molecular analysis. The three markers showed clear amplification and sequencing. However, the genetic variation and the resulting haplotype networks showed that only *Thymus capitatus* forms a well-defined ‘barcoding gap’ compared with the other taxa. The identification problems observed in the other *Thymus* spp. may be related to reduced gene flow among populations, resulting in high intraspecific and low interspecific genetic variation. This situation does not permit the definition of species-specific barcodes. A second hypothesis suggests that morphological traits used for the delimitation of *Thymus* spp. do not reflect real biological and molecular species boundaries. If this is the case, the taxonomy of *Thymus* should be revised through extensive sampling and analyses with different tools (i.e. molecular variability, morphology, geographical distribution, etc.) to define the natural units at the species level. © 2013 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2013, 171, 687–699.

**ADDITIONAL KEYWORDS:** *matK* – *rbcL* – species delimitation – *trnH-psbA.*

---

**INTRODUCTION**

The use of molecular identification techniques, such as DNA barcoding (Hebert *et al*., 2003), has recently taken on an important role in the definition of taxonomic status and evolutionary processes for almost all existing taxa (Newmaster, Ragupathy & Janovec, 2009). In particular, the promise of DNA barcoding is that it will provide a quick, simple and economic tool to identify and discover biological diversity (Casiraghi *et al*., 2010; De Mattia *et al*., 2012). In 2009, the Consortium for the Barcode of Life (CBOL) Plant

*Corresponding author. E-mail: massimo.labra@unimib.it*
This is evident in the case of taxonomically complex groups (TCGs) consisting of a genetic mixture of related individuals, often at more than one ploidy, in which biological diversity defies simple classification into discrete species (Ennos, French & Hollingsworth, 2005). Species in TCGs can be difficult to define and identify. This situation leads to practical difficulties in implementing conservation measures on endangered, rare or endemic taxa in TCGs. Secondly, a species-based approach based only on morphological analysis often fails to capture the whole diversity present in TCGs. Generally, the species belonging to TCGs are indeed hard to classify in stable and coherent taxa even for expert taxonomists (Ennos et al., 2012).

In such a context, the taxonomy of TCGs presents two problems: (1) the identification of stable and discrete characters of each species; and (2) the choice of effective tools with which to analyse these characters and to distinguish each species (Rieseberg, Wood & Baack, 2006).

Thymus is one of the TCGs in the European flora, consisting of 215 accepted species, distributed across the Eurasian continent, northern Africa and southern Greenland, with high levels of polymorphism as a result of post-glacial colonization dynamics and hybridization events occurring even between taxa that are not closely related (Bartolucci, 2010). Other important factors influencing the evolution of the genus are polyploidy, disloidal/aneuploidy (Morales, 2002) and gynodioecy (the co-existence of hermaphrodite and ‘female’ plants with much-reduced anthers and little or no viable pollen) in natural populations (Darwin, 1877; Thompson, 2002).

Jalas (1971), based on morphological characters, divided Thymus into two subgenera (Thymus, Coridothymus (Rchb.f.) Borbás) and eight sections: T. section Micantes Velen., T. section Mastichina (Mill.) Benth., T. section Piperella Willk., T. section Teucrioides Jalas, T. section Pseudothymbra Benth., T. section Thymus, Hypodromi (A. Kern.) Halácsy and T. section Serpillum (Mill.) Benth. In our work, nine species were analysed using the DNA core barcode markers (rbcL and matK) and the supplementary intergenic plastid region trnH-psbA. One species belongs to subgenus Coridothymus and the others belong to three sections of subgenus Thymus: T. section Thymus (one species), T. section Hypodromi (three species) and T. section Serpillum (four species). For one species (T. striatus Vahl), we also sampled three different intraspecific taxa to investigate their putative molecular variability with DNA barcoding markers. To evaluate the usefulness of the DNA barcoding approach in TCGs, multiple accessions for each taxon were sampled to investigate: (1) the utility of the three candidate loci for DNA barcoding; (2) the presence of a ‘barcoding gap’ (Meyer & Paulay, 2005) between the ranges of genetic interand intraspecific distances; and (3) the congruence between traditional taxonomic assessments (based on morphological data) and DNA barcoding data.

**MATERIAL AND METHODS**

**SAMPLING COLLECTION AND TAXONOMIC IDENTIFICATION BASED ON MORPHOLOGY**

Thirty-six samples representing nine Thymus spp. were collected for morphological and molecular (DNA barcoding) analyses (Tables 1, 2). We studied individuals, mostly from Italy, belonging to one species of subgenus Coridothymus and three sections of subgenus Thymus. To maximize the chance of observing intraspecific geographical variation, conspecifics were sampled from distant sites. At least three individuals per species were sampled, with the exception of T. paronychioides Čelak, an endemic species from Sicily (Italy) with only two known populations (Bartolucci & Peruzzi, in press); this was represented by only two individuals in our dataset. In addition, three different intraspecific taxa of T. striatus were considered to evaluate the genetic variability at the intraspecific level.

Taxonomic identification of samples was conducted according to dichotomous keys in floras and taxonomic papers concerning the Mediterranean area (Jalas, 1972; Baden, 1991; Morales, 2010; Bartolucci & Peruzzi, in press). The most distinctive traits of each accession are described in Table 1. In addition, sampled individuals were compared with herbarium specimens kept in the Floristic Research Centre of the Apennines (APP) to confirm their identification. Each specimen was deposited at the same herbarium and a sample for each individual was stored for DNA extraction and analysis. These samples were vouchered as ‘MIB-ZPL’ following the protocols specified by the Registry of Biological Repositories (http://www.biorepositories.org) and the data standards for BARCODE Records in INSDC (http://barcoding.si.edu/PDF/DWG_data_standards-Final.pdf). A list of samples and voucher names is included in Table 2 and the distribution of sampling localities is depicted in Figure 1.

**DNA BARCODING ANALYSIS**

DNA was isolated starting from 20 mg of plant material (young leaves) using the DNeasy Plant Mini kit (Qiagen, Milan, Italy). The concentration of extracted DNA for each sample was estimated both fluorometrically and by comparison of ethidium bromide-stained band intensities with a λ DNA standard. DNA barcoding analysis was performed with the rbcL and
<table>
<thead>
<tr>
<th>Taxon</th>
<th>Supraspecific classification</th>
<th>Habit</th>
<th>Stem indumentum (upper internode)</th>
<th>Leaf morphology</th>
<th>Leaf hairiness</th>
<th>Median cauline leaf (length $\times$ width) (mm)</th>
<th>Calyx length (mm)</th>
<th>Calyx morphology</th>
<th>Bract morphology</th>
<th>Inflorescence shape</th>
<th>Corolla colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. capitatus (L.) Hoffmanns. &amp; Link</td>
<td>Thymus subgenus Coridothymus (Rchb.f.) Borbas</td>
<td>Dwarf shrub</td>
<td>Holotrichous</td>
<td>Linear to lanceolate, rarely with revolute margins</td>
<td>Subglabrous or puberulent</td>
<td>(4)6–10 $\times$ 1.5–2</td>
<td>3.5–5.5</td>
<td>Cylindrical</td>
<td>Dorso-ventrally compressed with two lateral keels</td>
<td>Not similar to the leaves, ovate to lanceolate, greenish</td>
<td>Pink or purple, rarely white</td>
</tr>
<tr>
<td>T. vulgaris L. ssp. vulgaris</td>
<td>Thymus subgenus Thymus section Thymus subsection Thymus</td>
<td>Plant erect with woody branches</td>
<td>Holotrichous</td>
<td>Linear to elliptical, with revolute margins</td>
<td>Tomentose</td>
<td>3.5–7(9) $\times$ 1–3</td>
<td>2.5–4.5</td>
<td>Campanulate</td>
<td>Similar to the leaves, greenish</td>
<td>Capitula</td>
<td>Pink or purple</td>
</tr>
<tr>
<td>T. striatus Vahl ssp. acicularis (Waldst. &amp; Kit.) Ronninger</td>
<td>Thymus subgenus Thymus section Hypodromi (A. Kern.) Halácsy subsection Subbracteati (Kokkv.) Jalas</td>
<td>Plant with non-flowering long creeping branches, rarely with creeping branches ending in a terminal inflorescence</td>
<td>Holotrichous</td>
<td>Acicular to lanceolate, often falcate</td>
<td>Glabrous</td>
<td>Prominent, parallel</td>
<td>4–10 $\times$ 0.9–1.5</td>
<td>3.5–4.5</td>
<td>Cylindrical</td>
<td>Not similar to the leaves, trullate, purplish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td>T. striatus Vahl ssp. striatus</td>
<td>Thymus subgenus Thymus section Hypodromi (A. Kern.) Halácsy subsection Subbracteati (Kokkv.) Jalas</td>
<td>Plant with non-flowering long creeping branches, rarely with creeping branches ending in a terminal inflorescence</td>
<td>Holotrichous</td>
<td>Lanceolate to spatulate</td>
<td>Glabrous</td>
<td>Prominent, parallel</td>
<td>6.5–1314 $\times$ 1.4–2.6</td>
<td>4–5.5</td>
<td>Cylindrical</td>
<td>Not similar to the leaves, trullate, greenish to purplish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td>T. striatus var. ophioliticus (Lacaita) Fiori</td>
<td>Thymus subgenus Thymus section Hypodromi (A. Kern.) Halácsy subsection Subbracteati (Kokkv.) Jalas</td>
<td>Plant with non-flowering long creeping branches, rarely with creeping branches ending in a terminal inflorescence</td>
<td>Holotrichous</td>
<td>Acicular to lanceolate, often falcate</td>
<td>Glabrous</td>
<td>Prominent, parallel</td>
<td>5.5–1316 $\times$ 0.9–1.5</td>
<td>3.5–5</td>
<td>Cylindrical</td>
<td>Not similar to the leaves, trullate, purplish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td>T. spinulosus Ten.</td>
<td>Thymus subgenus Thymus section Hypodromi (A. Kern.) Halácsy subsection Subbracteati (Kokkv.) Jalas</td>
<td>Plant with non-flowering long creeping branches, rarely with creeping branches ending in a terminal inflorescence</td>
<td>Holotrichous</td>
<td>Lanceolate, Hirsut</td>
<td>Weak</td>
<td>6.5–15 $\times$ 1.1–2.2(2.7)</td>
<td>4–5(5.7)</td>
<td>Cylindrical</td>
<td>Not similar to the leaves, trullate, rarely caputla</td>
<td>Elongate to interrupted, rarely caputla</td>
<td>White, rarely pink</td>
</tr>
</tbody>
</table>
### Table 1. Continued

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Supraspecific classification</th>
<th>Habitus</th>
<th>Stem indumentum (upper internode)</th>
<th>Leaf morphology</th>
<th>Leaf hairiness</th>
<th>Median cauline leaf (length × width) (mm)</th>
<th>Calyx length (mm)</th>
<th>Calyx morphology</th>
<th>Bract morphology</th>
<th>Inflorescence shape</th>
<th>Corolla colour</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. paronychioides</em> Cˇelak</td>
<td><em>Thymus</em> subgenus <em>Thymus</em> section <em>Hypodromi</em> (A. Kern.) Halácsy subsection <em>Serpyllastrum</em> Villar</td>
<td>Plant with long, non-flowering, creeping branches</td>
<td>Holotrichous</td>
<td>Spathulate to lanceolate</td>
<td>Hirsute</td>
<td>Prominent, curved</td>
<td>(5)-6.5(-10.5) × (1.3)-1.6-2.6-3.2</td>
<td>4.5-6</td>
<td>Cylindrical</td>
<td>Not similar to the leaves, trullate, greenish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td><em>T. oenipontanus</em> Heinr. Braun</td>
<td><em>Thymus</em> subgenus <em>Thymus</em> section <em>Serpyllum</em> (Mill.) Benth. subsection <em>Isolepides</em> (Borbáš) Halácsy</td>
<td>Plant with creeping branches ending in a terminal inflorescence</td>
<td>Holotrichous</td>
<td>Elliptic-lanceolate</td>
<td>Glabrous</td>
<td>Prominent, curved</td>
<td>5–17 × 1.4–4.5/-7</td>
<td>3–5</td>
<td>Cylindrical</td>
<td>Similar to the leaves, greenish</td>
<td>Pink, rarely capitula white</td>
</tr>
<tr>
<td><em>T. longicaulis</em> C.Presl</td>
<td><em>Thymus</em> subgenus <em>Thymus</em> section <em>Serpyllum</em> (Mill.) Benth. subsection <em>Pseudomarginati</em> (Heinr. Braun ex Borbáš) Jalas</td>
<td>Plant with non-flowering, long, creeping branches</td>
<td>Amphitrichous</td>
<td>Linear-lanceolate to elliptical</td>
<td>Glabrous</td>
<td>Weak</td>
<td>(2)-6-12.15 × (1.5)-2-4.5/-6</td>
<td>2.5-3.5/-4.5</td>
<td>Campanulate</td>
<td>Similar to the leaves, greenish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td><em>T. praecox</em> Opiz ssp. <em>polytrichus</em> (A.Kern. ex Borbáš) Jalas</td>
<td><em>Thymus</em> subgenus <em>Thymus</em> section <em>Serpyllastrum</em> (Mill.) Benth. subsection <em>Pseudomarginati</em> (Heinr. Braun ex Borbáš) Jalas</td>
<td>Plant with long, non-flowering, creeping branches</td>
<td>Holotrichous or amphitrichous</td>
<td>Obovate, spathulate to suborbicular</td>
<td>Glabrous to hirsute</td>
<td>Prominent, curved, anastomosate from a marginal vein</td>
<td>5–10(15) × 2–5/-7</td>
<td>3-5(5.6)</td>
<td>Cylindrical</td>
<td>Similar to the leaves, usually purplish</td>
<td>Pink or purple</td>
</tr>
<tr>
<td><em>T. pulegioides</em> L. var. <em>pulegioides</em> Klokov</td>
<td><em>Thymus</em> subgenus <em>Thymus</em> section <em>Serpyllum</em> (Mill.) Benth. subsection <em>Alternantes</em> Klokov</td>
<td>Plant subrect to procumbent</td>
<td>Goniotrichous</td>
<td>Ovate to oblong-lanceolate</td>
<td>Glabrous</td>
<td>Weak to prominent, curved</td>
<td>4–15 × 1.5-10</td>
<td>3-4</td>
<td>Campanulate to cylindrical</td>
<td>Similar to the leaves, greenish</td>
<td>Pink or purple, rarely capitula white</td>
</tr>
</tbody>
</table>

Diagnostic traits for each taxon are provided according to Bartolucci (2010) and Bartolucci & Peruzzi (in press).
Table 2. Sampling table. List of the 36 samples analysed. For each sample, the morphospecies name, voucher number, geographical coordinates of the sampling locality and the haplotype for each barcode marker are provided.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Voucher sample</th>
<th>Latitude Longitude</th>
<th>Haplotype</th>
<th>rbcl</th>
<th>matK</th>
<th>trnH-psbA</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. capitatus (L.) Hoffmanns. &amp; Link</td>
<td>MIB:ZPL:04827 37°9′20.01″N 15°1′14.10″E</td>
<td>R5 M20 S33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. vulgaris L. var. vulgaris</td>
<td>MIB:ZPL:04819 41°17′31.51″N 13°15′35.33″E</td>
<td>R1 M16 S23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. striatus ssp. acicularis (Waldst. &amp; Kit.) Ronniger</td>
<td>MIB:ZPL:04797 42°13′33.10″N 12°58′26.89″E</td>
<td>R1 M1 S1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. striatus Vahl ssp. striatus</td>
<td>MIB:ZPL:04801 40°09′33.42″N 15°50′20.52″E</td>
<td>R1 M3 S4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. striatus var. ophioliticus (Lacaita) Fiori</td>
<td>MIB:ZPL:04804 43°55′15.30″N 11°04′14.20″E</td>
<td>R1 M1 S22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. paronychioides Čelak</td>
<td>MIB:ZPL:04848 46°39′21.98″N 13°46′26.79″E</td>
<td>R1 M17 S29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. longicaulis C.Presl</td>
<td>MIB:ZPL:04803 40°39′28.30″N 14°29′39.00″E</td>
<td>R1 M5 S6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. praecox Opiz ssp. polytrichus (A.Kern. ex Borbás) Jolas</td>
<td>MIB:ZPL:04846 42°27′2.86″N 13°33′20.59″E</td>
<td>R1 M17 S30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. pulegioides L. var. pulegioides</td>
<td>MIB:ZPL:04807 42°36′46.89″N 0°32′9.16″E</td>
<td>R1 M3 S11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

matK coding regions, and trnH-psbA intergenic spacer as an additional marker (Hollingsworth et al., 2011).

Polymerase chain reaction (PCR) for each marker was performed using puReTaq Ready-To-Go PCR beads (Amersham Biosciences, Milan, Italy) in a 25-μL reaction volume according to the manufacturer’s instructions. PCR cycles consisted of an initial denaturation (7 min at 94 °C), 35 cycles of denaturation (45 s at 94 °C), annealing (30 s at 50 °C for rbcl and 53 °C for matK and trnH-psbA) and extension (1 min at 72 °C), and a final extension for 7 min at 72 °C. The genes rbcl and matK were amplified using primers rbcl1F–rbcl724R (Fay et al., 1998) and matK950F–matK1326R (Sun, McLewin & Fay, 2001; Cuénoud et al., 2002), respectively, and the non-coding region trnH-psbA primers psbA and trnH (Newmaster et al., 2009) were used. Amplicons were bidirectionally sequenced using an ABI 155 3730XL automated sequencer at Macrogen Inc., Seoul, South Korea. Manual editing of raw traces and subsequent alignments of forward and reverse sequences using the Bioedit sequence alignment editor (version 7.0.5; Hall, 1999) enabled us to assign edited sequences for
Figure 1. Collection sites of the nine Thymus morphospecies. Three different symbols were used to distinguish the three intraspecific taxa of Thymus striatus.
RESULTS

DNA BARCODING SUCCESS

DNA extraction was successful for all the 36 samples with high DNA quality and good yield (i.e. 30–40 ng µL⁻¹). On amplification, all tested loci, rbcL, matK and trnH-psbA, yielded a single band and exhibited 100% amplification success with standard primers. All the PCR products were easily sequenced and high-quality bidirectional sequences were obtained. After primer trimming and alignment, the same sequence length was observed in all the analysed samples for rbcL (599 bp), whereas differences were observed in matK and trnH-psbA alignments, for which the sequence lengths ranged from 789 to 795 bp and from 394 to 431 bp, respectively, mainly as a result of insertions/deletions (indels).

Sequence alignment was used to evaluate genetic differences among samples; the highest overall K2P distance was shown by trnH-psbA (2.00%), followed by matK (0.60%) and rbcL (0.10%). Analysis performed with EINVERTED (Guindon & Gascuel, 2003) did not find any short inverted repeat regions, and therefore excluded their putative influence on genetic distance values (data not shown).

To evaluate the ‘DNA barcoding gap’, inter- and intraspecific genetic divergences were computed for each tested marker. The results suggested that trnH-psbA exhibited the highest interspecific variation (3.30%), followed by matK and rbcL with 1.70% and 0.30%, respectively (Table 3). Consistent intraspecific genetic distances were observed for trnH-psbA and matK among the analysed species, with average K2P distances of 1.85% (range, 0.70–2.60%) and 0.46% (range, 0.00–0.80%), respectively. For rbcL, the intraspecific genetic variation ranged from 0% [T. vulgaris L., T. praecox Opiz ssp. polytrichus (A.Kern. ex Borbás) Jalas, T. paronychioides, T. oenipontanus Heirn. Braun, T. longicaulis C.Presl and T. capitatus (L.) Hoffmanns. & Link] to 0.20% for T. striatus (Table 3).

Based on sequence diversity, five haplotypes were identified with rbcL and 20 with matK. The rbcL R1 haplotype was the most common and was shared

<table>
<thead>
<tr>
<th>Locus</th>
<th>Intraspecific comparison</th>
<th>Interspecific comparison</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean % variation (SE%)</td>
<td>Range (%)</td>
<td>Mean % variation (SE%)</td>
</tr>
<tr>
<td>matK</td>
<td>0.46 (0.17)</td>
<td>0.00–0.80</td>
<td>0.71 (0.21)</td>
</tr>
<tr>
<td>rbcL</td>
<td>0.04 (0.03)</td>
<td>0.00–0.20</td>
<td>0.11 (0.06)</td>
</tr>
<tr>
<td>trnH-psbA</td>
<td>1.85 (0.52)</td>
<td>0.70–2.60</td>
<td>2.05 (0.48)</td>
</tr>
</tbody>
</table>

Overall mean distances, intraspecific and interspecific comparisons are considered.

most species. The 3′ and 5′ terminals were clipped to generate consensus sequences for each taxon. The identification of short inverted repeat regions in the trnH-psbA spacer was performed as reported by Whitlock, Hale & Groff (2010), using the EMBOSS Software package (Rice, Longden & Bleasby, 2000). The EINVERTED algorithm (Guindon & Gascuel, 2003) was used with default parameters to detect the occurrence of inversions in the trnH-psbA region.

All sequences were deposited in GenBank. The accession numbers of the sequences are listed in Supporting Information Appendix S1. To assess the ability of the three barcode regions selected to identify the analysed samples unequivocally, genetic distances among different samples were calculated using MEGA 5.0 (Tamura et al., 2011) to generate Kimura two-parameter (K2P) distance matrices. Intra- and interspecific genetic distances were evaluated for all taxa.

Based on sequence similarity, samples were grouped into haplotypes for each barcode marker. Haplotype network analysis was carried out using TCS v. 1.21 (Clement, Posada & Crandall, 2000) which implements the Templeton, Crandall & Sing (1992) statistical parsimony procedure. The maximum number of mutational steps that constitutes a parsimonious connection between two sequence types was calculated with 95% confidence. According to Hart & Sunday (2007), a discrimination approach based on the parsimony connection limit can be useful in a DNA barcoding context. This approach might be useful to assign unknown specimens to known well-sampled taxa (Hart & Sunday, 2007). With this approach, taxa are correctly identified on the basis of correspondence between the number of resulting subnetworks and the number of taxa. A network analysis using this operational species definition might be particularly useful for DNA barcoding studies in which morphological or ecological species markers are labile (Hart & Sunday, 2007). Furthermore, network analysis clarifies the relationships among the haplotypes and defines the correlation between morphological species and genetic diversity.
amongst all *Thymus* morphospecies, with the exception of *T. capitatus* (haplotype R5). R2 is the rarest, and R3 and R4 are shared by two morphospecies, respectively (Table 3, Fig. 2A). These haplotype distributions resulted in a poor structured species network in which only *T. capitatus* was distinguished from the other taxa (Fig. 2A). Among the *matK* haplotypes, seven were shared by more than one population (Table 3, Fig. 2B). Six haplotypes were detected in the eight *T. striatus* accessions without any genetic structuring among different intraspecific taxa. Thirteen haplotypes (M1, M4, M5, M7, M8, M9, M10, M12, M13, M14, M16, M18 and M19) were exclusive haplotypes (Table 3, Fig. 2B).

In the *matK* network, each haplotype is separated by low genetic distances, with the exception of M15, M20 and the clade including the haplotypes M16, M13, M5, M6 and M8. However, only M20 is private for a single morphospecies (*T. capitatus*).

In the case of *trnH-psbA*, 34 haplotypes were detected for 36 analysed samples (Table 3), suggesting that almost all samples analysed had a distinct *trnH-psbA* sequence, with the exclusion of haplotypes S7 and S33, shared by two samples belonging to the same species (S33) or by two species (S7). This variability did not allow us to define an exhaustive network to explain species relationships, as also revealed by a neighbor-joining tree based on R2P distances provided in Supporting Information Figure S1.

**DISCUSSION**

Our data show that the three tested DNA barcoding markers can be easily amplified and sequenced for all the analysed samples. However, their discriminating power seems to be inadequate to distinguish the morphospecies considered. Only *T. capitatus* (previously placed in subgenus *Coridothymus*) showed private haplotypes and a well-defined ‘barcoding gap’. The separation between *T. capitatus* and the other taxa (subgenus *Thymus*) has also been emphasized by morphological data (i.e. calyx dorso-ventrally compressed with two lateral keels, ciliate), chemical profile (Figueiredo *et al*., 2008), isozymes (Ben El Hadj Ali, Guetat & Boussaid, 2012a) and random amplified polymorphic DNA (RAPD) profile (Ben El Hadj Ali, Guetat & Boussaid, 2012b). All this evidence led to a significant change in the taxonomic status of this species, which was therefore moved to the genus *Thymbra* L. Moreover, this consideration was also highlighted by recent phylogenetic analyses by Bräuchler, Meimberg & Heubl (2010) and Theodoridis *et al.* (2012) based on nuclear and plastid markers and considering different outgroups. Although the aims of our DNA barcoding approach did not include the clarification of phylogenetic issues, our molecular results further support the clear genetic distinction of *Thymbra capitata* (L.) Cav. (= *Thymus capitatus*) from the *Thymus* morphospecies in subgenus *Thymus* (Morales, 1986; Vila, 2002; Bartolucci, 2008; Govaerts *et al*., 2011).

The lower discrimination ability of DNA barcoding in the *Thymus* morphospecies (and intraspecific taxa) involved in this study agrees with the results obtained for different TCGs, such as *Euphrasia* L. (Ennos *et al*., 2005), *Crocus* L. (Seberg & Petersen, 2009), *Carex* L. (Starr *et al*., 2009) and *Dactylorhiza* Neck. ex Nevski (Ennos *et al*., 2005). In these studies, DNA barcoding could not resolve the relationships among closely related taxa resulting from recurrent ecotypic origins or arising through polyploidization or hybridization (Hollingsworth *et al*., 2011). In several cases, the core barcode markers and additional markers did not show any polymorphisms amongst most of the species of the TCGs, such as in the cases of some genera of Meliaceae (Muellner, Schaefer & Lahaye, 2011) and Lamiaceae (De Mattia *et al*., 2011). However, this is not true for *Thymus*, in which consistent intraspecific variability and a certain degree of haplotype diversity were detected using *matK* and *trnH-psbA*. The complex genetic structure observed in this genus could result from a high DNA mutation rate that characterizes *Thymus* populations, combined with interspecific hybridization and polyploidization events (Jalas & Kaleva, 1967; Morales, 1995; Stahl-Biskup & Sáez, 2002; Mahdavi & Karimzadeh, 2010). Another source of variability could also be caused by the conservation of ancestral polymorphisms or by recent speciation events among the morphospecies investigated here (Bräuchler *et al*., 2010).

Hollingsworth *et al.* (2011) have suggested that, for DNA barcoding to work successfully, it requires sufficient time since speciation for mutation to lead to a set of genetic characters grouping conspecific individuals together, separate from other species of the same genus. Although the *Thymus* samples in our study showed clear genetic mutations in the tested DNA barcode regions, sequence variation was not always shared among all the individuals of the same morphospecies. Thus, the genetic diversity showed by the three tested markers does not translate into a ‘barcoding gap’ situation or a species-specific barcode.

There are two possible explanations for these results. It could be that the analysed *Thymus* spp. are morphologically distinct, but genetic isolation between populations of each species has led to a reduced gene flow, so that mutations have not become fixed across the species. This situation can lead to the development of several population haplotypes not shared at the species level. This hypothesis is also
Figure 2. See caption on next page.
supported by population-level molecular analyses of different species of this genus, such as *Thymus herba-barona* Loesel., which has strong among-population differentiation (Molins et al., 2011), and *T. quinquecostatus* Čelak. (Quan et al., 2009), *T. vulgaris* (Belhassen et al., 1993; Tarayre et al., 1997) and *T. loscosii* Willk. (López-Pujol et al., 2004), in which the genetic diversity is consistent at the population level. If this is the case, our data suggest that the DNA mutation rate in the barcode regions is not the only key element to be required to distinguish species using DNA barcoding. When high intraspecific genetic diversity is accompanied by low gene flow among populations, it makes the DNA barcoding approach ineffective.

Alternatively, it could be that the morphological variation used for the delimitation of *Thymus* morphospecies does not reflect real species boundaries. Such a situation has been discussed by Ennos et al. (2005), who suggested that TCGs are sometimes represented by artificial entities identified on the basis of a few, weak morphological traits. This hypothesis is indeed supported by a large number of critical revisions of *Thymus*, with controversial discussion about the values of morphological characters (Morales, 1997; Diklic & Vasic, 2000; Aytas, 2003, 2006; Gomes Pinto et al., 2006; Ložiene, 2006; Blanco Salas, Vazquez Pardo & Ruiz Tellez, 2007; Dentant, 2007; Riera, Guemes & Rossello, 2007; Bartolucci, 2010; Molins et al., 2011; Bartolucci & Peruzzi, in press).

This scenario explains the sharing of the same haplotypes among different morphospecies of *Thymus* and the failure of DNA barcoding as a tool for species identification. As suggested by Blaxter & Floyd (2003), DNA markers can be useful in systems in which species limits are either subtle or cryptic, but nonetheless clear-cut. This is particularly true for DNA barcoding (Hebert et al., 2003), which cannot distinguish any molecular group if clear taxonomic boundaries are not determined.

Based on these results, we conclude that, before the application of DNA barcoding to elucidate the taxonomy of TCGs, an objective and quantitative analysis of taxonomic characters is required to identify natural units at the species level (Rieseberg et al., 2006; Jacobs et al., 2011). In this context, population genetic markers [e.g. amplified fragment length polymorphisms (AFLPs) or microsatellites or single nucleotide polymorphisms (SNPs)] should be used to investigate the population/species structuring in *Thymus* to verify classification problems, as shown in *Solanum* section *Petota*, in which almost half of the morphologically recognized taxa were found to be not genetically supported (Jacobs et al., 2011; see also Ovchinnikova et al., 2011). We believe that this approach could be useful in several genera of Lamioceae, in which relationships among close taxa are still unclear and boundaries among species are sometimes weak.

**ACKNOWLEDGEMENTS**

This work was supported by the Ministero dell’istruzione, dell’università e della Ricerca Italiano grant 2008BZYAH with the project entitled, ‘Tassonomia integrata per lo studio della biodiversità vegetale: DNA barcoding e analisi morfologiche’, and from RAS research grant cofinanced by PO Sardegna FSE 2007-2013 L.R.7/2007 with the project entitled, ‘Tassonomia integrata (morfologia e DNA barcoding) per la salvaguardia della biodiversità vegetale della Sardegna’. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**REFERENCES**


Darwin CF. 1877. \textit{The different forms of flowers on plants of the same species}. London: J. Murray.


Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Figure S1.** Neighbor-joining (NJ) tree based on the *trnH-psbA* dataset generated using MEGA version 5 [Tamura et al., 2011 – options: tree inference method, neighbor-joining; phylogeny test and options, bootstrap (100 replicates); gaps/missing data, pairwise deletion; codon positions, 1st + 2nd + 3rd + non-coding; substitution model, Kimura two-parameter (K2P); substitutions to include, transitions + transversions; pattern among lineages, same (homogeneous); rates among sites, uniform rates]. Bootstrap values of < 75% are not shown.

**Appendix S1.** Voucher and GenBank accession numbers (*rbcL, matK, trnH-psbA*).