Selection of DNA barcoding loci for *Nepeta deflersiana* Schweinf. ex Hedge from chloroplast and nuclear DNA genomes

F. Al-Qurainy, S. Khan, M. Nadeem, M. Tarroum and A.R.Z. Gaafar

Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia

Corresponding author: S. Khan
E-mail: salimkhan17@yahoo.co.in

Received July 17, 2013
Accepted November 28, 2013
Published February 21, 2014
DOI http://dx.doi.org/10.4238/2014.February.21.3

ABSTRACT. Molecular markers, mainly DNA-based are potential tools for DNA barcoding and phylogenetic study. The plant species belonging to the *Nepeta* genus have important medicinal value because of the presence of nepetalactones, and they have been used to treat human diseases. We amplified nuclear and chloroplast gene loci to develop a DNA barcode and phylogenetic study of *Nepeta deflersiana*. Among the studied loci, *psbA-trnH* and *rps16* showed less identity within the genus than the other loci using the Basic Local Alignment Search Tool of the National Center for Biotechnology Information GenBank database. These loci can be used for the development of a DNA barcode to identify and preserve the identity of this species. We also constructed the phylogram of *N. deflersiana* and other *Nepeta* species retrieved from the GenBank database (nonredundant DNA-internal transcribed spacer). *N. deflersiana* was placed in the same clade as *N. insaurica* with a 99% bootstrap value.

Key words: DNA barcode; Endemic plant; Chloroplast loci; Nepetalactone
INTRODUCTION

Nepeta (family: Lamiaceae) is a large genus that is composed of about 250-300 annual and perennial species (Tzakou et al., 2000; Miceli et al., 2005). The species that belong to this genus, known as catmint, are widely distributed in different areas including the Canary Islands, Eurasia, North Africa, and North and Central America (Dabiri and Sefidkon, 2003). This species is native to Europe, Asia, and Africa, and it has high species diversity mainly in the Mediterranean region and in mainland China (Mabberley, 1997). There are two species of the genus Nepeta in the Kingdom of Saudi Arabia. Nepeta deflersiana is found in Saudi Arabia and Yemen, whereas N. sheilae is found only in the northern Hijaz mountains of Saudi Arabia (Chaudhary, 2001).

The secondary metabolites produced in medicinal plants under biotic and abiotic conditions have many applications in the treatment of various human diseases. However, the medicinal value of this genus is due to the presence of essential oils and flavonoids (Jamzad et al., 2003), which have potential roles in the treatment of human diseases. Some Nepeta species contain cyclopentanoid iridoid derivatives such as monoterpenes, sesquiterpenes, and nepetalactones (McElvain et al., 1941; Javidnia et al., 2005; Letchamo et al., 2005). N. grandiflora and N. clarkei produce essential oils that contain (S)-germacrene D compound (Birkett et al., 2010) and other compounds. The chemical compounds present in the flowering tips of Nepeta spp and areal parts of N. persica both have sedative properties (Rabbani et al., 2008). Several species of Nepeta are extensively used in folk medicine as an antispasmodic, antiseptic, antiutisive, antiasthmatic, anti-inflammatory, diuretic, and expectorant (Bezanger-Beauquesnes et al., 1990; Bourrel et al., 1993; Miceli et al., 2005; Hussain et al., 2008). Some species of Nepeta have antibacterial, antioxidant, antiviral, anticandidal, angiotensin-converting enzyme-inhibitory activity, and fungicidal activities (Bourrel et al., 1993; Sattar et al., 1995; Işcan et al., 2011; Bisht et al., 2012; Salehi et al., 2012; Tundis et al., 2013). The essential oil of N. deflersiana showed antimicrobial and antioxidant activities because of the presence of monoterpenes (Mothana, 2012).

N. deflersiana is a woody aromatic perennial herb with many erect, tall stems from the base that can be 50-80 cm. Leaves are short petiolate, lamina triangular-ovate up to 4 x 1-3 cm, regularly crenate-dentate, gray-tomentose beneath, truncate-cordate at base, obtuse, or acute. Verticillasters are in lax spikes, floral leaves are leaf-like below, and they gradually become smaller toward the top. Bracteoles are setaceous, blue, and smaller than calyces and with white floccose. The calyx is about 6.5 mm long with white floccose, blue tubular, ± expanded and curved downward at mouth, and triangular and acute teeth. Corolla is blue, about 10 mm long, with a tube extended above the calyx; the upper lip is bilobed, and the lower lip is much longer than the upper and is spread. Nutlets are about 1.5 x 1 mm, dark brown, and minutely tuberculate (Chaudhary, 2001). However, these morphological characteristics are not sufficient for this species identification and more reliable molecular markers need to be developed.

Surface flavonoids were shown to provide useful taxonomic characteristics at various levels of classification (Tomas-Barberan and Gil, 1992). The external flavonoids were used to determine phylogenetic relationships in Nepeta at species and generic levels (Jamzad et al., 2003). Nepeta species were identified based on morphological markers (Bentham, 1832, 1848; Briquet, 1896). However, the study based on these markers, which
are very variable and do not reflect phylogenetic relationships among the species, and some ambiguities among these classifications are mainly the consequence of the paucity of reliable morphological characteristics. Because of the difficulty of using classical morphological and anatomical characteristics to establish relationships among the species of *Nepeta*, chemical and molecular characteristics were investigated to assess their potential value in phylogeny reconstruction (Jamzad, 2001).

Some gene loci of the chloroplast genome have been used to develop a DNA barcode based on their amplification rate and reproducibility. However, some loci have usefulness for identification at the family level rather than the species level. The chloroplast loci including *matK* and *trnH-psbA* are the most promising choices for the development of the DNA barcode. The *matK* locus is used for identification at the family, genus, and species levels. The variable loci of the chloroplast genome provide potential markers for the development of the DNA barcode. Therefore, it is necessary to determine the most variable loci in the chloroplast genome that can distinguish the closely related plant species. These genes, whose evolutionary rate and variability are high, are preferred for phylogenetic studies and DNA barcoding. In the chloroplast genome, some regions are mutationally active (Dong et al., 2012) and create problems in phylogeny reconstruction because of recombination and sequence convergence (Müller et al., 2006).

Many morphological characteristics in *Nepeta* are variable. Some of these, such as indumentum, leaf shape and size, and calyx and corolla characteristics, can vary among closely related species (Hedge and Lamond, 1968). As a result, the diagnostic use of such characteristics above the species level is problematic. Nutlet surface micromorphology was used for the taxonomic study of the *Nepeta* genus (Kaya and Dirmenci, 2008). The identification of the Lamiaceae family is very difficult based on morphological characteristics; however, chloroplast DNA markers can be used as discriminating DNA markers for this family (Theodoridis et al., 2012). However, the hybridization and introgression in the *Nepeta* genus is frequent in the natural habitat together with substantial age- or habitat-linked variation, making *Nepeta* a particularly complex genus.

**MATERIAL AND METHODS**

**Plant collection**

*N. deflersiana* was collected from the Shafa region, Taif, Saudi Arabia, and identified by a taxonomist based on morphological characteristics at King Saud University, Saudi Arabia (Figure 1).

**Chemicals and reagent preparation**

Extraction buffer consisted of 3% cetyltrimethylammonium bromide (CTAB) (w/v), 100 mM Tris-HCl, pH 8, 2 M NaCl, 25 mM ethylenediaminetetraacetic acid (EDTA), pH 8, 3% β-mercaptoethanol (v/v), and 3% polyvinylpyrrolidone (w/v) and was prepared fresh. Other reagents included 24:1 chloroform:isoamyl alcohol (v/v), Tris-EDTA, 10 mM Tris-HCl, pH 8, 1 mM EDTA, isopropanol, and 70% ethanol.
DNA barcoding of *Nepeta deflersiana*

**DNA extraction and sequence data generation**

The genomic DNA was isolated from the dried leaves (dried on silica gel) using the modified CTAB method (Khan et al., 2007). The purified genomic DNA was used in polymerase chain reaction (PCR) for the amplification of the nonredundant internal transcribed spacer (nr-ITS) locus using universal primers. The PCR bead (GE Healthcare, Spain) was employed for the amplification of the nr-ITS locus. The single reaction consists of 20 μL deionized sterile water, 25 ng DNA per reaction volume, and 10 pM of each forward and reverse primer. After mixing the PCR components, the reaction was performed in a Techne Thermal Cycler (UK) using the following conditions: denaturation at 94°C for 5 min; 40 cycles of denaturation at 94°C for 1 min, annealing at 49°C for 1 min, and extension at 72°C for 1 min; and a final extension step of 72°C for 5 min. The PCR products were purified using SolGent PCR Purification Kit prior to sequencing.

The amplified PCR product was directly sequenced at Macrogen Inc., South Korea, using dye terminator chemistry.

**Sequence alignment and phylogenetic analyses**

The sequenced chloroplast loci and nr-ITS were used in Basic Local Alignment Search Tool (BLAST) ([http://blast.ncbi.nlm.nih.gov/Blast.cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi)) to confirm the sequences of *N. deflersiana*. The related nr-ITS sequences of *Nepeta* species (32) were retrieved from the GenBank database. *Phlomis anisodonta* and *Scutellaria baicalensis* were used as outgroups for the phylogenetic reconstruction. The ITS sequence (682 bp) that was generated from *N. deflersiana* in this study has been submitted to the GenBank database (accession No. KF765442).

**RESULTS AND DISCUSSION**

There are numerous molecular markers that are available in the literature to identify and authenticate plant species, but they all have some limitations. Alone, one marker is not sufficient.
for the above study, and sequencing of more gene loci from the chloroplast genome is needed. Sequences from multiple genes give better results than a tree constructed from a single gene. However, it is possible that not all chloroplast loci of one genus will be present in all genera, and some of these loci may be very variable while others are not (Dong et al., 2012).

*N. deflersiana* is a medicinal and endemic plant of Saudi Arabia. Loci in the chloroplast genome are being used to develop a DNA barcode, and not all loci are successful for every plant. One gene locus gives better discrimination in one plant while it does not allow discrimination in another genus. We amplified various loci including *rps-16, rbcL, rpoB, rpoC1*, and *psbA-trnH* of the chloroplast genome from *N. deflersiana* to develop a DNA barcode. The sequences of all loci were used as queries in BLAST at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) to observe the sequence similarity in the same and different genera of the same family. All amplified sequences had variations in the taxa available from the GenBank database: nr-ITS of *N. deflersiana* was 97% similar to the same genus; *rbcL* and *rps16* were 99 and 97% similar to the same genus, respectively; *rpoC* had 99% similarity to the same genus and another genus of the same family; and *psbA-trnH* had 92% similarity in the same genus. The *rps16* and *psbA-trnH* loci showed less similarity to the same genus; hence, we could use these loci to develop a DNA barcode for *N. deflersiana*. In angiosperms, the chloroplast intron of *rps16* was used for identification at the genus, species, or even population levels (Nie et al., 2010; Xie et al., 2010; Xu et al., 2010: Khan et al., 2013). Similarly, *Meconopsis impedita* and *M. racemosa* were authenticated using the *rps16* intron region (Li and Dao, 2011). However, the sequences of chloroplast loci from the *Nepeta* genus were limited in the GenBank database; hence, molecular studies such as phylogeny reconstruction were not possible. We also studied the phylogeny of *N. deflersiana* by retrieving the nr-ITS sequences of the same genus from the NCBI (http://www.ncbi.nlm.nih.gov/). The exact size of nr-ITS and chloroplast gene sequences was determined after sequence alignment using ClustalX version 1.81 (Thompson et al., 1997). All sequences were edited and assembled using BIOEDIT v. 7.0.9.0 (Hall, 1999). Some sequences were edited manually with adjustments as needed. All characters were treated as equally weighted and unordered, and the gaps were treated as missing data. The branch support was evaluated using 1000 bootstrap replicates with random-sequence addition, equal weighting, and tree bisection and reconnection branch swapping, holding one tree at each replicate. There were a total of 540 positions in the final dataset. The bootstrap values that are greater than 50% are shown above the branches. The erroneous sequences were deleted from the final analysis after the sequence alignment and further processed for phylogenetic reconstruction. *N. deflersiana* was placed in a clade with *N. insaurica* as shown in the phylogenetic tree (Figure 2A, B and C), which may be assumed to be very close to this species. The phylogenetic tree was reconstructed using MEGA 5 (Tamura et al., 2011) with three methods: maximum likelihood (Jones et al., 1992), neighbor-joining (Saïtou and Nei, 1987), and maximum parsimony (Eck and Dayhoff, 1966). The tree that was constructed with these three methods was reproducible, and the phylogenetic relationship between *N. deflersiana* and *N. insaurica* was the same. The development of a DNA barcode is very important to identify plants and preserve their identity. Assessing genetic diversity is also very important for plant species that are endemic, rare, and endangered because it aids plant conservation. We are trying to collect populations of *N. deflersiana* to evaluate genetic diversity, which will also provide information about their genetic base.
Figure 2. A. Phylogram constructed using maximum likelihood method. B. Phylogram constructed using the maximum parsimony method. C. Phylogram constructed using the neighbor-joining method.

ACKNOWLEDGMENTS

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding the study through the research group project #RGP-VPP-014.

REFERENCES


DNA barcoding of *Nepeta deflersiana*


