ASSESSMENT OF GENETIC DIVERSITY IN SOME WILD PLANTS OF *ASTERACEAE* FAMILY BY RIBOSOMAL DNA SEQUENCE

M. H. AMAR[1], A. H. M. HASSAN[1] AND ESRAA A. M. EL SHERBENY[2]

1. *Egyptian Deserts Gene Bank, Desert Research Center, North Sinai, Egypt*
2. *Department of Plant Genetic Resources, Desert Research Center, Cairo, Egypt*

The *Asteraceae* (Compositae, alternate name) with its approximately 1,620 genera and more than 23,600 species is the largest family of flowering plants. The family is distributed worldwide except for Antarctica and is especially diverse in the tropical and subtropical regions of North America, the Andes, eastern Brazil, Southern Africa, the Mediterranean region, central Asia, and southwestern China. In Egypt, family *Asteraceae* is represented by 98 genera and 228 species, annuals, perennials, wide spread in all its phytogeographical regions. Many of these genera have economic potentialities, e.g. folk medicine (Bolous, 2002).

With the rapid development of molecular biology studies of *Asteraceae* germplasm identification and genetic diversity offer numerous reliable molecular marker information by means of Random Amplified Polymorphic DNA (RAPD) (Badr *et al*., 2012), Restriction Fragment Length Polymorphism (RFLP) (Ito *et al*., 2000), Inter-simple sequence repeat (ISSR) (Gharibi *et al*., 2011), Simple Sequence Repeat (SSR) (Simko, 2009), RAPD, ISSR and RFLP (Abd El-Tawab *et al*., 2010), amplified fragment length polymorphisms (AFLP) (Czarnecki *et al*., 2008) etc. Nevertheless, DNA-labeling techniques still have many problems particularly the wild medicinal herbs such as low repeatability, high subjectivity of experimental results and unshared study data from different laboratories (Wang *et al*., 1999). Thus, molecular sequence markers have become a significant tool in species classification, dependent on sequencing the differences in genome directly without environmental effects (Li *et al*., 2010).

Accordingly, advances in DNA sequencing techniques have allowed the extensive use of short DNA fragments, especially of the ribosomal DNA (rDNA). The utility of genes coding for ribosomal RNA (rDNAs) is found in the ubiquitous presence and relative conservation of many regions of their nucleotide sequences. A cluster of rDNA (Fig. 1) consists of the gene coding for 18S rRNA (small subunit ribosomal RNA, SSU rRNA), two internal transcribed spacers (ITS1 and ITS2) separated by the 5.8S rRNA gene, and the gene coding for the 28S rRNA (large subunit of ribosomal RNA, (LSU rRNA) (Hillis and Dixon, 1991). Additional two external transcribed spacers (ETS) are located upstream of the 18S

rDNA and downstream of the 28S rDNA. A nontranscribed spacer (NTS) separates adjacent copies of the rDNA repeat unit. Both spacers (ETS and NTS) are also called intergenic spacers (IGS) (Dabert *et al*., 2006).The major concerns with the use of the rDNA locus in taxonomic and phylogenetic analyses are the existence of polymorphisms among repeated units, which may cause extensive differentiation even within a single individual and provide useful tools for phylogenic studies (Wei *et al*., 2010). Accordingly, sequence comparison of the rDNA region is widely used in taxonomy and molecular phylogeny. It has typically been most constructive for variation between species, populations and even individuals (or inbred lines) as in tomato (Jo *et al*., 2009), rice (Chang *et al*., 2010), apple (Giaretta *et al*., 2010), and Compositae, Anthemideae (Sonboli *et al*., 2012).

In this study, we examined the molecular divergence of complete rDNA sequenced of ITS and IGS regions among four wild species of *Asteraceae* germplasm in Egypt. To our knowledge, even now there has been no report about the comparison of ITS and IGS sequences and their efficiency and utility as molecular markers in *Asteraceae* family.

MATERIALS AND METHODS

A. Plant Materials

A total of four wild and endemic species of Asteraceae family (*Echinops spinosus* L., *Achillea santolina* L., *Matricaria recutita* L. and *Artemisia monosperma* Delile) were collected from different eco-geographical localities of the natural habitat in Sinai protected area, Egypt. In Egypt, these species is recorded as rare, endemic and neglected wild medicinal taxon. Five plants from each species were collected together as bulk materials. The list of species and their collection sites (Latitude longitude and altitude) are presented in Table (1).

B. Methods

1. DNA Isolation

Total genomic DNA was isolated from fresh leaves following the procedure as previously described by Pirttila *et al.* (2001). Five DNA samples of each species were dissolved together as bulk DNA. The quality and concentration of the DNA samples were checked in a UV-1601 spectrophotometer (Shimadzu, Japan) and a portion of the DNA was diluted to 50 ng/µl for use in ITS and IGS analyses. Both the stock and diluted portions were stored at -20°C.

2. PCR amplification and sequencing of the ITS region

Complete ITS region of rDNA was amplified with universal primers ITS-1 (5' TCCGTAGGTGAACCTGCGG3') as forward primer and ITS-4 (5'TCCTCCGCTTATTGATATGC 3') as reverse primer as described by White *et al.* (1990). Final reaction volumes of 25 µl each contained 50 ng genomic DNA, 0.5 pmol of each primer, 0.2 mM dNTPs, 1U Taq DNA polymerase (Fermentas, Shenzhen, China), 2 µl of 10x PCR buffer sup-

plied by the manufacturer and about 2.5 mM $MgCL_2$. The amplification programmed consisted of pre denaturation at 94°C for 4 min; 35 cycles at 94°C for 45 s, 55°C for 60 s, 72°C for 90 s, and a final incubation at 72°C for 7 min. Then the PCR products were subjected to electrophoresis on a 1.5% agarose gel, stained with ethidium bromide, and visualized under ultraviolet (UV) light. The PCR fragments of each sample were excised and purified from the gels using E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek, Inc., Norcross, USA). The purified products of the PCR were ligated to pMD18-T Easy Vector using the appropriate kit (TaKaRa, Tokyo, Japan) and the ligation products were transformed into *Escherichia coli* DH5α competent cells. The recombinant clones were selected on Liquid Broth media plates containing ampicillin.

3. PCR amplification and sequencing of the IGS region

The primer pair 18S L (5'-GAACGCCTCTAAGTCAGAATCC-3') and 28S R (5'-ACTGGCAGAATCAACCAGGTA-3') was used to amplify across the IGS region of ribosomal DNA (White *et al.,* 1990). The reaction mixture (25 µl) containing 2 µl of 10x PCR buffer, 0.2 mM dNTPs, 0.5 pmol of each primer, 1U Taq DNA polymerase (Fermentas, Shenzhen, China), and 50 ng genomic DNA template. The PCR conditions were; 95°C for 5 min for initial genomic DNA denaturation; 35 cycles of 94°C for 1 min, 57°C for 45 s, 72°C for 1 min, and final extension at 72°C for 7 min. Then the PCR products of each sample were excised, purified and cloning were separated and visualized with the same procedure as for ITS. Three positive colonies from each amplified amplicons were selected for sequenced by the Uni-Gene Company (Shanghai, China). The Open Reading Frame Finder (http:// www.ncbi.nlm.nih.gov/gorf/gorf.) was used to verify the credibility of the results and their conformity. Sequence similarity was analyzed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/).

4. Sequence analysis

Vector sequences were cleaned and the sequences were aligned using Clustal X version 1.81 (Thompson *et al*., 1997) with manual adjustments wherever necessary. Gaps were positioned to minimize nucleotide mismatches. The MEGA program version 5.0 (Molecular Evolutionary Genetics Analysis, Tamura *et al.,* 2011) was employed to estimate GC and AT contents, nucleotide substitution, nucleotide diversity (π), estimated values of transition/transversion bias (R), substitutions (r) for each nucleotide pair, and cluster analysis among the four *Asteraceae* germplasm. We further computed Maximum Composite Likelihood (MCL) Estimate of the pattern of nucleotide substitution according to Tamura *et al.* (2004).

5. Phylogenetic analysis

Pair-wise evolutionary distance among four *Asteraceae* family was determined by Kimura 2-Parameter method

(K2P) (Kimura, 1980). The Maximum likelihood (ML) tree phylogenetic tree was conducted using MEGA version 5.

RESULTS AND DISCUSSION

1. Sequencing analysis for ITS and IGS

In this article, we used a comparative analysis approach using several parameters like nucleotide frequency, nucleotide substitution (r), nucleotide diversity (π), and the estimated values of transition/transversion bias (R) to provide better understanding of the genetic diversity and phylogenetic relations across the studied genotypes of the *Asteraceae* family. The results of the confrontation between DNA sequence analysis of the isolates and GenBank database ranged from 94 to 99% similarity, through BLAST search (Table 1), supporting good credibility for ITS and IGS. The length of variation for the entire ITS (650 to 750 bp) and IGS (800 to 950 bp) regions showed very distinctive sequences for individual species. Similarly, variation was observed in the nucleotide composition of the ITS and IGS, which may be due to the sequence length variation of the analyzed markers (Table 2).

With regard to ITS sequence divergence among taxa, the averages of nucleotide frequencies were A (25%), T (24%), C (26%), and G (25%) with an average of GC (51%) and AT (49%) contents (Table 2). The highest numbers of nucleotide frequency for ITS sequence was observed in *Artemisia monosperma* (729 bases), whereas the lowest one was recorded in *Achillea santolina* L. and *Matricaria recutita* L. (712 base). The maximum nucleotide percentage for GC content (53%) was observed in *Echinops spinosus* L. However, the lowest GC content (46%) was recorded in *Artemisia monosperma.* Within the analysis of IGS sequence divergence among taxa (Table 2), the averages of nucleotide length were A (25.1%), T (26.7%), C (26.1%), and G (22.1%) with an averages of GC (48.2%) and AT (51.8%) contents. The highest numbers of nucleotides for IGS sequence were observed in *Achillea santolina* L., *Artemisia monosperma*, and *Matricaria recutita* L., (640 bases). Whereas, *Echinops spinosus* L. recorded the minimum number of nucleotide frequency (365 bases). The maximum nucleotide percentage for GC content (53.6%) was observed in *Artemisia monosperma.* In contrast, the lowest GC content (46.3%) was recorded in *Achillea santolina* L. The Tajima's Neutrality test (Tajima, 1989) was performed to calculate the nucleotide diversity value (π). There were a total of 754 and 667 positions across the final dataset for ITS and IGS sequences, respectively. The nucleotide diversity rate (π) was observed higher in IGS (0.60) as compared to ITS sequence (0.49) (Table 3).

Within genomes, all organisms have DNA sequences that code for ribosomal RNA (rRNA), an essential component of cellular protein synthesis machinery (Kollipara *et al.*, 1997). Ribosomal RNA typically accounts for about 40% of all transcription within a cell, and ribosomal RNA makes up as much as 80% of cellular RNA (Moss and Stefanovsky,

1995). Owing to relatively rapid evolution, differences in sequence and/or length of rDNA are possible between closely related species of *Asteraceae* family (Zhao *et al.,* 2010). The IGS sequences, as an intergenic region, may bear functional sequences, such as promoter, enhancer, transcription stop signals, and reproduction start signals (Dutta and Verma, 1990). Meanwhile, the IGS sequences undergo conversion and concerted evolution to reach a homogenization within an array of repeats. Subsequently, the intergenic spacer of the rDNA cluster evolves quickly and is highly polymorphic sequence, providing a useful tool for assessing the sequence phylogeny and genetic variability studies (Singh *et al*., 2008).

2. Phylogenetic analysis

Based on the sequence data of the flanking regions of ITS or IGS sequence, a phylogenetic tree was constructed using Maximum likelihood (ML) method (Tamura *et al*., 2004) (Fig. 2 and 3). Maximum likelihood tree using Kimura two parameter distances (K2P) was created among the four Asteraceae germplasm, to provide a combined graphic representation of the patterns of divergences with ITS rDNA sequence data (Fig. 2). Within the group, two strongly supported clades were clearly distinguished among the four species of *Asteraceae* family. With regard to the first clade, *Achillea santolina* L. and *Artemisia monosperma* were grouped together in the first clade. Within the second clade, *Echinops spinosus* L. and *Matricaria recutita* L., were included in a sister clade. With respect to IGS rDNA sequence data (Fig. 3), *Achillea santolina* L., was closely related to *Matricaria recutita* L., in the first clade, while *Echinops spinosus* L. shared individually with the first clade. In contrast, *Artemisia monosperma* was placed independently in a separate clade (clade 4).

Taxonomic characterization leading to unambiguous identification of species and varieties is critically important for conservation and sustainable utilization of the *Asteraceae* germplasm. In *Asteraceae* family, molecular phylogeny at various taxonomic levels has been examined in several earlier studies through application of isozymes and RAPD (Ayers and Ryan, 1999), AFLP (Huang *et al.,* 2009), ISSR and RFLP (Abd El-Twab *et al.,* 2010), SSR (Iqbal *et al.,* 2011), as well as chloroplast DNA and rDNA markers (Sonboli *et al.,* 2012). In context, Dai *et al.* (2008) found that some closely related cultivars with identical ITS sequences in rice could be clearly discriminated based on the phylogenetic tree constructed by IGS sequences. In subsequent studies, (Plovanich and Panero, 2004; Dai *et al*., 2008; Li *et al*., 2010) confirmed that IGS sequences with the fastest rate of evolution could provide more hierarchical distinctions than ITS sequences. Therefore, it was concluded that the IGS region could be more suitable for measuring genetic relationship in different cultivars of subspecies, with more informational sites than ITS sequences in the *Asteraceae* germplasm.

3. Transition and transversion

It is a well-known fact that during DNA sequence evolution the rate of transitional changes differs from the rate of transversional changes, with transitions generally occurring more frequent than transversions. This difference is often referred to as transition bias, and estimation of the extent of transition bias may be of interest (Cortey *et al.,* 2011). In Table (4) the substation pattern and rates were estimated to compare the similarity matrix under the Tamura-Nei 93 test model (Tamura and Nei, 1993). The highest transition/ transversion rate ratios were recorded among IGS sequence data (k1 = 38.28, purines), (K2= 12.58, pyrimidines), respectively. Meanwhile, the lowest transition/ transversion rate ratios were observed among ITS sequence data (k1 = 2.983, purines), (K2= 2.746, pyrimidines), respectively. Moreover, the overall transition/transversion bias for IGS sequence data (R = 12.10) was superior compared to ITS sequence data (R = 1.43). This reflects that transitions are more dominant than transversion in *Asteraceae* germplasm across IGS sequence. This is compatible with the results of Wetzer (2001), who reported that transitions occur more frequently than trans-versions, even though for any given nucleotide position twice as many possible transversions may occur as transitions. In the context the results of Wang *et al.* (2011) elucidate that transitional substitutions at 3'UTR are more common than transversions and transitions are even more frequent than transversions at CpG sites compared with non-CpG sites. The Recent investigation by Kruger *et al.* (2012) elucidated that in a genome higher frequency of transition occurred than transversions substitutions.

In the existing study, our result from the IGS sequences confirmed the feasibility of utilizing these sequences for the study of species or intraspecies of *Asteraceae* germplasm than ITS sequence. Consequently, through previous results we can confirm that IGS sequence divergence seems to be the most appropriate regions as a significant molecular marker for classification, taxonomic and identification at the species level and beyond in *Asteraceae* germplasm.

In conclusion, the assessment of spacer length variation and rDNA polymorphisms in the rDNA genes in *Asteraceae* germplasm provides new insights in understanding the genetic variability among ecotypes and confirms that this is a useful region for genetic variability studies and phylogenetic relationships in *Asteraceae* germplasm. Despite the fact that *Asteraceae* germplasm is wild family, which has not yet been cultivated, its nutritional composition alone makes it an important resource. Therefore, focused research and development efforts are needed if this wild species can be raised from obscurity and improved sufficiently to contribute to the food supply in Egypt.

SUMMARY

Ribosomal DNA genes are organized in clusters of tandem repeated units, each of which consists of coding regions

(18S, 5.8S and 28S) and two internal transcribed spacers (ITS), in addition to intergenic spacer (IGS) region. Accordingly this article is focused on clarifying the sequence divergence of complete rDNA of ITS and IGS regions among four wild and endemic species of *Asteraceae* family in Egypt. Results indicated that there were a total of 754 and 667 positions across the final dataset for ITS and IGS sequences, respectively. IGS regions were superior compared to ITS region in several parameters like nucleotide diversity rate (π = 0.60), the estimated values of transition/transversion rate ratios (k1 = 38.28, purines), (K2 = 12.58, pyrimidines) and the overall transition/transversion bias (R = 12.10), respectively. This reflects that transitions are more dominant than transversion in *Asteraceae* germplasm across IGS markers. Thus, it was concluded that the IGS region could be more suitable for measuring genetic relationship in different subspecies of *Asteraceae*, with more informative sites than ITS sequences. Generally ribosomal DNA particularly intergenic spacer of the rDNA cluster evolves quickly and is highly polymorphic, providing a useful tool for assessing genetic diversity, taxonomic and phylogenetic studies in *Asteraceae* germplasm.

ACKNOWLEDGEMENTS

This work was financially supported by the open fund of the National Key Laboratory of Crop Genetic Improvement and the Bioversity International (Rome, Italy). Special thanks are given to Bioversity International for their valuable support. Special thanks to Mr. Wen-Fang Zeng, Mr. Nishawy, Mr. M. Eweas and Mr. Xin-Jian Zhang for their constructive comments and help.

REFERENCES

Abd El-Twab, M. H. and F. A Zahran (2010). RAPD, ISSR and RFLP analysis of phylogenetic relationships among congeneric species (*Anthemideae*, *Asteraceae*) in Egypt. Int. J. Bot., 6: 1-10.

Ayres, D. R. and F. J. Ryan (1999). Genetic diversity and structure of the narrow endemic *Wyethia reticulata* and its congener *W. bolanderi* (*Asteraceae*) using RAPD and allozyme techniques. Am. J. Bot., 86: 344-353.

Badr, A., W. Morsy, S. Abdelfattah, S. Shams and A. Shehab (2012). Genetic diversity in *Artemisia monosperma* and *Artemisia judaica* populations in Egypt based on morphological, karyological and molecular variations. J. Med. Plants Res., 6: 66-78.

Boulos, L. (2002). Flora of Egypt. 3: 146-230.

Chang, K. D., S. A. Fang, F. C. Chang and M. C. Chung (2010). Chromosomal conservation and sequence diversity of ribosomal RNA genes of two distant *Oryza* species. Genomics, 96: 181-190.

Cortey, M., L. Macera, J. Segale and T. Kekarainen (2011). Genetic variability and phylogeny of Torque teno sus virus 1 (TTSuV1) and 2 (TTSuV2) based on complete genomes. Vet. Microbiol., 148: 125-131.

Czarnecki, D. M., M. N. Rao, J. G. Norcini, Jr. F. G. Gmitter and Z. Deng (2008). Genetic diversity and differentiation among natural, production, and introduced populations of the narrowly endemic species coreopsis leaven worth (*Asteraceae*). J. Amer. Soc. Hort. Sci., 133: 234-241.

Dabert, M. (2006). DNA markers in the phylogenetics of the Acari. Biol. Lett., 43: 97-107.

Dai, X. J., L. J. Ou, W. J. Li, M. J. Liang and L. B. Chen (2008). Analysis of rDNA intergenic spacer (IGS) sequences in *Oryza sativa* L. and their phylogenetic implications. Acta Agro. Sin., 34: 1569-1573.

Dutta, S. K. and M. Verma (1990). Primary structure of the nontranscribed spacer region and flanking sequences of the ribosomal DNA of *Neuropora crassa* and comparison with other organisms. Biochem. Biophys. Res. Commun., 170: 187-193.

Gharibi, S., M. Rahimmalek, A. Mirlohi, M. M. Majidi and B. E. S. Tabatabae (2011). Assessment of genetic diversity in *Achillea millefolium* subsp. *millefolium* and *Achillea millefolium* subsp. *elbursensis* using morphological and ISSR markers. J. Med. Plant. Res., 5: 2413-2423.

Giaretta, D. R., A. Bogo, C. M. M. Coelho, A. F. Guidolin, A. C. M. Dantas and E. A. Gomes (2010). ITS-rDNA phylogeny of *Colletotrichum* spp. causal agent of apple glomerella leaf spot. Ciência Rural, Santa Maria., 40: 806-812.

Hillis, D. M. and M. T. Dixon (1991). Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol., 66: 411-453.

Huang, W. K., F. H. Wan, J. Y. Guo, B. D. Gao, B. Y. Xie and D. L. Peng (2009). AFLP analyses of genetic variation of *Eupatorium adenophorum* (Asteraceae) populations in China. Can. J. Plant Sci., 89: 119-126.

Iqbal, A., H. A. Sadaqat, A. S. Khan and M. Amjad (2011). Identification of sunflower (*Helianthus annuus*, *Asteraceae*) hybrids using simple-sequence repeat markers. Genet. Mol. Res., 10: 102-106.

Ito, M., T. Yahara, R. M. King, K. Watanabe, S. Oshita, J. Yokoyama and D. J. Crawford (2000). Molecular phylogeny of *Eupatorieae* (*Asteraceae*) estimated from cpDNA RFLP and its implieation

for the polyploid origin hypothesis of the tribe. J. Plant Res., 113: 91-96.

Jo, S. H., D. H. Koo, J. F. Kim, C. G. Hur, S. Lee, T. Yang, S. Y. Kwon and D. Choi (2009). Evolution of ribosomal DNA-derived satellite repeats in tomato genome. BMC Genet., 9: 42.

Kimura, M. A. (1980). Simple method for estimating rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol., 16: 111-120.

Kollipara, K. P., R. J. Singh and T. Hymowitz (1997). Phylogenetic and genomic relationships in the genus Glycine Willd, based on sequences from the ITS region of nuclear rDNA. Genome, 49: 57-68.

Kruger, D., D. Kapturska, C. Fischer, R. Daniel and T. Wubet (2012). Diversity measures in environmental sequences are highly dependent on alignment quality data from ITS and new LSU primers targeting Basidiomycetes. PLOS ONE, 7: e32139.

Li, Y., S. Shen, L. He, P. Xu and S. Lu (2010). Sequence analysis of rDNA intergenic spacer (IGS) of *Porphyra haitanensis*. J. Appl. Phycol., 22: 187-193.

Moss, T. and V. Y. Stefanovsky (1995). Promotion and regulation of ribosomal transcription in eukaryotes by RNA polymerase I. Prog. Nucleic Acid Res. Mol. Biol., 50: 25-66.

Pirttila, A. M., M. Hirsikorpi, T. Kamarainen, L. Jaakol and A. Hohtola (2001). DNA Isolation Methods for Medicinal and Aromatic Plants. Plant Mol. Biol. Rep., 19: 273.

Plovanich, A. E. and J. L. Panero (2004). A phylogency of the ITS and ETS for Montanoa (*Asteraceae*: *Heliantheae*). Mol. Phylogenet. Evol., 31: 815-821.

Simko, I. (2009). Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.). J. Hered. 100: 256-262.

Singh, A., R. M. Deverumath, S. RamaRao, V. P. Singh and S. N. Raina (2008). Assessment of genetic diversity, and phylogenetic relationships based on ribosomal DNA repeat unit length variation and Internal Transcribed Spacer (ITS) sequences in chickpea (*Cicer arietinum*) cultivars and its wild species. Genet. Resour Crop Evol., 55: 65-79.

Sonboli, A., K. Stroka, S. O. Kazempour and C. Oberprieler (2012). Molecular phylogeny and taxonomy of *Tanacetum* L. (Compositae, Anthemideae) inferred from

nrDNA ITS and cpDNA trnH-psbA sequence variation. Plant Syst. Evol., 298: 431-444.

Tajima, F. (1989). Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics, 123: 585-595.

Tamura, K. and M. Nei (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol., 10: 512-526.

Tamura, K., M. Nei and S. Kumar (2004). Prospects for inferring very large phylogenies by using the Neighbor-Joining method. Proc. Natl. Acad. Sci. USA., 101: 11030-11035.

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

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res., 25: 4876-4882.

Wang, J. B., W. J. Zhang and J. K. Chen (1999). Application of ITS sequences of nuclear rDNA in phylogenetic and evolutionary studies of angiosperms. Acta Phytotaxon Sin., 37: 407-416.

Wang, X. W., J. Luan, J. Li, Y. L. Su, J. Xia and S. Liu (2011). Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. BMC Genomics, 12: 458.

Wei, N. V., C. C. Wallace, C. F. Dai, K. R. M. Pillay and C. A. Chen (2010). Analyses of the ribosomal internal transcribed spacers (ITS) and the 5.8S gene indicate that extremely high rDNA heterogeneity is a unique feature in the *Scleractinian* coral genus *Acropora* (*Scleractinia*; *Acroporidae*). Zool. Studies., 45: 404-418.

Wetzer, R. (2001). Hierarchical analysis of mtDNA variation and the use of mtDNA for isopod (Crustacea: Peracarida: Isopoda) systematics. Contributions to Zoology 70: 23-39.

White, T. J., T. Bruns, S. Lee and J. Taylor (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols:A guide to methods and applications. (Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds). New York Academic Press, 315-322.

Zhao, H., F. Chen, S. Chen, G. S. Wu and W. Guo (2010). Molecular phylogeny of *Chrysanthenum*, *Ajania* and its allies (*Anthemideae*, *Asteraceae*) as inferred from nuclear ribosomal ITS and chloroplast *tm*L-F IGS sequences. Plant Syst. Evol., 284: 153-169.

Table (1): The eco-geographical localities of the four natural habitats of *Asteraceae* family.

Scientific name	Code	Location name	G. P. S. location			ITS similarity (%)	IGS similarity (%)
			Latitude	Longitude	Altitude		
Echinops spinosus L.	1	Rafah	31 15 57.1 N	034 08 57.7 E	270 F	AY538629 99	EU649670 96
Achillea santolina L.	2	El Gora	31 07 56.7 N	034 08 11.4 E	270 F	EU179212 99	AB359892 94
Matricaria recutita L.	3	Elsheik Zowaied	31 14 06.9 N	034 06 56.6 E	85 F	AY603253 99	GU818112 95
Artemisia monosperma L.	4	Elsheik Zowaied costal	31 13 59.8 N	034 06 55.3 E	97 F	GU724289 99	AB359871 94

Table (2): The evolutionary analyses using Tajima test among ITS and IGS sequences.

Scientific name	ITS nucleotide frequency					AT	CG	Scientific name	IGS nucleotide frequency					AT	CG
	T(U)	A	G	C	Total				T(U)	A	G	C	Total		
Achillea santolina L.	23.5	25.3	23.9	27.4	712	49.0	51.0	*Achillea santolina* L.	33.4	20.3	17.8	28.4	640	53.8	46.3
Artemisia monosperma L.	22.1	22.5	27.7	27.7	729	45.0	46.0	*Artemisia monosperma* L.	17.8	28.6	33.6	20.0	640	46.4	53.6
Avg.	24 %	25 %	25%	26 %	719	49 %	51 %	Avg.	26.7%	25.1%	22.1%	26.1%	571.3	51.8%	48.2%
Echinops spinosus L.	23.9	23.5	28.5	24.2	724	47.0	53.0	*Echinops spinosus* L.	21.9	31.5	19.2	27.4	365	53.4	46.6
Matricaria recutita L.	26.3	24.0	25.8	23.9	712	50.0	50.0	*Matricaria recutita* L.	33.6	20.0	17.8	28.6	640	53.6	46.4
Max.	26.3	25.3	28.5	27.7	729.0	50.0	53.0		33.6	31.5	33.6	28.6	640.0	53.8	53.6
Min.	22.1	22.5	23.9	23.9	712.0	45.0	46.0		17.8	20.0	17.8	20.0	365.0	46.4	46.3
Total					2877								2285		

Table (3): Nucleotide frequencies for ITS and IGS sequence among four natural habitats of *Asteraceae* family.

Sequence type	Number of Sites (M)	Number of positions (N)	Number of Segregating (S)	Nucleotide Diversity (π)
ITS	4	754	530	0.49
IGS	4	667	595	0.60

Table (4): Maximum composite likelihood estimate of the pattern of nucleotide substitution matrix for combined data of ITS and IGS sequences. Each entry is the probability of substitution (r) from one base (row) to another base (column). The rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics.

	ITS				IGS			
	A	T	C	G	A	T	C	G
A	-	*4.92*	*5.33*	**16.1**	-	*0.99*	*0.96*	**37.02**
T	*4.91*	-	**14.62**	*5.4*	*0.81*	-	**12.11**	*0.97*
C	*4.91*	**13.51**	-	*5.4*	*0.81*	**12.43**	-	*0.97*
G	**14.66**	*4.92*	*5.33*	-	**30.98**	*0.99*	*0.96*	-
	k1 = 2.983	*k2* = 2.746	*R* = 1.43		*k1* = 38.28	*k2* = 12.58	*R* = 12.10	
Over all mean distance = 286				*SE* = 8.15	Over all mean distance = 160			*SE* = 4.94

(*k1*) (purines) = Transition rate ratios, (*K2*) (pyrimidines) = Transversion rate ratios.
(*R*) = Transition/transversion bias.

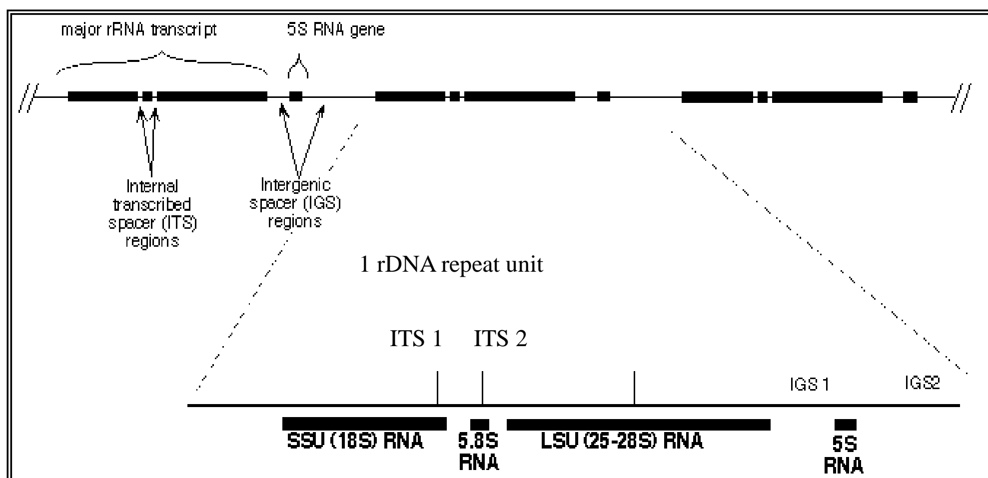

Fig. (1): The elementary structure of (rDNA) of plants.

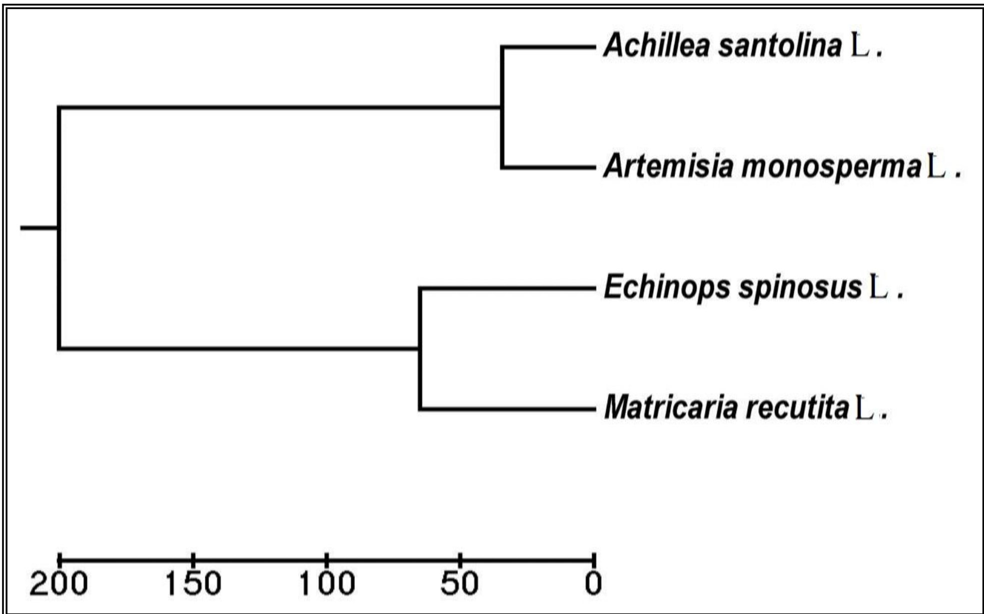

Fig. (2): ML tree generated among four *Asteraceae* genotypes based on ITS rDNA data.

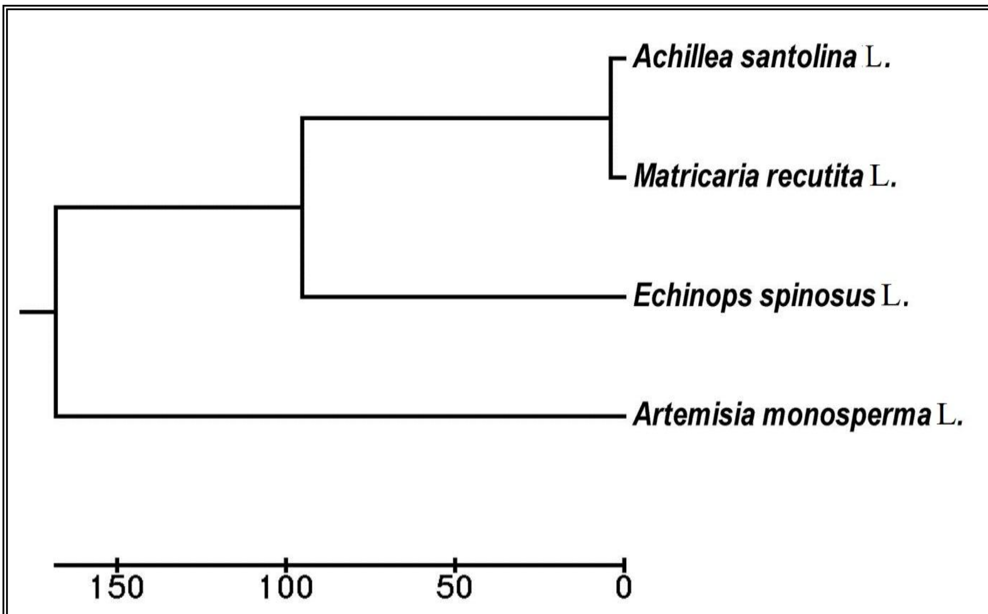

Fig. (3): ML tree generated among four *Asteraceae* genotypes based on IGS rDNA data.