Bangladesh J. Plant Taxon. **21**(1): 33-42, 2014 (June)

APPLICATION OF INTERNAL TRANSCRIBED SPACER OF NUCLEAR RIBOSOMAL DNA FOR IDENTIFICATION OF *ECHINOPS MANDAVILLEI* KIT TAN

FAHAD M.A. AL-HEMAID, M. AJMAL ALI[1], JOONGKU LEE[2], GÁBOR GYULAI[3] AND ARUN K. PANDEY[4]

Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

Keywords: Echinops mandavillei; Asteraceae; ITS; nrDNA; Endemic; Saudi Arabia.

Abstract

The present study explored the use of internal transcribed spacers (ITS) sequences (ITS1-5.8S-ITS2) of nuclear ribosomal DNA (nrDNA) for identification of *Echinops mandavillei* Kit Tan, an endemic species to Saudi Arabia. The sequence similarity search using Basic Local Alignment Search Tool (BLAST) and phylogenetic analyses of the ITS sequence of *E. mandavillei* Kit Tan showed high level of sequence similarity (98%) with *E. glaberrimus* DC. (section *Ritropsis*). The novel primary sequence and the secondary structure of ITS2 of *E. mandavillei* could have a potential use for molecular genotyping.

Introduction

The genus *Echinops* L. belonging to the subtribe Echinopsinae of Cynareae, of the family Asteraceae comprise about 120 species (Vidović, 2011), and distributed in tropical Africa, the Mediterranean basin, temperate regions of Eurasia, Central Asia, Mongolia and North-eastern China, with the maximum number of species occurring in the Caucasus and the Middle East (Susanna and Garcia-Jacas, 2007). The genus received considerable interest for establishing natural groups with infrageneric classification (Sánchez-Jiménez *et al*., 2010). Morphological characters, like the pappus, which is a key taxonomic character of Cynareae, the type and density of indumentum on stems, leaf shapes and phyllaries are considered least significance in dissemination of *Echinops* species (Mozaffarian, 2006; Sánchez-Jiménez *et al*., 2010). In Saudi Arabia, there are nine *Echinops* species, *viz*. *E. abuzinadianus* Chaudhary, *E. erinaceus* Kit Tan, *E. glaberrimus* DC., *E. hystrichoides* Kit Tan, *E. macrochaetus* Fresen., *E. mandavillei* Kit Tan, *E. sheilae* Kit Tan, *E. viscosus* DC. and *E. yemenicus* Kit Tan. Of them, *E. abuzinadianus, E. mandavillei* and *E. sheilae* are endemic to Saudi Arabia, while remaining species have been reported from different geographic locations of Arabian Peninsula. *E. mandavillei* was reported to occur in Dahna, Summan and Nafud sands (Chaudhary, 2000).

The DNA sequence technology provides series of new data for molecular phylogeny and DNA barcoding which has now-a-days changed the paradigm of species identification (Ali and Choudhary, 2011; Ali *et al*., 2014). From the first report of the utility of the internal transcribed spacers (ITS) sequence of nuclear ribosomal DNA (nrDNA) in plants (Baldwin, 1992), it has been

[1]Corresponding author. E-mail: majmalali@rediffmail.com
[2]International Biological Material Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305 806, South Korea
[3]Institute of Genetics and Biotechnology, St. István University, Gödöllo H-2103, Hungary
[4]Department of Botany, University of Delhi, Delhi 110007, India

extensively used to distinguish even very closely related species (Chen *et al*., 2010; Yao *et al.,* 2010). Moreover, in the last two decades, the ITS sequence technology has gained much attention, along with the smartest genes available for the molecular phylogeny and taxonomy (Ali *et al.,* 2013).

The ITS sequence technology has been used for molecular phylogeny of *Echinops* (Garnatje *et al*., 2005), and series of other genera of Cynareae (Susanna *et al*., 1999; Vilatersana *et al*., 2000; Wang *et al*., 2005, 2007; Hidalgo *et al*., 2006); however, these studies did not include systematics of *Echinops* species occurring in Saudi Arabia. Hence, the present study aims to establish molecular signature of *Echinops mandavillei* Kit Tan based on ITS sequence of nrDNA.

Materials and Methods

Plant materials:

The leaf material of *Echinops mandavillei* Kit Tan was collected from herbarium specimen (Saudi Arabia, Al-Nafud, 29.4'N, 39.58'E, 5 May 1985, H.O. Al-Hassan 195) housed at National Herbarium and Genebank, National Agriculture and Animal Resources Research Centre, Riyadh, Saudi Arabia (RIY). The taxonomic identification of specimen was confirmed with the aid of Flora of Saudi Arabia (Chaudhary, 2000).

ITS sequences of 39 species of *Echinops* (Table 1) were retrieved from the GenBank database of NCBI (National Centre for Biotechnology Information; www.ncbi.nlm.nih.gov). *Brachylaena discolor* DC., from the tribe Tarchonantheae Kostel and *Cardopatium corymbosum* (L.) Pers. from the subtribe Cardopatiinae Less. were chosen as outgroups (Table 1) according to previous report based on molecular characters (Susanna *et al*., 2006; Sánchez-Jiménez *et al*., 2010).

Table 1. List of *Echinops* species used in the present study along with accession numbers.

	Taxa	Accession number
Ingroup		
sect. ***Acantholepis*** (Less.) Jaub. & Spach		
1.	*Echinops acantholepis* Jaub. & Spach	AY8262223
sect. ***Chamaechinops*** Bunge		
2.	*E. fastigiatus* Kamelin & Tscherneva	GU116503
3.	*E. humilis* M. Bieb	GU116514
4.	*E. integrifolius* Kar. & Kir.	GU116517
sect. ***Echinops***		
5.	*E. arachniolepis* Rech. f.	GU116486
6	*E. dahuricus* Fisch.	GU116493
7.	*E. freitagii* Rech. f.	GU116504
8.	*E. kotschyi* Boiss.	GU116520
9.	*E. latifolius* Tausch	GU116521
10.	*E. nizvanus* Rech. f.	GU116530
11.	*E. parviflorus* Boiss. & Buhse	GU116533
12.	*E. przewalskyi* Iljin	GU116535
13.	*E. ritrodes* Bunge	GU116539
14.	*E. setifer* Iljin	GU116540
15.	*E. sphaerocephalus* L.	GU116541
16.	*E. spiniger* Iljin	GU116542
17.	*E. transcaucasicus* Iljin	GU116546

Table 1 contd.

Taxa	Accession number
sect. ***Hamolepis*** R. E. Fr.	
18. *E. hoehnelli* Schweinf	GU116506
sect. ***Hololeuce*** Rech. f.	
19. *E. hololeucus* Rech. f.	GU116513
sect. ***Nanechinops*** Bunge	
20. *E. gmelini* Turcz.	GU116510
sect. ***Oligolepis*** Bunge	
21. *E. cephalotes* DC.	GU116487
22. *E. cornigerus* DC.	GU116552
23. *E. echinatus* Roxb.	GU116497
24. *E. ghoranus* Rech. f.	GU116508
25. *E. griffithianus* Boiss.	GU116512
26. *E. ilicifolius* Bunge	GU116516
27. *E. leucographus* Bunge	GU116522
28. *E. lipskyi* Iljin	GU116523
sect. ***Phaeochaete*** Bunge	
29. *E. longifolius* A. Rich	GU116524
sect. ***Psectra*** Endl.	
30. *E. strigosus* L.	AY5386532
sect. ***Ritropsis*** Greuter & Rech. f.	
31. *E. chardinii* Boiss. & Buhse	GU116490
32. *E. dichrous* Boiss. & Hausskn.	GU116495
33. *E. endotrichus* Rech. f.	GU116500
34. *E. gaillardotii* Boiss.	GU116507
35. *E. glaberrimus* DC.	GU116509
36 *E. mandavillei* Kit Tan	KJ187107
37. *E. orientalis* Trautv.	GU116532
38. *E. spinosissimus* Turra	HE687348
39. *E. tenuisectus* Rech. f.	GU116551
sect. ***Terma*** Endl.	
40. *E. exaltatus* Schrad.	GU116501
Outgroup	
41. *Brachylaena discolor* DC.	AY8262363
42. *Cardopatium corymbosum* (L.) Pers.	AY8262383

DNA isolation and amplification:

Genomic DNA was extracted from 10 mg silica gel-dried leaves using the protocol of DNeasy Plant Mini kit (QIAGEN, Valencia, CA, USA). The ITS regions were amplified using the primers ITS1 and ITS4 as described by White *et al*. (1990). Double-stranded polymerase chain reaction (PCR) products were produced through 35 cycles of 95°C for 1 min, 48°C for 1 min and 72°C for 1 min, with a 10 min final extension cycle at 72°C. PCR products were purified with SolGent PCR Purification kit-Ultra (SolGent, Daejeon, South Korea), and forwarded to sequencing using the same primers, 2L BigDye, 1µl primer (20 pM), template DNA and purified water to reach a 10µl reaction volume. Cycle sequencing used was 25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min.

DNA sequencing and data analysis:

DNA sequencing was performed by ABI Prism 377 automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). Each sample was sequenced in the sense and anti-sense direction. The nucleotide sequences of both DNA strands were obtained and analyzed by Sequence Navigator (Perkin-Elmer/Applied Biosystems) to ensure accuracy of the base pair sequences. The sequence was submitted to GenBank (accession number KJ187107).

Sequence alignments were performed using CLUSTAL X, version 1.81 (Thompson *et al*., 1997), and sequence alignments were subsequently adjusted manually using BioEdit (Hall, 1999). Gaps were treated as missing data in phylogenetic analyses. The maximum parsimony and neighbour-joining analyses with 1,000 bootstrap replicates (Felsenstein, 1985) were performed using PAUP* 4.0b10 (Swofford, 2002). The boundaries between ITS1, 5.8S and ITS2 gene were determined according to the ITS sequences of *Echinops* available in GenBank. The ITS2 database (http://its2.bioapps.biozentrum.uni-wuerzburg.de/) was used to predict the secondary structures (Koetschan *et al.*, 2012).

Results and Discussion

The ITS region (ITS1-5.8S-ITS2) of *Echinops mandavillei* Kit Tan sequenced in the present study was found 634 bp, where ITS1 region 252 bp (GC content 54%), 5.8S gene 164 bp (GC content 53%), and ITS2 region 218 bp (GC content 50%). The BLAST search of ITS sequence of *E. mandavillei* Kit Tan showed maximum identity (98%) with *E. glaberrimus* DC. Parsimony analysis of the entire ITS region resulted in 431 maximally parsimonious trees with consistency index of 0.691, homoplasy index of 0.459, and retention index of 0.763. The phylogenetic tree constructed by the present analyses shows *Echinops* to be monophyletic (bootstrap support 100%; Fig. 1). The tree also provides a clear resolution at the sectional level and the result confirms an earlier report (Sánchez-Jiménez *et al.,* 2010), and *E. mandavillei* Kit Tan nested within the clade of the section *Ritropsis* (Fig. 1). Figure 2 illustrates specific nucleotide differences between *E. mandavillei* Kit Tan and *E. glaberrimus* DC., in total seven SNPs (four nucleotides in ITS1 region, i.e. at the alignment position 11, 81, 226 and 234, and three nucleotides in ITS2 region, i.e. at the alignment position 4, 58 and 165) were observed.

Table 2. Loci of SNPs (single nucleotide polymorphism) ITS sequences of *E. mandavillei* compared to *E. glaberrimus.*

Region	Position in sequence alignment	*E. mandavillei* → *E. glaberrimus*
ITS1	11th	T → C
	81th	G → R
	226th	T → C
	234th	C → T
ITS2	4th	A → C
	58th	A → G
	165th	T → C

The secondary structures of ITS2 region of *E. mandavillei* Kit Tan and *E. glaberrimus* DC. were constructed and compared (Fig. 3 A-B), which contained a central ring (primary ring) and four helices. However, the two structures differed in the four helical regions, in stem loop numbers, sizes, position, and screw angle. On the basis of the ITS2 secondary structure, *E. mandavillei* Kit Tan could be discriminative from other species of the genus.

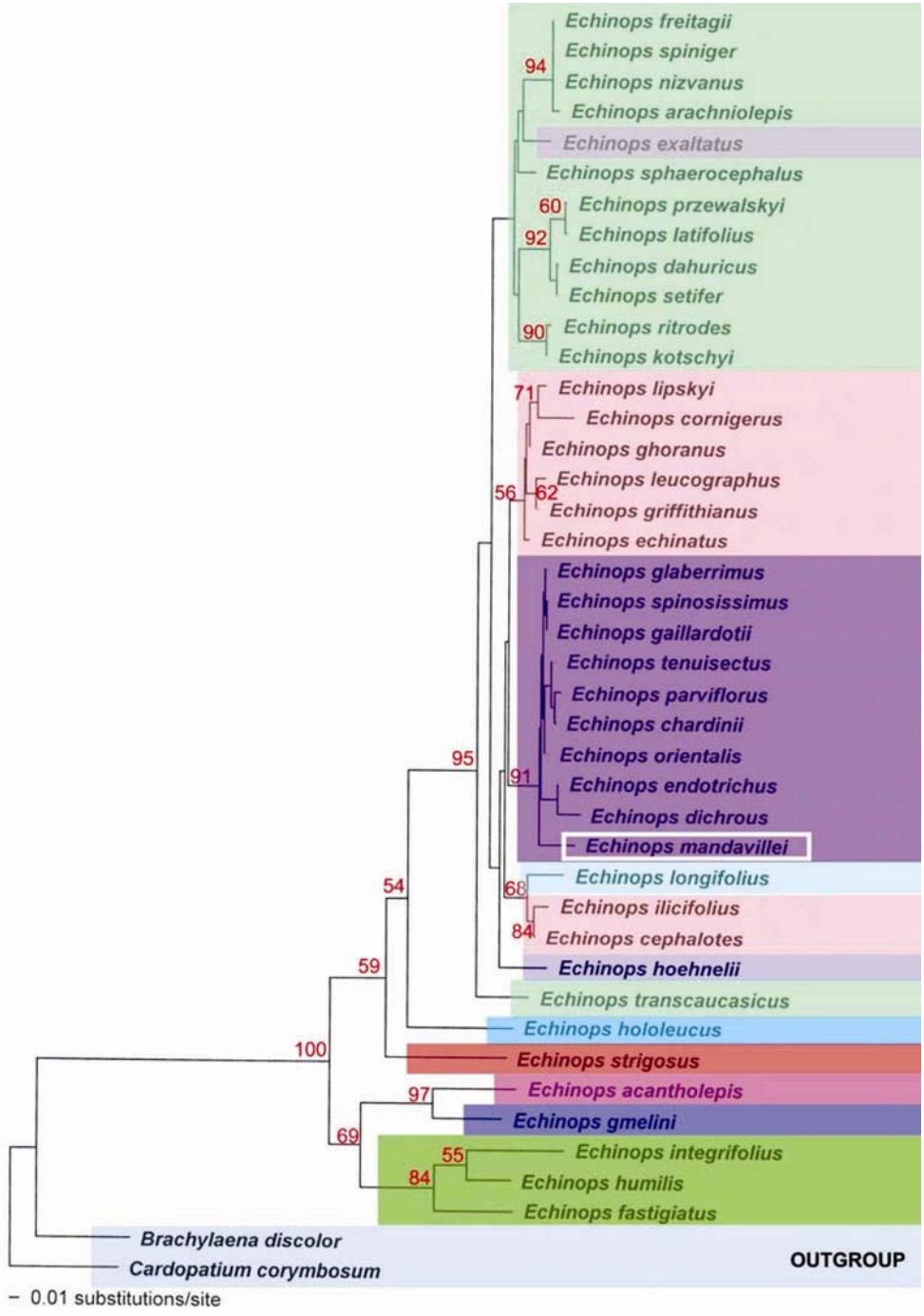

Fig. 1. Neighbour joining tree of *Echinops* species including *E. mandavillei* inferred from ITS sequences of nrDNA. Bootstrap values greater than 50% in 1,000 bootstrap replicates are shown above lines.

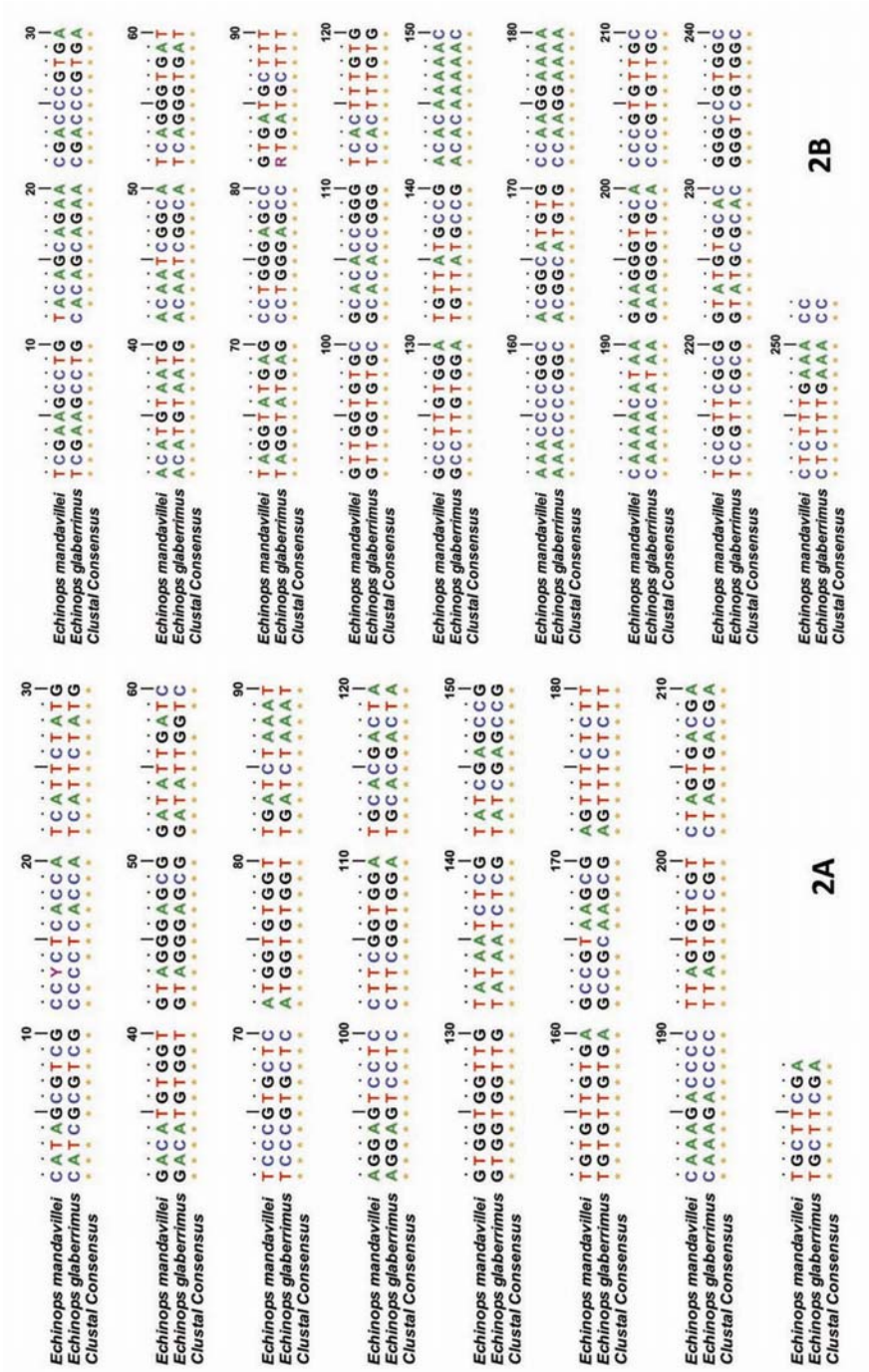

Fig. 2A. Alignments of ITS1 sequences of *E. mandavillei* compared to *E. glaberrimus*, B. Alignments of ITS2 sequences of *E. mandavillei* compared to *E. glaberrimus*. Gaps in clustal line indicate nucleotide differences.

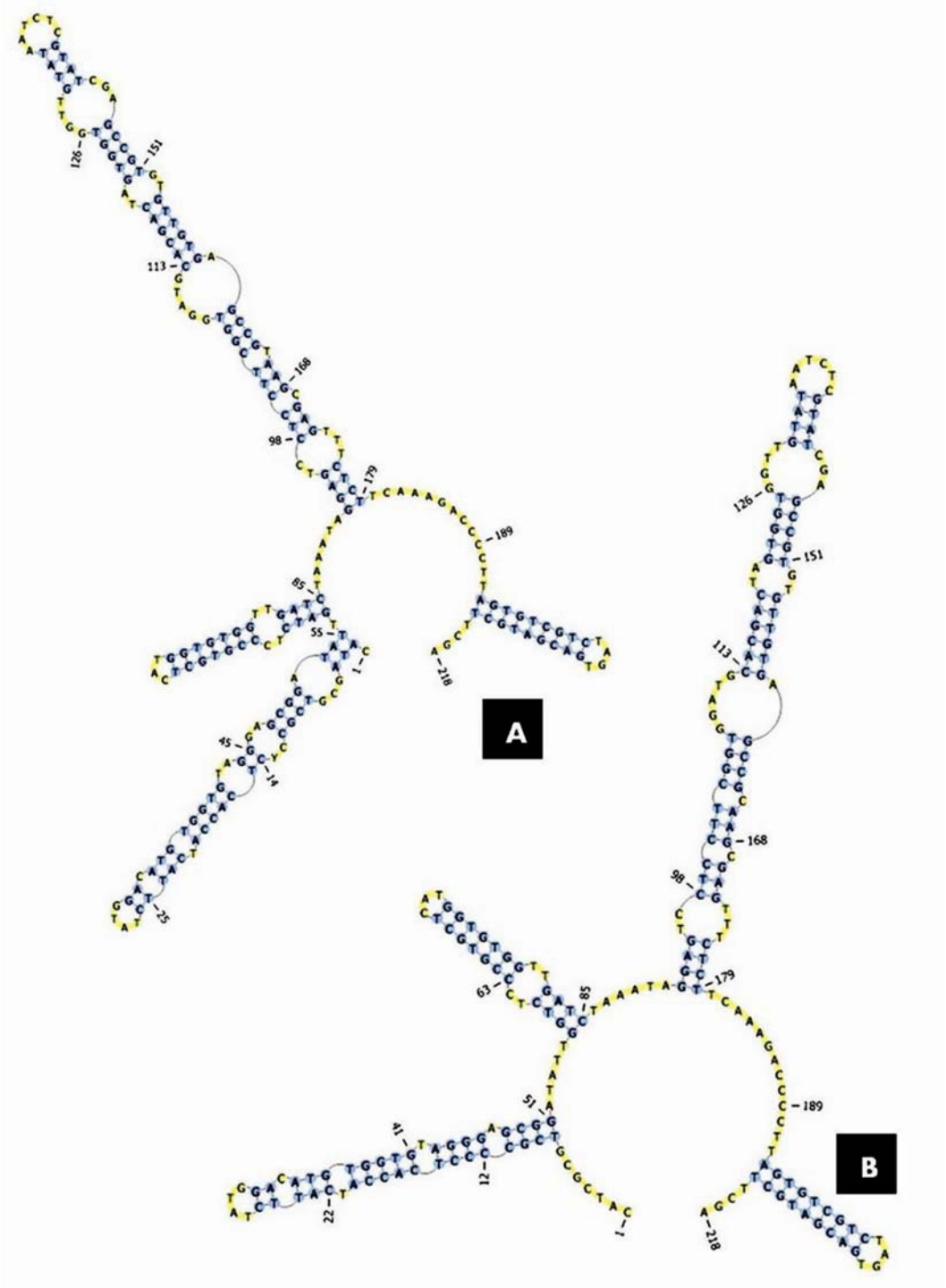

Fig. 3. The secondary structures of the ITS2 regions of *E. mandavillei* (A) compared to *E. glaberrimus* (B).

The morphological identification depends on sufficient experience and can easily be affected by the geographical environment and biocoenosis (Marcon *et al*., 2005; Rai *et al*., 2012). In contrast, DNA sequence is hardly influenced by environmental characteristics and developmental stages (Liu *et al*., 2011); and therefore, the DNA barcoding may be an effective supplement to traditional/classical morphological methods (see Hebert *et al*., 2003). The species identification using DNA barcodes has been successfully used across the algae, fungi, plants, and animals, hence; the DNA barcoding has now been proven useful in biodiversity assessment, biomonitoring, forensics, illegal trade of endangered species and their products, ecology, medicinal and poisonous plants and conservation genetics (see Hebert *et al*., 2003; Fišer Pečnikar and Buzan, 2014; Ali *et al*., 2014).

DNA barcoding efforts worldwide have resulted in the formation of the Consortium for the Barcode of Life (CBOL), and the Barcode of Life Database (BOLD), which contain more than 2.7 million records, with 2 million barcodes belonging to over 170,000 species (Ratnasingham and Hebert, 2007; BOLD Systems, 2013). The China Plant BOL Group has proposed that ITS1/ITS2 should be incorporated into the core of barcode for seed plants (Li *et al*., 2011). In the present study, we supplied the ITS barcode of *E. mandavillei* Kit Tan which is new for GenBank databases. An increasing number of studies also suggest that DNA secondary structures are crucial for genomic stability and cellular processes, such as transcription (Bochman *et al*., 2012; Salvi and Mariottini, 2012), and our study has also provided new data of *E. mandavillei* Kit Tan for this purposes.

Acknowledgements

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for the funding of this research through the Research Group Project No. RGP-VPP-195.

References

Ali M.A, Al-Hemaid, F.M., Choudhary, R.K., Lee, J., Kim, S.Y. and Rub, M.A. 2013. Status of *Reseda pentagyna* Abdallah & A.G. Miller (Resedaceae) inferred from combined nuclear ribosomal and chloroplast sequence data. Bangladesh J. Plant Taxon. **20**(2): 233-238.

Ali, M.A., Gábor, G., Norbert, H., Balázs, K., Al-Hemaid, F.M.A., Pandey, A.K. and Lee, J. 2014. The changing epitome of species identification- DNA barcoding. Saudi J. Biol. Sci. http://dx.doi.org/10.1016/j.sjbs.2014.03.003

Ali, M.A. and Choudhary R.K. 2011. India needs more plant taxonomists. Nature **471**: 37.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Phylogenet. Evol. **1**: 3-16.

Bochman, M.L., Paeschke, K. and Zakian, V.A. 2012. DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet. **13**: 770-780.

BOLD Systems 2013. BOLD Systems v3. http://www.boldsystems.org/ (Retrieved on 6 May 2014)

Chaudhary, S. 2000. *Echinops*. *In:* Chaudhary, S. (Ed.), Flora of the Kingdom of Saudi Arabia. Ministry of Agriculture and Water, National Herbarium, National Agriculture and Water Research Center, Riyadh, Saudi Arabia, Vol. **II** (3), pp. 194-199.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y. and Leon, C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE **5**(1): e8613.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**: 783-791.

Fišer Pečnikar, Z. and Buzan, E.V. 2014. 20 years since the introduction of DNA barcoding: from theory to application. J. Appl. Genet. **55**: 43-52.

Garnatje, T., Susanna, A., Garcia-Jacas, N., Vilatersana, R. and Vallès, J. 2005. A first approach to the molecular phylogeny of the genus *Echinops* L. (Asteraceae): sectional delimitation and relationships with the genus *Acantholepis* Less. Folia Geobot. **40**: 407-419.

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nuc. Acids Symp. Ser. **41**: 95-98.

Hebert, P.D.N., Cywinska, A., Ball, S.L. and deWaard, J.R. 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. **270**: 313-321.

Hidalgo, O., Garcia-Jacas, N., Garnatje, T. and Susanna, A. 2006. Phylogeny of *Rhaponticum* (Asteraceae, Cardueae-Centaureinae) and related genera inferred from nuclear and chloroplast DNA sequence data: Taxonomic and biogeographic implications. Ann. Bot. **97**: 705-714.

Koetschan, C., Hackl, T., Müller, T., Wolf, M., Förster, F. and Schultz, J. 2012. ITS2 database IV: Interactive taxon sampling for Internal Transcribed Spacer 2 based phylogenies. Mol. Phylogenet. Evol. **63**: 585-588.

Li, D.Z., Gao, L.M., Li, H.T., Wang, H., Ge, X.J., Liu, J.Q., Chen, Z.D., Zhou, S.L., Chen, S.L., Yang, J.B., Fu, C.X., Zeng, C.X., Yan, H.F., Zhu, Y.J., Sun, Y.S., Chen, S.Y., Zhao, L., Wang, K., Yang, T. and Duan, G.W. 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA **108**: 19641-19646.

Liu, C., Liang, D., Gao, T., Pang, X., Song, J., Yao, H., Han, J., Liu, Z., Guan, X., Jiang, K., Li, H. and Chen, S. 2011. PTIGS-IdIt, a system for species identification by DNA sequences of the *psb*A-*trn*H intergenic spacer region. BMC Bioinformatics **12**: S4.

Marcon, A.B., Barros, I.C. and Guerra, M. 2005. Variation in chromosome numbers, CMA bands and 45S rDNA sites in species of *Selaginella* (Pteridophyta). Ann. Bot. **95**: 271-276.

Mozaffarian, V. 2006. A taxonomic survey of *Echinops* L. tribe Echinopeae (Asteraceae) in Iran: 14 new species and diagnostic keys. Iranian J. Bot. **11**: 197-239.

Rai, P.S., Bellampalli, R., Dobriyal, R.M., Agarwal, A., Satyamoorthy, K. and Narayana, D.A. 2012. DNA barcoding of authentic and substitute samples of herb of the family Asparagaceae and Asclepiadaceae based on the ITS2 region. J. Ayurveda Integr. Med. **3**: 136-140.

Ratnasingham, S. and Hebert, P.D.N. 2007. Bold: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes **7**: 355-364.

Salvi, D. and Mariottini, P. 2012. Molecular phylogenetics in 2D: ITS2 rRNA evolution and sequence-structure barcode from Veneridae to Bivalvia. Mol. Phylogenet. Evol. **65**: 792-798.

Sánchez-Jiménez, I., Lazkov, G.A., Hidalgo, O. and Garnatje, T. 2010 Molecular systematics of *Echinops* L. (Asteraceae, Cynareae): A phylogeny based on ITS and *trn*L-*trn*F sequences with emphasis on sectional delimitation. Taxon **59**(3): 698-708.

Susanna, A. and Garcia-Jacas, N. 2007. The tribe Cardueae. *In*: Kadereit, J.W. and Jeffrey, C. (Eds), The families and genera of vascular plants, Vol. **8**, Flowering plants; Eudicots; Asterales. Heidelberg, Springer, pp. 135-158.

Susanna, A., Garcia-Jacas, N., Hidalgo, O., Vilatersana, R. and Garnatje, T. 2006. The Cardueae (Compositae) revisited: Insights from ITS, *trn*L-*trn*F, and *mat*K nuclear and chloroplast DNA analysis. Ann. Missouri Bot. Gard. **93**: 150-171.

Susanna, A., Garnatje, T. and Garcia-Jacas, N. 1999. Molecular phylogeny of *Cheirolophus* (Asteraceae: Cardueae-Centaureinae) based on ITS sequences of nuclear ribosomal DNA. Plant Syst. Evol. **214**: 147-160.

Swofford, D.L. 2002. PAUP* (v. 4.0b10). Phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**: 4876-4882.

Vidović, B. 2011. A new *Aceria* species (Acari: Eriophyoidea) on *Echinops ritro* L. subsp. *ruthenicus* (M. Bieb.) Nyman (Asteraceae) from Serbia and a supplement to the original description of *Aceria brevicincta* (Nalepa 1898). Zootaxa **2796**: 56-66.

Vilatersana, R., Susanna, A., Garcia-Jacas, N. and Garnatje, T. 2000. Generic delimitation and phylogeny of the *Carduncellus*-*Carthamus* complex (Asteraceae) based on ITS sequences. Plant Syst. Evol. **221**: 89-105.

Wang, Y.J., Liu, J.Q. and Miehe, G. 2007. Phylogenetic origins of the Himalayan endemic *Dolomiaea*, *Diplazoptilon* and *Xanthopappus* (Asteraceae: Cardueae) based on three DNA regions. Ann. Bot. **99**: 311-322.

Wang, Y.J., Pan, J.T., Liu, S.W. and Liu, J.Q. 2005. A new species of *Saussurea* (Asteraceae) from Tibet and its systematic position based on ITS sequence analysis. Bot. J. Linn. Soc. **147**: 349-356.

White, T.J., Bruns, T., Lee, S. and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In:* Innis, M.A., Gelfand, D.H., Sninksky, J.J. and White, T.J. (Eds), PCR protocols: a guide to method and amplifications. Academic Press, San Diego, California, pp. 315-322.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P. and Chen, S. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS ONE **5**(10): e13102.

(Manuscript received on 7 February 2014; revised on 9 May 2014)