Bangladesh J. Plant Taxon. **22**(2): 111-118, 2015 (December)

# MOLECULAR EVOLUTIONARY RELATIONSHIPS OF *EUPHORBIA SCORDIFOLIA* JACQ. WITHIN THE GENUS INFERRED FROM ANALYSIS OF INTERNAL TRANSCRIBED SPACER SEQUENCES

FAHAD M.A. AL-HEMAID, M. AJMAL ALI[1], JOONGKU LEE[2], SOO-YONG KIM[3] AND M. OLIUR RAHMAN[4]

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

*Keywords*: *Euphorbia scordifolia*; Euphorbiaceae; ITS; Genotyping.

**Abstract**

The present study explored molecular phylogenetic analysis of 28 species of *Euphorbia* L. for the identification and establishment of molecular evolutionary relationships of *Euphorbia scordifolia* Jacq. within the genus based on the internal transcribed spacers (ITS) sequences (ITS1-5.8S-ITS2) of nuclear ribosomal DNA (nrDNA). The sequence similarity search using Basic Local Alignment Search Tool (BLAST) of the ITS sequence of *E. scordifolia* showed the closest sequence similarity to *E. supina* Raf. The analysis of ITS sequence data revealed four major clades consistent with subgeneric classifications of the genus. Molecular data support placement of *E. scordifolia* in the subgenus *Chamaesyce*.

## Introduction

The genus *Euphorbia* L. (Euphorbiaceae) comprising ca. 2000 species, which is one of the largest genera of the flowering plants (Frodin, 2004; Riina *et al*., 2013). The main molecular phylogenetic studies of *Euphorbia* species have addressed the overall phylogeny of the genus, with its four subgeneric clades of *Rhizanthium*, *Esula*, *Euphorbia*, and *Chamaesyce* (Steinmann and Porter, 2002; Bruyns *et al*., 2006; Park and Jansen, 2007; Zimmermann *et al*., 2010). In Saudi Arabia, the genus *Euphorbia* is represented by 38 species. Of them, *E. scordifolia* Jacq. is distributed in Cape Verde Island, Ethiopia, Somalia, Sudan, Yemen and also in western region of Saudi Arabia (Abedin *et al*., 2001). The morphological characters of *E. scordifolia* overlap with *E. supina* Raf. (Abedin *et al*., 2001).

The internal transcribed spacers (ITS) sequence of nuclear ribosomal DNA region including the 5.8S gene is the most widely used molecular marker to infer phylogenetic relationships among plant species (Baldwin *et al*., 1995; Ali *et al*., 2014). Although reliance on nrDNA ITS sequence as the sole source of phylogenetic evidence has come under criticism because of certain features of its evolution; however, it remains the most efficient locus for generating species-specific phylogenetic inferences and genotyping in most groups of plants (Ali *et al*., 2013, 2014, 2015). While searching for DNA sequences of *E. scordifolia* in GenBank as a part of a research for genotyping of unresolved taxonomic status of flowering plants of Saudi Arabia, it was found that *E. scordifolia* have not previously been sequenced. A perusal of taxonomic literature revealed that

[1]Corresponding author. Email: majmalali@rediffmail.com
[2]Department of Environment and Forest Resources, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, South Korea.
[3]International Biological Material Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305 806, South Korea.
[4]Department of Botany, University of Dhaka, Dhaka 1000, Bangladesh.

the molecular evolutionary relationships of *E. scordifolia* distributed in Saudi Arabia is also unknown. Therefore, the present study aims at molecular genotyping of *E. scordifolia* based on ITS sequence of nrDNA.

## Materials and Methods

*Taxon sampling*:

Leaf materials of *E. scordifolia* were collected from the herbarium specimens [voucher- Al-Rawshan, altitude 1122 m, 19.08.1978, Don Bermant 146] housed at National Herbaium & Genebank, National Agriculture & Animal Resources Research Center, Ministry of Agriculture, Riyadh, Saudi Arabia (RIY); and the taxonomic identification was confirmed through consultation of Flora of Saudi Arabia (Abedin *et al*., 2001).

*DNA extraction, amplification and sequencing*:

Total genomic DNA was extracted using Qiagen DNeasy Plant Mini Kit (Valencia, CA, USA). ITS sequences of nuclear ribosomal DNA were amplified using AccuPower HF PCR PreMix (Bioneer, Daejeon, South Korea) and primer ITS1 (5′-GTCCACTGAACCTTATCATTT AG-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) of White *et al*. (1990) via polymerase chain reaction (PCR). Each 20 µl volumes of PCR premix contained 2 µl of 10x buffer, 300 µM dNTPs, 1 µl of a 10 pM solution of each primer and 1 unit of HF DNA polymerase. One round of amplification consisted of denaturation at 94 °C for 5 min, followed by 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 for 5 min at 72 °C. PCR products were purified with the SolGent PCR Purification Kit-Ultra (SolGent, Daejeon, South Korea) prior to sequencing. The sequencing reaction was performed in a 10 µl final volume with the BigDye Terminator cycle sequencing kit (Perkin-Elmer, Applied Biosystems). Cycling conditions included an initial denaturation at 94 °C for 5 min, followed by 30 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 4 min. The sequenced products were precipitated with 17 µl of deionized sterile water, 3 µl of 3 M NaOAc, and 70 µl of 95% EtOH. The capillary gel electrophoresis was conducted with Long Ranger Single Packs (FMC BioProducts) by an ABI 3100 automated DNA sequencer (Perkin-Elmer, Applied Biosystems). The sequences were analyzed by ABI Sequence Navigator (Perkin-Elmer/Applied Biosystems). Nucleotide sequences of both DNA strands were analyzed to ensure accuracy. The sequences were subjected to BLAST-searched (Altschul *et al*., 1990) by NCBI server (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

*Phylogenetic analysis*:

ITS sequences of nrDNA of 28 species of *Euphorbia* (Table 1) were retrieved from GenBank database of National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). *Neoguillauminia cleopatra* and *Dichostemma glaucescens* were chosen as outgroup taxa according to previous work (Barres *et al*., 2011) and were retrieved from GenBank (Table 1). Sequence alignment was performed using CLUSTAL X version 1.81 (Thompson *et al*., 1997). Sequence alignment was subsequently adjusted manually using BioEdit (Hall, 1999). Gaps were treated as missing data in phylogenetic analyses. The generated sequences were submitted to GenBank (Table 1). The boundaries between the ITS1, 5.8S and ITS2 gene for *E. scordifolia* were determined in the aligned data matrix, and were exported as a Nexus file and subsequently analysed using Maximum Parsimony (MP) and Maximum Likelihood (ML) methods by MEGA5 (Tamura *et al*., 2011). The distribution and pattern of nucleotide substitution in all sequences was investigated using HYPERMUT (Rose and Korber, 2000).

**Table 1. Plant accessions used for the molecular phylogenetic analysis of *Euphorbia scordifolia.***

| Group | Subgenus | Taxon | GenBank Accession No. |
|---|---|---|---|
| Ingroup | *Rhizanthium* | *Euphorbia antso* Denis | AF537579 |
| | | *Euphorbia atrispina* N.E. Br. | AF537568 |
| | | *Euphorbia balsamifera* Ait. | AF537571 |
| | | *Euphorbia clava* Jacq. | AF537569 |
| | | *Euphorbia namuskluftensis* L.C. Leach | AF537562 |
| | | *Euphorbia obesa* Hook. f. | AF537566 |
| | *Esula* | *Euphorbia aphylla* Brouss. | AF537540 |
| | | *Euphorbia dendroides* L. | AF537539 |
| | | *Euphorbia peplus* L. | AF537532 |
| | | *Euphorbia schimperi* C. Presl | AF537537 |
| | | *Euphorbia schimperiana* Hochst. *ex* A. Rich. | JN207816 |
| | *Euphorbia* | *Euphorbia abdelkuri* Balf. f. | AF537458 |
| | | *Euphorbia beharensis* Leandri | AJ508983 |
| | | *Euphorbia cylindrifolia* Marn.-Lap. & Rauh | AJ508955 |
| | | *Euphorbia drupifera* Thonn. | AF537480 |
| | | *Euphorbia epiphylloides* Kurz | AF537484 |
| | | *Euphorbia milii* Des Moul. | AJ508974 |
| | | *Euphorbia ramipressa* Croizat | AF537481 |
| | | *Euphorbia supina* Raf. | EU659773 |
| | | *Euphorbia teke* Schweinf. *ex* Pax | AF537485 |
| | *Chamaesyce* | *Euphorbia fulgens* Karw. *ex* Klotzsch | AF537404 |
| | | *Euphorbia graminea* Jacq. | AF537410 |
| | | *Euphorbia heterophylla* L. | GU214931 |
| | | *Euphorbia ipecacuanhae* L. | AF537397 |
| | | *Euphorbia leucocephala* Lotsy | GU214932 |
| | | *Euphorbia misera* Benth. | AF537383 |
| | | *Euphorbia pulcherrima* Willd. *ex* Klotzsch | GU214943 |
| | | *Euphorbia scordifolia* Jacq. | KR704890 |
| | | *Euphorbia sphaerorhiza* Benth. | AF537412 |
| Outgroup | | *Neoguillauminia cleopatra* (Baill.) Croizat | AF537581 |
| | | *Dichostemma glaucescens* Pierre | AF537584 |

## Results and Discussion

The combined length of ITS region (ITS1-5.8S-ITS2) in *E. scordifolia* was 642 bp. The ITS1 region was 266 bp (GC content 53%), the 5.8S gene was 162 bp long (GC content 56%), and the ITS2 region was 213 bp (GC content 58%). The BLAST search of ITS sequence of *E. scordifolia* showed high identity level (95%) with *E. humifusa* Willd. followed by *E. glyptosperma* Engelm., *E. maculata* L., *E. tettensis* Klotzsch and *E. meganaesos* Featherm. Parsimony analysis of the entire ITS region resulted in five maximally parsimonious trees, the consistency index was 0.491, the retention index was 0.709, and the composite index was 0.367 (0.348) for all sites and parsimony-informative sites (in parentheses). There were a total of 499

positions in the final dataset, of which 223 were parsimony informative. The phylogenetic tree recovered by the analyses provided a clear resolution of taxon included in the analysis at the subgeneric level. *Eupphorbia scordifolia* nested within the clade of the subgenus *Chamaesyce.* The ML analyses recovered tree topology similar to MPT; and therefore, only the ML topology is presented here (Fig. 1). A total of 36 specific nucleotide differences, i.e. 19 in ITS1 and 17 in ITS2 region were detected between *E. scordifolia* and *E. supina* (Table 2).

**Table 2. Differences of DNA base pairs between the ITS sequences of *Euphorbia supina* and *E. scordifolia.***

| Specific nucleotide differences | | | | | |
|---|---|---|---|---|---|
| Position in sequence alignment | ITS1 | | Position in sequence alignment | ITS2 | |
| | *E. supina* | *E. scordifolia* | | *E. supina* | *E. scordifolia* |
| 18 | G | A | 3 | T | C |
| 41 | T | T | 22 | C | T |
| 45 | C | G | 25 | T | C |
| 56 | G | T | 37 | - | G |
| 93 | C | T | 49 | C | T |
| 106 | T | - | 56 | A | R |
| 113 | C | T | 74 | T | C |
| 135 | A | C | 94 | T | C |
| 136 | A | T | 126 | T | C |
| 137 | A | T | 146 | A | G |
| 147 | T | C | 151 | C | A |
| 148 | G | T | 163 | C | T |
| 149 | C | T | 170 | T | A |
| 208 | C | T | 173 | G | A |
| 212 | C | T | 174 | A | T |
| 215 | C | T | 191 | T | C |
| 232 | T | C | 192 | G | A |
| 254 | G | A | | | |
| 258 | G | A | | | |

The Tandem Repeats Finder (Benson, 1999) was used to detect repeats in the ITS sequences. Differences in substitution rates can discriminate functional forms of pseudogenes (Buckler and Holtsford, 1996a,b). The analysis using the program HYPERMUT showed excessive levels of G =>A mutations which indicates that all differences arose from a single substitution sequence. The result was compared to the reference sequences and their physical locations along the sequences were graphically illustrated (Fig. 2).

The use of DNA sequences to identify organisms has been proposed as a more efficient approach than traditional and morphological taxonomic parameters (Tautz *et al.*, 2003). In fact, the recent development in DNA molecular systematic techniques including molecular hybridization, cloning, restriction endonuclease digestions and DNA sequencing and phylogenetic theory have changed the epitome of species identification as well as our understanding of the relationships among organisms at various levels in the tree of life which has been advanced greatly

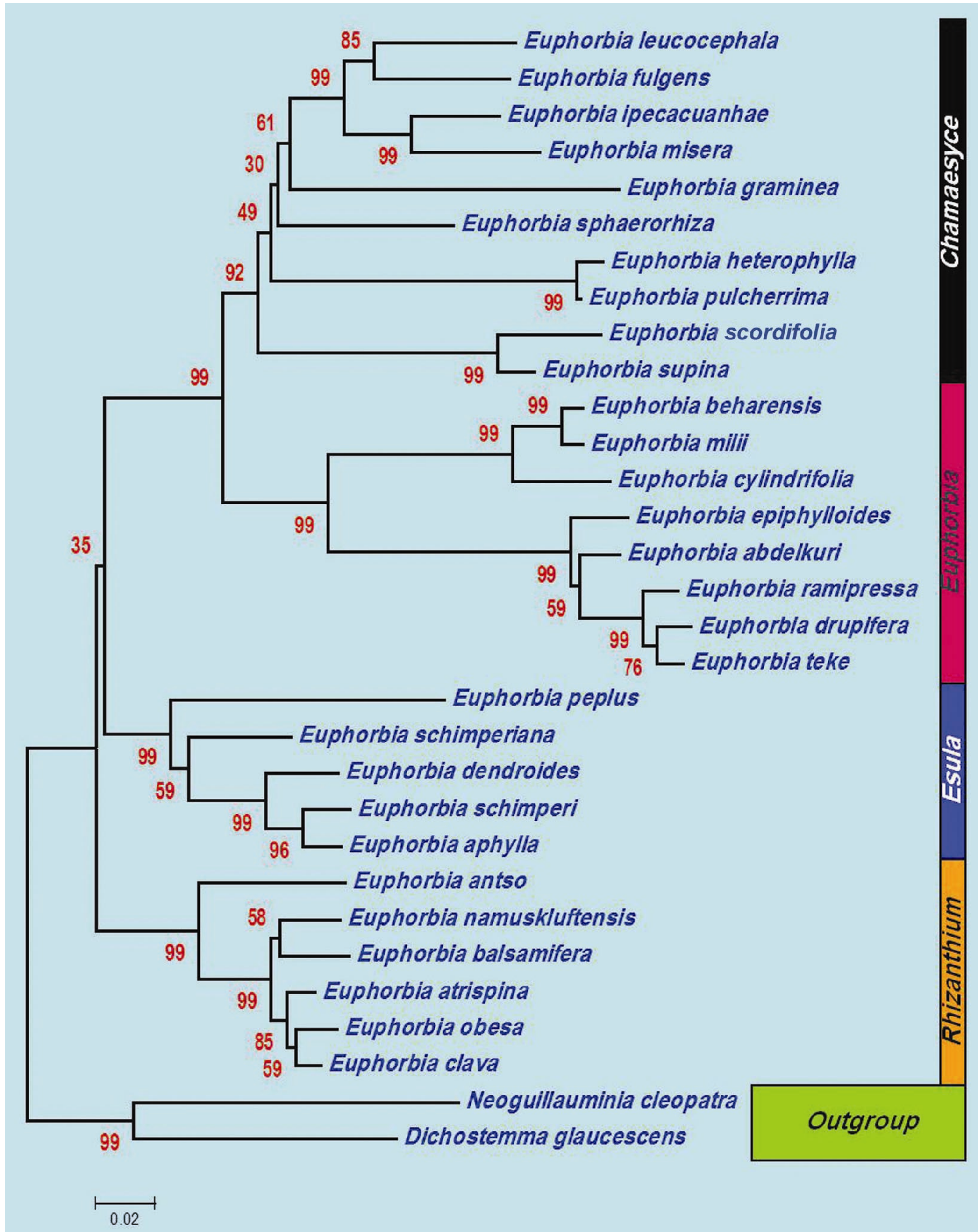


Fig.1. A maximum likelihood (ML) tree inferred from analysis of sequence data of internal transcribed spacer (ITS) region of nuclear ribosomal DNA. Bootstrap values (1000 × replicates) are indicated.

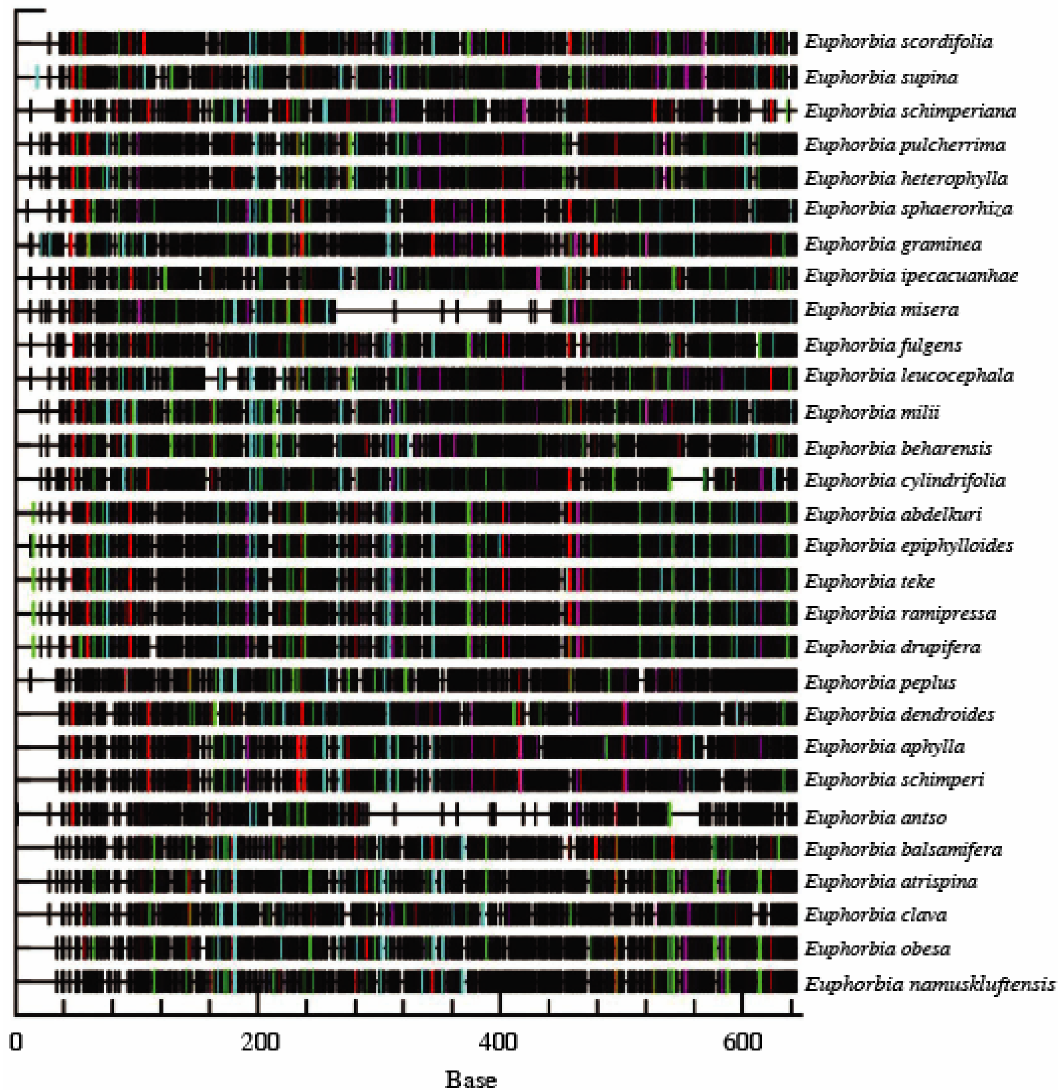


Fig. 2. Schematic illustration of the distribution of substitution sites across the ITS region obtained from 29 species of *Euphorbia*, using *Dichostemma glaucescens* as reference (red = GG > AG, cyan = GA > AA, green = GC > AC, magenta = GT > AT, black = not G > A transition, yellow = gap).

(Ali *et al.*, 2014). From the first report of the utility of the nrDNA ITS sequence in plants (Baldwin, 1992), it has been extensively used to distinguish even very closely related species (Chen *et al*., 2010; Yao *et al*., 2010). Moreover, during the last two decades, the nrDNA ITS sequence has gained much attention as smartest gene available for the molecular signature of a taxon (Ali *et al*., 2013).

The present study is the first report of inferring the nrDNA ITS based molecular genotyping of the *E. scordifolia*. Since, the majority of the species of the genus *Euphorbia* have to be sequenced; the present study will nevertheless help in DNA barcoding / molecular identification of *E.*

*scordifolia* as well as it will also participate in addressing the complete phylogeny of the genus *Euphorbia*. The DNA barcodes show promise in providing a practical, standardized, species-level identification tool that can be used for biodiversity assessment, life history, ecological studies and forensic analysis (Szabó *et al*., 2005; Mansour *et al*., 2009; Gyulai *et al*., 2012; Ali *et al*., 2014, 2015). Hence, the nrDNA ITS sequence of *E. scordifolia* will be of immense importance in barcoding of the genus *Euphorbia* in particular, and in the analysis of plant biodiversity of Saudi Arabia in general.

**Acknowledgement**

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

**References**

Abedin, S., Mossa, J.S., Al-Said, M.S. and Al-Yahya, M.A. 2001. Euphorbiaceae. *In*: Chaudhary, S. (Ed.), Flora of Saudi Arabia, Ministry of Agriculture and Water, National Herbarium, National Agriculture and Water Research Center, Riyadh, Saudi Arabia, **II**(1): 291–395.

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

Ali, M.A., Gyulai, 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. **21**(3): 204–231.

Ali, M.A., Gyulai, G. and Al-Hemaid, F. 2015. Plant DNA Barcoding and Phylogenetics. LAP Lambert Academic Publishing, Germany, pp. 1−298.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**: 403–410.

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.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S. and Donoghue, M.J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Miss. Bot. Gard. **82**: 247–277.

Barres, L., Vilatersana, R., Molero, J., Susanna, A. and Galbany-Casals, M. 2011. Molecular phylogeny of *Euphorbia* subg. *Esula* sect. *Aphyllis* (Euphorbiaceae) inferred from nrDNA and cpDNA markers with biogeographic insights. Taxon **60**(3): 705–720.

Benson, G. 1999. Tardem repeats finder: a program to analyze DNA sequence. Nucleic Acids Res. **27**: 573–580.

Bruyns, P.V., Mapaya, R.J. and Hedderson, T. 2006. A new subgeneric classification of *Euphorbia* (Euphorbiaceae) based on ITS and *psb*A–*trn*H sequence data. Taxon **55**: 397–420.

Buckler, E.S. and Holtsford, T.P. 1996a. *Zea* ribosomal repeat evolution and mutation patterns. Mol. Biol. Evol. **13**: 623–632.

Buckler, E.S. and Holtsford, T.P. 1996b. *Zea* systematics: ribosomal ITS evidence. Mol. Biol. Evol. **13**: 612–622.

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.

Frodin, D. 2004. History and concepts of big plant genera. Taxon **53**: 753–776.

Gyulai, G., Horváth, L., Lágler, R. and Holly, L. 2012. The Hungarian gene bank collections of common millet (*Panicum miliaceum*) and the application for conservation genetics. European J. Plant Sci. Biotech. **6**(SI2): 69–102.

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

Mansour, A., Ismail, H.M., Ramadan, M.F. and Gyulai, G. 2009. Comparative genotypic and phenotypic analysis of tomato (*Lycopersicon esculentum*) cultivars grown under two different seasons in Egypt. African J. Plant Sci. Biotechnol. **3**(1): 73–79.

Park, K.-R. and Jansen, R.K. 2007. A phylogeny of Euphorbieae subtribe Euphorbiinae (Euphorbiaceae) based on molecular data. J. Plant Biol. **50**: 644–649.

Riina, R., Peirson, J.A., Geltman, D.V., Molero, J., Frajman, B., Pahlevani, A., Barres, L., Morawetz, J.J., Salmaki, Y., Zarre, S., Kryukov, A., Bruyns, P.V. and Berry, P.E. 2013. A worldwide molecular phylogeny and classification of the leafy spurges, *Euphorbia* subgenus *Esula* (Euphorbiaceae). Taxon **62**(2): 316–342.

Rose, P.P. and Korber, B.T. 2000. Detecting hypermutations in viral sequences with an emphasis on G A hypermutation. Bioinformatics **16**: 400−401.

Szabó, Z., Gyulai, G., Humphreys, M., Horváth, L., Bittsánszky, A., Lágler, R. and Heszky, L. 2005. Genetic variation of melon (*C. melo*) compared to an extinct landrace from the middle ages (Hungary) I. rDNA, SSR and SNP analysis of 47 cultivars. Euphytica **146**: 87–94.

Steinmann, V.W. and Porter, J.M. 2002. Phylogenetic relationships in Euphorbieae (Euphorbiaceae) based on ITS and *ndh*F sequence data. Ann. Miss. Bot. Gard. **89**: 453–490.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**(10): 2731–2739.

Tautz, D., Arctander, P., Minelli, A., Thomas, R.H. and Vogler, A.P. 2003. A plea for DNA taxonomy. Trends Ecol. Evol. **18**(2): 70–74.

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.

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, 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.

Zimmermann, N.F.A., Ritz, C.M. and Hellwig, F.H. 2010. Further support for the phylogenetic relationships within *Euphorbia* L. (Euphorbiaceae) from nrITS and *trn*L–*trn*F IGS sequence data. Plant Syst. Evol. **286**: 39–58.

*(Manuscript received on 11 September 2015; revised on 22 October 2015)*