IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)
ISSN: 2455-264X, Volume 6, Issue 6 (Nov. – Dec. 2020), PP 12-23
www.iosrjournals.org

Molecular Identification and Phylogenetic Relationship of some *Origanum* and *Thymus* Species Based on ITS Regions

Mohamed Zoelfakar Sayed Ahmed[1], El-Shaimaa Saad El-Demerdash[1] and Shafik Darwish Ibrahim[2]

[1](*Genetic Resources Department, Desert Research Center (DRC), 1, Mathaf El-Matariya Street, El-Matariya B.O.P 11753 El-Matariya, Cairo, Egypt.)*

[2](*Molecular Genetics and Genome Mapping Lab., AGERI, ARC, 9 Gamma St., Giza – Cairo 12619, Egypt.)*

Abstract:

Background: *Lamiaceae or Labiatae family is one of the major important plant families of multiple usesin aromatic purposes, medicine and food. Oregano (Origanum) and thyme (Thymus) the scope of our study are belonging to familyLamiaceae. Butin spite of their importance, they are poorly identified on the basis of molecular levels three Thymus species (T. vulgaris, T. capitatus and T. decussatus)and two Origanum species (O. vulgareand O. syriacum L., subsp. sinaicum) were chosen for the preset study. Molecular identification and characterization studies based on DNA molecular marker (ITS region) are more precise, reliable and powerful tool to assess the phylogenetic relationships between studied plant species with 17 genera in Lamiaceae family.*

Materials and Methods:*Specific one fragment ofPCR product about 710±15 bp from was produced using the universal primer (ITS1 and ITS4) with highly conserved of ITS regions of rDNA through the 5 samples under study and sequencing of the obtained fragment was conducted.*

Results:*The sequence lengths of the ITS region of three Thymus species were 685bp, 681bp and 680bp with T. vulgaris (MT974175.1), T. capitatus(MT974176.1) and T. decussatus (MT974177.1) and two Origanum species were 676bp and 680bp nucleotides with O. vulgare (MT974178.1) and O. syriacum L., subsp. sinaicum (MT974179.1), respectively. Each sequence was aligned individually at BLAST to confirm each species and to determine its relation to other sequences. Multiple sequence alignment (MSA) involved 56 nucleotide sequences of nuclear (ITS region) was done using MEGA7 software by Maximum Likelihood method with the highest log likelihood (-3348.45) and phylogenetic relationship showed that 3 Thymus and 2 Origanum species were closely related to Thymus and Origanum genus and they were grouped in one main cluster with 17 generabelonging to the family (Lamiaceae). These results indicated thatmolecular studies based on thesequence of ITS regions are important for confirmation of the phenotypic characterization of 3 Thymus and 2 Origanum species inEgypt.*

Conclusion:*Molecular sequence markers (DNA barcoding) is a reliable technique at molecular levelin the field of Identification, classification and differentiation between wild Thymus genus, Origanum genus and other generabelonging to family Lamiaceae as medicinal plants and thus, it is an endangered plant that needs a strict conservation plan.The alignments of 56 nucleotide sequences of the nuclear ITS region (ITS1, 5.8S and ITS2) with: T. vulgaris, T. capitatus, T. decussatus, O. vulgare and O. syriacum L. subsp. sinaicum were used and able to distinguish species under research with high similarities and to the closely related species of the Thymusgenus, Origanum genus and other 17 genera belonging to family Lamiaceae and take them away from four different plant species of different families as out group of the family Lamiaceae. Furthermore, phylogenetic trees were constructed based on 56 nucleotide sequences encoded from the ITS region usingMEGA 7 program, includingunder investigation species that showed highest percentages of similarity with species from 17 genera of family Lamiaceae clustered together, while they showed low percentages of similarities towards relatively distantly related species as out group of the family.*

Key Word*:Thymus vulgaris; Thymus capitatus;Thymus decussatus; Origanum vulgare; Origanum syriacum L. subsp. sinaicum;nuclear ITS region; DNA barcoding.*

Date of Submission: 10-12-2020 Date of Acceptance: 25-12-2020

I. Introduction

The genus *Origanum* and *Thymus*, are important multipurpose as aromatic perennial herbs used in folk medicine food additives since ancient times. They are belonging to the mint family (*Lamiaceae* or *Labiatae*)and they are considered one of the most economically important herbs [18]. The *Lamiaceae* family is one of the largest families among the dicotyledons. It has been classified into seven subfamilies;Nepetoideae is the largest subfamily of *Lamiaceae*, including numerous familiar plants such as oregano (*Origanum*) and thyme (*Thymus*), the scope of our study [12]. In the past few years,several molecular marker types showed usefulness in studying

genetic diversity and phylogenetic relationship for several plant species such as Restriction fragment length polymorphism (RFLP); Randomly Amplified Polymorphic DNA(RAPD);Simple Sequence Repeats(SSR); Inter Simple Sequence Repeat(ISSR); start codon targeted markers (SCoT) andAmplified Fragment Length Polymorphism (AFLP). These traditional approaches generally require high quality DNA for amplification, which can lead to problems with reproducibility and accuracy [3]. In addition, experimental error and subjective factors such as scoringPCR bands on a gel are eliminated or minimized in sequence based protocols.Thus, molecular sequence markers (DNA barcoding)have become wieldy used in species classification and identification, dependent on sequencing from gene to genomeproving to be an effective tool that enables rapid and accurate identification of plant species using standardized DNA markers [7, 15].

In Eukaryotic cells, Ribosomal RNA organizes into two ribosomal subunits: the large ribosomal subunit (LSU) and small ribosomal subunit (SSU) [19].The Internal Transcribed Region (ITS) one of the most used polymorphic regions it is a space of noncoding RNA sequence on a common of primary transcript of RNA and has been removed by splicing during RNA processing [4]. Ribosomal RNA genes (known as ribosomal DNA or rDNA) are a type of non-coding RNA and parts of repeat units that are arranged in tandem arrays. Nuclear ribosomal DNA has ITS1, ITS2and the 5.8S subunit are collectively known as the internal transcribed spacer (ITS) region [9.10.11]. The ITS region of rDNA (600-700bp) repeats are highly conserved and are one of the most successfully used to amplify using universal PCR primersof the nuclear genome in studying phylogenetic genomic relationships of plants at lower taxonomic levels [1,2]. Furthermore, DNA barcoding is one of the widely accepted technologies based on nucleotide sequence for the identification of multiple species according to CBOL[5]. ITS1 and ITS2 locusas a single individual ofpolymorphisms among repeated units are widely used universal primers for various kinds of organisms and phylogeny reconstruction to compare the maternally inherited chloroplast and mitochondrial markers for evolutionary studies in taxonomic and phylogenetic analyses[16, 21].The ITS2 locuswas widely used as a universal DNA barcode and has been tested as the optimal DNA barcode for identifying plant species [23].DNA barcoding is an efficient, cheap and fast standard method for evaluation and identification of different plant species [13]. The development of DNA-based markers became important for the authentication of medicinal plants [20]. DNA barcoding system used to give correct authentication and identification of plant systematic at the molecular level and ITS2 sequence-structure information could be used as an alternative mini-barcode when a full-length ITS is not present [24]. New universal primers of plant-specificwere designed for PCR amplifying the whole ITS region to test 335 samples[6].The levels of ITS2 sequence variation from 17,203 sequences, representing 5,439 species in 113 genera of seed plants were examined, to ascertain the association between species status and their molecular divergence [17]. The ITS2 region is a powerful barcoding region for identifying Chinese *Zanthoxylum* species, and will be useful for phylogenetic analyzing and managing Chinese *Zanthoxylum* germplasm collections [25].Theviability and the potentiality of the ITS and *rbcL* markers were tested for the identification of *Solanumnigrum*and *Solanumvillosum*, as medicinal plants in Egypt [8].

The present study aims to amplify and sequence nuclear ITS regions among three wild types (two Thymus species and one Origanum species), two cultivated types (one Thymus and one Origanum species).Furthermore, to study functional annotation and homology modeling of sequences ITS regions among these species using Basic Local Alignment Search Tool (BLAST).Moreover,to identify plant species under study at molecular level. In addition, the main objective of this study is to utilize the sequences of ITS region as useful molecular markers to determine the genetic relationships with 17 genera belonging to family *Lamiaceae* using MEGA 7 software and exploring the evolutionary relationships.Also, to detect the efficiency of ITS markers in identification and discrimination of the studied plant species. A comprehensive study was designed to test ITS regions for authentication and phylogenetic relationship investigation of 3 *Thymus* and 2 *Origanum* species as medicinal plants from Egypt.

II. Material and Methods

2.1 Plant Materials collection:

In Egypt, three *Thymus* species (*T. vulgaris*Benth., *T. capitatus*, *T. decussatus* Benth.)and two *Origanum* species (*O. vulgare* L. and *O. syriacum* L., subsp. *sinaicum* Boiss.) were chosen for the present investigation. The fresh young leaves were collected as bulk for each species in spring of 2018, transferred into liquid nitrogen, and frozen at - 80 °C till use. *Thymus vulgaris* Benth. and*Origanum vulgare* L. (cultivated types) were collected from private farms at Kirdasa region, Giza Governorate, Egypt. For two wild types: *Thymus capitatus* (L.) Hoffmanns& Link was collected from 26 Km (26 kilometers) from the North Coast, MarsaMatrouh Governorate, Egypt and the other wild types(*Thymus decussatus* Benth.) and *Origanum syriacum* L. subsp. *sinaicum* (Boiss.) were collected from the top one of the mountain of Saint Catherine Protectorate, South Sinai Governorate, Egypt.

2.2 DNA extraction, primers design and PCR amplification:

Genomic DNAs were extracted from100mg fine powdered tissue from 5 samples by DNeasy Plant Mini Kit (Qiagen Inc., Cat.no.69104, USA) this was performed following the manufacturer's instructions. The final concentrations of 35-40 ng/µl DNA were amplified via standard polymerase chain reaction (PCR). One pair primers (ITS1 and ITS4) were selected from the literature and the BOLD primer dataset platform (http://www.boldsystems.org/index.php/Public_Primer_PrimerSearch), The entire noncoding nuclear(ITS1: (5`-TCC GTA GGT GAA CCT GCG G-3`) and ITS4: (5`-TCC TCC GC TTA TTG ATA TGC-3`) PCR fragment a 710±15 bp of the ITS region according to [22] and CBOL [5].

Phusion®*Taq*, the High-Fidelity DNA polymerase (Thermo Scientific, Product codes: F-530L, 500 Unit) was used. For ITS region master mix: The amplification reaction was carried out in a 50 µl total volume containing; 4 µl cDNA, 10 µl 5X Phusion HF Buffer, 1 µl 10mM dNTP mix, 2.5 µl Fwd_primer ITS1(10 µM),2.5 µl Rev_primer ITS4(10 µM), 0.5 µl PhusionDNA polymerase, 29.5 µl DEPC H_2O and spin for 15 Sec. The reaction mixture was incubated in a Perkin Elmer thermocycler 9700. PCR program for ITSregions, the temperature profile in different cycles was as follows: an initial strand separation cycle at 98°C for 3 min followed by 35 cycles comprised of a denaturation step at 98°C for 30 seconds, an annealing step at 58°C for 30 seconds and an extension step at 72°C for 45 seconds. The final cycle was a polymerization cycle for 7 min at 72°C. PCR product purification was used QIAquick gel extraction kit (Qiagen, Cat. No. 28704) from the agarose gel up to 10 µg can bind to each QIAquick column.

2.3 Sequence editing, alignment and phylogenetic inference:

Sequencing chromatograms 5 ITS regions obtained were analyzed by Macrogen, Seoul, south of Korea for each studied Three *Thymus* species and two *Origanum* species. All nucleotide sequences of 5 ITS regions were searched in The National Center for Biotechnology Information GenBank Database (NCBI), (http://www.ncbi.nlm.nih.gov). The homology searches were performed with Basic Local Alignment Search Tool of several sequences (BLASTn online program) their homologies with sequences published in DDBJ/EMBL/GenBank database which are available using the NCBI.

Constructing of the phylogenetic tree and investigation of the relationships were carried out using alignment among the sequences of 56 ITS regions from GenBank NCBI database including 5 ITS regions of *Thymus* and *Origanum* species, other species belonging to *Lamiaceae* family and different species with *Linum usitatissimum* (MK066769.1), *Glycine max* (MT230614.1), *Chenopodium album* (MK953683.1) and *Arabidopsis thaliana* (MG886683.1) from different families as out group. Phylogenetic tree analysis tree was performed using the Maximum Likelihood (ML) tree methods to detect positive selection under site models by the MEGA 7.0 software program [14].

III. Result and Discussion

The ITS region (ITS1, 5.8S and ITS2) as DNA barcodeshas been found to be an effective universal region andhas proved to be easily amplified and sequenced with the currently established primers for accurate identification of plant systematic at the molecular level and in the verification of herbal medicines [7, 15, 24] as shown in Figure 1.Total genomic DNA was extracted from5 samples.With universal primers ITS1 and ITS4 on conserved regions (ITS region), we observed good results of one fragment of PCR amplification about 710±15 bp of the ITS region according to CBOL [5] and [22].PCR products werepurified and sequenced. The isolated nucleotide sequences have been uploaded to the GenBankdatabase NCBI (National Center for Biotechnology Information) by BankIt, online from the website (https://www.ncbi.nlm.nih.gov/WebSub/) after it was processed by email (gb-admin@ncbi.nlm.nih.gov).

The sequence length of theITS region available at the DDBJ/EMBL/GenBank database of three *Thymus* species were 685, 681 and 680 bp with *Thymus vulgaris* (MT974175.1), T*hymuscapitatus* (MT974176.1) and *Thymusdecussatus*(MT9741777.1), respectively. In the same context, sequenceslength of the ITS region available forthe two *Origanum* species 676bp and 680bp with *Origanum vulgare*L. (MT974178.1) and *Origanum syriacum* L. subsp. *Sinaicum* (MT974179.1), respectively as shown in Table 1 and Figure (2 & 3).

Molecular phylogenetic analysis based on sequences of ITS region:

Sequence homology of each nucleotide of the nuclear (ITS region) for 56 sequences in NCBI including 5 sequences of current study against the sequences of 51 other species from 17 genera belonging to family *Lamiaceae* or *Labiatae* such as: (Code1) for*Thymus vulgaris*, (Code2) for*Thymus capitatus*, (Code3) for*Thymus decussatus*, (Code4) for*Origanum vulgare*, (Code5) for*Origanum syriacum* L. subsp. *Sinaicum* as current study, (Code 6 to 14) for*Thymus genus*, (Code 15 to 22) for*Origanum* genus, (Code 23 to 24) for*Conradina* genus,(Code 25 to 26), for*Dicerandra* genus, (Code 27 to 28)for*Monarda* genus, (Code 29 to 30) for*Pycnanthemum*genus, (Code 31 to 32) for*Bystropogon*genus, (Code 33 to 34) for*Clinopodium*genus, (Code 35 to 36) for *Satureja* genus, (Code 37 to 38) for *Mentha* genus, (Code 39 to 40) for *Agastache* genus, (Code 41

to 42) for*Dracocephalum* genus, (Code 43 to 44) for*Nepeta*genus,(Code 45 to 46) for *Prunella* genus,(Code 47 to 48) for*Lepechinia* genus, (Code 49 to 50) for*Rosmarinus*genus, (Code 51 to 52) for *Saliva* genusand (Code 53 to 56) for 4 different species *Linum usitatissimum* (MK066769.1), *Glycine max* (MT230614.1), *Chenopodium album* (MK953683.1) and *Arabidopsis thaliana* (MG886683.1) as out group of the family were tested by top-scoring hits through NCBI database using Basic Local Alignment Search Tool (BLASTn) as shown in Table 2. Multiple sequence alignments (MSA) of 56 ITS region were carried out between the selected nucleotide sequences of two genera (*Origanum* and *Thymus* genus)including the 3 *Thymus* and 2 *Origanum* species, *(T. vulgaris, T.capitatus, T.decussatus, O.* vulgare, *O. syriacum* subsp. *sinaicum*)and 4 species from different families as an out-group available in GenBank databases, we will discuss it in brief.

The results revealed that *Thymus vulgaris* (MT974175.1, current study) was closely related with high identity and similarity (99.27-93.24%), to 11 genus *Thymus*, (92.33-90.99%) to 10 genus *Origanum*, withE-value = Zero and low similarity with other 15 genera (93.61-82.57%) withE-value =Zero, While similarity 85.02% *Dracocephalummoldavica* (MH710906.1) withE-value =8e-169, 82.57% *Dracocephalum grandiflorum*(AJ420999.1) with E=5e-166, 85.60% *Nepetalaxiflora*(AJ420995.1) withE-value =3e-178, 83.38% *Prunella vulgaris* (MH711205.1) withE-value =1e-177, 85.80 *Prunella grandiflora* (HQ228231.1) withE-value =6e-159, 83.74% *Salvia rosmarinus* (MH645776.1) withE-value =2e-180. In contrast, it showed low similarities with the other 4 different species as out group: 94.86% *Linum usitatissimum* (MK066769.1) withE-value =3e-77, 96.99% Glycine max (MT230614.1) withE-value =2e-79, 87.46% *Chenopodium album* (MK953683.1) withE-value =1e-92, 94.92% *Arabidopsis thaliana* (MG886683.1) withE-value =3e-78.

The results revealed that *Thymus capitatus* (MT974176.1, current study) was closely related with high identity and similarity (99.05-92.59%) to 11 genus *Thymus*, (94.55-93.10%) to 10 genus *Origanum*, withE-value = Zero and showed similarity (98.37-84.37%) with the other 15 genera withE-value = Zero. While similarity 85.69% to*Nepetalaxiflora* (AJ420995.1) withE-value =2e-180. On the other hand, it showed similarities with the other 4 different species as out group: 93.55% *Linum usitatissimum* (MK066769.1) withE-value =9e-78, 96.99% *Glycine max* (MT230614.1) withE-value =2e-79, 86.21% *Chenopodium album* (MK953683.1) withE-value =3e-88, 94.92% *Arabidopsis thaliana* (MG886683.1) withE-value =3e-78.

The results revealed that *Thymus decussatus* (MT974177.1, current study) revealed high similarity (99.11-93.44%) to 11 genus *Thymus*, (93.24-91.94%) 10 genus *Origanum*, withE-value = Zero and similarity (93.93-83.12%) with other 15 genera withE-value = Zero. While similarity 83.38% *Dracocephalum moldavica*(MH710906.1) withE-value =2e-175, 83.12% *Dracocephalum grandiflorum*(AJ420999.1) with E=5e-171, 86.05% *Nepetamirzayanii*(AJ515309.1) withE-value =7e-180, 83.64% *Nepetalaxiflora*(AJ420995.1) withE-value =3e-179, 86.04% *Prunella grandiflora* (HQ228231.1) withE-value =4e-161. However, it showed similarities with the other 4 different species as out group: 94.41% *Linum usitatissimum* (MK066769.1) withE-value =3e-77, 96.99% *Glycine max* (MT230614.1) withE-value =2e-79, 87.46% *Chenopodium album* (MK953683.1) withE-value =1e-92, 94.92% *Arabidopsis thaliana* (MG886683.1) withE-value =3e-78.

The results revealed that *Origanum vulgare* L.(MT974178.1, current study) showed high similarity (94.64 - 90.64%) to 12 genus *Thymus*, (100.00 - 97.35%) to 9 genus *Origanum*, withE-value = Zero and similarity (94.61-83.92%) with other 15 genera withE-value = Zero. While, 83.53% *Dracocephalum moldavica* (MH710906.1) with E-value =1e-176, 82.87% *Dracocephalum grandiflorum* (AJ420999.1) E-value =3e-168, 85.59%*Nepetamirzayanii* (AJ515309.1) E-value =2e-169, 82.34%*Nepetalaxiflora* (AJ420995.1) E-value =3e-164, 86.39% *Prunella grandiflora* (HQ228231.1) E-value =2e-164. On the other hand, it showed similarities with the other 4 different species as out group: 94.97% *Linum usitatissimum* (MK066769.1) withE-value =7e-79, 96.77% *Glycine max* (MT230614.1) withE-value =1e-77, 86.01% *Chenopodium album* (MK953683.1) withE-value =5e-86, 94.35% *Arabidopsis thaliana* (MG886683.1) withE-value =1e-76.

The results revealed that *O. syriacum* L. subsp. *sinaicum* (MT974179.1, current study) showed high similarity (93.37-89.37%) to 12 genus*Thymus*, (100.00-92.40%) to 9 genus *Origanum*, withE-value = Zero and similarity with other 15 genera (92.99-82.42%) withE-value =Zero. While, 85.02%*Conradinabrevifolia* (AY943461.1) with E-value= 7e-174, 84.85%*Conradinaetonia*(AF369165.1) E-value= 9e-173, 88.14%*Dicerandrachristmanii* (DQ860350.1) E-value= 4e-176, 83.07% *Dracocephalum moldavica* (MH710906.1) with E-value= 3e-173, 82.42% *Dracocephalum grandiflorum* (AJ420999.1) with E-value= 9e-164, 85.97%*Nepetamirzayanii* (AJ515309.1) with E-value= 9e-169, 85.38%*Nepetalaxiflora* (AJ420995.1) with E-value= 6e-161, 82.97% *Prunella vulgaris* (MH711205.1) with E-value= 1e-177, 85.71% *Prunella grandiflora* (HQ228231.1)with E-value= 2e-158, 82.93% *Salvia rosmarinus* (MH645776.1) with E-value 1e-176. In contrast, it showed similarities with the other 4 different species as out group: 94.97% *Linum usitatissimum* (MK066769.1) withE-value =7e-79, 96.39% *Glycine max* (MT230614.1) withE-value =1e-77, 86.81% *Chenopodium album* (MK953683.1) withE-value =2e-89, 94.35% *Arabidopsis thaliana* (MG886683.1) withE-value =2e-76.

The sequences alignment of the 3 *Thymus* species and 2 *Origanum* species for ITS1, 5.8S and ITS2 regions were obtained. The results revealed that ITS1 and ITS2 parts withconserved and with high conserved

with 5.8S region of two genera as shown in Figures (4, 5 & 6).The ITS region barcode sequencing has been used for the identification of medicinal plants at the molecular level. Our results were in agreement with several authors such as [20, 25]. The internal transcribed spacer (ITS) region and/or its subregion (ITS2) have been proposed as a standard DNA barcode marker in plants [15]. The most useful for discrimination in the ITS2 region with a correct identification rate of 100%, also exhibited significantly higher intra and interspecific divergence in Chinese *Zanthoxylum* [25]. The efficiency of the ITS region more than *rbcL*, where *rbcL*was confirmed the identification of *Solanumnigrum*and *Solanumvillosum*, as medicinal plants in Egypt at the generic level, while ITS region at the species level[8].

Weobserved themotif which described in other groups was found in our species under study such as 5`AAGGAA3` in the central region of ITS1 sequence previously described for flowering plants was also present in oursequences,in location 163 and 168as shown in Figures 4 and was helpful in aligning the 5 ITS sequences for phylogenetic analysis.The nuclear ribosomal internal transcribed spacer 1 (ITS1) region presented a higher discrimination power, higher values of variable and informative characters than the other datasets (ITS2 region) showed higher numbers of conserved sites in Figure 6 and also this region should be used as a starting point to identify *Thymus and Origanum* species according to [10] and [11].High conservative patterns in ITS2 sequences have been previously reported and related to structural constraints that are present at very deep phylogenetic scales in eukaryotes [23, 25].We also observed about 155 base pair (nucleotide) withthe highly conserved motif in the central region of the 5.8S sequences has been reported (Figure5). In molecular biology, the 5.8S ribosomal RNA (5.8S rRNA) is a non-coding RNA component of the large subunit of the eukaryotic ribosome and so plays an important role in protein translation by RNA polymerase I. The 5.8S rRNA can be used as a reference gene for miRNA detection. The 5.8S ribosomal RNA is used to better understand other rRNA processes and pathways in the cell. The 5.8S rRNA is initially transcribed along with the 18S and 28S rRNA in the 45S preribosomal RNA, along with the ITS 1 and ITS 2 and a 5' and 3' ETS (External transcribed spacer).

Based on the results, phylogenetic relationships analyses enable us to check the close relationships of species from 17 different genera belonging to the family *Lamiaceae* which included 3 *Thymus* species in this investigation (*T. vulgaris*, *T. capitatus* and *T. decussatus*) and2 *Origanum* species (*O. vulgare*and *O. syriacum* subsp.sinaicum) were conducted in MEGA 7.0 software program [14] by Maximum Likelihood (ML) tree with the highest log likelihood (-3348.45).The phylogenetic trees were constructed based on 56 nucleotide sequences encoded from the ITS regionused, including our speciesthat have the highest percentages of similarity with species from 17 different genera, showing closely related species clustering together from the family *Lamiaceae* and they showed low percentages of similarities towards relatively distantly related species(4species) as out group of the family (Figure 7).

IV. Figures and Tables

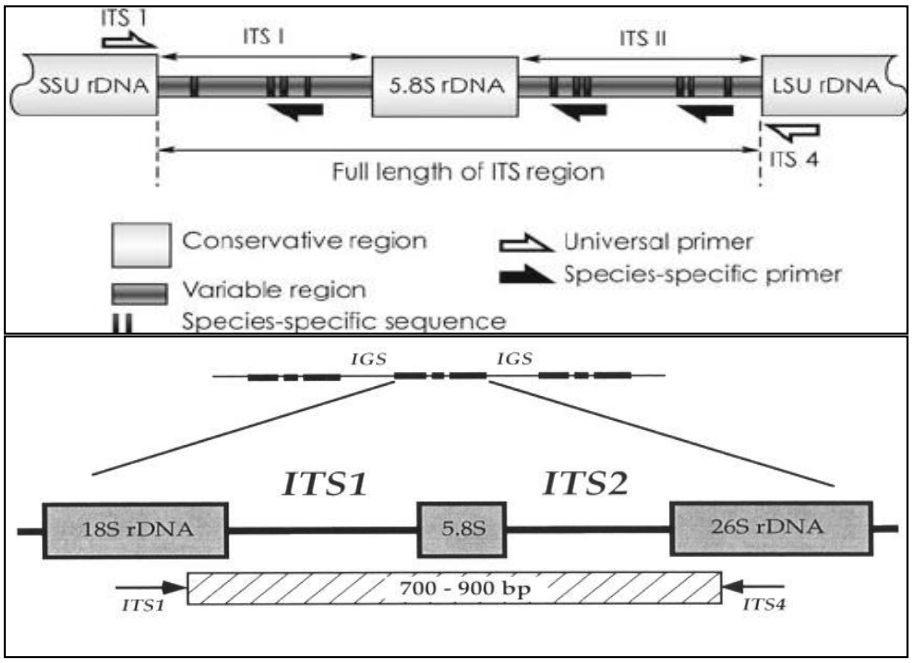

Figure 1: The elementary structure of (rDNA) of plant. Forward primer was ITS1and ITS4 was the reverse primer. Size length of ITS region was ITS1: (208± 56bp), 5.8S (83± 2bp) and ITS2 310± 67bp).

Table 1: Nucleotide sequences length of the nuclear ITS region (ITS1, 5.8S and ITS2) with: *T. vulgaris, T.capitatus,T. decussatus,O. vulgare* and *O. syriacum* L. subsp. *sinaicum* were used in the current study.

	Scientific name	ITS1	5.8S	ITS2	Total length (ITS region)
I. *Thymus* genus					
1-	Seq1_*Thymus vulgaris* (MT974175.1)	238bp	155bp	292bp	685bp
2-	Seq2_*Thymus capitatus* (MT974176.1)	233bp	155bp	293bp	681bp
3-	Seq3_*Thymus decussatus* (MT974177.1)	240bp	155bp	285bp	680bp
II. *Origanum* genus					
4-	Seq4_*Origanum vulgare* (MT974178.1)	237bp	155bp	284bp	676bp
5-	Seq5_*O. syriacum subsp. sinaicum* (MT974179.1)	250bp	155bp	275bp	680bp

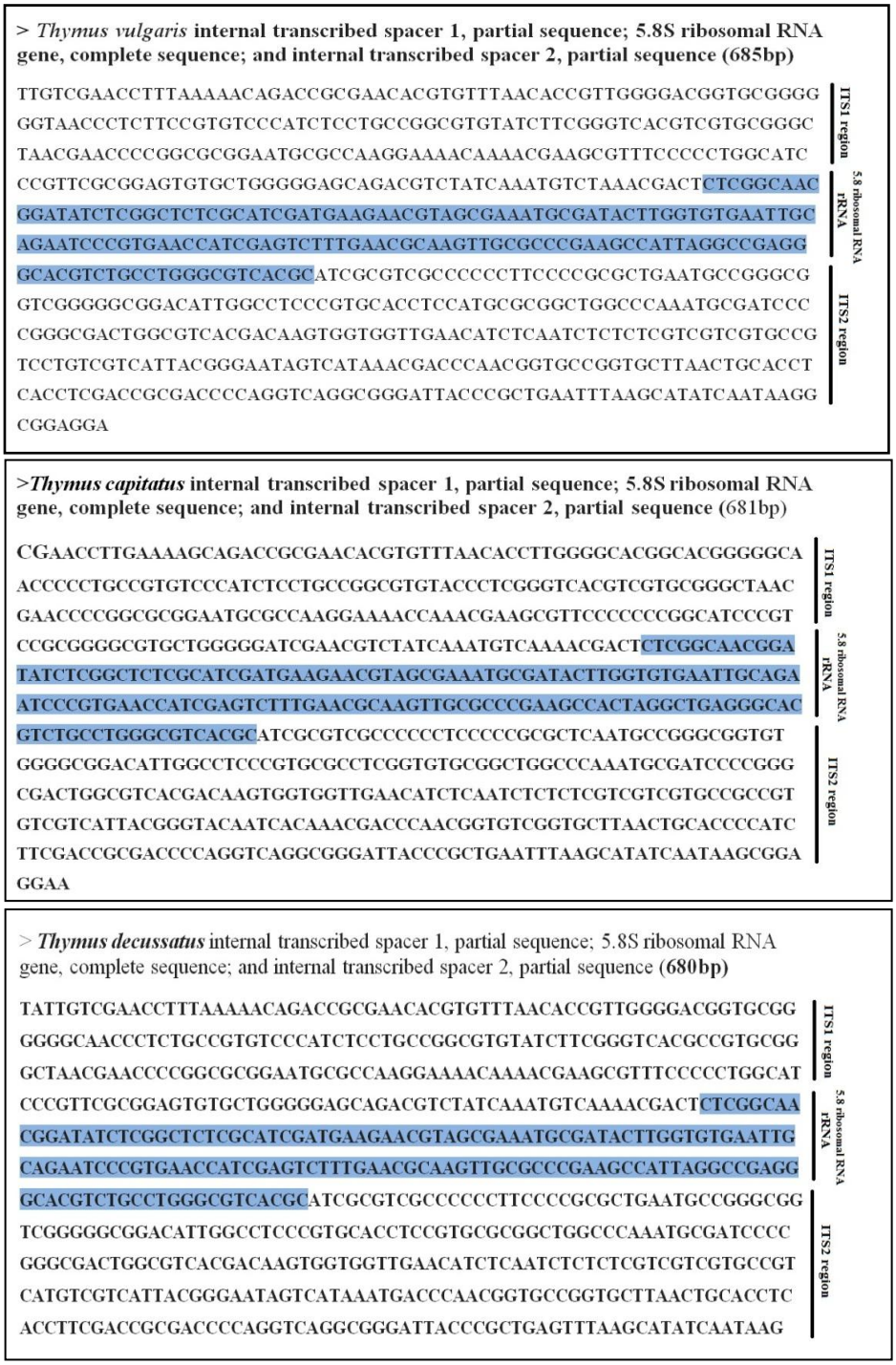

Figure 2: The nucleotide sequence of the nuclear ITS region (ITS1, 5.8S, ITS2) from *Thymus vulgaris* (685bp), *Thymus capitatus* (681bp) and *Thymus decussatus* (680bp) with GenBank accession no. MT974175.1, MT974176.1 and MT974177.1, respectively.

>*Origanum vulgare* **internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence (676bp)**

TTGTCGAACCCTGAAAAGTAGACCGCGAACACGTGTTTAACATCATGGGGGACGGTGCGGG
GGCAACCCTCTGCCGTGACCCATCTCCTGCGGGCGTGTATCTTCGGGTCACGTCTTGCGGGC
TAACGAACCCCGGCGCGGAATGCGCCAAGGAAAACTAAACGAAGCGTTTCCCCCCAGCATC
CCGTCCGCGGAGCGTGTTGGGGGATCGAACGTCTATCAAATGTCAAAACGACTCTTGGCAAC
GGATATCTCGGCTCTCGCATCGATTAAAAACGTAACGAAATGCGATACTTGGTGTGAATTGC
AGAATCCCGTGAACCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCCTTAGGCTGAGGG
CACGTCTGCCTGGGCGTCACGCATCGCGTCGCCCCCCTTCCCCGCGCTCAAAGCCGGGTGTT
AGGGGGCGGACATTGGCCTCCCGTGTACTTCGGTGTGCGGCTGGCCCAAATGCGATCCCCG
GGCGACTAGCGTCACGACAAGTGGTGGTTGAACATCTCAATCTCTCTCGTAGTCGTGCAGCT
GTGTCGTCATTACGGGCACAATCACAAATGACCCAACGGTGTCGGTGCGTAACTGCACCCCA
TCTTCGACCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAGTTTAAGCATATCAATAA

ITS1 region | 5.8 ribosomal RNA rRNA | ITS2 region

>*Origanum syriacum* subsp. *sinaicum* **internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence (680bp)**

TGCGGAAGGACATTGTCGAACCTCGAAAAGTAGACTGCGAACACGTGTTTAACATCATGGGG
GGCGGTGCGGGGGCAACCCTCTGCCGTGACCCATTTCCTGCGGACGTGTATCTTTGGGTCAC
GTCTTGCGGGCTAACGAACCCCGGCGCGGAATGCGCCAAGGAAAACAAAACGAAGCGTTC
CCCCCTAGCATCCCGTCCGCGGAGCGTGTTTGGGGGATCGAACGTCTATCAAATGT
CAAAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCG
AAATGCGATACTTGGTGTGAATTGCAGAATCCCGTGAACCATCGAGTCTTTGAACG
CAAGTTGCGCCCGAAGCCATTAGGCTGAGGGCACGTCTGCCTGGGCGTCACGCATC
GCGTCGCCCCCCTTCCCCGCGCTCAAAGCCGGGCGATAGGGGGCGGACATTGGCC
TCCCGTGTACTTCGGTGTGCGGCTGGCCCAAATGCGATCCCCGGGCGACTAGCGTC
ACGACAAGTGGTGGTTGAACATCTCAATCTCTGTCGTAGTCGTGCAGCTGAGTCGC
CATTACGGGCACAATCACAAATGACCCAACGGTGTCGGTGCGTAACTGCACCCCAT
CTTCGACCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAGTTTAAGCA

ITS1 region | 5.8 ribosomal RNA rRNA | ITS2 region

Figure 3: The nucleotide sequence of nuclear ITS region (ITS1, 5.8S, ITS2) from *Origanum vulgare* L. (676bp) and *Origanum syriacum* L. subsp. *sinaicum* (680bp) with GenBank accession no. MT974178.1 and MT974179.1, respectively.

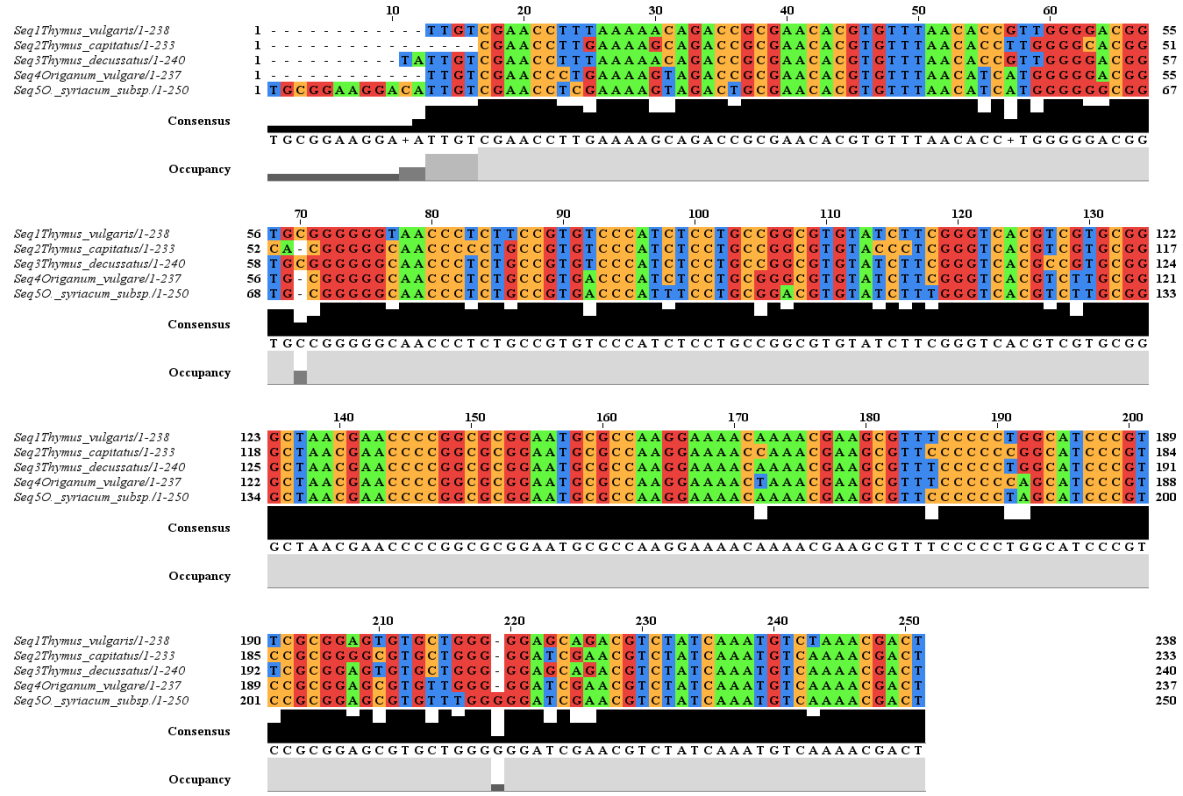

Figure 4: The nucleotide sequence alignment of the nuclear ITS1 region from*T. vulgaris* (238bp),*T. capitatus* (233bp), *T. decussatus* (240bp),*O. vulgare* L. (237bp) and *O. syriacum* L. subsp. *sinaicum* (250bp)were used in the current study.

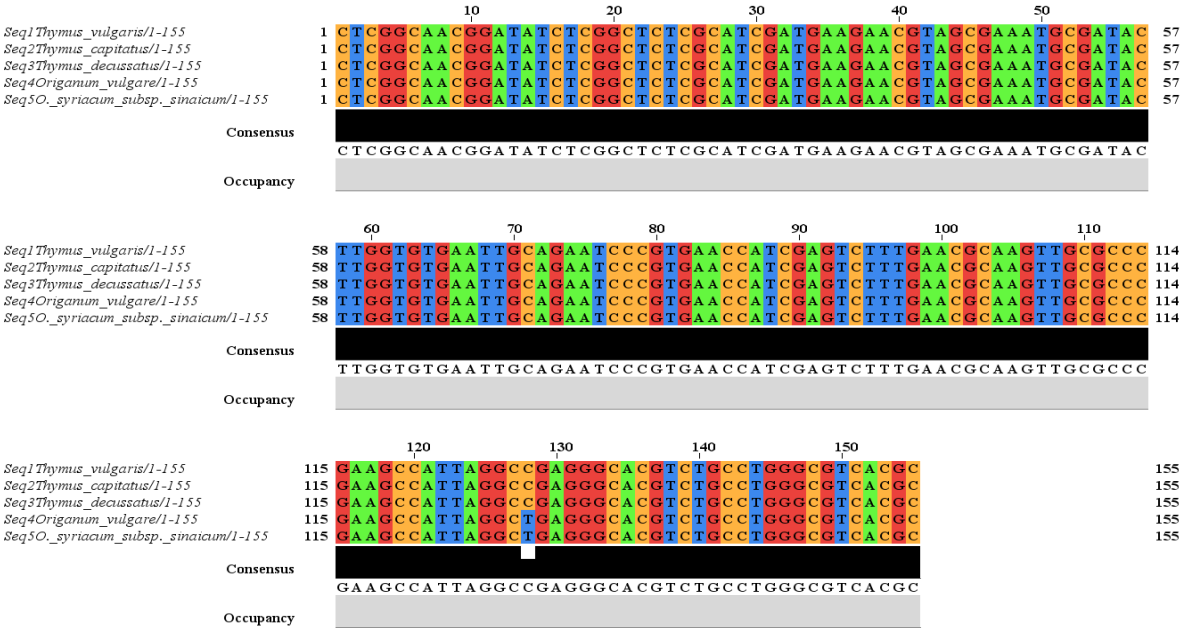

Figure 5: The nucleotide sequence alignment of the nuclear 5.8S region (155bp) with high conserves region from *Thymus* species (*T. vulgaris*, *T. capitatus*, *T. decussatus*) and *Origanum* species(*O. vulgare* and *O. syriacum* L. subsp. *sinaicum*)were used in the current study.

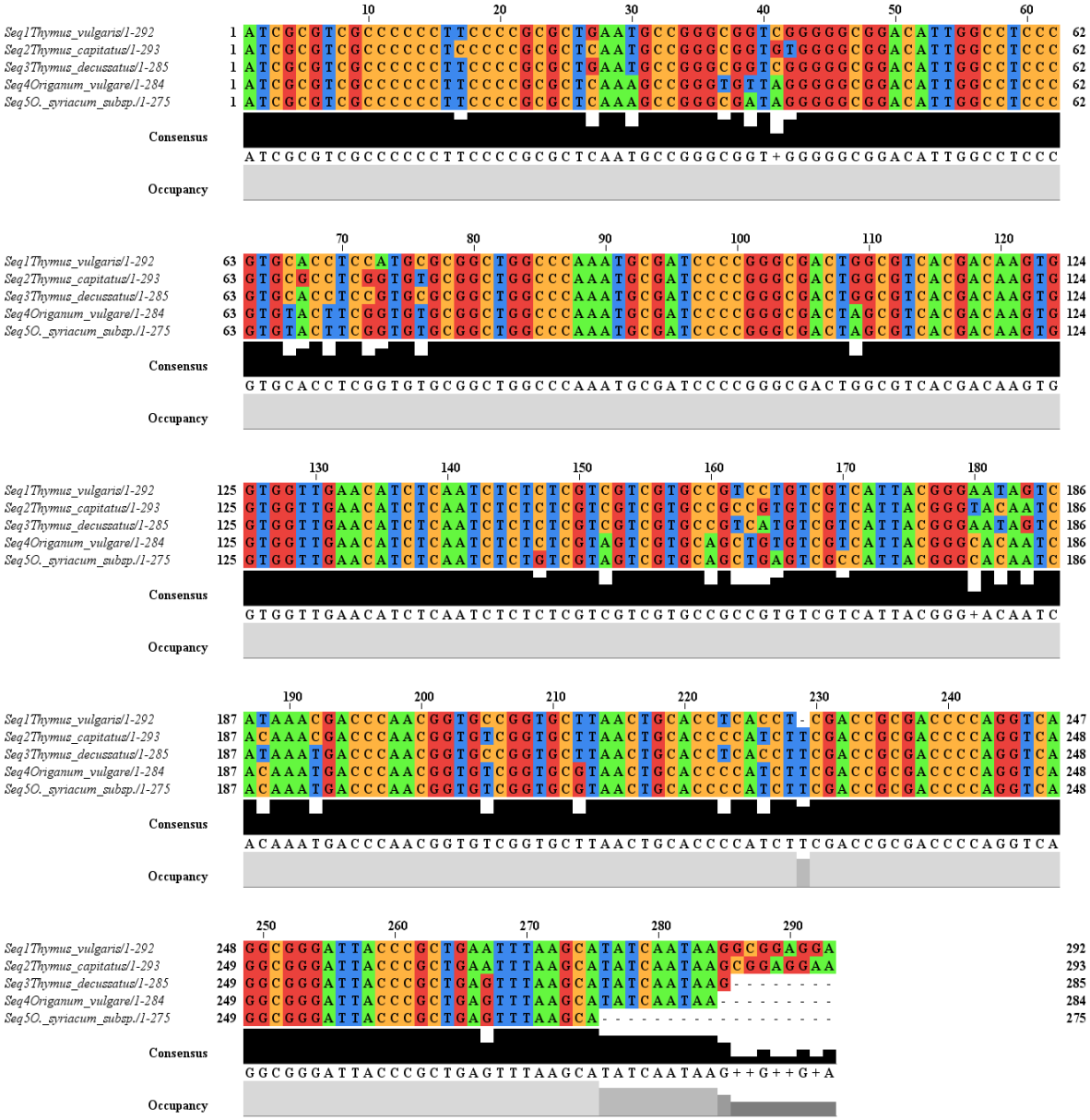

Figure 6: The nucleotide sequence alignment of the nuclear ITS2 region from *Thymus vulgaris* (292bp),*T. capitatus* (293bp), *T. decussatus* (285bp), *O. vulgare* (284bp) and *O. syriacum* L. subsp. *sinaicum* (275bp)were used in thecurrent study.

Table 2: Homology of nucleotide sequences for 56 selected accession lists for the nuclear (ITS region) sequenced in this study, BLAST top hits against GenBank database, similarity score and accession no.

Family: *Lamiaceae* Tribe: *Mentheae*	**Code**	**Scientific name of Nuclear ITS region(length bp)**	**GenBank Accession No.**	**Similarity% (bp)**				
				Seq1 (685bp)	**Seq2 (681bp)**	**Seq3 (680bp)**	**Seq4 (676bp)**	**Seq5 (680bp**
Current study	1	*Seq1_Thymus vulgaris* (685bp)	MT974175.1	------	94.00	98.67	92.47	91.34
	2	Seq2_*Thymus capitatus* (681bp)	MT974176.1	94.00	------	94.37	94.64	93.37
	3	Seq3_*Thymus decussatus* (680bp)	MT974177.1	98.67	94.37	------	93.35	92.40
	4	Seq4_*Origanum vulgare* (676bp)	MT974178.1	92.47	94.64	93.35	------	97.60
	5	Seq5_*O. syriacum* subsp. *sinaicum* (680bp)	MT974179.1	91.34	93.37	92.40	97.60	------
G1: *Thymusgenus*	6	*Thymus vulgaris* (643bp)	AY506646.1	97.97	92.59	97.04	90.64	89.37
	7	*Thymus quinquecostatus*(696bp)	EU556517.1	99.13	94.58	99.12	92.91	91.76
	8	*Thymus quinquecostatus*(625bp)	AY443435.1	98.08	93.10	98.08	91.52	90.43

	9	*Thymus mongolicus*(682bp)	EU556518.1	99.27	94.44	99.11	92.86	91.73
	10	*Thymbracapitata*(695bp)	JQ669137.1	93.24	99.05	93.56	94.32	93.06
	11	*Thymbracapitata*(612bp)	GU381453.1	93.28	98.86	93.44	93.95	92.66
	12	*Thymus eigii* (700bp)	KX257372.1	99.11	94.33	99.26	93.16	92.23
	13	*Thymus caespititius*(613bp)	GU381457.1	94.78	93.66	95.27	92.85	91.09
	14	*Thymus daenensis*(656bp)	MF465192.1	98.76	93.47	98.45	92.11	91.06
G2: *Origanum genus*	15	*Origanum vulgare* (693bp)	MH645777.1	92.33	94.55	93.24	100.00	97.49
	16	*Origanum vulgar* (704bp)	DQ667243.1	92.13	94.36	92.93	99.70	97.19
	17	*Origanum vulgare* (612bp)	GU381468.1	92.12	94.12	92.78	99.84	97.38
	18	*O. syriacum var. sinaicum* (644bp)	JX163026.1	90.99	93.10	91.94	97.35	99.84
	19	*Origanum syriacum* (644bp)	JX163033.1	90.99	93.10	91.94	97.35	99.84
	20	*Origanum syriacum* (644bp)	JX163037.1	90.99	93.10	91.94	97.35	99.84
	21	*O. syriacum var. bevanii*(644bp)	JX163046.1	90.99	93.10	91.94	97.35	99.84
	22	*O. syriacum var. sinaicum* (644bp)	JX163030.1	91.15	93.26	92.09	97.51	100.00
G3: *Conradina genus*	23	*Conradinabrevifolia*(617bp)	AY943461.1	86.69	89.90	87.50	87.25	85.02
	24	*Conradinaetonia*(618bp)	AF369165.1	86.55	89.21	87.20	86.89	84.85
G4: *Dicerandra genus*	25	*Dicerandrafrutescens*(644bp)	AY506642.1	86.80	88.23	87.60	85.36	84.73
	26	*Dicerandrachristmanii*(579bp)	DQ860350.1	89.17	90.36	89.54	87.18	88.14
G5: *Monarda genus*	27	*Monardadidyma*(686bp)	MN636289.1	87.57	89.55	88.47	87.22	98.41
	28	*Monardafistulosa*(629bp)	MG677129.1	87.32	89.18	88.11	86.61	85.76
G6: *Pycnan. genus*	29	*Pycnanthemummuticum*(615bp)	AY943494.1	87.95	89.98	88.76	87.03	98.13
	30	*Pycnanthemum beadle* (621bp)	MK355996.1	87.86	89.60	88.67	86.40	85.65
G7:*Bystropogon genus*	31	*Bystropogonoriganifolius*(649bp)	AY506635.1	86.98	89.17	87.77	86.51	85.54
	32	*Bystropogonwildpretii*(638bp)	AY704584.1	87.20	88.58	88.15	85.85	84.91
G8:*Clinopodium genus*	33	*Clinopodiumtomentosum*(649bp)	DQ017559.1	89.51	91.06	90.48	88.72	88.17
	34	*Clinopodiumcylindristachys*(651bp)	DQ017562.1	89.52	91.23	90.34	88.40	87.56
G9: *Satureja genus*	35	*Satureja thymbrifolia*(612bp)	GU381454.1	92.79	97.88	93.11	93.96	92.35
	36	*Satureja linearifolia*(612bp)	GU381455.1	93.61	98.37	93.93	94.61	92.99
G10: *Mentha genus*	37	*Mentha Canadensis* (715bp)	KC473228.1	88.56	90.52	89.00	88.22	87.01
	38	*Mentha spicata*(686bp)	KY072948.1	88.20	90.28	88.64	88.00	87.15
G11 *Agastache genus*	39	*Agastache barberi*(814bp)	AJ421001.1	84.76	86.79	85.04	84.88	85.12
	40	*Agastache urticifolia*(728bp)	DQ667247.1	84.01	86.83	84.20	85.07	84.97
G12: *Dracoc. genus*	41	*Dracocephalum moldavica*(717bp)	MH710906.1	85.02	84.64	83.38	83.53	83.07
	42	*Dracocephalum grandiflorum*(784bp)	AJ420999.1	82.57	84.37	83.12	82.87	82.42
G13:*Nepeta*	43	*Nepetamirzayanii*(784bp)	AJ515309.1	86.14	86.98	86.05	85.59	85.97
	44	*Nepetalaxiflora*(830bp)	AJ420995.1	85.60	85.69	83.64	82.34	85.38
G14: *Prunella genus*	45	*Prunella vulgaris* (734bp)	MH711205.1	83.38	85.80	83.84	83.92	82.97
	46	*Prunella grandiflora*(534bp)	HQ228231.1	85.80	88.01	86.04	86.39	85.71
G15:*Lepechinia genus*	47	*Lepechinialancifolia*(711bp)	DQ667306.1	85.80	88.10	86.03	86.38	85.23
	48	*Lepechiniaconferta*(672bp)	DQ667307.1	85.80	87.21	85.65	85.95	84.79
G16: *Rosmarinus genus*	49	*Salvia rosmarinus*(684bp)	MH645776.1	83.74	85.05	83.74	84.27	82.93
	50	*Rosmarinusofficinalis*(668bp)	EU796893.1	84.64	86.12	84.66	85.32	83.91
G17: Saliva *genus*	51	*Salvia wardii*(684bp)	MG824290.1	85.53	87.17	85.80	86.28	85.42
	52	*Salvia prattii*(685bp)	MG824253.1	85.53	86.88	85.84	86.24	85.59
Family *linaceae*	53	*Linum usitatissimum* (578bp)	MK066769.1	94.86	93.55	94.41	94.97	94.97
Family *Fabaceae*	54	*Glycine max* (666bp)	MT230614.1	96.99	96.99	96.99	96.39	96.39
Family*Amarnthaceae*	55	*Chenopodium album* (655bp)	MK953683.1	87.46	86.21	87.46	86.01	86.81
Family *Brassicaceae*	56	*Arabidopsis thaliana*(722bp)	MG886683.1	94.92	94.92	94.92	94.35	94.35

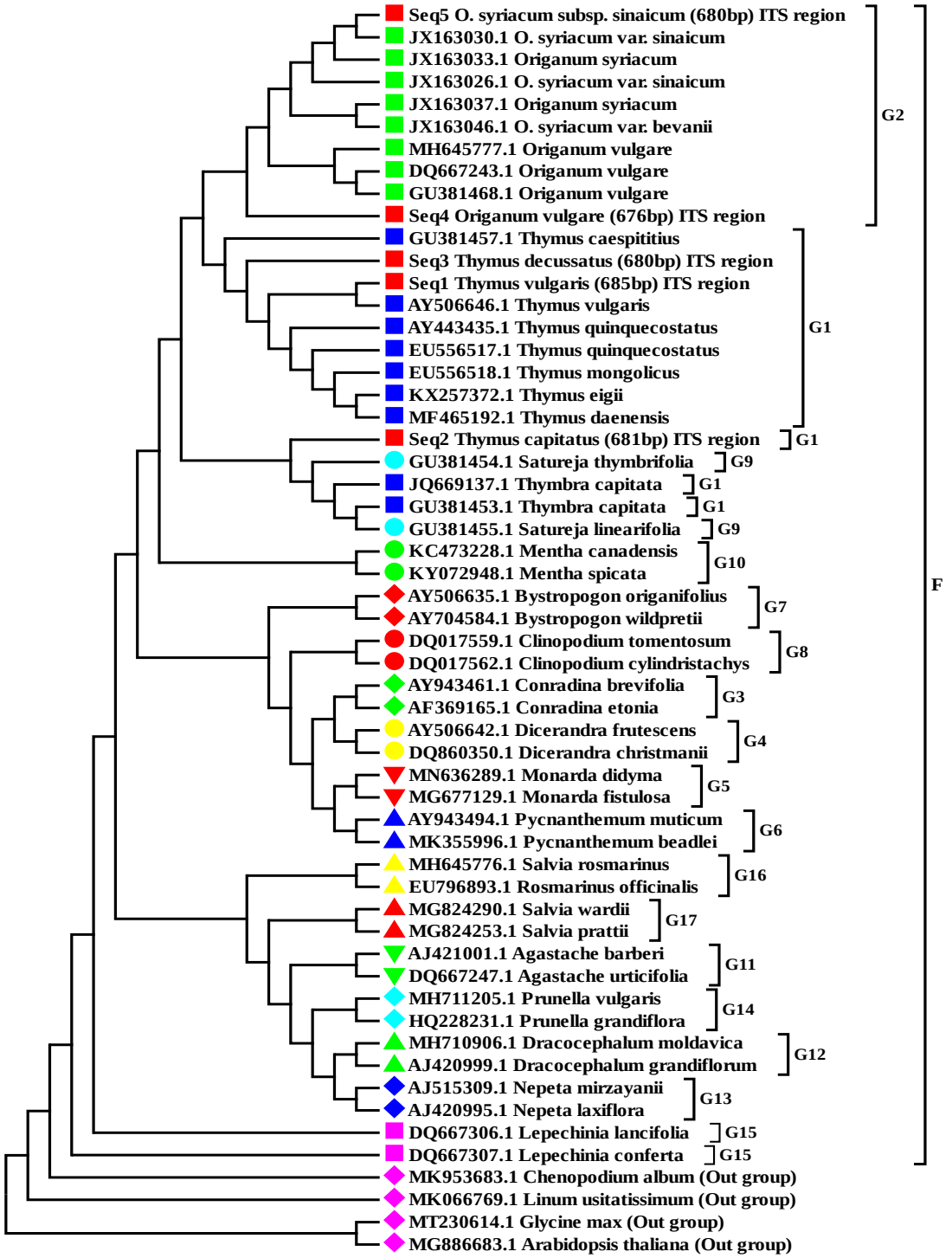

Figure 7: Molecular Phylogenetic analysis involved 56 nucleotide sequences of the nuclear (ITS region) were conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-3348.45) is shown [14]. Seq1: *T. vulgaris*, Seq2: *T. capitatus*, Seq3: *T. decussatus*, Seq4: *Origanum vulgare* and Seq5: *Origanum syriacum* L. subsp.*sinaicum* were used as current study. G1: *Thymu*, G2: *Origanum*, G3: *Conradina*, G4: *Dicerandra*, G5: *Monarda*, G6: *Pycnanthemum*, G7: *Bystropogon*, G8: *Clinopodium*,G9: *Satureja*, G8: *Clinopodium*, G10: *Mentha*, G11 *Agastache*, G12: *Dracocephalum*, *G13*:*Nepeta*, G14: *Prunella*, G15:*Lepechinia*, G16: *Rosmarinus*, G17: Saliva and F: Family *Lamiaceae*. *Linum usitatissimum*(MK066769.1),*Glycine max* (MT230614.1), *Chenopodium album* (MK953683.1) and*Arabidopsis thaliana* (MG886683.1) were used as out group. F: family *Lamiaceae*.

Reference

[1]. Baldwin BG. Molecular phylogenetics of *Calcydenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: Chromosomal and morphological evolution reexamined. Am. J. Bot., 1993; 80:222-238.

[2]. Baldwin BG, Sanderson MJ, Porter JM,*et al*. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden, 1995; 82: 247-277.

[3]. Caesar RM, Sorensson M, CognatoAI. Integrating DNA data and traditional taxonomy to streamline biodiversity assessment: An example from edaphic beetles in the Klamath ecoregion, California, USA," *Diversity and Distributions*, 2006; (12)5:483–489.

[4]. Calonje M, Martín–Bravo S, Dobes C, Gong W, *et al*. Non– coding nuclear DNA markers in phylogenetic reconstruction. Plant Syst. Evol., 2009; 282(3): 257– 280.

[5]. CBOL Plant Working Group. A DNA Barcode for Land Plants. Proceedings of the National Academy of Sciences of the United States of America, 2009; 106, 12794-12797. http://dx.doi.org/10.1073/pnas.0905845106

[6]. Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S. Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Molecular Ecology Resources. 2016; 16, 138–149. doi: 10.1111/1755-0998.12438.

[7]. China Plant BOL Group. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporates into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA, 2011; 108, 19641–19646.

[8]. El-Atroush H. DNA Barcoding of Two Medicinal Plants Using Molecular Markers. Egypt. Acad. J. Biolog. Sci., 2020; 12(1):83-92.

[9]. Ferreira-Cerca S, Poll G, Gleizes PE, TschochnerH, MilkereitP. Roles of Eukaryotic ribosomal protein in maturation and transport of pre 18S rRNA and ribosome function. Mol. Cell. 2005; 20(2): 263-275.

[10]. Giudicelli GC, Mäder G, FreitasLB. Efficiency of ITS sequences for DNA barcoding in *Passiflora* (Passifloraceae). Int. J. Mol. Sci., 2015; 16:7289-7303.

[11]. Giudicelli GC, Mäder G, Silva-Arias GA, ZaberlanPM, BonattoSL, FreitasL. Secondary structure of nrDNA Internal Transcribed Spacers as a useful tool to align highly divergent species in phylogenetic studies. Genetics and Molecular Biology, 2017; 40, 1(suppl), 191-199.DOI: http://dx.doi.org/10.1590/1678-4685-GMB-2016-0042

[12]. Harley RM,Atkins S,Budantsev AL, et al.In: *Labiatae*. PP. 167–275 in: Kadereit, J.W. (ed.), The families and genera of vascular plants, 2004; vol. 7, Flowering plants: Dicotyledons (*Lamiales* except *Acanthaceae* including *Avicenniaceae*). Berlin and Heidelberg: Springer.

[13]. Khan MQ, Khalil AT, RehmanK, ShinwariZK. Searching for DNA barcodes in plants. Global J. of Biotech. andBioch., 2015; 10, 1–10.

[14]. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary Genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 2016; 33:1870-1874.

[15]. Li XW, Yang Y, Henry RJ, Rosetto M, Wang YT, Chen SL. Plant DNA barcoding: from gene to genome. Biol. Rev., 2015; 90:157–166.

[16]. Poczai P,HyvnenJ. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. Mol. Biol. Rep. 2010; 37(4):1897–1912. DOI: 10.1007/s11033-009-9630-3

[17]. Qin Y, Li M, Cao Y, GaoY, Zhang W. Molecular thresholds of ITS2 and their implications for molecular evolution and species identification in seed plants. Scientific Reports, 2017; 7: 17316. https://doi.org/10.1038/s41598-017-17695-2

[18]. Raja RR. In: medicinally potential plants of *Lamiaceae* family: An overview, RJMP, 2012; Vol. 6(3),pp. 203-213 ISSN 1819-3455. DOI: 10.3923/rjmp.2012.203.213

[19]. Richard GF, Kerrest A,B Dujon. Comparative genomics andmolecular dynamics of DNA repeats in Eukaryotes. Microbiology and MolecularBiology Reviews.2008; 72(4): 686-727.DOI:10.1128/MMBR.00011-08

[20]. Techen N, Parveen I, Pan Z, Khan IA. DNA barcoding of medicinal plant material for identification. Curr. Opin. Biotechnol. 2014; 25: 103-110. https://doi.org/10.1016/j.copbio.2013.09.010

[21]. Wei NV, Wallace CC, Dai C-F,PillayKRM, ChenCA. Analyses of the ribosomalinternal transcribed spacers (ITS)and the 5.8S gene indicate that extremely high rDNA heterogeneityis a unique feature in the*Scleractinian*coral genus *Acropora*(*Scleractinia*; *Acroporidae*). ZoologicalStudies, 2006; 45(3): 404-418.

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

[23]. Yao H, Song J, Liu C, Luo K, Han J, et al. Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS ONE.2010; 5(10): e13102. doi:10.1371/journal.pone.0013102

[24]. Zhang W, Yang S, Zhao H, Huang L. Using the ITS2 sequence-structure as a DNA mini-barcode: a case study in authenticating the traditional medicine "Fang Feng". Biochem. Syst. Ecol. 2016;69, 188–194. doi: 10.1016/j.bse.2016.10.007

[25]. Zhao L-L, Feng S-J, Tian J-Y, Wei A-Z, Yang T-X. Internal transcribed spacer 2 (ITS2) barcodes: A useful tool for identifying Chinese *Zanthoxylum*. Applications in Plant Sciences2018; 6 (6): e1157. doi:10.1002/aps3.1157

Mohamed Zoelfakar Sayed Ahmed, et. al. "Molecular Identification and Phylogenetic Relationship of some Origanum and Thymus Species Based on ITS Regions." *IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)*, 6(6), (2020): pp. 12-23.