Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2020 17(4): 3078-3099. **OPEN ACCESS**

Authentication and Phylogenetic Analyses of some Endemic *Thymus* Species in Egypt Using DNA Barcodes *rbcL* and *matK* Gene Sequence

Mohamed Z.S. 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**

***Correspondence:** saadelshaimaa1@yahoo.com **Received 07-09-2020, Revised: 04-12-2020, Accepted: 10-12-2020 e-Published: 20-12-2020**

Genus *Thymus* is one of the largest genera of family Lamiaceae, it comprises about 250 species represented globally. Several important uses of *Thymus* species were reported in previous studies such as their medicinal, pharmaceutical and culinary uses. In spite of their importance, they are poorly identified on the basis of molecular levels. DNA barcoding is a reliable technique at molecular level in the field of authentication, classification and differentiation between plant species. Therefore, the current investigation was conducted to identify, discriminate and authenticate three endemic *Thymus* species in Egypt (*Thymus vulgaris, Thymus capitatus and Thymus decussatus*) using two chloroplast genes (coding sequence) of the most common DNA barcodes; ribulose 1, 5-biphosphate carboxylase large subunit (*rbcL*) and maturase K (*matK*) gene. The partial sequence lengths of *rbcL* gene of the three *Thymus* species were 577, 528 and 544 nucleotides for *T. vulgaris* (MN972464.1), *T. capitatus* (MN972465.1) and *T. decussatus* (MN972466.1), respectively. Similarly, for *matK* gene, they were 643, 800 and 788 nucleotides for *Thymus vulgaris* (MN972467.1), *Thymus capitatus* (MN972468.1) and *Thymus decussatus* (MN972469.1), respectively. The alignment of isolated sequences with the recorded species in NCBI database was conducted using basic local alignment search tool for several sequences (BLASTn and BLASTp). The obtained results revealed that *rbcL* and *matK* genes nucleotides sequence isolated from the three *Thymus* species under investigation were closely related and showed high similarities to the NCBI recoded 17 genera belonging to family Lamiaceae while they showed low similarities towards recorded species which considered as an out group of family Lamiaceae. Furthermore, a phylogenetic tree analysis was constructed based on amino acid sequences of 61 *rbcL* genes and 52 *matK* genes using MEGA 7 program by Maximum Likelihood (ML) method. The DNA barcodes (*rbcL* and *matK*) were able to distinguish the Egyptian *Thymus* species under research successfully from species of the same family and also from those belonging to different families.

Keywords: DNA barcoding, *rbcL* gene, *matK* gene, MEGA 7 program, family Lamiaceae, *Thymus vulgaris*, *Thymus capitatus* and *Thymus decussatus*

INTRODUCTION

Genus Thymus belongs to the tribe Mentheae, subfamily Nepetoideae in family Lamiaceae (Labiatae). It is a family of flowering plants commonly known as the mint or dead nettle or sage family. It contains about 236 genera and has been stated to contain 6.900 to 7.200 species (Harley et al., 2004). Many plants are aromatic in all parts and include widely used culinary herbs. Jalas, (1971) classified Thymus into two subgenera and eight sections based on morphological characters. Thymus is one of the largest genera including about 250 species represented globally (Raja 2012).

Thyme is the general name for the many herb varieties of the Thymus species. It is a tiny perennial shrub, with a semi evergreen ground cover that seldom grows quite to (40 cm) with horizontal and upright habits (Saleh et al., 2015). Common or garden Thyme is considered the principal type, and is utilized commercially for flowering and ornamental purposes (Nadiya et al., 2016) It grows in mountainous areas, used as a beverage instead of or with tea and added to some food to give it an acceptable flavour,

Thymus species are considered as medicinal plants due to their pharmacological and biological properties. their properties are due to their main components, thymol and carvacrol. Fresh Thyme has the highest level of antioxidants among all herbs. The plant is used in folk medicine frequently where it is prescribed to treat mouth infections, stomach, intestine and airways, coughing and gastroenteritis and expel intestinal worms, as well as to strengthen the heart (Mohamed et al., 2013). Extracts from Thyme have been used in traditional medicine for the treatment of several respiratory diseases like asthma and bronchitis and for the treatment of other pathologies. Also, several properties were reported such as antiseptic, antispasmodic, antitussive, antimicrobial, antifungal, antioxidative, and antiviral (Ocana and Reglero, 2007). Additionally, the amazing Thyme benefits can be attributed to its rich nutritional value. The nutrients in Thyme have disease-preventing and health-promoting properties (Vineetha, 2014). Ancient Egyptians used Thyme for embalming.

When there is an insufficient morphological or anatomical data for the identification of the sample, a stretch of DNA sample might be helpful in identifying the species, samples with many fragments can provide multiple species identification, giving a clear picture of habitats, offers critical clue to the investigators (Ferri et al., 2015). Therefore, DNA barcoding technique was chosen for the present study to distinguish between three Thymus species grown in Egypt. Schori and Showalter, (2011) used DNA barcodes to identify 14 species of medicinal plants from Pakistan and distinguished them from other similar species. DNA barcoding is a technique for identifying organisms based on sequence data obtained from one to several gene regions (http://barcoding.si.edu/DNABarCoding.htm). Several chloroplast gene regions were typically used as DNA barcodes such as the two chloroplast gene fragments ribulose 1, 5-biphosphate carboxylase large subunit (rbcL) and maturase K; (matK) genes which considered core barcodes (Hollingsworth et al., 2009). Researchers have chosen a combination of two regions (matK and rbcL) as a satisfactory compromise that likely agrees with DNA barcoding criteria. In combination, these genes could identify species in 72 % of the cases and almost in all cases at the genus level (Hebert 2003; Chase et al. 2005; De Vere et al., 2012; Fazekas et al., 2012).

The Consortium for the Barcode of Life (CBOL) Plant Working Group (CBOL 2009, 2010) considered all the criteria for barcodes when evaluating the seven potential candidates, including four coding regions (matK, rbcL, rpoC1 and rpoB) and three non-coding regions (psbA-trnH, atpF-atpH and psbK-psbI).

Moreover, taxonomic expertise for plant species identification as well traditional morphological characterization are not sufficient for identification and authentication of plant species samples such as flowers, roots, seeds, and pollen (Chase and Fay, 2009).Therefore, the large data obtained from DNA barcodes can be used beyond taxonomic studies to ascertain the results obtained. Furthermore, DNA barcoding techniques play an important role in the identification of plants having a problematic taxonomic identity. For biodiversity screening, the DNA barcoding technique has a highly discriminating power between organisms and has the ability to identify polymorphic plant species **(**Ajmal et al., 2014). Also, it has employed for detection of adulteration and verification of plant materials included in several industries such as pharmacuticals (Asahina et al., 2010; Xue and Li, 2011). Several studies have been conducted to evaluate the efficiency of DNA barcodes in identification and verification of plant species such as De Mattia et al., (2011; 2012) who identified 64 plant species of six different genera (i.e. Mentha, Ocimum, Origanum, Salvia, Thymus and Rosmarinus) using four barcode regions (rpoB, rbcL, matK and trnH-psbA). They pointed to the benefit of using DNA barcoding technique in saving time and costs consumed during technique implementation.

Further study has been carried out by Federici et al. 2013 to distinguish between 36 samples of Thymus spp. using the core barcode regions matK, rbcL and the plastid intergenic spacer trnH-psbA**.** Nithaniyal et al. 2014 utilized DNA barcodes as a first attempt to assemble a reference barcode library for the 429 trees of India. DNA barcoding was completed using rbcL and matK markers to inventory species for applications in biodiversity monitoring, conservation impact assessment, monitoring of illegal trading, authentication of traded medicinal plants etc. In addition, Theodoridis et al. (2012) evaluated the utility of DNA barcoding in discriminating Labiatae species in Chios Island (Greece) and the adjacent Çeşme-Karaburun Peninsula (Turkey) using three cpDNA regions (matK, rbcL and trnH-psbA). They tested them either as single region or as multi region barcodes and found that matK and trnH-psbA taken as useful in discriminating species of the Labiatae, for the species they examined, as any multi region combination according to consortium for the Barcode of Life (CBOL) for medicinal plants of the Labiatae (Lamiaceae) family. They also added that matK and trnH-psbA could serve as single-region barcodes for Labiatae species contributing to the conservation and the trade control of valuable plant resources. For more than a decade, other several applications has been widely used in the DNA barcoding, molecular systematics, authentication identification and community phylogenetics of some medicinal plants (Sahare and Srinivasu, 2012; Vohra and Khera, 2013; Techen et al. 2014; Zhou et al. 2014; Parveen et al. 2016; Bezeng et al. 2017; Chen et al. 2019; Skuza et al. 2019).

Since, importance of Thymus species is due to active ingredients and essential oils which confer them their medicinal, culinary, pharmaceutical properties. Moreover, threatens such as over collecting of plants and human constructions in plants natural habitats which put them in their way to extinction. Therefore the current investigation was carried out to identify and authenticate the two endemic wild Thymus species and one cultivar species in Egypt (*T. capitatus, T. decussatus* and *T. vulgaris*) using two of the most important barcode regions of chloroplast genes (rbcL and matK gene). In addition, to distinguish between the three Thymus species under study and the other available species on NCBI database to fulfill a strict conservation plan and maintenance of the studied species.

MATERIALS AND METHODS

Plant Materials collection

Three available Thymus species in Egypt were chosen for the present investigation (*T. vulgaris, T. capitatus* and *T. decussatus*). Three samples representative to each plant species were collected (Fig.1), identified and authenticated by Dr. Yousry Abd-Elhady, Ecology and Range Management Department, Desert Research Center (DRC), Egypt (no voucher specimen of this material has been deposited in a publicly available herbarium). The fresh young leaves were collected as bulk for each species in spring of 2018, transferred into liquid nitrogen, and kept frozen at - 80 °C till use. *Thymus vulgaris* Benth. (cultivated type) which was collected from private farms at Kirdasa region, Giza Governorate, Egypt. For two wild types: *Thymus capitatus* (L.) Hoffmanns & Link was collected from the naturally grown rocky ridge habitats, especially wet sites distributed in small spot at Kilo 26 from the North Coast, Marsa Matrouh Governorate, Egypt and the other wild type is *Thymus decussatus* (Benth.) which was collected from the top one of mountain of Saint Catherine Protectorate, South Sinai Governorate, Egypt (Fig 1).

DNA extraction, primers design and PCR amplification

Genomic DNAs were extracted from 100mg fine powdered tissue using DNeasy Plant Mini Kit (Qiagen Inc., Cat.no.69104, USA). The concentrations and quality of the genomic DNA samples were estimated on spectrophotometer ND-2000 (Nanodrop, USA). In this investigation, successful isolation of pure DNA with high molecular weight is quite a challenge to reduce degraded DNA of plant materials especially in medicinal plants. The resulted DNA concentration for Thymus species (25–30 ng /ML) was checked at 260 and 280 nm.

DNA fragments were amplified via standard polymerase chain reaction (PCR). The entire coding plastid *rbcLa* (first part of *rbcL* gene ~ 700bp or less) was amplified using the primer pairs *rbcLa*_1Fwd (5'-ATG TCA CCA CAA ACA GAG ACT AAA GC-3') and *rbcLa*_599Rev (5’-GTA AAA TCA AGT CCA CCR CG-3’) and PCR fragment a 599 bp of the *rbcL* gene as previously described by Levin *et al*., (2003) and Kress and Erickson, 2007). For *matK* primer design, 13 ORF full length sequence *matK* genes from different species belong to Lamiaceae family were retrievable from the National Center for Biotechnology Information (NCBI) database (GenBank). Accession numbers GU381790.1 (*T. pulegioides*), AY840173.1 (*T. serpyllum*), GU381789.1 (*T. caespititius*), GU381791.1 (*T. broussonetii* subsp. hannonis), GU381792.1 (*T. vulgaris*), GU381802.1 (*O. vulgare*), GU381799.1 (*O. elongatum*), GU381798.1 (*O. dictamnus*), GU381797.1 (*O.*

rotundifolium), GU381801.1 (*O. microphyllum*), GU381800.1 (*O. dayi*), AY840165.1 (*O. vulgare*), and MG256495.1 (*Mentha spicata*) were used for multiple sequences alignment of nucleotide (BLASTn) to design specific primers pair of matK gene (Fig. 2). Then, the downloaded full ORF matK gene sequences with these accession numbers were saved in format (FASTA), files then aligned by MEGA 7 software program (Kumar et al., 2016), (www.megasoftware.net). The sequence from location base 466 - 488 to location base 1327-1349 was commonly present in all aligned sequences. The entire coding plastid maturase gene (*matK*) was amplified ~ 884bp or less using the primer pairs *matK*_466Fwd (5`- GTC CAT GTG GAA ATC TTG ATT C -3`) and *matK*_1349Rev (5`- CGT ACA GTA CTT TTG TGT TTA CG -3`) and PCR fragment ~ 884 bp or less 850 bp according to (CBOL Plant Working Group 2009). Phusion® *Taq*, the High-Fidelity DNA polymerase (Thermo Scientific, Product codes: F-530L, 500 Unit) was used. The amplification reaction mixture for *rbcl* and *matK* was done in a 50 µl total volume. Reaction mixture containing; 4 µl cDNA, 10 µl 5X Phusion HF Buffer, 1 µl 10mM dNTP mix, 2.5 µl primer 1 (10 µM), 2.5 µl primer 2 (10 µM), 0.5 µl Phusion DNA polymerase, 29.5 µl DEPC H_2O and spin for 15 Sec. The reaction mixture was incubated in a Perkin Elmer thermo cycler 9700. PCR program for *rbcL* regions, the temperature profile in different cycles was as follows: an initial strand separation cycle at 98°C for 3 min followed by 35 cycles including a denaturation step at 98°C for 30 seconds, an annealing step at 55°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 program for *matK* regions, 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 55°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.

Gel electrophoresis

PCR amplification samples were resolved by electrophoresis in a 1.5% agarose gel containing ethidium bromide (0.5ug/ml) in 1X TAE buffer at 95 volts. A 1Kbp DNA ladder was used as DNA standard size marker. PCR products were visualized on UV light and photographed using a Gel Documentation System (BIO-RAD 2000).

QIAquick® gel extraction kit protocol for purification

QIAquick gel extraction kit (Qiagen, Cat. No. 28704 - 28706) was used for PCR product purification or fast clean-up from the agarose gel up to 10 µg can bind to each QIAquick column. This was performed following the manufacturer's instructions.

Sequence editing, alignment and phylogenetic inference

Sequencing chromatograms of obtained rbcL and matK regions were analyzed by Macrogen, Seoul, south of Korea and translated into amino acid sequences by the ExPASy online program (https://web.expasy.org/translate) for each studied three Thymus species. All nucleotide sequences of rbcL and matK gene, Open Reading Frame (ORF) were searched in NCBI database, The National Center for Biotechnology Information GenBank Database (http://www.ncbi.nlm.nih.gov).

(a) *Thymus vulgaris* L. (Cultivar plant)

(b) *Thymus capitatus* L. (Ecotype of wild species)

(c) *Thymus decussatus* Benth. (Ecotype of wild species)

Figure 1: The three *Thymus* species under study: (a) *Thymus vulgaris* L. (Cultivar plant), *Thymus capitatus* L. (Ecotype of wild species) and (c) *Thymus decussatus* Benth. (Ecotype of wild species).

Species/Abbrv	Gr	
1. GU381790.1 Thymus pulegioides		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCGTTGGGTAAAAGATGT
2. GU381792.1 Thymus vulgaris		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCACTGGGTAAAAGATGT
3. AY840173.1 Thymus serpyllum		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
4. GU381789.1 Thymus caespititius		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
5. GU381791.1 T. broussonetii subsp. hannonis		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCACTGGGTAAAAGATGT
6. GU381802.1 Origanum vulgare		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
7. GU381799.1 Origanum elongatum		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
8. GU381798.1 Origanum dictamnus		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
9. GU381797.1 Origanum rotundifolium		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
10. GU381801.1 Origanum microphyllum		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
11. GU381800.1 Origanum dayi voucher		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
12. AY840165.1 Origanum vulgare		TACTAATACCCCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT
13. MG256495.1 Mentha spicata		TACTAATACCTCGCTCTGTCCATGTGGAAATCTTGATTCAAACTCTTCGCCATTGGGTAAAAGATGT

*matK*_466Fwd : 5'- GTC CAT GTG GAA AT CTT GAT TC -3'

Species/Abbrv	Gr	
1. GU381790.1 Thymus pulegioides		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
2. GU381792.1 Thymus vulgaris		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
3. AY840173.1 Thymus serpyllum		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
4. GU381789.1 Thymus caespititius		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
5. GU381791.1 T. broussonetii subsp. hannonis		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
6. GU381802.1 Origanum vulgare		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
7. GU381799.1 Origanum elongatum		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
8. GU381798.1 Origanum dictamnus		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
9. GU381797.1 Origanum rotundifolium		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
10. GU381801.1 Origanum microphyllum		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
11. GU381800.1 Origanum dayi voucher		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
12. AY840165.1 Origanum vulgare		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT
13. MG256495.1 Mentha spicata		TCTTGTGCTAGAACTTTGGCACGTAAACACAAAAGTACTGTACGTGCTTTTTTAAAAAGATTAGGTT

*matK*_1349Rev :5'- CGT ACA GTA CTT TTG TGT TTA CG -3'

Figure 2: Multiple alignment nucleotide sequences of 13 *matK* gene from 13 accession no. of different species belong to *Lamiaceae* family to design forward (matK_466Fwd) and reverse (matK_1349Rev) primer *matK* gene.

The homology searches were performed with Basic Local Alignment Search Tool of several sequences (BLASTn/BLASTp program) on the basis of their homologies with sequences published in DDBJ/EMBL/GenBank database which are available through NCBI database (Altschul et al. 1990).

The phylogenetic trees were constructed using MEGA 7.0 software program following the Maximum Likelihood (ML) tree methods to detect positive selection under site models (Kumar et al., 2016) (http://megasoftware.net/)

RESULTS AND DISCUSSION

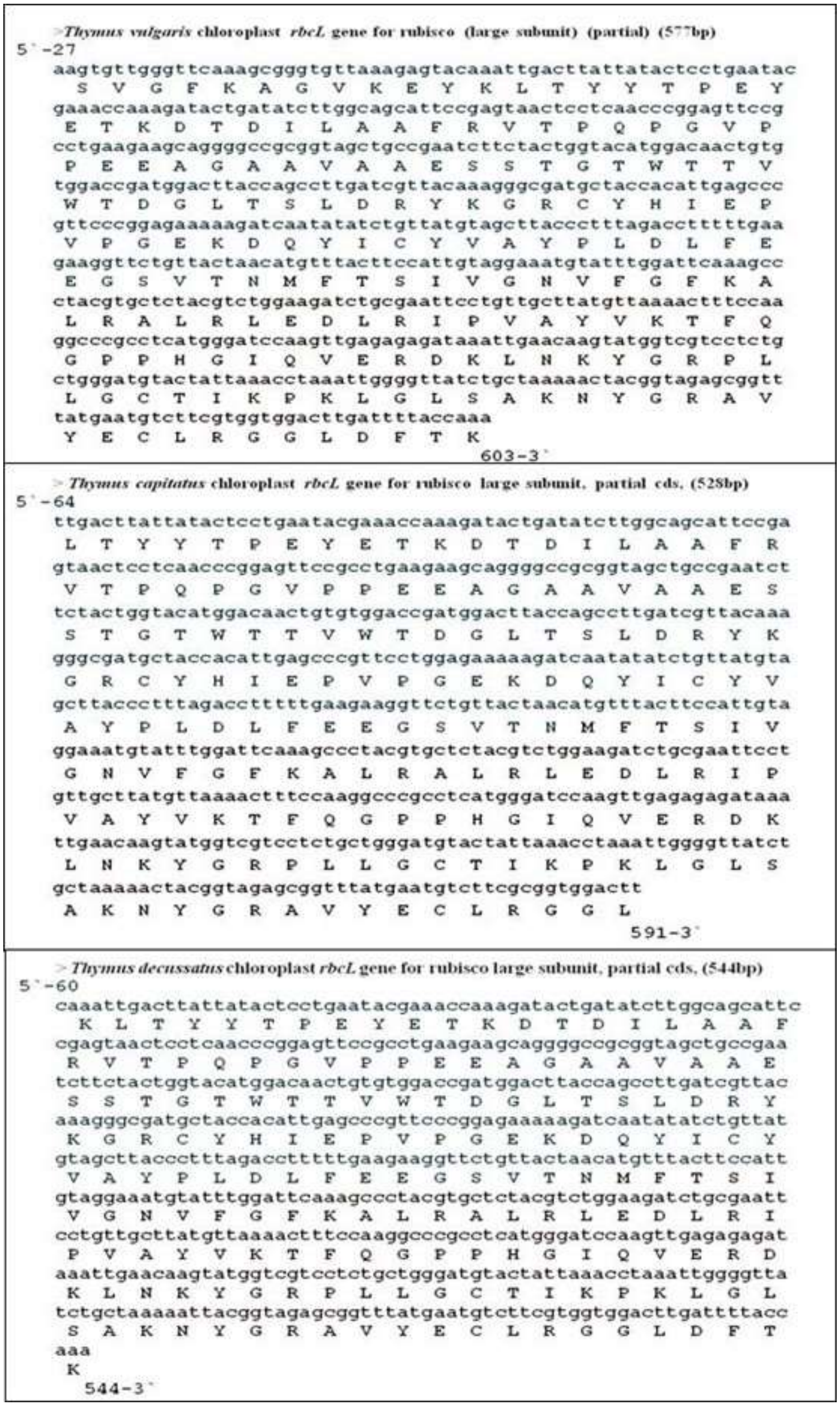
The DNA barcoding technique involved several consecutive steps to achieve identification and documentation of the studied *Thymus* species. These steps included DNA extraction, PCR amplification of *rbcL* and *matK* conserved regions, sequencing of the obtained fragments and alignment of the examined sequences with others included on NCBI database. PCR amplification of conserved regions (*matK* and *rbcL*) was carried out using both universal primers *rbcL* and *matK* genes, PCR amplification has resulted in amplified fragments sized about 700 or less 600bp for the first part of partial fragment length *rbcL* gene while about 900 to 850bp for partial *matK* gene in the three *Thymus* species (*Thymus vulgaris* L. (cultivar plant), *Thymus capitatus* L. (Ecotype of wild species) and *Thymus decussatus* Benth. (Ecotype of wild species). Specific single fragment was obtained, purified and sequenced for each sample, size and reading sequence were determined by Macrogen analysis, and then aligned by BLAST web sites to identify sequences similarities. The nucleotide sequences isolated in this paper have been recorded in the GenBank database NCBI (National Center for Biotechnology Information) by BankIt, online from website (https://www.ncbi.nlm.nih.gov/WebSub/) after it was processed by email (gb-admin@ncbi.nlm.nih.gov), The three *rbcL* gene sequences are available in DDBJ/EMBL/GenBank database with accession no. (MN972464.1/QKE50850.1) from *Thymus vulgaris* (577 bp) which predicted to encode a protein of 192 amino acids, from *Thymus capitatus*, (528bp), which predicted to encode a protein of 176 amino acids (accession no. MN972466.1/QKE50852.1) and from *Thymus decussatus* (544bp), which predicted to encode a protein of 181 amino acids as shown in Figure (3). In the same context, the three sequenced *matK* genes are available in

GenBank (accession no. MN972467.1/QKE50853.1) from *Thymus vulgaris* (643bp) which predicted to encode a protein of 214 amino acids, (accession no. MN972468.1/QKE50854.1), from *Thymus capitatus*, (800bp), which predicted to encode a protein of 266 amino acids (accession no. MN972469.1/QKE50855.1) and from *Thymus decussatus* (788bp), which predicted to encode a protein of 261 amino acids as shown in Figure (4).

Molecular phylogenetic analysis based on sequences of *rbcL* gene

Sequence homology of each nucleotide of 48 *rbcL* genes represented by (Code 1 to 3) for *Thymus* species (current study), (Code 4 to 36) for *Thymus* genus, (Code 37 to 44) for *Origanum* genus, and (Code 45 to 48) for 4 different species from different families as an out group of the family, all were tested by top-scoring hits through NCBI database using basic local alignment search tool (BLASTn) as shown in Table (1).

Multiple sequence alignments (MSA) of 48 *rbcL* genes were carried out between sequences of selected nucleotides of two genera (*Thymus* and *Origanum*) belonging to the family Lamiaceae or *Labiatae* including the 3 *Thymus* species (*T. vulgaris, T. capitatus and T. decussatus)* under study and four different species from different families as an out-group available in GenBank database will be discussed briefly. The results revealed that *Thymus vulgaris* (current study) was closely related with high identity and similarity (100%) to *T. vulgaris* ISr29 with accession no. (KU856611.1). Moreover, *T. vulgaris* showed similarity less than 100% with 34 accessions of *Thymus* genus, 8 accessions of *Origanum* genus and with 4 different species as an out group of the family, we will discuss it briefly. Firstly, *T. vulgaris* showed high similarity (100%, 99.82%, 99.82%, 99.81% and 99.65%,) with *T. vulgaris* ISr29 (KU856611.1), *Thymus* sp. AT-2018 (MF695015.1), *Thymus* sp. AT-2018 (MF694789.1) *T. vulgaris* (Z37471.1) and *T. praecox* subsp. *polytrichus* (JN890697.1), it showed 98.60% similarity with *T. serpyllum* (MK105914.1), 99.98% - 99.43% with 8 *Origanum* genus which belongs also to family Lamiaceae with E-value = Zero. On the other hand, it showed less similarities with the 4 different species from different families other than Lamiaceae (89.81%, 91.06%, 90.77% and 87.55%) with *Linum usitatissimum* (MG946893.1), *Helianthus annuus* (AB530968.1), *Arabidopsis thaliana* (AB917053.1) and *Ginkgo biloba* (HQ619753.1), respectively with E-value = Zero as an out group of the family Lamiaceae or Labiatae. Secondly, the Multiple sequence alignments (MSA) of *Thymus capitatus rbcL* gene (current study) showed high relatedness to *Thymus sp.* SH-2010 (AB586374.1), *Thymus sp.* MIB zpl (FR720585.1), and *T. praecox* (KT695419.1) with high identity and similarity (100%).



>*Thymus vulgaris* chloroplast *rbcL* gene for rubisco (large subunit) (partial) (577bp)

```
5'-27
   aagtgttgggttcaaagcgggtgttaaagagtacaaattgacttattatactcctgaatac
     S  V  G  F  K  A  G  V  K  E  Y  K  L  T  Y  Y  T  P  E  Y
   gaaaccaaagatactgatatcttggcagcattccgagtaactcctcaacccggagttccg
    E  T  K  D  T  D  I  L  A  A  F  R  V  T  P  Q  P  G  V  P
   cctgaagaagcaggggccgcggtagctgccgaatcttctactggtacatggacaactgtg
    P  E  E  A  G  A  A  V  A  A  E  S  S  T  G  T  W  T  T  V
   tggaccgatggacttaccagccttgatcgttacaaagggcgatgctaccacattgagccc
    W  T  D  G  L  T  S  L  D  R  Y  K  G  R  C  Y  H  I  E  P
   gttcccggagaaaaagatcaatatatctgttatgtagcttaccctttagaccttttgaa
    V  P  G  E  K  D  Q  Y  I  C  Y  V  A  Y  P  L  D  L  F  E
   gaaggttctgttactaacatgtttacttccattgtaggaaatgtatttggattcaaagcc
    E  G  S  V  T  N  M  F  T  S  I  V  G  N  V  F  G  F  K  A
   ctacgtgctctacgtctggaagatctgcgaattcctgttgcttatgttaaaactttccaa
    L  R  A  L  R  L  E  D  L  R  I  P  V  A  Y  V  K  T  F  Q
   ggcccgcctcatgggatccaagttgagagagataaattgaacaagtatggtcgtcctctg
    G  P  P  H  G  I  Q  V  E  R  D  K  L  N  K  Y  G  R  P  L
   ctgggatgtactattaaacctaaattggggttatctgctaaaaactacggtagagcggtt
    L  G  C  T  I  K  P  K  L  G  L  S  A  K  N  Y  G  R  A  V
   tatgaatgtcttcgtggtggacttgattttaccaaa
    Y  E  C  L  R  G  G  L  D  F  T  K
                                       603-3'
```

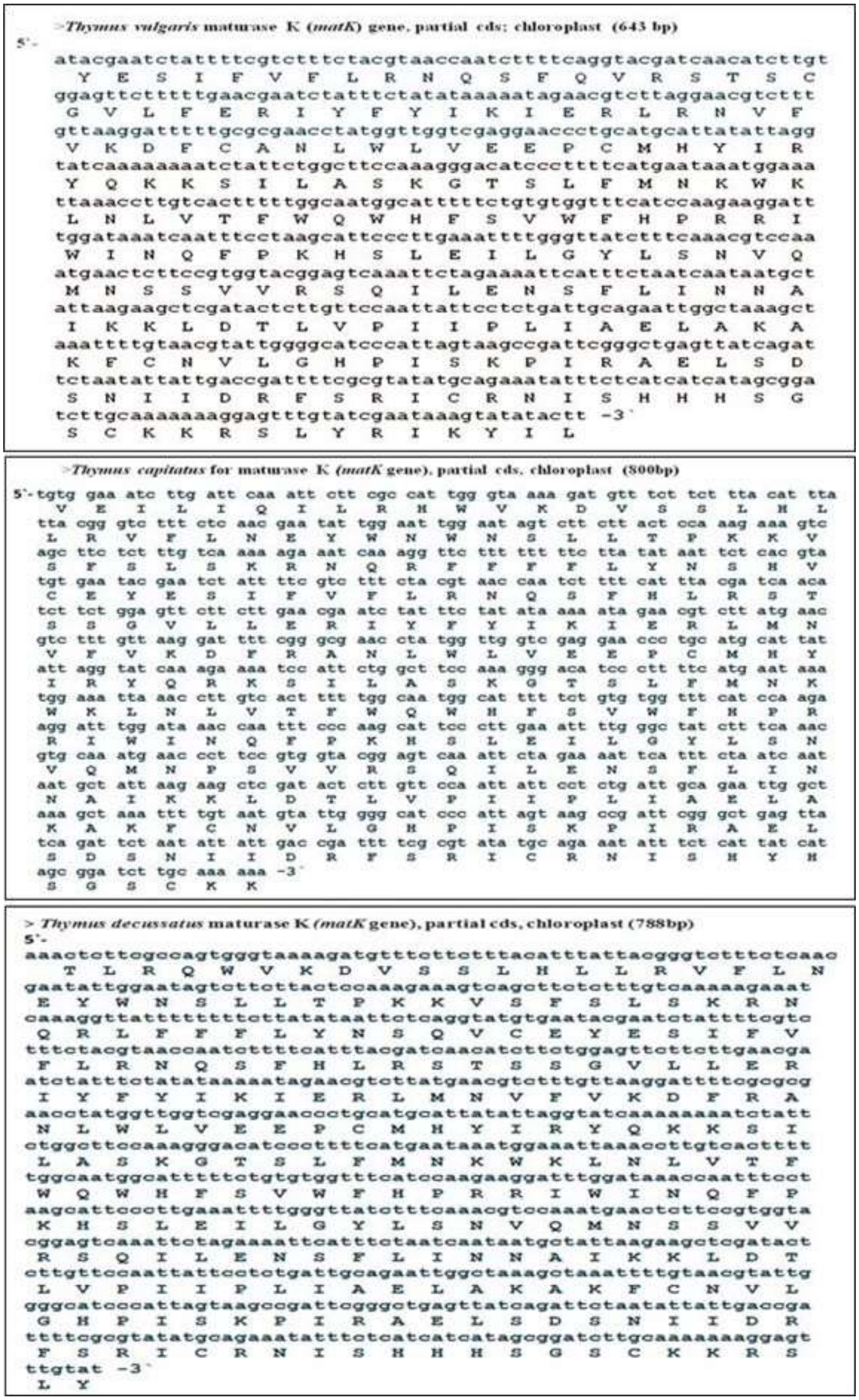
>*Thymus capitatus* chloroplast *rbcL* gene for rubisco large subunit, partial cds, (528bp)

```
5'-64
   ttgacttattatactcctgaatacgaaaccaaagatactgatatcttggcagcattccga
    L  T  Y  Y  T  P  E  Y  E  T  K  D  T  D  I  L  A  A  F  R
   gtaactcctcaacccggagttccgcctgaagaagcaggggccgcggtagctgccgaatct
    V  T  P  Q  P  G  V  P  P  E  E  A  G  A  A  V  A  A  E  S
   tctactggtacatggacaactgtgtggaccgatggacttaccagccttgatcgttacaaa
    S  T  G  T  W  T  T  V  W  T  D  G  L  T  S  L  D  R  Y  K
   gggcgatgctaccacattgagcccgttcctggagaaaaagatcaatatatctgttatgta
    G  R  C  Y  H  I  E  P  V  P  G  E  K  D  Q  Y  I  C  Y  V
   gcttaccctttagaccttttgaagaaggttctgttactaacatgtttacttccattgta
    A  Y  P  L  D  L  F  E  E  G  S  V  T  N  M  F  T  S  I  V
   ggaaatgtatttggattcaaagccctacgtgctctacgtctggaagatctgcgaattcct
    G  N  V  F  G  F  K  A  L  R  A  L  R  L  E  D  L  R  I  P
   gttgcttatgttaaaactttccaaggcccgcctcatgggatccaagttgagagagataaa
    V  A  Y  V  K  T  F  Q  G  P  P  H  G  I  Q  V  E  R  D  K
   ttgaacaagtatggtcgtcctctgctgggatgtactattaaacctaaattggggttatct
    L  N  K  Y  G  R  P  L  L  G  C  T  I  K  P  K  L  G  L  S
   gctaaaaactacggtagagcggtttatgaatgtcttcgcggtggactt
    A  K  N  Y  G  R  A  V  Y  E  C  L  R  G  G  L
                                                  591-3'
```

>*Thymus decussatus* chloroplast *rbcL* gene for rubisco large subunit, partial cds, (544bp)

```
5'-60
   caaattgacttattatactcctgaatacgaaaccaaagatactgatatcttggcagcattc
     K  L  T  Y  Y  T  P  E  Y  E  T  K  D  T  D  I  L  A  A  F
   cgagtaactcctcaacccggagttccgcctgaagaagcaggggccgcggtagctgccgaa
    R  V  T  P  Q  P  G  V  P  P  E  E  A  G  A  A  V  A  A  E
   tcttctactggtacatggacaactgtgtggaccgatggacttaccagccttgatcgttac
    S  S  T  G  T  W  T  T  V  W  T  D  G  L  T  S  L  D  R  Y
   aaagggcgatgctaccacattgagcccgttcccggagaaaaagatcaatatatctgttat
    K  G  R  C  Y  H  I  E  P  V  P  G  E  K  D  Q  Y  I  C  Y
   gtagcttaccctttagaccttttgaagaaggttctgttactaacatgtttacttccatt
    V  A  Y  P  L  D  L  F  E  E  G  S  V  T  N  M  F  T  S  I
   gtaggaaatgtatttggattcaaagccctacgtgctctacgtctggaagatctgcgaatt
    V  G  N  V  F  G  F  K  A  L  R  A  L  R  L  E  D  L  R  I
   cctgttgcttatgttaaaactttccaaggcccgcctcatgggatccaagttgagagagat
    P  V  A  Y  V  K  T  F  Q  G  P  P  H  G  I  Q  V  E  R  D
   aaattgaacaagtatggtcgtcctctgctgggatgtactattaaacctaaattggggtta
    K  L  N  K  Y  G  R  P  L  L  G  C  T  I  K  P  K  L  G  L
   tctgctaaaaattacggtagagcggtttatgaatgtcttcgtggtggacttgattttacc
    S  A  K  N  Y  G  R  A  V  Y  E  C  L  R  G  G  L  D  F  T
   aaa
    K
     544-3'
```

Figure 3: The partial fragment *rbcL* gene and deduced amino acid sequence from *Thymus vulgaris*, *Thymus capitatus and Thymus decussatus* (577, 528 and 544bp) with GenBank accession numbers; MN972464.1/QKE50850.1, MN972465.1/QKE50851.1 and MN972466.1/QKE50852.1, respectively. Nucleotide sequence and deduced sequence of amino acid residues of *rbcL* gene are indicated by a single letter code used by https://web.expasy.org/translate/.



>*Thymus vulgaris* maturase K (*matK*) gene, partial cds; chloroplast (643 bp)

```
5'-
atacgaatctattttcgtctttctacgtaaccaatcttttcaggtacgatcaacatcttgt
 Y   E   S   I   F   V   F   L   R   N   Q   S   F   Q   V   R   S   T   S   C
ggagttcttttgaacgaatctatttctatataaaaatagaacgtcttaggaacgtcttt
 G   V   L   F   E   R   I   Y   F   Y   I   K   I   E   R   L   R   N   V   F
gttaaggattttgcgcgaacctatggttggtcgaggaaccctgcatgcattatattagg
 V   K   D   F   C   A   N   L   W   L   V   E   E   P   C   M   H   Y   I   R
tatcaaaaaaaatctattctggcttccaaagggacatcccttttcatgaataaatggaaa
 Y   Q   K   K   S   I   L   A   S   K   G   T   S   L   F   M   N   K   W   K
ttaaaccttgtcacttttggcaatggcattttctgtgtggtttcatccaagaaggatt
 L   N   L   V   T   F   W   Q   W   H   F   S   V   W   F   H   P   R   R   I
tggataaatcaatttcctaagcattcccttgaaattttgggttatctttcaaacgtccaa
 W   I   N   Q   F   P   K   H   S   L   E   I   L   G   Y   L   S   N   V   Q
atgaactcttccgtggtacggagtcaaattctagaaaattcatttctaatcaataatgct
 M   N   S   S   V   V   R   S   Q   I   L   E   N   S   F   L   I   N   N   A
attaagaagctcgatactcttgttccaattattcctctgattgcagaattggctaaagct
 I   K   K   L   D   T   L   V   P   I   I   P   L   I   A   E   L   A   K   A
aaatttgtaacgtattgggcatcccattagtaagccgattcgggctgagttatcagat
 K   F   C   N   V   L   G   H   P   I   S   K   P   I   R   A   E   L   S   D
tctaatattattgaccgattttcgcgtatatgcagaaatatttctcatcatcatagcgga
 S   N   I   I   D   R   F   S   R   I   C   R   N   I   S   H   H   H   S   G
tcttgcaaaaaaaggagtttgtatcgaataaagtatatactt -3'
 S   C   K   K   R   S   L   Y   R   I   K   Y   I   L
```

>*Thymus capitatus* for maturase K (*matK* gene), partial cds, chloroplast (800bp)

```
5'-tgtg gaa atc ttg att caa att ctt cgc cat tgg gta aaa gat gtt tct tct tta cat tta
     V   E   I   L   I   Q   I   L   R   H   W   V   K   D   V   S   S   L   H   L
   tta cgg gtc ttt ctc aac gaa tat tgg aat tgg aat agt ctt ctt act cca aag aaa gtc
    L   R   V   F   L   N   E   Y   W   N   W   N   S   L   L   T   P   K   K   V
   agc ttc tct ttg tca aaa aga aat caa agg ttc ttt ttt ttc tta tat aat tct cac gta
    S   F   S   L   S   K   R   N   Q   R   F   F   F   F   L   Y   N   S   H   V
   tgt gaa tac gaa tct att ttc gtc ttt cta cgt aac caa tct ttt cat tta cga tca aca
    C   E   Y   E   S   I   F   V   F   L   R   N   Q   S   F   H   L   R   S   T
   tct tct gga gtt ctt ctt gaa cga atc tat ttc tat ata aaa ata gaa cgt ctt atg aac
    S   S   G   V   L   L   E   R   I   Y   F   Y   I   K   I   E   R   L   M   N
   gtc ttt gtt aag gat ttt cgg gcg aac cta tgg ttg gtc gag gaa ccc tgc atg cat tat
    V   F   V   K   D   F   R   A   N   L   W   L   V   E   E   P   C   M   H   Y
   att agg tat caa aga aaa tcc att ctg gct tcc aaa ggg aca tcc ctt ttc atg aat aaa
    I   R   Y   Q   R   K   S   I   L   A   S   K   G   T   S   L   F   M   N   K
   tgg aaa tta aac ctt gtc act ttt tgg caa tgg cat ttt tct gtg tgg ttt cat cca aga
    W   K   L   N   L   V   T   F   W   Q   W   H   F   S   V   W   F   H   P   R
   agg att tgg ata aac caa ttt ccc aag cat tcc ctt gaa att ttg ggc tat ctt tca aac
    R   I   W   I   N   Q   F   P   K   H   S   L   E   I   L   G   Y   L   S   N
   gtg caa atg aac cct tcc gtg gta cgg agt caa att cta gaa aat tca ttt cta atc aat
    V   Q   M   N   P   S   V   V   R   S   Q   I   L   E   N   S   F   L   I   N
   aat gct att aag aag ctc gat act ctt gtt cca att att cct ctg att gca gaa ttg gct
    N   A   I   K   K   L   D   T   L   V   P   I   I   P   L   I   A   E   L   A
   aaa gct aaa ttt tgt aat gta ttg ggg cat ccc att agt aag ccg att cgg gct gag tta
    K   A   K   F   C   N   V   L   G   H   P   I   S   K   P   I   R   A   E   L
   tca gat tct aat att att gac cga ttt tcg cgt ata tgc aga aat att tct cat tat cat
    S   D   S   N   I   I   D   R   F   S   R   I   C   R   N   I   S   H   Y   H
   agc gga tct tgc aaa aaa -3'
    S   G   S   C   K   K
```

> *Thymus decussatus* maturase K (*matK* gene), partial cds, chloroplast (788bp)

```
5'-
aaactcttcgccagtgggtaaaagatgtttcttctttacatttattacgggtctttctcaac
   T   L   R   Q   W   V   K   D   V   S   S   L   H   L   L   R   V   F   L   N
gaatattggaatagtcttcttactccaaagaaagtcagcttctctttgtcaaaaagaaat
 E   Y   W   N   S   L   L   T   P   K   K   V   S   F   S   L   S   K   R   N
caaaggttatttttttcttatataattctcaggtatgtgaatacgaatctattttcgtc
 Q   R   L   F   F   F   L   Y   N   S   Q   V   C   E   Y   E   S   I   F   V
tttctacgtaaccaatcttttcatttacgatcaacatcttctggagttcttcttgaacga
 F   L   R   N   Q   S   F   H   L   R   S   T   S   S   G   V   L   L   E   R
atctatttctatataaaaatagaacgtcttatgaacgtctttgttaaggattttcgcgcg
 I   Y   F   Y   I   K   I   E   R   L   M   N   V   F   V   K   D   F   R   A
aacctatggttggtcgaggaaccctgcatgcattatattaggtatcaaaaaaaatctatt
 N   L   W   L   V   E   E   P   C   M   H   Y   I   R   Y   Q   K   K   S   I
ctggcttccaaagggacatcccttttcatgaataaatggaaattaaaccttgtcactttt
 L   A   S   K   G   T   S   L   F   M   N   K   W   K   L   N   L   V   T   F
tggcaatggcattttctgtgtggtttcatccaagaaggatttggataaaccaatttcct
 W   Q   W   H   F   S   V   W   F   H   P   R   R   I   W   I   N   Q   F   P
aagcattcccttgaaattttgggttatctttcaaacgtccaaatgaactcttccgtggta
 K   H   S   L   E   I   L   G   Y   L   S   N   V   Q   M   N   S   S   V   V
cggagtcaaattctagaaaattcatttctaatcaataatgctattaagaagctcgatact
 R   S   Q   I   L   E   N   S   F   L   I   N   N   A   I   K   K   L   D   T
cttgttccaattattcctctgattgcagaattggctaaagctaaattttgtaacgtattg
 L   V   P   I   I   P   L   I   A   E   L   A   K   A   K   F   C   N   V   L
gggcatcccattagtaagccgattcgggctgagttatcagattctaatattattgaccga
 G   H   P   I   S   K   P   I   R   A   E   L   S   D   S   N   I   I   D   R
ttttcgcgtatatgcagaaatatttctcatcatcatagcggatcttgcaaaaaaaggagt
 F   S   R   I   C   R   N   I   S   H   H   H   S   G   S   C   K   K   R   S
ttgtat -3'
 L   Y
```

Figure 4: The partial fragment *matK* gene and deduced amino acid sequence from *Thymus vulgaris*, *Thymus capitatus* and *Thymus decussatus* (643, 800 and 788bp) with GenBank accession numbers; MN972467.1/QKE50853.1, MN972468.1/QKE50854.1 and MN972469.1/QKE50855.1, respectively. Nucleotide sequence and deduced sequence of amino acid residues of *matK* gene are indicated by a single letter code used by https://web.expasy.org/translate/.

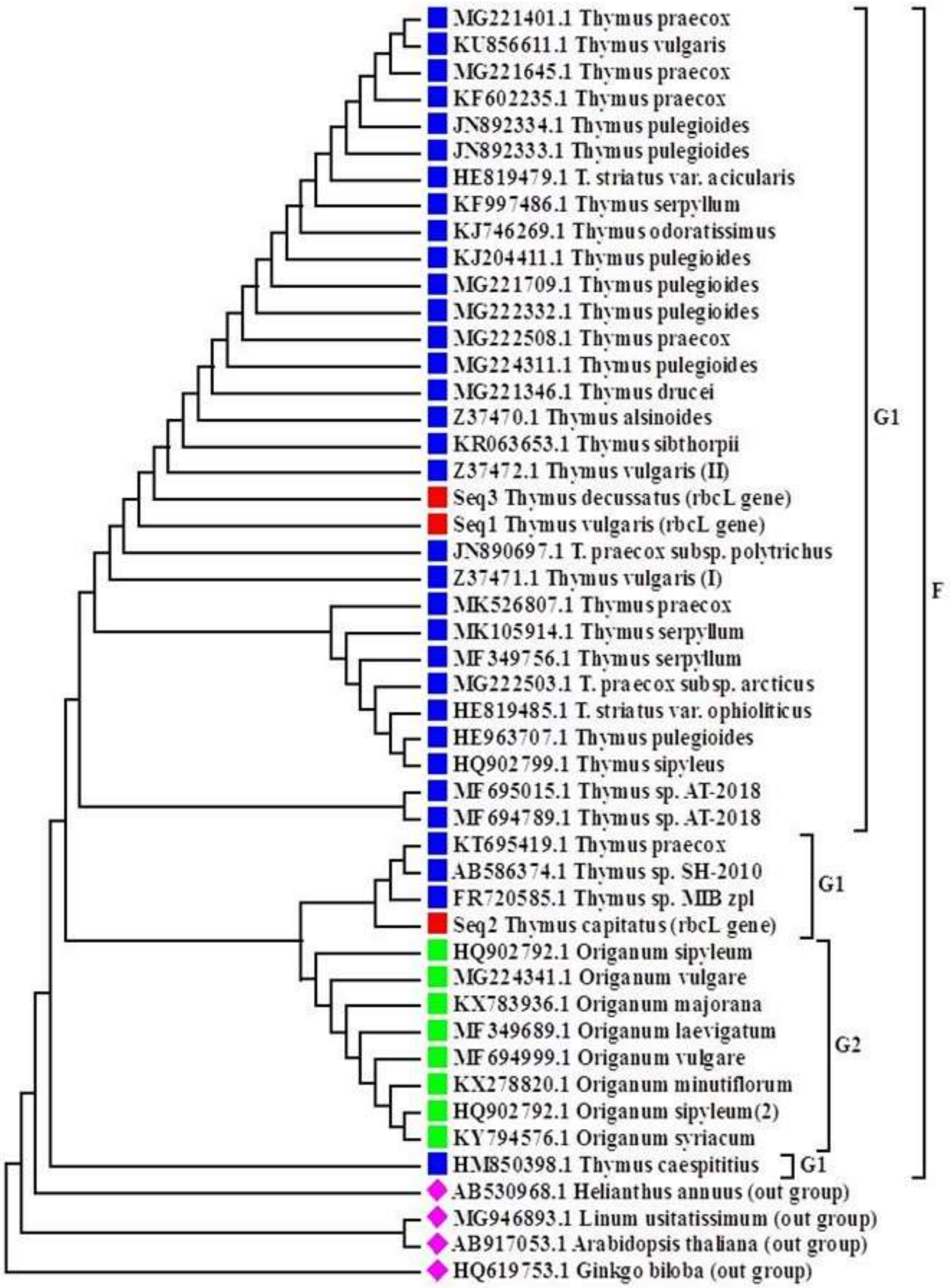


Figure 5: Molecular Phylogenetic analysis involved 48 amino acid sequences of *rbcL* genes was conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-2092.26).
(G1: *Thymus* genus - G2: *Origanum* genus - F: family *Lamiaceae* - Seq1: *Thymus vulgaris*, Seq2: *Thymus capitatus* and Seq3: *Thymus decussatus.*

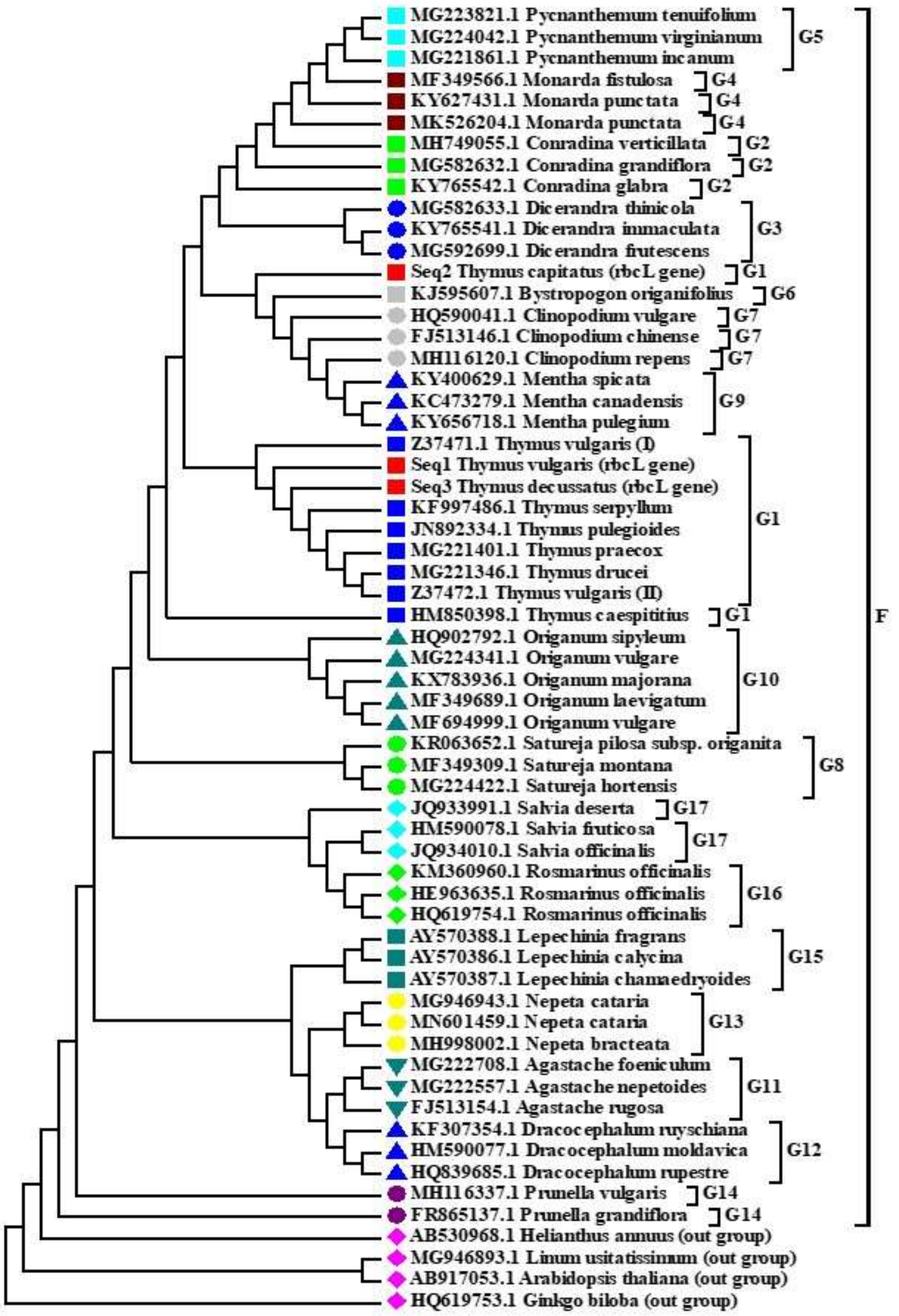


Figure 6: Molecular Phylogenetic analysis involved 61 amino acid sequences of *rbcL* gene was conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-2610.87). G1: *Thymus* species, G2: *Conradina* species, G3: *Dicerandra* species, G4: *Monarda* species, G5: *Pycnant.* Species, G6: *Bystropogon* species, G7: *Clinopodium* species, G8: *Satureja* species, G9: *Mentha* species, G10: *Origanum* species, G11: *Agastache* species, G12: *Dracocephalum* species, G13: *Nepeta* species, G14: *Prunella* species, G15: *Lepechinia* species, G16: *Rosmarinus* species, G17: *Salvia* species and 4 different species from different families as out group. F was family *Lamiaceae.* Seq1: *Thymus vulgaris*, Seq2: *Thymus capitatus* and Seq3: *Thymus decussatus*.

Table 1: Homology of nucleotide sequences for 48 selected accession lists and its related _Thymus vulgaris_, _Thymus capitatus_ and _Thymus decussatus_ for chloroplast, _rbcL_ gene partial sequenced in this study, BLAST top hits against GenBank database, similarity score and GenBank accession number.

Code	Scientific name of chloroplast _rbcL_ gene (length bp)	GenBank Accession No.	Similarity% (bp)		
			*T. vulgaris	_*T. capitatus_	_*T. decussatus_
1	_Thymus vulgaris_ (577bp)	MN972464.1	-------	99.62	99.82
2	_Thymus capitatus_ (528bp)	MN972465.1	99.62	-------	99.43
3	_Thymus decussatus_ (544bp)	MN972466.1	99.82	99.43	-------
4	_Thymus vulgaris_ (1420bp)	Z37471.1	99.65	99.43	99.63
5	_Thymus vulgaris_ (II) (1420bp)	Z37472.1	99.83	99.81	99.63
6	_Thymus vulgaris_ ISr29 (532bp)	KU856611.1	100.00	99.80	100.00
7	_Thymus sibthorpii_ (629bp)	KR063653.1	99.65	99.62	99.82
8	_Thymus alsinoides_ (1420bp)	Z37470.1	99.65	99.62	99.82
9	_Thymus drucei_ (552bp)	MG221346.1	99.82	99.61	99.81
10	_Thymus serpyllum_ (553bp)	MF349756.1	99.82	99.61	99.62
11	_Thymus serpyllum_ (1349bp)	KF997486.1	99.64	99.62	99.82
12	_Thymus serpyllum_ (573bp)	MK105914.1	98.60	99.24	99.07
13	_Thymus sp._ AT-2018 (552bp)	MF694789.1	99.82	99.61	99.62
14	_Thymus_ sp. AT-2018 (553bp)	MF695015.1	99.82	99.61	99.62
15	_Thymus pulegioides_ (552bp)	MG224311.1	99.82	99.61	100.00
16	_Thymus pulegioides_ (552bp)	MG222332.1	99.82	99.61	100.00
17	_Thymus pulegioides_ (552bp)	MG221709.1	99.82	99.61	100.00
18	_Thymus pulegioides_ (549bp)	KJ204411.1	99.82	99.61	100.00
19	_Thymus pulegioides_ (528bp)	HE963707.1	99.62	99.62	99.44
20	_Thymus pulegioides_ (541bp)	JN892333.1	99.82	99.60	100.00
21	_Thymus pulegioides_ (539bp)	JN892334.1	99.81	99.60	100.00
22	_Thymus caespititius_ (568bp)	HM850398.1	99.29	99.62	99.08
23	_Thymus odoratissimus_ (658bp)	KJ746269.1	99.64	99.62	99.82
24	_T. striatus_ var. _ophioliticus_ (599bp)	HE819485.1	99.64	99.62	99.82
25	_T. striatus_ var. _acicularis_ (599bp)	HE819479.1	99.64	99.62	99.82
26	_T. praecox subsp. arcticus_ (552 bp)	MG222503.1	99.82	99.61	99.62
27	_Thymus praecox_ (473bp)	KT695419.1	99.79	100.00	99.77
28	_Thymus praecox_ (1377bp)	KF602235.1	99.63	99.62	100.00
29	_Thymus praecox_ (535bp)	MG221645.1	99.81	99.61	100.00
30	_Thymus praecox_ (549bp)	MK526807.1	99.26	99.01	99.02
31	_Thymus praecox_ (528bp)	MG221401.1	99.81	99.59	100.00
32	_Thymus praecox_ (552bp)	MG222508.1	99.82	99.61	100.00
33	_T. praecox subsp. polytrichus_ (524bp)	JN890697.1	99.81	99.60	100.00
34	_Thymus sipyleus_ (523bp)	HQ902799.1	99.62	99.62	99.43
35	_Thymus sp._ MIB zpl (551bp)	FR720585.1	98.80	100.00	99.40
36	_Thymus sp._ SH-2010 (1301bp)	AB586374.1	99.64	100.00	99.45
37	_Origanum laevigatum_ (1420bp)	Z37426.1	99.98	99.81	99.26
38	_Origanum laevigatum_ (553bp)	MF349689.1	99.64	99.81	99.42
39	_Origanum sipyleum_ (523bp)	HQ902792.1	99.43	99.81	99.42
40	_Origanum vulgare_(552bp)	MG224341.1	99.64	99.81	99.42
41	_Origanum vulgare_ (553bp)	MF694999.1	99.64	99.81	99.42
42	_Origanum majorana_ (552bp)	KX783936.1	99.64	99.81	99.42
43	_Origanum minutiflorum_ (555bp)	KX278820.1	99.64	99.80	99.42
44	_Origanum syriacum_ (926bp)	KY794576.1	99.46	99.81	99.26
45	_Linum usitatissimum_ (1419bp)	MG946893.1	89.81	89.43	90.28
46	_Helianthus annuus_ (1244bp)	AB530968.1	91.06	91.05	90.87
47	_Arabidopsis thaliana_ (1326bp)	AB917053.1	90.77	90.17	90.44
48	_Ginkgo biloba_ (553bp)	HQ619753.1	87.55	88.01	87.91

*_Thymus vulgaris_ (MN972464.1), _Thymus capitatus_ (MN972465.1) and _Thymus decussatus_ (MN972466.1) (Species under study)

.

Table 2. Homology of nucleotide sequences for 61 selected accession lists and its related *T. vulgaris*, *T. capitatus* and *T. decussatus* for chloroplast, *rbcL* gene partial sequenced in this study, BLASTn top hits against GenBank database, similarity score and GenBank accession number.

Family *Lamiaceae* Tribe: *Mentheae*	Code	Scientific name of chloroplast *rbcL* gene (length bp)	GenBank Accession No.	Similarity% (bp)		
				T. vulgaris	*T. capitatus*	*T. decussatus*
1. Genus *Thymus*	1	**Thymus vulgaris* (577bp)	MN972464.1	-------	99.62	99.82
	2	**Thymus capitatus* (528bp)	MN972465.1	99.62	-------	99.43
	3	**Thymus decussatus* (544bp)	MN972466.1	99.82	99.43	-------
	4	*Thymus vulgaris*	Z37471.1	99.65	99.43	99.63
	5	*Thymus vulgaris* (II)	Z37472.1	99.83	99.81	99.63
	6	*Thymus drucei*	MG221346.1	99.82	99.61	99.81
	7	*Thymus serpyllum*	KF997486.1	99.64	99.62	99.82
	8	*Thymus pulegioides*	JN892334.1	99.81	99.60	100.00
	9	*Thymus caespititius*	HM850398.1	99.29	99.62	99.08
	10	*Thymus praecox*	MG221401.1	99.81	99.59	100.00
2. Genus *Conradina*	11	*Conradina glabra*	KY765542.1	99.60	99.78	99.57
	12	*Conradina grandiflora*	MG582632.1	99.24	99.41	99.02
	13	*Conradina verticillata*	MH749055.1	99.64	99.81	99.42
3. Genus *Dicerandra*	14	*Dicerandra thinicola*	MG582633.1	99.26	99.41	99.02
	15	*Dicerandra immaculata*	KY765541.1	99.19	99.34	99.13
	16	*Dicerandra frutescens*	MG592699.1	99.24	99.34	99.02
4. Genus *Monarda*	17	*Monarda punctata*	MK526204.1	99.63	99.80	99.41
	18	*Monarda punctata*	KY627431.1	99.63	99.80	99.42
	19	*Monarda fistulosa*	MF349566.1	99.64	99.81	99.42
5. Genus *Pycnant.*	20	*Pycnanthemum incanum*	MG221861.1	99.62	99.81	99.42
	21	*Pycnanthemum tenuifolium*	MG223821.1	99.64	99.81	99.42
	22	*Pycnanthemum virginianum*	MG224042.1	99.64	99.81	99.42
6. Genus *Bystropogon*	23	*Bystropogon origanifolius*	KJ595607.1	99.65	100.00	99.45
7. Genus *Clinopodium*	24	*Clinopodium vulgare*	HQ590041.1	99.65	100.00	99.45
	25	*Clinopodium chinense*	FJ513146.1	99.65	100.00	99.45
	26	*Clinopodium repens*	MH116120.1	99.64	100.00	99.45
8. Genus *Satureja*	27	*Satureja pilosa subsp. origanita*	KR063652.1	99.31	99.62	99.08
	28	*Satureja montana*	MF349309.1	99.46	99.61	99.23
	29	*Satureja hortensis*	MG224422.1	99.46	99.61	99.23
9. Genus *Mentha*	30	*Mentha spicata*	KY400629.1	99.65	100.00	99.44
	31	*Mentha canadensis*	KC473279.1	99.65	100.00	99.45
	32	*Mentha pulegium*	KY656718.1	99.65	100.00	99.45
10. Genus *Origanum*	33	*Origanum sipyleum*	HQ902792.1	99.43	99.81	99.42
	34	*Origanum vulgare*	MG224341.1	99.64	99.81	99.42
	35	*Origanum majorana*	KX783936.1	99.64	99.81	99.42
	36	*Origanum laevigatum*	MF349689.1	99.64	99.81	99.42
	37	*Origanum vulgare*	MF694999.1	99.64	99.81	99.42
11. Genus *Agastache*	38	*Agastache rugosa*	FJ513154.1	97.92	98.48	97.79
	39	*Agastache foeniculum*	MG222708.1	98.19	98.45	98.08
	40	*Agastache nepetoides*	MG222557.1	98.19	98.45	98.08
12. Genus *Dracocephalum*	41	*Dracocephalum moldavica*	HM590077.1	97.05	98.45	96.88
	42	*Dracocephalum rupestre*	HQ839685.1	97.23	97.54	97.06
	43	*Dracocephalum ruyschiana*	KF307354.1	97.39	97.73	97.23
13. Genus *Nepeta*	44	*Nepeta cataria*	MG946943.1	97.57	97.92	97.43
	45	*Nepeta bracteata*	MH998002.1	97.23	97.54	97.06
	46	*Nepeta cataria*	MN601459.1	97.57	97.92	97.43
14. Genus *Prunella*	47	*Prunella vulgaris*	MH116337.1	98.03	98.30	97.79
	48	*Prunella grandiflora*	FR865137.1	97.21	97.53	97.03
15. Genus	49	*Lepechinia chamaedryoides*	AY570387.1	98.38	98.67	98.16

Lepechinia	50	*Lepechinia fragrans*	AY570388.1	98.25	98.67	98.16
	51	*Lepechinia calycin*	AY570386.1	98.26	98.67	98.16
16. Genus *Rosmarinus*	52	*Rosmarinus officinalis*	HE963635.1	98.87	99.23	98.69
	53	*Rosmarinus officinalis*	HQ619754.1	98.92	99.22	98.85
	54	*Rosmarinus officinalis*	KM360960.1	99.12	99.43	98.90
17. Genus *Salvia*	55	*Salvia deserta*	JQ933991.1	99.13	99.62	99.08
	56	*Salvia fruticosa*	HM590078.1	99.13	99.62	99.08
	57	*Salvia officinalis*	JQ934010.1	99.13	99.62	99.08
Family *linaceae*	58	*Linum usitatissimum*	MG946893.1	89.81	89.43	90.28
Family *Asteraceae*	59	*Helianthus annuus*	AB530968.1	91.06	91.05	90.87
Family Brassicaceae	60	*Arabidopsis thaliana*	AB917053.1	90.77	90.17	90.44
Family Ginkgoaceae	61	*Ginkgo biloba*	HQ619753.1	87.55	88.01	87.91

* *Thymus vulgaris* (MN972467.1), *Thymus capitatus* (MN972468.1) and *Thymus decussatus* (MN972467.1) were (Species under study)

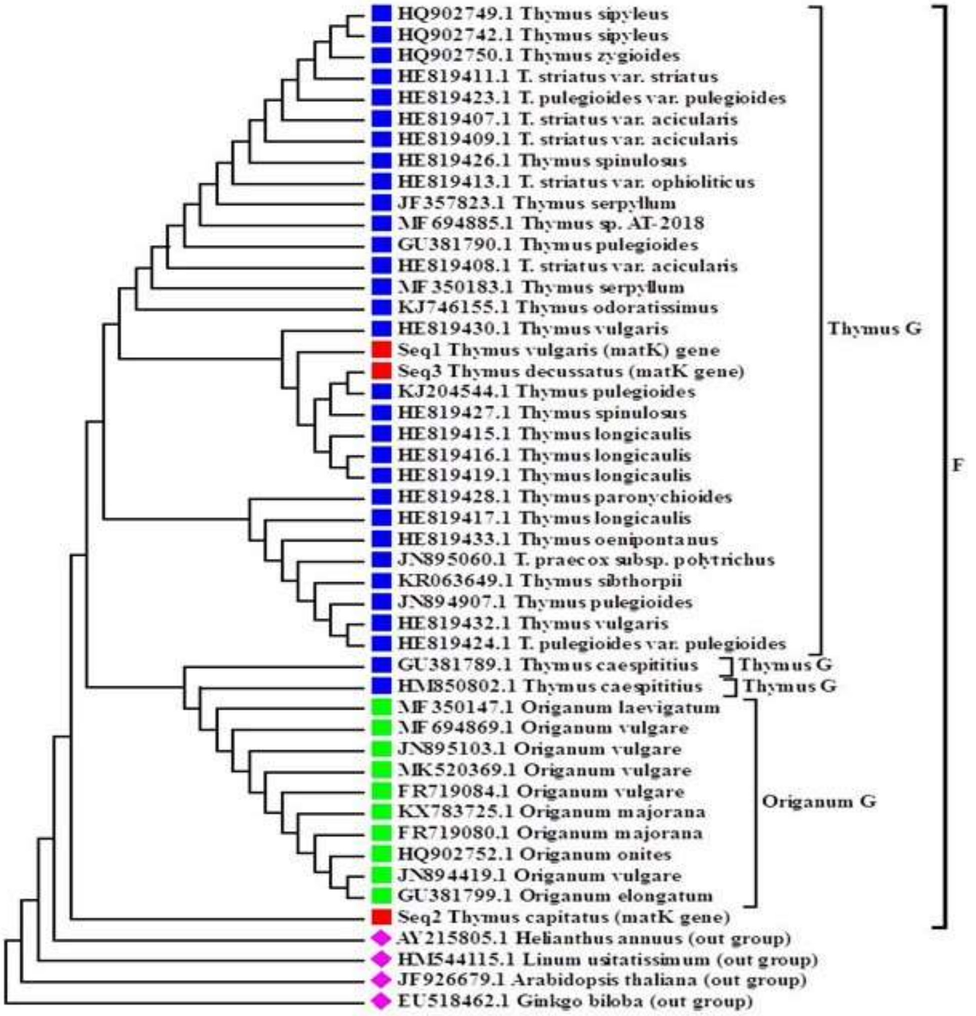


Figure 7: Molecular Phylogenetic analysis involved 48 amino acid sequences of maturase gene (matK gene) was conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-3275.04). (*Thymus* G: *Thymus* genus - *Origanum* G: *Origanum* genus -F: family *Lamiaceae* - Seq1: *Thymus vulgaris*, Seq2: *Thymus capitatus* and Seq3: *Thymus decussatus* .

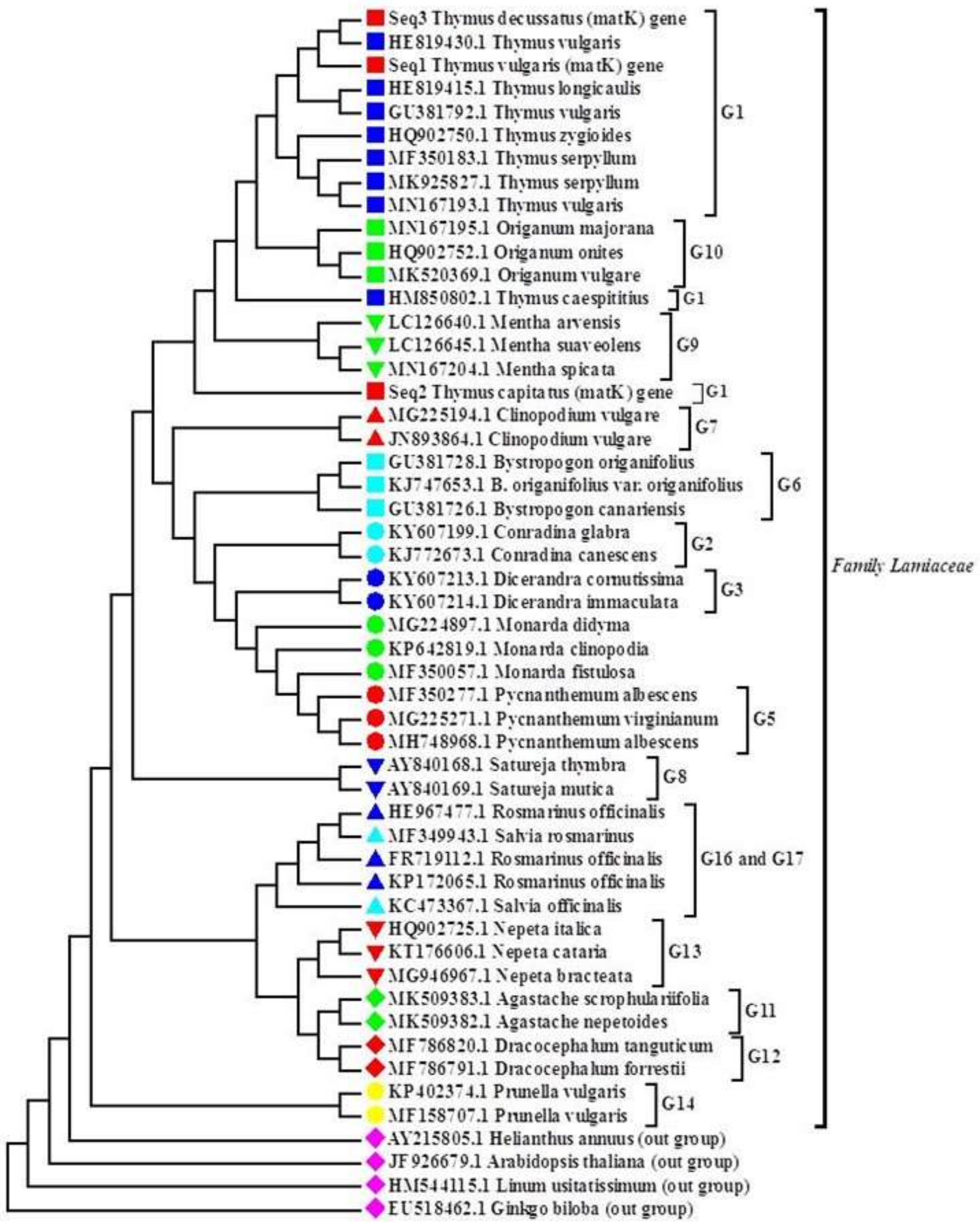


Figure 8: Molecular Phylogenetic analyses involved 52 amino acid sequences of maturase gene (matK) from 17 genus belong to the family Lamiaceae were conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-3908.76) is shown. G1: *Thymus* species, G2: *Conradina* species, G3: *Dicerandra* species, G4: *Monarda* species, G5: *Pycnant.* Species, G6: *Bystropogon* species, G7: *Clinopodium* species, G8: *Satureja* species, G9: *Mentha* species, G10: *Origanum* species, G11: *Agastache* species, G12: *Dracocephalum* species, G13: *Nepeta* species, G14: *Prunella* species, G15: *Lepechinia* species (No. record), G16: *Rosmarinus* species, G17: *Salvia* species and 4 different species from different families as out group.

Table 3: Homology of nucleotide sequences for 48 selected accession lists and its related _Thymus vulgaris_, _Thymus capitatus_ and _Thymus decussatus_ for chloroplast, partial maturase gene (_matK_) sequenced in this study, BLAST top hits against GenBank database, similarity score and GenBank accession number.

Code	Scientific name of chloroplast _rbcL_ gene (length bp)	GenBank Accession No.	Similarity% (bp)		
			*T. vulgaris	_*T. capitatus_	_*T. decussatus_
1	_*Thymus vulgaris_ (643bp)	MN972467.1	-------	97.39	98.88
2	_*Thymus capitatus_ (800bp)	MN972468.1	97.39	-------	97.57
3	_*Thymus decussatus_ (788bp)	MN972469.1	98.88	97.57	------
4	_Thymus vulgaris_ (789bp)	HE819432.1	98.30	97.81	99.34
5	_Thymus vulgaris_ (789bp)	HE819430.1	98.64	97.68	99.60
6	_T. praecox subsp. Polytrichus_ (795bp)	JN895060.1	98.29	96.50	98.07
7	_Thymus sibthorpii_ (1014bp)	KR063649.1	98.29	98.41	98.56
8	_Thymus pulegioides_ (1542bp)	GU381790.1	98.71	93.00	93.20
9	_Thymus pulegioides_ (808bp)	JN894907.1	98.33	97.71	99.34
10	_Thymus pulegioides_ (748bp)	KJ204544.1	98.46	97.48	99.47
11	_Thymus sp. AT-2018_ (789bp)	MF694885.1	98.24	97.57	99.24
12	_Thymus serpyllum_ (829bp)	JF357823.1	98.37	97.62	99.23
13	_Thymus serpyllum_ (810bp)	MF350183.1	98.44	98.55	98.70
14	_T. striatus var. ophioliticus_ (795bp)	HE819413.1	98.64	98.71	98.68
15	_Thymus spinulosus_ (789bp)	HE819427.1	98.47	97.55	99.47
16	_Thymus spinulosus_ (789bp)	HE819426.1	98.47	97.81	99.34
17	_Thymus longicaulis_ (789bp)	HE819419.1	98.47	97.55	99.47
18	_Thymus longicaulis_ (789bp)	HE819415.1	98.47	97.55	99.47
19	_Thymus longicaulis_ (789bp)	HE819417.1	97.96	97.55	99.07
20	_Thymus longicaulis_ (789bp)	HE819416.1	98.64	97.68	99.60
21	_T. striatus var. acicularis_ (789bp)	HE819409.1	98.47	97.94	99.47
22	_T. striatus var. acicularis_ (789bp)	HE819407.1	98.47	97.81	99.34
23	_T. striatus var. acicularis_ (789bp)	HE819408.1	98.30	97.81	99.34
24	_T. striatus var. striatus_ (789bp)	HE819411.1	98.30	98.19	99.20
25	_T. pulegioides var. pulegioides_ (789bp)	HE819423.1	98.30	97.68	99.20
26	_T. pulegioides var. pulegioides_ (789bp)	HE819424.1	98.13	97.68	99.20
27	_Thymus oenipontanus_ (789bp)	HE819433.1	98.13	97.68	99.20
28	_Thymus paronychioides_ (789bp)	HE819428.1	98.13	97.68	99.20
29	_Thymus zygioides_ (802bp)	HQ902750.1	98.61	98.29	98.38
30	_Thymus sipyleus_ (802bp)	HQ902749.1	98.44	97.56	98.52
31	_Thymus sipyleus_ 789bp)	HQ902742.1	98.44	97.90	99.46
32	_Thymus odoratissimus_ (693bp)	KJ746155.1	98.26	97.57	99.28
33	_Thymus caespititius_ (789bp)	HM850802.1	97.85	97.94	97.66
34	_Thymus caespititius_ (1494bp)	GU381789.1	98.05	98.64	99.16
35	_Origanum laevigatum_ (817bp)	MF350147.1	98.44	98.69	98.58
36	_Origanum vulgare_ (785)	MF694869.1	98.38	98.60	98.47
37	_Origanum vulgare_ (724bp)	JN895103.1	98.37	98.62	99.30
38	_Origanum vulgare_ (782bp)	MK520369.1	98.34	98.59	98.44
39	_Origanum vulgare_ (810bp)	FR719084.1	98.30	98.71	98.42
40	_Origanum majorana_ (793bp)	KX783725.1	98.30	98.71	98.41
41	_Origanum majorana_ (810bp)	FR719080.1	98.13	98.45	98.15
42	_Origanum onites_ (769bp)	HQ902752.1	98.26	98.69	98.39
43	_Origanum vulgare_ (773bp)	JN894419.1	98.04	98.53	98.22
44	_Origanum elongatum_ (1530bp)	GU381799.1	98.39	98.65	99.50
45	_Linum usitatissimum_ (1116bp)	HM544115.1	71.13	70.67	70.50
46	_Helianthus annuus_ (1503bp)	AY215805.1	79.00	79.10	78.42
47	_Arabidopsis thaliana_ (1041bp)	JF926679.1	71.61	72.43	71.32
48	_Ginkgo biloba_ (1539bp)	EU518462.1	67.36	66.80	66.80

***** _Thymus vulgaris_ (MN972467.1), _Thymus capitatus_ (MN972468.1) and _Thymus decussatus_ (MN972467.1) were used as current study.

Table 4: Homology of nucleotide sequences for 52 selected accession lists and its related *T. vulgaris*, *T. capitatus* and *T. decussatus* for chloroplast partial maturase gene (*matK*) sequenced in this study, BLAST top hits against GenBank database, similarity score and GenBank accession.

Family: *Lamiaceae* Tribe: *Mentheae*	**Code**	**Scientific name of chloroplast *rbcL* gene (length bp)**	**GenBank Accession No.**	**Similarity% (bp)**		
				T. vulgaris	***T. capitatus***	***T. decussatus***
1. Genus *Thymus*	1	**Thymus vulgaris* (643bp)	MN972467.1	-------	97.39	98.88
	2	**Thymus capitatus* (800bp)	MN972468.1	97.39	-------	97.57
	3	**Thymus decussatus* (788bp)	MN972469.1	98.88	97.57	------
	4	*Thymus caespititius*	HM850802.1	97.85	97.94	97.66
	5	*Thymus zygioides*	HQ902750.1	98.61	98.29	98.38
	6	*Thymus longicaulis*	HE819415.1	98.47	97.55	99.87
	7	*Thymus vulgaris*	HE819430.1	98.64	97.68	99.60
	8	*Thymus vulgaris*	GU381792.1	98.70	98.30	99.83
	9	*Thymus serpyllum*	MF350183.1	98.44	97.55	98.70
	10	*Thymus serpyllum*	MK925827.1	98.60	97.83	99.49
	11	*Thymus vulgaris*	MN167193.1	98.60	97.95	99.49
2. Genus *Conradina*	12	*Conradina glabra*	KY607199.1	96.19	97.94	97.14
	13	*Conradina canescens*	KJ772673.1	96.32	98.16	97.24
3. Genus *Dicerandra*	14	*Dicerandra cornutissima*	KY607213.1	95.52	97.63	96.84
	15	*Dicerandra immaculata*	KY607214.1	95.84	97.94	97.14
4. Genus *Monarda*	16	*Monarda didyma*	MG224897.1	96.02	98.23	96.63
	17	*Monarda clinopodia*	KP642819.1	95.89	98.14	96.54
	18	*Monarda fistulosa*	MF350057.1	96.26	98.30	96.77
5. Genus *Pycnant.*	19	*Pycnanthemum albescens*	MF350277.1	96.17	98.47	96.88
	20	*Pycnanthemum virginianum*	MG225271.1	96.20	98.48	96.90
	21	*Pycnanthemum albescens*	MH748968.1	96.35	98.54	96.99
6. Genus *Bystropogon*	22	*Bystropogon canariensis*	GU381726.1	96.85	99.01	97.89
	23	*Bystropogon origanifolius*	GU381728.1	96.85	99.01	97.89
	24	*Bystropogon origanifolius*	KJ747653.1	96.86	98.95	97.41
7. Genus *Clinopodium*	25	*Clinopodium vulgare*	MG225194.1	96.87	98.81	97.27
	26	*Clinopodium vulgare*	JN893864.1	96.88	98.75	97.23
8. Genus *Satureja*	27	*Satureja thymbra*	AY840168.1	96.75	98.98	97.82
	28	*Satureja mutica*	AY840169.1	96.90	99.14	97.98
9. Genus *Mentha*	29	*Mentha arvensis*	LC126640.1	97.36	98.96	97.45
	30	*Mentha suaveolens*	LC126645.1	97.36	99.22	97.69
	31	*Mentha spicata*	MN167204.1	97.36	99.36	97.72
10. Genus *Origanum*	32	*Origanum onites*	HQ902752.1	98.26	98.69	98.39
	33	*Origanum vulgare*	MK520369.1	98.34	98.59	98.44
	34	*Origanum majorana*	MN167195.1	98.29	98.59	98.36
11. Genus *Agastache*	35	*Agastache scrophulariifolia*	MK509383.1	94.00	96.15	94.67
	36	*Agastache nepetoides*	MK509382.1	94.01	96.16	94.68
12. Genus *Dracocephalum*	37	*Dracocephalum tanguticum*	MF786820.1	91.64	93.68	92.73
	38	*Dracocephalum forrestii*	MF786791.1	92.22	94.34	93.31
13. Genus *Nepeta*	39	*Nepeta italica*	HQ902725.1	93.24	95.41	93.97
	40	*Nepeta cataria*	KT176606.1	92.39	94.88	93.08
	41	*Nepeta bracteata*	MG946967.1	92.86	94.50	93.08
14. Genus *Prunella*	42	*Prunella vulgaris*	KP402374.1	94.04	96.86	95.44
	43	*Prunella vulgaris*	MF158707.1	93.71	95.99	95.35
15. Genus *Lepechinia*	---	*No record*	*No record*	----	----	----
16. Genus *Rosmarinus*	44	*Rosmarinus officinalis*	KP172065.1	94.08	95.24	95.28
	45	*Rosmarinus officinalis*	FR719112.1	94.06	95.23	95.21
	46	*Rosmarinus officinalis*	HE967477.1	93.96	95.30	95.12
17. Genus *Salvia*	47	*Salvia rosmarinus*	MF349943.1	94.25	95.44	95.34
	48	*Salvia officinalis*	KC473367.1	94.09	96.25	94.58
Family *linaceae*	49	*Linum usitatissimum*	HM544115.1	71.13	70.67	70.50

Family *Asteraceae*	50	*Helianthus annuus*	AY215805.1	79.00	79.10	78.42
Family Brassicaceae	51	*Arabidopsis thaliana*	JF926679.1	71.61	72.43	71.32
Family Ginkgoaceae	52	*Ginkgo biloba*	EU518462.1	67.36	66.80	66.80

.

***** *Thymus vulgaris* (MN972467.1), *Thymus capitatus* (MN972468.1) and *Thymus decussatus* (MN972467.1) (Species under study).

It also revealed similarity less than 100% with the other 32 accessions of genus *Thymus* and 8 accessions of genus *Origanum,* out of which 99.01% belong to the family *Lamiaceae* with E-value = Zero and showed low similarity with the other different species (89.43%, 91.05% and 90.17%) with *Linum usitatissimum* (MG946893.1), *Helianthus annuus* (AB530968.1) and *Arabidopsis thaliana* (AB917053.1), respectively with E-value = Zero. In addition, it showed low similarity 88.01% to *Ginkgo biloba* (HQ619753.1) with E-value = 5e-179 as an out group of the family *Lamiaceae*. Thirdly, the Multiple sequence alignments (MSA) of nucleotide partial *rbcL* gene from *Thymus decussatus* (current study) exhibited high identity and similarity 100% to 12 accessions no. namely, *Thymus vulgaris* ISr29 (KU856611.1), *T. pulegioides* (MG224311.1), *T. pulegioides* (MG222332.1), *T. pulegioides* (MG221709.1), *T. pulegioides* (KJ204411.1), *T. pulegioides* (JN892333.1), *T. pulegioides* (JN892334.1), *T. praecox (*KF602235.1), *T. praecox* (MG221645.1), *T. praecox* (MG221401.1), *T. praecox* (MG222508.1), *T. praecox* subsp*.* and *polytrichus* (JN890697.1). It also exhibited similarity less than 100% with 23 accessions of genus *Thymus* and 8 accessions of genus *Origanum*, of which 99.07 % belong to family *Lamiaceae* with E-value = Zero whereas *T. decussatus* showed low similarity with the other different species (90.28%, 90.87% and 90.44%) *Linum usitatissimum* (MG946893.1), *Helianthus annuus*, (AB530968.1) and *Arabidopsis thaliana* (AB917053.1), respectively with E-value = Zero. Moreover, it showed low similarity 87.91% with *Ginkgo biloba* (HQ619753.1) with E-value = 1e-179 as an out group of family *Lamiaceae*. Based on the previous results, phylogenetic relationships analyses enable us to check the species relatedness to plant species under investigation (*T. vulgaris, T. capitatus and T. decussatus*).The phylogenetic trees were constructed based on the 48 amino acids sequence encoded from the *rbcL* gene including 3 *Thymus* species (*T. vulgaris, T. capitatus and T. decussatus*) that have the highest percentages of similarity with species from genus *Thymus* and genus *Origanum*, showing that, closely related species from family *Lamiaceae* were clustered together and they showed low percentages of similarities towards relatively distant related species belonging to different families (out group of family Lamiaceae) as shown in Figure 5.

Sequence homology of each nucleotide of the 61 chloroplast *rbcL* genes searches (Code 1 to 3) for *Thymus* species (current study), (Code 4 to 10) *for Thymus* genus, (Code 11 to 13) for *Conradina* genus, (Code 14 to16) for *Dicerandra* genus, (Code 17 to 19) *for Monarda* genus, (Code 20 to 22) for *Pycnanthemum* genus, (Code 23) *for Bystropogon* genus, (Code 24 to 26) for *Clinopodium* genus, (Code 27 to 29) for *Satureja* genus, (Code 30 to 32) for *Mentha* genus, (Code 33 to 37) for *Origanum* genus, (Code 38 to 40) *for Agastache* genus, (Code 41 to 43) for *Dracocephalum* genus, (Code 44 to 46) for *Nepeta* genus, (Code 47 to 48) for *Prunella* genus, (Code 49 to 51) for *Lepechinia* genus, (Code 52 to 54) for *Rosmarinus* genus, (Code 55 to 57) for *Salvia* genus, and (Code 58 to 61) for 4 different species *Linum usitatissimum* (MG946893.1), *Helianthus annuus*, (AB530968.1) *Arabidopsis thaliana* (AB917053.1) and *Ginkgo biloba* (HQ619753.1) as an 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).

The Phylogenetic relatioships analyses were conducted by MEGA 7.0 software program using Maximum Likelihood (ML) tree with the highest log likelihood (-2610.87) (Kumar et al. 2016). The phylogenetic trees were constructed based on the 61 amino acids sequences encoded from the *rbcL* genes that have the highest similarity with species from 17 different genera, showing closely related species clustering together from family Lamiaceae and they showed low percentages of similarities towards relatively distant related species from families other than family Lamiaceae (Figure 6).

Molecular phylogenetic analysis based on sequences of *matK* gene

Sequence homology of each nucleotide sequence of the 48 chloroplast *matK* genes (Code 1 to 3) for *Thymus* species (current study), (Code 4 to 34) for *Thymus* genus, (Code 35 to 44) for *Origanum* genus, and (Code 45 to 48) for 4 different species as an out group of the family were tested by top-scoring hits through NCBI database using Basic Local Alignment Search Tool (BLASTn). Multiple sequence alignments (MSA) of 48 *matK* genes were carried out between nucleotides sequences of two genera (34 *Thymus* and 10 *Origanum* genus) belonging to the family *Lamiaceae* or *Labiataet* of which, the 3 *Thymus* species, *T. vulgaris, T. capitatus and T. decussatus* (current study) and 4 different species from different families as an out group available in GenBank database. The results revealed that *Thymus vulgaris* (current study) in this investigation was closely related to 43 accession numbers sharing high identity and similarity as shown in Table (3). For convenience, *T. vulgaris* revealed high similarity to 33 accessions (98.88 – 97.85%) with *Thymus decussatus* (MN972469.1), *T. pulegioides* (GU381790.1), *T. vulgaris* (HE819430.1), *T. striatus var. ophioliticus* (HE819413.1), *T. vulgaris* (HE819432.1), *T. zygioides* (HQ902750.1), *T. sipyleus* (HQ902742.1), *T. sipyleus* (HQ902749.1), *T. caespititius* (GU381789.1), *T. caespititius* (HM850802.1) and 99.44% - 99.04% with 10 *Origanum* genera which belongs also to family *Lamiaceae* with E-value = Zero. Moreover, it showed low similarity with the other different species 71.13% *Linum usitatissimum* (HM544115.1) with E-value = 1e-68, 79.00% *Helianthus annuus* (AY215805.1) with E-value =2e-160, 71.61% *Arabidopsis thaliana* (JF926679.1) with E-value = 2e-54 and 67.36% *Ginkgo biloba* (EU518462.1) with E-value = 7e-24 as an out group of family *Lamiaceae*. Additionally, *Thymus capitatus* (current study) was closely related and shared high identity and similarity to 33 accessions (98.71 – 93.00%) of which, *T. striatus var. ophioliticus* (HE819413.1), *T. caespititius* (GU381789.1), *T. sibthorpii* (KR063649.1), *T. vulgaris* (HE819432.1), *T. vulgaris* (HE819430.1), *T. vulgaris* (MN972467.1), *T. serpyllum* (MF350183.1), *T. caespititius* (HM850802.1), *T. pulegioides* (GU381790.1), and the other listed *Thymus* species with E-value = Zero, 98.71% - 98.45% with 10 *Origanum* genus which belongs also toamily *Lamiaceae* with E-value = Zero, while it showed low similarity with the other different species 70.67% *Linum usitatissimum* (HM544115.1) with E-value = 3e-78, 79.10% *Helianthus annuus* (AY215805.1) with E-value = Zero, 72.43% *Arabidopsis thaliana* (JF926679.1) with E-value = 1e-89 and 66.80% *Ginkgo biloba* (EU518462.1) with E-value = 6e-19 as an out group of the family *Lamiaceae*. Finally, the MSA of nucleotide partial *matK* gene for *Thymus decussatus* (current study) showed its close relatedness to 33 accessions (99.60 – 93.20 %) with high identity and similarity out of which, *T. vulgaris* (HE819430.1), *T. longicaulis* (HE819419.1), *T. longicaulis* (HE819415.1), *T. longicaulis* (HE819416.1), *T. spinulosus* (HE819427.1), *T. longicaulis* (HE819417.1), *T. longicaulis* (HE819415.1), *T. vulgaris* (HE819432.1), *T. vulgaris* (MN972467.1), *T. striatus var. ophioliticus* (HE819413.1), *Thymus pulegioides* (GU381790.1) and the other listed *Thymus* species with E-value = Zero, 99.50% - 98.15% with 10 *Origanum* genera which belong also to family *Lamiaceae* with E-value = Zero, while it showed low similarity with the other different species 70.50% *Linum usitatissimum* (HM544115.1) with E-value = 2e-75, 78.42% *Helianthus annuus* (AY215805.1) with E-value = Zero, 71.32% *Arabidopsis thaliana* (JF926679.1) with E-value = 4e-76 and 66.80% *Ginkgo biloba* (EU518462.1) with E-value = 6e-19 as an out group of the family *Lamiaceae*. Phylogenetic relationships analyses were conducted in MEGA 7.0 software program by Maximum Likelihood (ML) tree with the highest log likelihood (-3275.04) (Kumar *et al*., 2016).They enable us to check and clarify the relatedness between plant species. The phylogenetic trees were constructed based on the 48 amino acid sequences encoded from the maturase gene (*matK*) including 3 *Thymus* species (*T. vulgaris, T. capitatus and T. decussatus*) that have the highest percentages of similarity with species from *Thymus* genus and *Origanum* genus, showing closely related species clustering together from family *Lamiaceae* and they showed low similarities towards relatively distant related species belonging to 4 different families as an out group of the family (Figure 7).

Sequence homology of each nucleotide of the 52 chloroplast maturase gene (matK) searches (Code 1 to 3) for Thymus species (current study), (Code 4 to 11) for Thymus genus, (Code 12 to 13) Conradina genus, (Code 14 to15) for Dicerandra genus, (Code 16 to 18) for Monarda genus, (Code 19 to 21) Pycnanthemum genus, (Code 22 to 24) for Bystropogon genus, (Code 25 to 26) for Clinopodium genus, (Code 27 to 28) for Satureja genus, (Code 29 to 31) for Mentha genus, (Code 32 to 34) for Origanum genus, (Code 35 to 36) for Agastache genus, (Code 37 to 38) for Dracocephalum genus, (Code 39 to 41) for Nepeta genus, (Code 42 to 43) for Prunella genus and Lepechinia genus (No record), (Code 44 to 46) for Rosmarinus genus, (Code 47 to 48) for Salvia genus, and (Code 49 to 52) for 4 different species Linum usitatissimum (HM544115.1), Helianthus annuus (AY215805.1), Arabidopsis thaliana (JF926679.1) and Ginkgo biloba (EU518462.1) as an out group of family were tested by top-scoring hits through NCBI database using Basic Local Alignment Search Tool (BLASTn) as shown in Table (4). The phylogenetic trees were conducted using MEGA 7.0 software program by Maximum Likelihood (ML) tree with the highest log likelihood (-3908.76) (Kumar et al., 2016). The phylogenetic trees were constructed based on the 52 amino acid sequences encoded from maturase gene (matK) including 3 Thymus species (*T. vulgaris, T. capitatus* and *T. decussatus*) that have the highest percentages of similarity with species from 16 genus, no record with genus Lepechinia, showing closely related species clustering together from family Lamiaceae and they showed low similarities towards relatively distant related species from different families as an out group of the family (Figure 8).

Our data analysis were in agreement with the Plant Working Group CBOL, 2009; Fazekas et al., 2008; and Newmaster et al., 2008 for examined the suitability of different leading candidate markers and proposed the two-locus combination of matK and rbcL as the core plant barcodes. Other authors Bafeel et al., (2011) used universal matK primer for matK as a barcode. Schori and Showalter, (2011) analyzed 14 species from Labiatae in Pakistan, the rbcL, matK, psbA-trnH loci, could serve as single-region barcodes depending on plant to be identified, one region was preferred over the other to aid with species identification. Bafeel et al., (2012) tested the potential of the rbcL marker for the identification of wild plants belonging to diverse families of arid regions. Maximum likelihood tree analysis was performed to evaluate the discriminatory power of the rbcL gene. Federici et al., (2013) showed clear amplification and sequencing of 36 samples of Thymus spp. using molecular analysis based on the core barcode regions (matK and rbcL) and the plastid intergenic spacer trnH-psbA. Skuza et al., (2019) observed that nucleotide sequences had a high variability within matK and rbcL regions and the matK region is suitable for differentiation and discrimination between the studied species.

CONCLUSION

The results revealed that the two partial genes (*rbcL* and *matK*) as a multi locus barcoding markers are the most suitable for the molecular authentication and identification of the three Egyptian *Thymus* species under study. They also enabled us to define the closest 17 different genera of family Lamiaceae to Thymus species and could be helpful in the identification and monitoring of the adulteration of seed varieties under national and international trade.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

AUTHOR CONTRIBUTIONS

M.Z.S participated in collection of plant samples from their natural habitats in Egypt and also in Experimental work, data analysis and drafting of the manuscript and discussion.

ELS.S participated in drafting manuscript and discussion, reviewing manuscript

Sh.D.I participated in experimental work and data analysis. All authors read and approved the final version.

REFERENCES

Ajmal AM, Gyulai G, Hidwegi N, Kerti B, Al Hemaid F, Pandey AK, Lee J (2014) The changing epitome of species identification–DNA barcoding. Saudi J Biol Sci 21:204–231

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol., 215*(5), 403-410.

Asahina H, Shinozaki J, Masuda K, Morimitsu Y, Satake M (2010) Identification of medicinal *Dendrobium* species by phylogenetic analyses using *matK* and *rbcL* sequences. Journal of Natural Medicines 64: 133–138.

Bafeel, S. O., Arif, I. A., Bakir, M. A., Al Homaidan, A. A., Al Farhan, A. H., & Khan, H. A. (2012). DNA barcoding of arid wild plants using rbcL gene sequences. Genet Mol Res, 11(3), 1934-1941. doi:10.4238/2012.July.19.12

Bafeel, S.; Arif, I.; Bakir, M.; Khan, H.; Farhan, A.; Al-Homaidan, A.; Ahamed, A.; Thomas, J. Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics J. 2011, *4*, 195–198.

Bezeng, B.S., Davies, T.J., Daru, B.H., Kabongo, R.M., Maurin, O., Yessoufou, K., *et al*. 2017. Ten years of barcoding at the African Centre for DNA Barcoding. Genome, 60 (7): 629–638. doi:10.1139/gen-2016 0198. PMID:28340301.

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the United States of America 106: 12794–12797

CBOL: Consortium for the Barcode of Life. Available, http://www.barcoding. si.edu/. Accessed 2010 June 1.

Chase M.W., and M.F. Fay, (2009) Barcoding of Plants and Fungi. Science 325: 682–683.

Chase, M.W., N. Salamin, M. Wilkinson, J.M. Dunwell, R.P. Kesanakurthi, N. Haider and V. Savolainen, 2005. Land plants and DNA barcodes: short-term and long-term goals. Philos. Trans., Ser. B 360: 1889–1895.

Chen, J., Zhao Y-B, Wang Y-J, and Li X-G. 2019. Identification of species and material medical within *Saussurea* subg. *Amphilaena* based on DNA barcodes. Peer J 7:e6357 http://doi.org/10.7717/peerj.6357

De Mattia F, Bruni I, Galimberti A, Cattaneo F, Casiraghi M, and Labra M. (2011). A comparative study of different DNA barcoding markers for the identification of some members of *Lamiaceae*. Food Research International 44: 693–702.

De Mattia, F., Gentili R. Bruni I. Galimberti A. Sgorbari S. Casiraghi M. and Labrai M. (2102). A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. Botanical Journal of *the Linnean Society,* 169: 518–529.

De Vere, N., T.C.G. Rich, C.R. Ford, S.A. Trinder, C. Long, *et al*. (2012) DNA barcoding the native flowering plants and conifers of Wales. PLoS ONE 7(6): e37945. doi:10.1371/journal.pone.0037945

Fazekas AJ, Burgess KS, Kesanakurti PR et al. (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS ONE, 3, e2802.

Fazekas AJ, Kuzmina ML, Newmaster SG, and Hollingsworth PM (2012) DNA barcoding methods for land plants. In: Kress JW, Erickson DL (eds) DNA barcodes: methods and protocols. Methods Mol Biol 858:223–252. doi:10.1007/978-1-61779-591-6_11

Federici, S. Galimberti, A. Bartolucci, F. Bruni, I. De Mattia, F. Cortis, P. and Labra, M. (2013). DNA barcoding to analyse taxonomically complex groups in plants: the case of *Thymus* (*Lamiaceae*) Botanical Journal of the Linnean Society, 171: 687–699.

Ferri, G., Corradini, B., Ferrari, F., Santunione, A.L., Palazzoli, F. and Alu, M. (2015) Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci. Int. Genet. 15, 131–136. doi: 10.1016/j.fsigen.2014.10.005

Harley, R.M., Atkinson, S., Budantsev, A.L., Cantino, P.D., Conn, B.J., Grayer, R., Harley, M.M., De Kok, R., Krestovskaja, T., Morales, R., Paton, A.J., Ryding, O., Upson, T. (2004). *Labiatae*. In: J.W. Kadereit (ed.). The Families and Genera of Vascular Plants, vol. 7: (167–275). Berlin: Springer.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc. Biol. Sci. 270:313–321

Hollingsworth M, Clark AA, Forrest LL et al. (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. Molecular Ecology Resources, 9, 439–457.

Jalas J. 1971. Notes on *Thymus* L. (*Labiatae*) in Europe. I. Supraspecific classification and nomenclature. *Botanical Journal of the Linnean Society* 64: 199–235.

Kress J, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS One 6: 1-10

Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Levin, R.A., W.L. Wagner, P.C. Hoch, M. Nepokroeff, J.C. Pires, E.A. Zimmer and K.J. Sytsma, 2003. Family-level relationships of

Onagraceae based on chloroplast *rbcL* and *ndhF* data. American Journal of Botany, 90: 107-115.

Mohamed A, Mohamed A, Omar AA. A study to find thyme oil dose that kill 50% of mice and minimal dose that kill all mice and maximum nonlethal Dose. Nature and Science. 2013; 11(12): 52-53.

Nadiya RM, Yadav KC, Anurag V. Optimization of process parameters in extraction of Thyme oil using response surface methodology (RSM). International Journal of Science, Engineering and Technology. 2016; 4(1): 248-256.

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. Molecular Ecology Resources, 8, 480–490.

Nithaniyal, S., Newmaster, S.G., Ragupathy, S., Krishnamoorthy, D., Vassou, S.L. and Parani, M. (2014) DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. PLoS ONE, 9, e107669.

Ocana A and Reglero G. Effects of Thyme extract oils (from Thymus vulgaris, Thymuszygis, and Thymus hyemalis) on cytokine production and gene expression of oxLDL-Stimulated THP-1-Macrophages. Journal of Obesity. 2012; 1-11, Doi:10.1155/2012/104706.

Parveen, I., Gafner, S., Techen, N., Murch, S. J., and Khan, I. A. (2016). DNA barcoding for the identification of botanicals in herbal medicine and dietary supplements: strengths and limitations. *Planta Med.* 82, 1225–1235. doi: 10.1055/s-0042-111208

Raja R.R., (2012). Medicinally Potential Plants of *Labiatae* (*Lamiaceae*) Family: An Overview. Research Journal of Medicinal Plants, 6, 203-213.

Sahare, P. and Srinivasu, T. (2012) Barcoding for authentic identification of medicinal plants. Int. J. Eng. Sci. 1, 33–36.

Saleh H, Azizollah JK, Ahmadreza H, Raham A. The Application of Thymus vulgaris in traditional and modern medicine: A Review. Global Journal of Pharmacology. 2015; 9 (3): 260-266. etrieved 2016-06-09.

Schori, M. and Showalter, A.M. (2011) DNA barcoding as a means for identifying medicinal plants of Pakistan. Pak. J. Bot.,43(SI): 1-4.

Skuza L, Szućko I, Filip E, Adamczyk A (2019) DNA barcoding in selected species and subspecies of Rye (*Secale*) using three chloroplast loci (*matK*, *rbcL*, *trnH-psbA*). Notulae Botanicae Horti Agrobotanici Cluj-Napoca 47(1):54–62 DOI:47.15835

Techen, N., I. Parveen, Z. Pan and A.I. Khan, 2014. DNA barcoding of medicinal plant material for identification. Current Opinion in Biotechnology, 25: 103-110.

Theodoridis S, Stefanaki A, Tezcan M, Aki C, Kokkini S, Vlachonasios KE: 2012 DNA barcoding in native plants of the *Labiatae* (*Lamiaceae*) family from Chios Island (Greece) and the adjacent Cesme-Karaburun Peninsula (Turkey). Mol. Ecol. Resour., 12:620-633.

Vineetha. 22 Amazing benefits and uses of Thyme. Health Beckon. 2014. http://www.healthbeckon.com/thyme-benefits/ 2

Vohra, P. and Khera, K.S. (2013) DNA barcoding: current advance and future prospects a review. AJBLS, 3, 185–189.

Xue C-Y, and Li D-Z (2011) Use of DNA barcode sensu lato to identify traditional Tibetan medicinal plant *Gentianopsis paludosa* (Gentianaceae). Journal of Systematics and Evolution 49: 267–270.

Zhou J, Wang WC, Liu MQ, Liu ZW. 2014. Molecular authentication of the traditional medicinal plant *Peucedanum praeruptorum* and its substitutes and adulterants by DNA - barcoding technique**.** Pharmacogn Mag., 10(40):385-390.