

## MOLECULAR IDENTIFICATION OF *LA V E N D U L A D E N T A T A* L., *M E N T H A L O N G I F O L I A* (L.) H U D S. A N D *M E N T H A* × *P I P E R I T A* L. B Y D N A B A R C O D E S

SHAWKAT MAHMOUD AHMED<sup>1</sup>

*Biology Department, Faculty of Education, Ain Shams University, Cairo, Egypt*

*Keywords: Lavendula; Mentha; ITS; ITS2; rbcL; matK; trnH.*

### Abstract

Five DNA barcodes were tested for identification and discrimination of *Lavendula dentata* L., *Mentha longifolia* (L.) Huds. and *Mentha* × *piperita* L. New DNA barcodes have been registered for *L. dentata* from Taif, Saudi Arabia. The separate clading of *L. dentata* and *M. longifolia* through the phylogenetic analyses proved their endemism to Saudi Arabia. The phylogenetic trees revealed from the ITS2, *matK* and *trnH* data demonstrated that all *Mentha* species formed monophyletic clusters except hybrid *M. × piperita* from Taif which formed separate clades distinguishing it from the two parents; *M. aquatica* L. and *M. spicata* L. DNA barcoding could be considered as a good approach for distinguishing and identifying the mint plants, though it was not possible to confirm the relationship between hybrids and their putative parents.

### Introduction

The family Lamiaceae comprising about 7,173 species under 236 genera possesses medicinal and aromatic herbs such as lavender, basil, mint, rosemary and thyme, that have been widely utilized as teas, spices, traditional medicines or raw material for the food and pharmaceutical industries (Theodoridis *et al.*, 2012).

*Lavendula dentata* is one of five naturally growing lavender species in Saudi Arabia that has been known as the main center of origin of the genus (Miller, 1985). Lavender species as medicinal plants, are distributed in highlands of Albaha, Asir, and Taif and are exploited for the production of high-quality lavender honeys. Locally known as Habak, Al-Madinah mint or wild mint, *Mentha longifolia* and peppermint, *Mentha* × *piperita* are present in the spontaneous flora of Saudi Arabia but also under cultivation. Traditionally, they have been used as medicinal agents to treat colds, cough, headaches, asthma and digestive disorders. Recent studies proved the antiviral, antimicrobial, antioxidant, anti-inflammatory and anticancer characteristics as therapeutic activities for the extracts derived from *Mentha* species (Anwar *et al.*, 2017). Hybridization and polyploidy play an important role in the speciation of the members belonging to genus *Mentha* such as *M. × piperita* that is considered as a hybrid of the two mints; *M. spicata* and *M. aquatica* (Mogosan *et al.*, 2017) making them good targets for molecular studies.

Various studies have been performed to identify and classify species of Lamiaceae collected from Saudi Arabia based on anatomical and cytological studies (Abdel Khalik, 2016) and biochemical analyses (Kasem, 2016), however, very little is known about DNA barcoding information. DNA sequences for the species under study will be compared in a database against retrieved sequences of identified individuals from the GenBank. If the query sequence matches with one in the database, this will help in identification, discrimination or gaining a new barcodes for these species (Hajibabaei *et al.*, 2007). Therefore, the objectives of this research include:

---

<sup>1</sup>Present address: Biology Department, Faculty of Science, Ta'if University, Ta'if, 5700, Saudi Arabia.  
Email: shamahmoh@gmail.com

i) utility of specific DNA regions, two nuclear internal transcribed spacers (ITS and ITS2) and the plastid DNA regions (*rbcL*, *matK* and *trnH*) for developing DNA barcodes and subsequently identification for the three species; *Lavendula dentata*, *Mentha longifolia* and *M. × piperita* occurred in Taif highlands of Saudi Arabia; ii) discriminating between species under study and those retrieved from the GenBank and iii) exploring the interspecific variation between *M. longifolia* and *M. × piperita*.

## Materials and Methods

### Plant materials

Two wild species, namely *Lavendula dentata* and *Mentha longifolia*, and the hybrid species, *M. × piperita* belonging to family Lamiaceae were collected from Taif highlands, Saudi Arabia. Species identification was confirmed following Colletete (1999).

### DNA extraction and amplification

DNA of fresh young leaves was extracted using CTAB method as described by Doyle and Doyle (1987). The purified DNA was amplified for ITS, ITS2, *rbcL*, *matK* and *trnH* barcodes using universal primers.

### PCR sequencing

The PCR products of the three Lamiaceae species for the five DNA barcodes were purified and sequenced at Macrogen Inc., South Korea. All sequences of the three species generated in this research were deposited in GenBank (accession numbers are listed in Table 1).

### Sequences alignment and phylogenetic analyses

The sequences of ITS, ITS2, *matK*, *rbcL* and *trnH* of *L. dentata*, *M. longifolia* and *M. × piperita* were subjected to BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) to confirm them from the other related Lamiaceae species existing in the GenBank database. Sequence alignments were performed by MUSCLE algorithm (Edgar, 2004; Tamura *et al.*, 2013). The equality of evolutionary rate parameters between sequences of the three species under study and the retrieved species from GenBank were calculated by Tajima's relative rate test (Tajima, 1993). Nucleotide substitution rates and Transition/Transversion bias (R) were estimated using Maximum Likelihood method. The phylogenetic trees were constructed by the Maximum likelihood bootstrap (MLB) analysis. A total of 1,000 bootstrap replicates were performed. The software of MEGA6 was used for all operations (Tamura *et al.*, 2013).

**Table 1. Accession numbers in GenBank of sequences of *Lavendula dentata*, *Mentha longifolia* and *M. × piperita* generated in this study.**

Taxa	ITS	ITS2	<i>matK</i>	<i>rbcL</i>	<i>trnH</i>
<i>L. dentata</i>	LC373552.1	LC373553.1	-	LC373554.1	LC373555.1
<i>M. longifolia</i>	-	LC378378.1	-	LC378379.1	-
<i>M. × piperita</i>	-	LC374287.1	LC374288.1	LC374289.1	LC374290.1

## Results and Discussion

### Identification of *Lavendula dentata*

Sequences of *L. dentata* for ITS, ITS2, *rbcL* and *trnH* barcoding loci were submitted to BLAST at the GenBank database, however, any sequence of *L. dentata* was detected in the database, thus the present study succeeded in registering new DNA barcodes for *L. dentata* from

Taif. Sequences of species belonging to the genus *Lavendula* showing high similarities to those of *L. dentata* were retrieved for the statistical analyses. ITS showed the highest sequence length (775 bp) followed by *rbcL* (537 bp), ITS2 (358 bp) and *trnH* (346 bp), whereas, the variable sites percentage after alignment was higher in *trnH* (24%) than those of ITS, ITS2 and *rbcL*. The GC ratios scored in loci ITS and ITS2 (60.1 and 65.9) was found greater than those of *rbcL* and *trnH* (Table 2). In comparison with the retrieved *Lavendula* species, the rates of transitions to transversions showed notable substitution changes in the sequences of *L. dentata* (Table 2). Transitions generally occurred more than transversions. Transition/transversion bias (R) was found relatively high and ranged from 1.19 to 2.82 demonstrating a molecular evolution within *Lavendula* genome. This putative evolution in *L. dentata* was confirmed through tests of Tajima relative evolutionary rate that displayed an accelerated rates of evolution ( $P$ -values  $<0.05$ ) for all loci under study. The results revealed that ITS, ITS2, *rbcL* and *trnH* have sufficient efficiency in sequence quality as well as in species identification across the genome of the genus *Lavendula*. For further identification of *L. dentata*, sequences of the four loci were used to reconstruct four phylogenetic trees (Fig. 1). Except the tree revealed from ITS2, the separate clustering of *L. dentata* in the phylogenetic trees of ITS, *rbcL* and *trnH* proved its endemism to Saudi Arabia. The development of different DNA barcodes is better than single locus for more accurate results (Khan *et al.*, 2013). The identification of species within a community through DNA barcodes contributes to the construction of the barcode library for terrestrial plants (Burgess *et al.*, 2011).

**Table 2. Statistics derived from the sequencing, alignment and BLAST processes for all loci employed in the present investigation.**

Parameters	Loci				
	ITS	ITS2	<i>rbcL</i>	<i>matK</i>	<i>trnH</i>
% Variable sites after alignment for <i>Lavendula dentata</i>	0.01	0.06	0.01	-	0.24
% Variable sites after alignment for <i>Mentha longifolia</i>	0.11	0.12	0.04	-	-
% Variable sites after alignment for <i>M. × piperita</i>	-	0.12	0.04	0.29	0.31
Sequence length of <i>L. dentata</i>	775	358	537	-	346
Sequence length of <i>M. longifolia</i>	362	347	528	-	-
Sequence length of <i>M. × piperita</i>	-	349	540	810	403
GC ratio in <i>L. dentata</i>	60.1	65.9	43.6	-	28.6
GC ratio in <i>M. longifolia</i>	51.1	66.6	44.1	-	-
GC ratio in <i>M. × piperita</i>	-	67.9	43.7	34.8	31.2
Number of the retrieved <i>Lavendula</i> species from the GenBank	3	2	4	-	2
Number of the retrieved <i>Mentha</i> species from the GenBank	4	9	9	11	14

#### Identification of *Mentha longifolia*

Sequences of ITS, ITS2 and *rbcL* were used to identify *M. longifolia*. ITS2 recorded the lowest sequence length, whereas, the variable sites (%) and GC ratio of it were greater than those of ITS and *rbcL* (Table 2). Transitions were found to be more than transversions leading to substitution changes in the sequences of *M. longifolia* (Table 3). An evolution within *M. longifolia* genome was noticed through the high transition/transversion bias (R) that ranged from 3.51 in *rbcL* to 1.81 in ITS2. Except data of *rbcL*, Tajima relative evolutionary rate displayed an accelerated rates of evolution ( $P$ -values  $<0.05$ ) in *M. longifolia* (Table 4). Sequences of *M. longifolia* for ITS, ITS2 and *rbcL* that submitted to BLAST at the GenBank retrieved 4, 9 and 9 *Mentha* species, respectively (Table 2). *M. longifolia* and the retrieved *Mentha* species reconstructed three phylogenetic trees (Fig. 2) which revealed that *M. longifolia* was represented in separate clade demonstrating variability between it and other *Mentha* species, and proved its

endemism to Saudi Arabia. Similar result was obtained by Khan *et al.* (2013) in *Senecio asirensis* using nrDNA ITS.

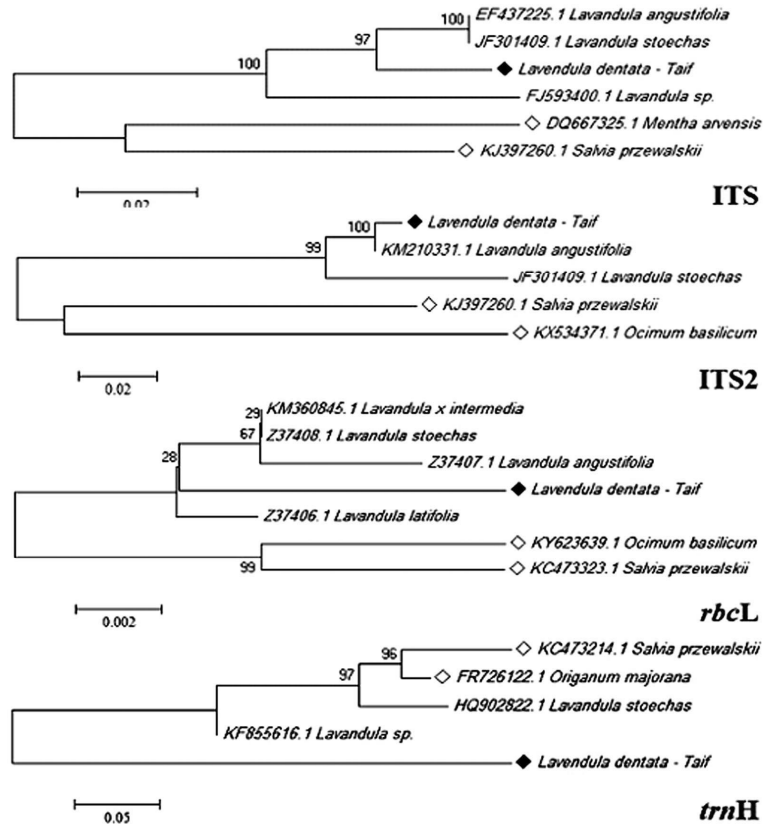


Fig. 1. Phylogenetic trees of *Lavendula dentata* and the retrieved species based on four loci. (◇) refers to the outgroup.

#### Identification of *M. × piperita*

Sequences of ITS2, *rbcL*, *matK* and *trnH* were used to identify *M. × piperita*. As found in *M. longifolia*, ITS2 showed the lowest sequence length (349 bp) and the highest GC ratio (67.9). Whereas, the percentage of variable sites (31%) was detected in *trnH* locus (Table 2). An obvious evolution was also observed within *M. × piperita* genome through the high transition/transversion bias (R) that ranged from 0.79 in *matK* to 7.01 in *rbcL* (Table 3). The previous result was supported by Tajima relative evolutionary rate that displayed an accelerated rates of evolution ( $P$ -values <0.05) in *M. × piperita* except that of *rbcL* (Table 4). The null hypothesis of equal evolution rates between *M. × piperita* from Taif and its ancestors; *M. spicata* and *M. aquatica* from one hand, and the retrieved *M. × piperita* from the other hand, was rejected because the  $P$ -values were lower than 0.05 in ITS2, *matK* and *trnH* revealing the accelerated evolutionary rate of *M. × piperita* from Taif and subsequently reflecting the variance among them. *M. × piperita* and the retrieved *Mentha* species from the GenBank library were analyzed to form four phylogenetic trees (Fig. 2). The phylogenetic trees from the ITS2, *matK* and *trnH* data demonstrated that all the



*Mentha* species formed monophyletic clusters except the hybrid *M. × piperita* from Taif which formed separate clades. The differences between *M. × piperita* under study and the other retrieved *Mentha* species could be explained due to an evolutionary process.

Little divergence in *rbcL* tree (Fig. 2) and the acceptance of the null hypothesis of equal evolutionary rates among *Mentha* species through *rbcL* data (Table 4) could be due to the symmetry in *rbcL* sequence of *Mentha* species. Kshirsagar *et al.* (2015) reported the same limitation of *rbcL* gene in closely related species of the two genera *Ardisia* Sw. and *Swertia* L. These results were in accordance with those of Theodoridis *et al.* (2012) who showed that *matK* and *trnH* were more useful in discriminating Lamiaceae species than *rbcL*. It was noticed that ITS2, *rbcL*, *matK* and *trnH* distinguished *M. × piperita* from the two parents, *M. aquatica* and *M. spicata* through the phylogenetic trees. These genetic differences might be due to most commercial

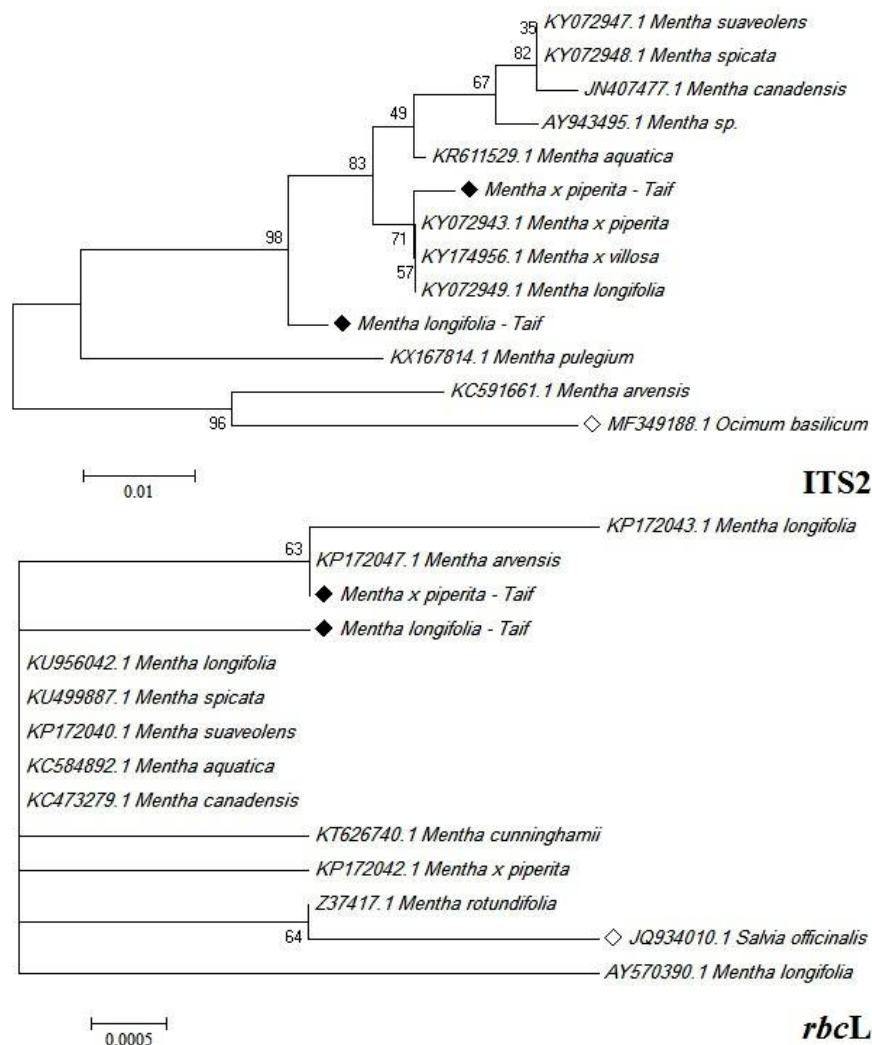


Fig. 2. Phylogenetic tree diverges between *Mentha longifolia* and *M. × piperita* based on ITS2 and *rbcL* sequences. (◇) refers to the outgroup.

hybrids, i.e. *M. × piperita* is sterile or subfertile, therefore, crossing with parental or nonparental species is expected. They may also form complex hybrid populations through vegetative propagation and polyploidy. These possibilities lead to great genetic diversity and subsequently to several taxonomic problems (De Mattia *et al.*, 2011).

**Table 4. Tajima relative rate tests of loci for *L. dentata*, *M. longifolia* and *M. × piperita*.**

Loci	Outgroup	Testing group		RI	RD	RA	RB	$\chi^2$	P value
		(A)	(B)						
ITS	<i>L. angustifolia</i>	<i>L. dentata</i> -Taif	<i>L. stoechas</i>	512	0	36	0	36.0	<0.05
ITS2	<i>L. angustifolia</i>	<i>L. dentata</i> -Taif	<i>L. stoechas</i>	280	0	2	17	11.8	<0.05
<i>rbcL</i>	<i>L. angustifolia</i>	<i>L. dentata</i> -Taif	<i>L. stoechas</i>	529	0	5	0	5.00	<0.05
<i>trnH</i>	<i>L. angustifolia</i>	<i>L. dentata</i> -Taif	<i>L. stoechas</i>	126	6	60	16	25.5	<0.05
ITS	<i>M. suaveolens</i>	<i>M. longifolia</i> -Taif	<i>M. spicata</i>	70	0	33	0	33.0	<0.05
	<i>M. spicata</i>	<i>M. longifolia</i> -Taif	<i>M. suaveolens</i>	70	0	33	1	30.12	<0.05
ITS2	<i>M. × piperita</i>	<i>M. longifolia</i> -Taif	<i>M. longifolia</i>	297	0	9	0	9.00	<0.05
	<i>M. longifolia</i>	<i>M. longifolia</i> -Taif	<i>M. × piperita</i>	297	0	9	0	9.00	<0.05
<i>rbcL</i>	<i>M. × piperita</i>	<i>M. longifolia</i> -Taif	<i>M. longifolia</i>	524	0	1	0	1.00	>0.05
	<i>M. longifolia</i>	<i>M. longifolia</i> -Taif	<i>M. × piperita</i>	524	0	1	1	0.00	>0.05
ITS2	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. aquatica</i>	272	0	7	0	7.00	<0.05
	<i>M. aquatica</i>	<i>M. × piperita</i> -Taif	<i>M. spicata</i>	272	0	7	0	7.00	<0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. × piperita</i>	296	0	4	0	4.00	<0.05
<i>rbcL</i>	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. aquatica</i>	525	0	1	0	1.00	>0.05
	<i>M. aquatica</i>	<i>M. × piperita</i> -Taif	<i>M. spicata</i>	525	0	1	0	1.00	>0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. × piperita</i>	524	0	1	1	0.00	>0.05
<i>matK</i>	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. aquatica</i>	411	1	263	0	263	<0.05
	<i>M. aquatica</i>	<i>M. × piperita</i> -Taif	<i>M. spicata</i>	411	1	263	0	263	<0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. × piperita</i>	400	1	250	0	250	<0.05
<i>trnH</i>	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. aquatica</i>	177	0	114	2	108.1	<0.05
	<i>M. aquatica</i>	<i>M. × piperita</i> -Taif	<i>M. spicata</i>	177	0	114	2	108.1	<0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. × piperita</i>	178	1	114	1	111.0	<0.05
ITS2	<i>M. spicata</i>	<i>M. longifolia</i> -Taif	<i>M. × piperita</i> -Taif	320	0	3	6	1.0	>0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. longifolia</i> -Taif	320	0	3	6	1.0	>0.05
<i>rbcL</i>	<i>M. spicata</i>	<i>M. longifolia</i> -Taif	<i>M. × piperita</i> -Taif	524	0	1	1	0.00	>0.05
	<i>M. spicata</i>	<i>M. × piperita</i> -Taif	<i>M. longifolia</i> -Taif	524	0	1	1	0.00	>0.05

The Tajima relative rate test was used to examine the equality of evolutionary rate for *L. dentata*, *M. longifolia* and *M. × piperita* and other relative species with different outgroups.

RI is the identical sites in all three sequences

RD is the divergent sites in all three sequences

RA is the number of unique differences in the sequence A

RB is the number of unique differences in the sequence B

$\chi^2$  test statistic more than 3.841 ( $P < 0.05$ ) indicates accelerated evolution

P value greater than 0.05 is often used to accept the null hypothesis of equal rates between lineages

#### Discrimination between *M. longifolia* and *M. × piperita*

Sequences of ITS2 and *rbcL* were used to discriminate between *M. longifolia* and *M. × piperita*. Statistics in Table 1 revealed slight differences between them. Results of mean nucleotide substitution rates, transition/transversion bias (R) and Tajima relative evolutionary rate were similar in these two taxa (Tables 3 & 4). ITS2 and *rbcL* trees were used to assess genetic divergences between *M. longifolia* and *M. × piperita*. A suitable divergence was detected in the

two phylogenetic trees displaying the efficacy of the two barcodes in distinguishing between them through the Maximum Likelihood method (Fig. 2). Thakur *et al.* (2016) stated that the convenient barcode exhibits large interspecific but little intraspecific divergence and its sequence length must be short enough to be available in a single amplification. This comparability of interspecific sequence variation is a significant aspect for barcoding identification of species in local floras. Establishing a local barcode data will be useful in several ecological applications, such as the reconstruction of community phylogenies, palaeoecological studies of ecosystems and analyzing the diets of human and other animals (Valentini *et al.*, 2009). Depending on these data, the DNA barcoding could be considered as a good approach for distinguishing and identifying the mint plants, however, it was not possible to confirm the relationship between hybrids and their putative parents.

Finally, it could be concluded that the identification and discrimination of *L. dentata*, *M. longifolia* and *M. × piperita* were necessary and valuable for their great economic importance. ITS, *matK* and *trnH* were found to be more effective barcodes than ITS2 and *rbcL* for the authentication of these species and hybrid. DNA barcoding provided new insight that will contribute to the taxonomy of Lamiaceae taxa around the world and the conservation of the genetic resources of these valuable taxa occurring in Saudi Arabia.

#### Acknowledgement

The technical support from Macrogen Inc., South Korea is gratefully acknowledged.

#### References

- Abdel Khalik, K.N. 2016. A Systematic revision of the genus *Plectranthus* L. (Lamiaceae) in Saudi Arabia based on morphological, palynological and micromorphological characters of trichomes. *Am. J. Plant Sci.* **7**: 1429–1444.
- Anwar, F., Alkharfy, K.M., Rehman, N., Adam, E.H.K. and Gilani, A. 2017. Chemo-geographical variations in the composition of volatiles and the biological attributes of *Mentha longifolia* L. essential oils from Saudi Arabia. *Int. J. Pharmacol.* **13**: 408–424.
- Burgess, K.S., Fazekas, A.J., Kesanakurti, P.R., Graham, S.W., Husband, B.C., Newmaster, S.G., Percy, D.M., Hajibabaei, M. and Barrett, S.C.H. 2011. Discriminating plant species in a local temperate flora using the *rbcL+matK* DNA barcode. *Methods Ecol. Evol.* **2**: 333–340.
- Collenette, S. 1999. Wild Flowers of Saudi Arabia. National Commission for Wildlife Conservation and Development (NCWCD), Riyadh.
- De Mattia, F., Bruni, I., Galimberti, A., Cattaneo, F., Casiraghi, M. and Labra, M. 2011. A comparative study of different DNA barcoding markers for the identification of some members of Lamiaceae. *Food Res. Int.* **44**: 693–702.
- Doyle, J.J. and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**: 11–15.
- Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**(5): 1792–1997.
- Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N. and Hickey, D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* **23**: 167–172.
- Kasem, W.T. 2016. Pollen grains and seed morphology as related to biochemical patterns in five species of genus *Ocimum* L. (Lamiaceae Juss.) of Saudi Arabia. *J. Phytol.* **8**: 42–48.
- Khan, S., Al-Qurainy, F., Nadeem, M. and Tarroum, M. 2013. Selection of chloroplast DNA markers for the development of DNA barcode and reconstruction of phylogeny of *Senecio asirensis* Boulos and J.R.I. wood. *Pak. J. Bot.* **45**(2): 703–710.



- Kshirsagar, P., Umdale, S., Chavan, J. and Gaikwad, N. 2015. Molecular authentication of medicinal plant, *Swertia chirayita* and its adulterant species. Proc. Natl. Acad. Sci. India Sect. B, Biol. **87**(1): 1–7.
- Miller, A.G. 1985. The genus *Lavendula* in Arabia and Tropical NE Africa. Notes Roy. Bot. Gard. Edinburgh **42**(3): 503–528.
- Mogosan, C., Vostinaru, O., Oprean, R., Heghes, C., Filip, L., Balica, G. and Moldovan, R. 2017. A comparative analysis of the chemical composition, anti-inflammatory, and antinociceptive effects of the essential oils from three species of *Mentha* cultivated in Romania. Molecules **22**: 263–274.
- Tajima, F. 1993. Simple methods for testing molecular clock hypothesis. Genetics **135**: 599–607.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. and Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. **30**: 2725–2729.
- Thakur, V., Tiwari, S., Tripathi, N., Tiwari, G. and Sapre, S. 2016. DNA barcoding and phylogenetic analyses of *Mentha* species using *rbcL* sequences. Ann. Phytomed. **5**(1): 59–62.
- Theodoridis, S., Stefanaki, A., Tezcan, M., Aki, C., Kokkini, S. and Vlachonassios, K.E. 2012. DNA barcoding in native plants of the Labiatae (Lamiaceae) family from Chios Island (Greece) and the adjacent Cesme-Karaburun Peninsula (Turkey). Mol. Ecol. Resour. **12**: 620–633.
- Valentini, A., Miquel, C., Nawaz, M., Bellemain, E., Coissac, E., Pompanon, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., Swenson, J. and Taberlet, P. 2009. New perspectives in diet analyses based on DNA barcoding and parallel pyrosequencing: the *tmL* approach. Mol. Ecol. Resour. **9**: 51–60.

(Manuscript received on 14 January 2018; revised on 6 November 2018)