

## DNA-Barcoding of Some Medicinal Plant Species in Saudi Arabia Using *rbcL* and *matK* Genes

Abdulrahman Alasmari\*

Department of Biology, Faculty of Science, University of Tabuk, Tabuk, 71491, Saudi Arabia

\*Corresponding Author: Abdulrahman Alasmari. Email: ab.alsmari@ut.edu.sa

Received: 09 April 2020; Accepted: 22 June 2020

**Abstract:** In the Kingdom of Saudi Arabia (KSA), thousands of plants are considered to have therapeutic value. The ambiguous use of identification mainly morphological characteristics of many plants has resulted in the adulteration and displacement of plant products which undermine their therapeutic value and weak documentation of plant resources. The aims of this study were therefore to evaluate genetic variability and explore the phylogeographic architecture for Saudi medicinal plant samples using *rbcL* and *matK* genes as barcodes for genomic identification. The *matK* and *rbcL* sequences collected for these samples were used as key markers for examining the relationship between Saudi medicinal plant species based on genetic diversity. During our study we were successful in identifying and documenting 4 different species (*Foeniculum vulgare*, *Nitraria retusa*, *Dodonaea viscosa*, and *Rumex nervosus*) located in Saudi Arabia using DNA barcoding technique. A total number of 8 sequences were obtained with a total sequence length of 6176 bp, where it ranged from 617 bp to 878 bp with an average length of 772 bp. The total number of *rbcL* sequences length is 2801 bp, where it ranges from 617 bp to 807 bp with an average length of 700.2 bp. Out of the 4 plant samples used, only three samples were identified correctly on the species level with an identity percentage higher than 95% using *rbcL* gene. Additionally, 4 *matK* sequences have been retrieved belong to 4 species. The total number of *matK* sequences length is 3375 bp, where it ranges from 819 bp to 878 bp with an average length of 843.8 bp. Out of the 4 plant samples used, only two samples were identified correctly on the species level with an identity percentage higher than 98% using *matK* gene. Both *rbcL* and *matK* have been able to identify most of our collected plant samples by genus, and some by species. Using only one DNA-barcoding technique was not reliable for plant identification, where *matK* and *rbcL* must be used as a dual DNA-barcoding procedure.

**Keywords:** Phylogeographic architecture; genomic identification; DNA barcoding

### 1 Introduction

Ethnic knowledge of ancient and yet still usable conventional Saudi medicine among locals and medical healers. In the Kingdom of Saudi Arabia (KSA), higher than 1200 (more than 50%) of the total plants (2250)



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

are considered to be of therapeutic value [1]. Additionally, this indigenous knowledge and cultural experiences were passed orally without recording and the traditional healers die without passing on their information [2]. Additionally, today medicinal plants are promoted as herbal medicines or natural health products. They are often considered as safe but adulterated, falsified and low-quality products pose serious security threats to consumers and the current markets [3]. Inaccurate identification use predominantly morphological characteristics of many plants has resulting in adulteration and displacement of plant products which undermine their therapeutic value [4].

Morphological characterization continues to be the foundation of plant taxonomic diagnosis. Unfortunately there are some substantial drawbacks to relying solely on morphology. This has raised a need for more advanced and sufficient molecular tools could be used recursively to identify medical plant taxonomic profiles with lower margin of error. The evolution of molecular genetics has resulted in a new approach based on the diversity of nucleotide sequences across species called “DNA barcoding” [5,6]. Morphology-based identification of medicinal plants is a classic approach, and molecular techniques based on DNA have progressively been used for taxonomic identification. Short uniform segment of the genome that serves as a “barcode” pattern has been suggested as a technology which provides accurate identification of species [7]. The plant committee Barcode of Life (CBOL) consortium has proposed *matK* and *rbcL* as the main plant barcode regions [8]. Barcoding is helpful in identifying the origin species of medicines and has also been used to identify contamination of foreign material in natural medicines, particularly those in small pieces or forms of powder [9]. Development within DNA-barcoding systematic has led to different approaches in herbal phylogenetic being introduced and the use of recently developed molecular markers expanded.

The plastid genome is uniparently acquired, non-re-combined and a structurally stable genome which is the most motivating rival genes for plant DNA barcoding [10]. The RuBisCO large subunit (*rbcL*) plastid marker was commonly used to study unknown taxonomic position of species to elucidate taxonomic connections between different species [11]. It was supported by the stable exon structure of the *rbcL* gene and high amino acid sequence similarity as a reliable marker for these research fields [12]. Maturase K is a chloroplast-encoding gene nested in the large single copy region of the chloroplast genome [13]. The *matK* gene is regarded to evolve quickly, as the gene seems to have a high degree of mutation rate and its sequence is much more varied than most other genes [14,15]. For a reliable plant DNA barcode including *rbcL* and *matK*, not fewer than two genes are essential. *rbcL* can be recognized by genus and family level whereas *matK* is the variable part of the code for species confirmation [16]. Moreover, bioinformatics analysis techniques have provided more insight on genetic diversity of species populations and offer more comprehensive about genetic sequence analysis [17].

The integration of such technology into the identification and phylogenetic analysis of Saudi medicinal plants could therefore be key to the conservation of natural resources. The aim of this study was therefore to evaluate genetic variability and explore the phylogeographic architecture for Saudi medicinal plant samples using *rbcL* and *matK* genes as barcodes for genomic identification. The *matK* and *rbcL* sequences collected for these samples were used as key markers for examining the relationship between Saudi medicinal plant species based on genetic diversity. Additionally, after aligning new sequences against GenBank databases, the efficacy of *rbcL/matK* based DNA barcoding was assessed for species identification and differentiation.

## 2 Materials and Methods

### 2.1 Plant Material

Four medicinal plant samples were obtained from different geographical locations in KSA. These samples belong to four different medicinal plant species (*Dodonaea viscosa*, *Foeniculum vulgare*, *Nitraria retusa*, and *Rumex nervosus*).

## 2.2 DNA Barcode Analysis

### 2.2.1 Extraction and Purification of DNA

Four different medicinal plant leaf specimens were collected and ground to a fine powder in liquid nitrogen using a sterile mortar and pestle. For DNA extraction and purification reported by Alshehri et al. [18]. The concentration and quality of extracted DNA was estimated by running on 1% agarose gel electrophoresis, using a DNA size marker (100 bp DNA Ladder from New England Bio-labs).

### 2.2.2 PCR Amplification

For *rbcL* and *matK* PCR primers the reverse and forward primers were, respectively, *rbcL*-F (5'-ATGTCACCACAAACAGAGACTAAAGC-3') and *rbcL*-R (5'-TCGCATGTACCTGCAGTAGC-3'), *matK*-F (5'-ATCCATCTGGAAATCTTAGTTC-3') and *matK*-R (5'-CTTCCTCTGTAAAGAATTC-3'). For *rbcL* and *matK* genes the PCR the reaction mixture consisted of 15 mM MgCl<sub>2</sub>, 1× buffer (Promega), 0.2 mM dNTPs, 1 μ of Taq DNA polymerase (GoTaq, Promega), 20 pcoml of each primers, 40 ng DNA and distilled deionized water to a volume of 25 μL. PCR amplification was performed in 40 cycles (94°C for 30 s, 50°C for 30 s and 72°C for 30 s). Electrophoresis of the amplification of PCR products was screened in an agarose gel (1.5%) containing ethidium bromide (0.5 μg/ml). PCR products have been viewed and photoed using a Gel Documentation System (BIO-RAD 2000) on UV light. The PCR was carried out with a Perkin-Elmer/GeneAmp<sup>®</sup> PCR System 9700 (PE Applied Biosystems, USA) programmed to fulfill 40 cycles after an initial denaturation cycle for 5 min at 94°C. Each cycle consisted of a denaturation step at 94°C for 30 s, an annealing step at 50°C for 30 s and an elongation step at 72°C for 30 s. The primer extension segment was extended to 7 min at 72°C in the final cycle. The amplification products were determined by electrophoresis in a 1.5% agarose gel using ethidium bromide (0.5 μg/ml) in 1X Tris borate Edita (TBE) buffer at 95 volts. For PCR product sizes determination, a marker 100 bp DNA ladder (Promega, USA), was used as a molecular size standard. Gel images were visualized using UV transilluminator and photographed using a Gel Documentation System (BIO-RAD 2000, USA).

### 2.2.3 PCR Purification and Sequencing

GeneJET PCR Purification Kit (Catalog number: K0701) was used for the purification of amplified PCR products using manufacturer's manual. The extracted PCR products have been incubated for 2 min at ambient temperature, or at -20°C when stored. Macrogen Inc., Korea used Sanger DNA sequencing tech to perform the *rbcL* and *matK* PCR products DNA sequencing.

### 2.2.4 Computational Analysis

The product sequences *rbcL* and *matK* PCR were used to identify species utilizing NCBI-BLAST to search GenBank databases [4]. Species identifications were adopted if they gave more than 98% similarity to the database comparison sequences. NCBI-Blast disclosed the most similar 50 *rbcL* and *matK* sequences to be used for more phylogenetic analysis. Sequence Alignments were conducted using CLUSTALW [19]. The phylogenetic analysis was performed using iTOL [3]. Multivariate phylogenetic visualizing and clustering of data of similarity was shown using ClustVis web tool [20]. Extra sequences of medicinal plants obtained from GenBank have been used in the geographical distribution screening shown using the "rworldmap" package through R [21].

## 3 Result and Discussion

### 3.1 Species Identification

A total number of 8 sequences were obtained with a total sequence length of 6176 bp, where it ranged from 617 bp to 878 bp with an average length of 772 bp (Supplementary 1). Among these sequences, 4 *rbcL* sequences have been retrieved belong to 4 different plant species. The total number of *rbcL* sequences length

is 2801 bp, where it ranges from 617 bp to 807 bp with an average length of 700.2 bp. The NCBI-BLAST system successfully identified 4 *rbcL* sample sequences by genus. Out of the 4 plant samples used, only three samples were identified correctly on the species level with an identity percentage higher than 95% (Tab. 1). Although *rbcL* was successful to identify sample 4 genus, it did not identify the correct species of the sample belong to *Rumex nervosus* (Tab. 1). Through further investigation it has been concluded that, there is only one DNA-barcoding sequence related to this species in NCBI database until now (3/3/2020), which belongs to *ITS2* gene. This could confirm the lack of information in the NCBI database regarding this plant, and could shine a light for further studies. The sequence coverage for most sequenced *rbcL* genes was 100%, which

**Table 1:** The NCBI blast results for retrieved *matK* and *rbcL* sequences. The scoring results of sequences alignment, where sample species (SS), sequence coverage (QC), *E*-value (EV), sequences identity score (SI) reveals genes with high similarity to our retrieved sequences

S. No.	SS	gene	TS	QC	EV	SI	Accession
1	<i>Foeniculum vulgare</i>	<i>matK</i>	<i>Foeniculum vulgare</i>	97.00%	0	99.65%	YP_009235860.1
		<i>matK</i>	<i>Foeniculum vulgare</i>	96.00%	0	100.00%	AWM67414.1
		<i>matK</i>	<i>Foeniculum vulgare</i>	96.00%	0	100.00%	AFC82721.1
2	<i>Nitraria retusa</i>	<i>matK</i>	<i>Nitraria roborowskii</i>	99.00%	6E-177	95.96%	QHE65566.1
		<i>matK</i>	<i>Nitraria roborowskii</i>	88.00%	1E-169	95.85%	AEQ47266.1
		<i>matK</i>	<i>Nitraria roborowskii</i>	87.00%	7E-169	95.83%	AEQ47267.1
3	<i>Dodonaea viscosa</i>	<i>matK</i>	<i>Dodonaea viscosa</i>	91.00%	1E-170	97.60%	AKJ76991.1
		<i>matK</i>	<i>Dodonaea viscosa</i>	92.00%	3E-169	96.85%	YP_009437361.1
		<i>matK</i>	<i>Dodonaea viscosa</i>	88.00%	9E-168	98.36%	QAV55715.1
4	<i>Rumex nervosus</i>	<i>matK</i>	<i>Commicarpus plumbagineus</i>	91.00%	1E-174	99.62%	ATD50701.1
		<i>matK</i>	<i>Commicarpus scandens</i>	91.00%	1E-174	99.62%	ATD50705.1
		<i>matK</i>	<i>Commicarpus boissieri</i>	91.00%	2E-174	99.62%	ATD50686.1
		<i>matK</i>	<i>Commicarpus coctoris</i>	91.00%	2E-174	99.62%	ATD50689.1
5	<i>Foeniculum vulgare</i>	<i>rbcL</i>	<i>Rhodiola bupleuroides</i>	91.00%	9E-133	95.42%	AII70523.1
		<i>rbcL</i>	<i>Primula japonica</i>	98.00%	6E-132	94.77%	AAM50479.1
		<i>rbcL</i>	<i>Rhodiola himalensis</i>	91.00%	1E-131	95.42%	AII70548.1
		<i>rbcL</i>	<i>Fenerivia emarginata</i>	91.00%	1E-131	94.12%	AER51997.1
6	<i>Nitraria retusa</i>	<i>rbcL</i>	<i>Lysimachia christinae</i>	97.00%	2E-147	96.89%	AFD30872.2
		<i>rbcL</i>	<i>Rosaceae sp.</i>	97.00%	2E-147	97.33%	QAT81573.1
		<i>rbcL</i>	<i>Crypteronia paniculata</i>	97.00%	3E-147	97.33%	AOO78041.1
7	<i>Dodonaea viscosa</i>	<i>rbcL</i>	<i>Acer negundo</i>	97.00%	6E-113	89.05%	CBK52928.1
		<i>rbcL</i>	<i>Euodia fraxinifolia</i>	97.00%	7E-113	89.05%	AHI17504.1
		<i>rbcL</i>	<i>Carapa procera</i>	97.00%	1E-112	89.05%	ACN42354.1
8	<i>Rumex nervosus</i>	<i>rbcL</i>	<i>Rumex nepalensis</i>	99.00%	2E-148	96.92%	AEM05112.1
		<i>rbcL</i>	<i>Rumex nepalensis</i>	99.00%	2E-148	96.92%	AEM05113.1
		<i>rbcL</i>	<i>Rumex acetosa</i>	99.00%	8E-148	96.46%	ANW81115.1

indicate high PCR amplification performance. Our findings are not at variance with previous results recorded *rbcL* is the most successful sequence quality and recovered loci [22]. The identification power of *rbcL* on species level was lower than expected, which could indicate small difference between inter-species and intra-species variation was observed. This *rbcL* gene limitation has also been reported in similar species of *Ardisia*, *Sweretia chirayita* and *Mentha* [23].

Additionally, 4 *matK* sequences have been retrieved belong to 4 species (*Foeniculum vulgare*, *Nitraria retusa*, *Dodonaea viscosa*, and *Rumex nervosus*). The total number of *matK* sequences length is 3375 bp, where it ranges from 819 bp to 878 bp with an average length of 843.8 bp. The NCBI-BLAST system successfully identified 3 *matK* sample sequences by genus. Out of the 4 plant samples used, only two samples were identified correctly on the species level (*Foeniculum vulgare*, and *Dodonaea viscosa*) with an identity percentage higher than 98% (Tab. 1). Although *matK* was successful to identify sample 3 genus, it did not identify the correct species of the samples of *Nitraria retusa*, and *Rumex nervosus* (Tab. 1). The sequence coverage for most sequenced *matK* genes was 98%, which indicate high PCR amplification performance (Tab. 1). The ability of *matK* to differentiate between species is not always reliable, *matK* has been reported to have significant variation and can be used in the nutmeg family News master for DNA barcode [24], while it cannot be used to distinguish *Myristica* species. Therefore, it was suggested that minimum requirement for species differentiation through *matK* gene sequencing is 99.64% similarity (Figs. 1 and 2) [23,25,26].

### 3.2 Sequence Alignment and Phylogenetic Analysis

The sequence alignment of *rbcL* sequences revealed mutation rate of sequence structure, where sequence mutation in middle sequence is quite high (Fig. 3). The heatmap of sequence similarities revealed high abundance of sequences similarity in inter-species and high evolution rate within intra-species sequences (Fig. 4). Where the *Foeniculum* genus revealed high sequence differentiation compared to other studied species (Fig. 4). Additionally, both Figs. 3 and 4 indicate high genetic variability of sample 3 (*Dodonaea viscosa*) compared to other studied species. The phylogenetic tree built using *rbcL* sequences was grouped into 3 different groups, each group containing one genus (Fig. 5). Except for the *rbcL* of *Dodonaea viscosa* all *rbcL* was able to differentiate between all species, where *Foeniculum vulgare* sequences were clustered at 0.01 of sequence difference, *Rumex nervosus* at 0.18, and *Nitraria retusa* at 0.24.

The sequence alignment of *matK* sequences revealed high conservation of sequence structure among same genus and high mutation rate between different genus, where sequence mutation is not restricted to sequence area (Fig. 6). The heatmap of sequence similarities revealed low ability of *matK* gene sequences to differentiate between some genus (Fig. 7). The phylogenetic tree built using *matK* sequences was grouped into 3 different groups, only two groups contain one genus (*Nitraria retusa*, and *Foeniculum vulgare*) (Fig. 8). Except for *Dodonaea viscosa* and *Rumex nervosus* the *matK* was able to differentiate between species depending on phylogentic analysis, however *Foeniculum vulgare* sequences were clustered at 0.75 of sequence difference, and *Nitraria retusa* at 0.018.

(A) *Foeniculum vulgare*

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene, partial cds; chloroplast</a>	1465	1465	99%	0.0	98.62%	<a href="#">MG946837.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare chloroplast rbcL gene for Rubisco large subunit</a>	1465	1465	99%	0.0	98.62%	<a href="#">LT576823.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare chloroplast complete genome</a>	1465	1465	99%	0.0	98.62%	<a href="#">KR011054.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare voucher MoSEL-270 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast</a>	1465	1465	99%	0.0	98.62%	<a href="#">KP974256.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare voucher MoSEL-268 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast</a>	1465	1465	99%	0.0	98.62%	<a href="#">KP974255.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; plastid</a>	1465	1465	99%	0.0	98.62%	<a href="#">KM360787.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum sp. AP-2010 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene, partial cds; chloroplast</a>	1465	1465	99%	0.0	98.62%	<a href="#">GQ162782.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare voucher OTA061332 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbc) gene, partial cds; chloroplast</a>	1465	1465	99%	0.0	98.62%	<a href="#">GQ120445.1</a>
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare chloroplast gene for Ribulosebiphosphate Carboxylase large subunit, partial cds</a>	1465	1465	99%	0.0	98.62%	<a href="#">D44567.1</a>
<input checked="" type="checkbox"/> <a href="#">Anethum graveolens ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene, partial cds; chloroplast</a>	1460	1460	99%	0.0	98.50%	<a href="#">MG946945.1</a>

**Foeniculum vulgare ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene, partial cds; chloroplast**

Sequence ID: [MG946837.1](#) Length: 1407 Number of Matches: 1

Range 1: 149 to 948 [GenBank](#) [Graphics](#)

[Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1465 bits(762)	0.0	789/800(99%)	1/800(0%)	Plus/Plus
Query 1	CGGTAGCTGCCGAACACAACTGGTACATGGACCACTGTGTGGACCGATGGACTTACCA	60		
Sbjct 149	CGGTAGCTGCCGAATCTCTACTGGTACATGGACCACTGTGTGGACCGATGGACTTACCA	208		
Query 61	GCTCTTGATCGTTACAAAGGGCGCTGCTACGGAATCGAGCCCGTTGCTGGAGAAGAAATC	120		
Sbjct 209	GCTCTTGATCGTTACAAAGGGCGCTGCTACGGAATCGAGCCCGTTGCTGGAGAAGAAATC	268		
Query 121	AATATATCGCTTATGTAGCTTACCCATTAGACCTTTTGAAGAAGTTCTGTTACTAACCA	180		
Sbjct 269	AATATATCGCTTATGTAGCTTACCCATTAGACCTTTTGAAGAAGTTCTGTTACTAACCA	328		
Query 181	TGTTTACTTCCATTGTAGGTAATGATTTGGGTTCAAGCCCTGCGGCTCTACGTCTGG	240		
Sbjct 329	TGTTTACTTCCATTGTAGGTAATGATTTGGGTTCAAGCCCTGCGGCTCTACGTCTGG	388		
Query 241	AAGATCTGCGAATCCCGTTGCTTATGTTAAACTTTTCAAGGACCGCTCATGGCATCC	300		
Sbjct 389	AAGATCTGCGAATCCCGTTGCTTATGTTAAACTTTTCAAGGACCGCTCATGGCATCC	448		
Query 301	AAGTTGAGAGAGATAAATGAACAAGTATGGTCTCCCTGTTGGATGACTATTAAC	359		
Sbjct 449	AAGTTGAGAGAGATAAATGAACAAGTATGGTCTCCCTGTTGGATGACTATTAAC	508		
Query 360	CTAAATGGGGTTATCCGCTAAAACTACGGTAGAGCGGTTATGAATGCTCCCGGGTG	419		
Sbjct 509	CTAAATGGGGTTATCCGCTAAAACTACGGTAGAGCGGTTATGAATGCTCCCGGGTG	568		
Query 420	GACTTGATTTTACCAAGACGATGAGAATGTGAACCTCCCACTTATGCGTTGGAGAG	479		
Sbjct 569	GACTTGATTTTACCAAGACGATGAGAATGTGAACCTCCCACTTATGCGTTGGAGAG	628		
Query 480	ATGGTTCTTATTTTGTGCCGAAGCAATTTATAAGCACAGGCTGAACTGGTGAATCA	539		
Sbjct 629	ATGGTTCTTATTTTGTGCCGAAGCAATTTATAAGCACAGGCTGAACTGGTGAATCA	688		
Query 540	AAGGGCATTACTTGATGCTACTGCGGGTACATGTGAAGMATGATGAAAAGGCTATAT	599		
Sbjct 689	AAGGGCATTACTTGATGCTACTGCGGGTACATGTGAAGMATGATGAAAAGGCTATAT	748		
Query 600	TTGCAGAGAATGGGAGTTCCATCGTAATGCATGATTACTTAACAGGGGATTCACTG	659		
Sbjct 749	TTGCAGAGAATGGGAGTTCCATCGTAATGCATGATTACTTAACAGGGGATTCACTG	808		
Query 660	CAAACTAGCTTAGCCATTATTGCCGAGATAATGGCTACTTCTTCACATCCACCGTG	719		
Sbjct 809	CAAACTAGCTTAGCCATTATTGCCGAGATAATGGCTACTTCTTCACATCCACCGTG	868		
Query 720	CAATGATGACGTTATGATAGACAGAAATCATGGTATACACTTCCGTTACTAGCTA	779		
Sbjct 869	CAATGATGACGTTATGATAGACAGAAATCATGGTATACACTTCCGTTACTAGCTA	928		
Query 780	AAGCGTTACGATGTCTGGT	799		
Sbjct 929	AAGCGTTACGATGTCTGGT	948		

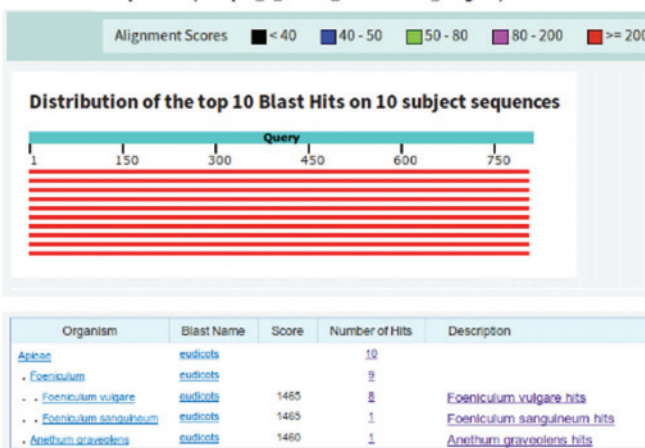


Figure 1: (continued)

(B) *Nitraria retusa*

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	<a href="#">Nitraria retusa ribulose 1,5-bisphosphate carboxylase (rbcl) gene, partial cds; chloroplast gene for chloroplast product</a>	1286	1286	99%	0.0	99.27%	U39278.2
<input checked="" type="checkbox"/>	<a href="#">Nitraria tangutorum chloroplast, complete genome</a>	1275	1275	99%	0.0	98.98%	MK347423.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria sibirica chloroplast, complete genome</a>	1275	1275	99%	0.0	98.98%	MK347422.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria roborowskii chloroplast, complete genome</a>	1275	1275	99%	0.0	98.98%	MK347421.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria praevisia ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-like (rbcl) gene, partial sequence; chloroplast</a>	1275	1275	99%	0.0	98.98%	DQ267162.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria tangutorum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-like (rbcl) gene, partial sequence; chloroplast</a>	1275	1275	99%	0.0	98.98%	DQ267158.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria billardierei voucher N1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcl) gene, partial cds; chloroplast</a>	1273	1273	98%	0.0	98.98%	KT377258.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria roborowskii ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-like (rbcl) gene, partial sequence; chloroplast</a>	1273	1273	98%	0.0	98.98%	DQ267161.1
<input checked="" type="checkbox"/>	<a href="#">Tetraena mongolica ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-like (rbcl) gene, partial sequence; chloroplast</a>	1273	1273	98%	0.0	98.98%	DQ267163.1
<input checked="" type="checkbox"/>	<a href="#">Nitraria sibirica ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-like (rbcl) gene, partial sequence; chloroplast</a>	1258	1258	97%	0.0	98.82%	DQ267160.1

**Nitraria retusa ribulose 1,5-bisphosphate carboxylase (rbcl) gene, partial cds; chloroplast gene for chloroplast product**

Sequence ID: [U39278.2](#) Length: 1333 Number of Matches: 1

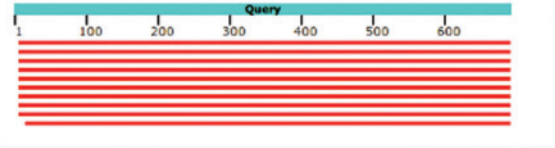
Range 1: 16 to 704 GenBank Graphics

Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
1286 bits(669)	0.0	684/689(99%)	1/689(0%)	Plus/Plus
Query 6	AGCCGGTGTT	-AAGATTATAAATTGACTTATTATACTCTGAATATGAAACCAAGATAC	64	
Sbjct 16	AGCCGGTGTT	AAAGATTATAAATTGACTTATTATACTCTGAATATGAAACCAAGATAC	75	
Query 65	TGATATCTTGGCAGCATTCCGAGTAACCTCAACCCGGAGTCCACCCGAAGAAGCAGG	124		
Sbjct 76	TGATATCTTGGCAGCATTCCGAGTAACCTCAACCCGGAGTCCACCCGAAGAAGCAGG	135		
Query 125	GGCTGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACGGTATGGACCGATGGGCT	184		
Sbjct 136	GGCTGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACGGTATGGACCGATGGGCT	195		
Query 185	TACCAGCCTTGATCGTTACAAAGGACGATGCTACAACATCGAGCCCTGGTGGAGAAGA	244		
Sbjct 196	TACCAGCCTTGATCGTTACAAAGGACGATGCTACAACATCGAGCCCTGGTGGAGAAGA	255		
Query 245	AAATCAATATATGTTATGTAGCTTATCCTTAGACCTTTTGAAGAAGGTTCTGTTAC	304		
Sbjct 256	AAATCAATATATGTTATGTAGCTTATCCTTAGACCTTTTGAAGAAGGTTCTGTTAC	315		
Query 385	TAAATGTTTACTTCCATTGTGGGTAAATGATTTGGGTTCAAAGCCCTGCGTGTCTACG	364		
Sbjct 316	TAAATGTTTACTTCCATTGTGGGTAAATGATTTGGGTTCAAAGCCCTGCGTGTCTACG	375		
Query 365	TCTAGAGGATCTACGAATCCCTCTGCGTATATTAATACTTCCAAAGCCCGCCTCACGG	424		
Sbjct 376	TCTAGAGGATCTACGAATCCCTCTGCGTATATTAATACTTCCAAAGCCCGCCTCACGG	435		
Query 425	TATCCAGGTTGAGAGAGATAAATTGAACAAGTATGGCCGTCCTTATTGGGATGACTAT	484		
Sbjct 436	TATCCAGGTTGAGAGAGATAAATTGAACAAGTATGGCCGTCCTTATTGGGATGACTAT	495		
Query 485	TAAACCTAAATGGGGTTATCCGCTAAGAACTACGGTAGAGCAGTTATGAATGTCTACG	544		
Sbjct 496	TAAACCTAAATGGGGTTATCCGCTAAGAACTACGGTAGAGCAGTTATGAATGTCTACG	555		
Query 545	TGGTGGACTTGACTTTACCAAGATGATGAGAACGTGAACCTGCAACATTATATGCGTTG	604		
Sbjct 556	TGGTGGACTTGACTTTACCAAGATGATGAGAACGTGAACCTGCAACATTATATGCGTTG	615		
Query 605	GAGAGACCGTTTCTTATTTTGGCGGAAGCTATTATAAAGTGCAGGCTGAACAAGTGG	664		
Sbjct 616	GAGAGACCGTTTCTTATTTTGGCGGAAGCTATTATAAAGTGCAGGCTGAACAAGTGG	675		
Query 665	AATCAAAGTCACTTAAATGCTACTG	693		
Sbjct 676	AATCAAAGTCACTTAAATGCTACTG	704		

Alignment Scores ■ < 40 ■ 40 - 50 ■ 50 - 80 ■ 80 - 200 ■ >= 200

**Distribution of the top 10 Blast Hits on 10 subject sequences**



Organism	Blast Name	Score	Number of Hits	Description
<a href="#">rosids</a>	<a href="#">eudicots</a>		10	
• <a href="#">Nitraria</a>	<a href="#">eudicots</a>		9	
• <a href="#">Nitraria retusa</a>	<a href="#">eudicots</a>	1286	1	<a href="#">Nitraria retusa hits</a>
• <a href="#">Nitraria tangutorum</a>	<a href="#">eudicots</a>	1275	2	<a href="#">Nitraria tangutorum hits</a>
• <a href="#">Nitraria sibirica</a>	<a href="#">eudicots</a>	1275	2	<a href="#">Nitraria sibirica hits</a>
• <a href="#">Nitraria roborowskii</a>	<a href="#">eudicots</a>	1275	2	<a href="#">Nitraria roborowskii hits</a>
• <a href="#">Nitraria praevisia</a>	<a href="#">eudicots</a>	1275	1	<a href="#">Nitraria praevisia hits</a>
• <a href="#">Nitraria billardierei</a>	<a href="#">eudicots</a>	1273	1	<a href="#">Nitraria billardierei hits</a>
• <a href="#">Tetraena mongolica</a>	<a href="#">eudicots</a>	1273	1	<a href="#">Tetraena mongolica hits</a>

Figure 1: (continued)

(C) *Rumex nervosus*

Sequences producing significant alignments Download Manage Columns Show 10

select all 10 sequences selected

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<a href="#">Rumex acetosa chloroplast, complete genome</a>	1256	1256	100%	0.0	98.54%	<a href="#">NC_042390.1</a>
<a href="#">Rumex acetosa chloroplast, complete genome</a>	1256	1256	100%	0.0	98.54%	<a href="#">MH359405.1</a>
<a href="#">Rumex acetosella ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1256	1256	100%	0.0	98.54%	<a href="#">HM850316.1</a>
<a href="#">Rumex acetosa ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1256	1256	100%	0.0	98.54%	<a href="#">AY395058.1</a>
<a href="#">Rumex induratus ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1244	1244	100%	0.0	98.24%	<a href="#">HM850321.1</a>
<a href="#">Rumex induratus ribulose-1,5-bisphosphate carboxylase (rbcL) gene, partial cds, chloroplast gene for chloroplast product</a>	1244	1244	100%	0.0	98.24%	<a href="#">AF297122.1</a>
<a href="#">Rumex acetosa isolate SM05 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, plastid</a>	1242	1242	99%	0.0	98.38%	<a href="#">KY378051.1</a>
<a href="#">Rumex acetosa isolate SM04 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, plastid</a>	1242	1242	99%	0.0	98.38%	<a href="#">KY378050.1</a>
<a href="#">Rumex acetosa isolate SM10 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, plastid</a>	1242	1242	99%	0.0	98.38%	<a href="#">KY378049.1</a>
<a href="#">Rumex acetosa isolate SM12 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, plastid</a>	1242	1242	99%	0.0	98.38%	<a href="#">KY378048.1</a>

**Rumex acetosa chloroplast, complete genome**  
 Sequence ID: [NC\\_042390.1](#) Length: 160269 Number of Matches: 1

Range 1: 56587 to 57269 [GenBank](#) [Graphics](#) Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
1256 bits(653)	0.0	673/683(99%)	0/683(0%)	Plus/Plus

```

Query 1      GTTAGAGAATACAAATTGACTTATTACTCCTGACTATGAACCCCATGACCATGATATC 60
Sbjct 56587   GTTAAAGAATACAAATTGACTTATTACTCCTGACTATGAACCCCATGACCATGATATC 56646
Query 61     TTGGCAGCATTTTCGAGTAACCTCAACCTGGAGTTCACCCAGAAGAAGCAGGGCCGCG 120
Sbjct 56647   TTGGCAGCATTTTCGAGTAACCTCAACCTGGAGTTCACCCAGAAGAAGCAGGGCCGCG 56706
Query 121    GTAGCTCCGAATCTTCTACTGGTACATGGACAACCTGTGTGGACCGATGGCTTACCAGC 180
Sbjct 56707   GTAGCTCCGAATCTTCTACTGGTACATGGACAACCTGTGTGGACCGATGGCTTACCAGC 56766
Query 181    CTTGATCGTTACAAGGACGATGCTACCCATCGAGCCTGTTCTCTGGAGAAGAAAGTCAG 240
Sbjct 56767   CTTGATCGTTACAAGGACGATGCTACCCATCGAGCCTGTTCTCTGGAGAAGAAAGTCAG 56826
Query 241    TTTAATGCTTATGTAGCTTACCCATTAGACCTTTTGAAGAAGGTTCTGTACTAACATG 300
Sbjct 56827   TTTAATGCTTATGTAGCTTACCCATTAGACCTTTTGAAGAAGGTTCTGTACTAACATG 56886
Query 301    TTTACTTCCATTGTGGGTAATGATATTGGGTTCAAAGCCCTGCGTGTCTACGTTGGAG 360
Sbjct 56887   TTTACTTCCATTGTGGGTAATGATATTGGGTTCAAAGCCCTGCGTGTCTACGTTGGAG 56946
Query 361    GATTGCGAATCCCTCCTGCTTATACGAAAACCTTCCAAGGCCCGCCTCATGGTATCAA 420
Sbjct 56947   GATTGCGAATCCCTCCTGCTTATACGAAAACCTTCCAAGGCCCGCCTCATGGTATCAA 57006
Query 421    GTTGAGAGAGATAAATTGAACAAATATGGACGTCCTTATTGGGATGACTATTAACCG 480
Sbjct 57007   GTTGAGAGAGATAAATTGAACAAATATGGACGTCCTTATTGGGATGACTATTAACCG 57066
Query 481    AAATTGGGTTGTCCGCTAAGAACTACGGCCGAGCAGTTTATGAATGCTCTCGTGGCGGA 540
Sbjct 57067   AAATTGGGTTGTCCGCTAAGAACTACGGCCGAGCAGTTTATGAATGCTCTCGTGGCGGA 57126
Query 541    CTTGATTTTACCAAAGATGATGAAAACGTGAACTCCCAACATTTATGCGTTGGAGAGAC 600
Sbjct 57127   CTTGATTTTACCAAAGATGATGAAAACGTGAACTCCCAACATTTATGCGTTGGAGAGAC 57186
Query 601    CGGGTCTTATCTGTGGGGAAGCTATTTTAAATCACAGTCTGAAACACGATTTTAAA 660
Sbjct 57187   CGGTCTTATCTGTGGGGAAGCTATTTTAAATCACAGTCTGAAACACGATTTTAAA 57246
Query 661    TGACATTACTTGAATGCTACTGC 683
Sbjct 57247   GGACATTACTTGAATGCTACTGC 57269
    
```

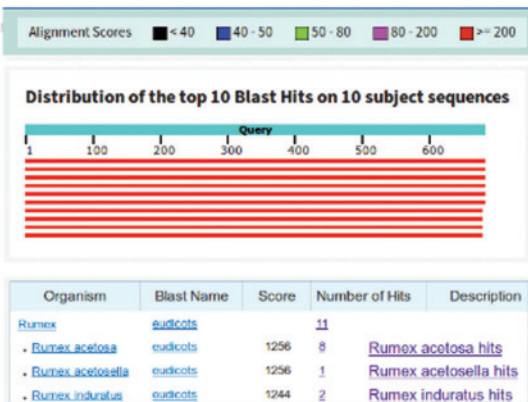


Figure 1: (continued)



(D) *Dodonaea viscosa*

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/> <a href="#">Dodoniaea viscosa chloroplast, complete genome</a>	1006	1006	100%	0.0	95.15%	<a href="#">MF155892.1</a>
<input checked="" type="checkbox"/> <a href="#">Harpullia cupanioides isolate J096 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1006	1006	100%	0.0	95.15%	<a href="#">KR329416.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodoniaea viscosa ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1006	1006	100%	0.0	95.15%	<a href="#">GU935451.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodoniaea viscosa chloroplast partial rbcL gene for ribulose bisphosphate carboxylase large subunit, specimen voucher FFR1-YU004</a>	1006	1006	100%	0.0	95.15%	<a href="#">FN599447.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodoniaea viscosa ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1006	1006	100%	0.0	95.15%	<a href="#">DQ978445.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodoniaea viscosa subsp. angustifolia chloroplast partial rbcL gene for ribulose bisphosphate carboxylase large subunit, specimen voucher Forest J</a>	1006	1006	100%	0.0	95.15%	<a href="#">AM235129.1</a>
<input checked="" type="checkbox"/> <a href="#">Harpullia cupanioides isolate C001 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1000	1000	100%	0.0	94.98%	<a href="#">KR529412.1</a>
<input checked="" type="checkbox"/> <a href="#">Conchopetalum brachysepalum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	1000	1000	100%	0.0	94.98%	<a href="#">DQ978442.1</a>
<input checked="" type="checkbox"/> <a href="#">Sinoradikofera minor voucher CPG20218 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds, chloroplast</a>	996	996	98%	0.0	95.21%	<a href="#">JX027258.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodoniaea triquetra ribulose-1,5-bisphosphate carboxylase (rbcL) gene, partial cds, chloroplast gene for chloroplast product</a>	992	992	100%	0.0	94.96%	<a href="#">U30922.2</a>

**Dodonaea viscosa chloroplast, complete genome**

Sequence ID: [MF155892.1](#) Length: 159375 Number of Matches: 1

Range 1: 58504 to 59121 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1006 bits(523)	0.0	588/618(95%)	1/618(0%)	Plus/Plus
Query 1	AGCCGGTGT -AAGATTATAAATGACTTATTATACTCCTGACTATGTAACCAAGATAC	59		
Sbjct 58504	AGCCGGTGTAAAGATTATAAATGACTTATTATACTCCTGACTATGTAACCAAGATAC	58563		
Query 60	TGATATCTTGGCAGCATTC CAGTAACCTCAACCCGGGGTTCGGCTGAGGAAGCAGG	119		
Sbjct 58564	TGATATCTTGGCAGCATTC CAGTAACCTCAACCCGGGGTTCGGCTGAGGAAGCAGG	58623		
Query 120	GGCCCGGTAGCTCGGAATCTTCTACTGGTACATGGACAACCTGTGGACCGATGGGCT	179		
Sbjct 58624	GGCCCGGTAGCTCGGAATCTTCTACTGGTACATGGACAACCTGTGGACCGATGGGCT	58683		
Query 180	TACCAGCCTTGATCGTTATAAAGGACGATGCTACAACATTGAGCCTGTGCTGGAGAAGA	239		
Sbjct 58684	TACCAGCCTTGATCGTTATAAAGGACGATGCTACAACATTGAGCCTGTGCTGGAGAAGA	58743		
Query 240	AAATCAATATATGTTATGTAGCTTATCCCTTAGACCTTTTGAAGAAGGTTCTGTTAC	299		
Sbjct 58744	AAATCAATATATGTTATGTAGCTTATCCCTTAGACCTTTTGAAGAAGGTTCTGTTAC	58803		
Query 300	TAACATGTTACTCCATTGTGGTAATGTATTTGGGTTTAAAGCCCTGCGGCCTCACC	359		
Sbjct 58804	TAACATGTTACTCCATTGTGGTAATGTATTTGGGTTTAAAGCCCTGCGGCCTCACC	58863		
Query 360	TCTAGAGGATCTACGAATCCCTCCCGGTATTCGAAAACCTTCGAAGGGCCCTCAGG	419		
Sbjct 58864	TCTAGAGGATCTACGAATCCCTCCCGGTATTCGAAAACCTTCGAAGGGCCCTCAGG	58923		
Query 420	TATCCAAGTTGAGAGAGATAAATTGGACAAGTATGGACGTCCCTATTGGGATGGGCTAT	479		
Sbjct 58924	CATCCAAGTTGAGAGAGATAAATTGAACAAGTATGGACGTCCCTATTGGGATGACTAT	58983		
Query 480	TAAACCTAAATGGGATATCTCTAAGAAGTACTGTAGAATTTATGAATGTCAGG	539		
Sbjct 58984	TAAACCTAAATGGGATATCTCTAAGAAGTACTGTAGAATTTATGAATGTCAGG	59043		
Query 540	TCGTGGACTCACGTGGCAACCATGGTGGAGGTTCACTCCCAACCGCTTATACGTTT	599		
Sbjct 59044	TGGTGGACTTACTTACCAAGATGATGAGAAGCTAACTCCCAACCATTTATGCGTTG	59103		
Query 600	GAGAAACCGTTTTTGTGTT	617		
Sbjct 59104	GAGAGACCGTTCTTGT	59121		

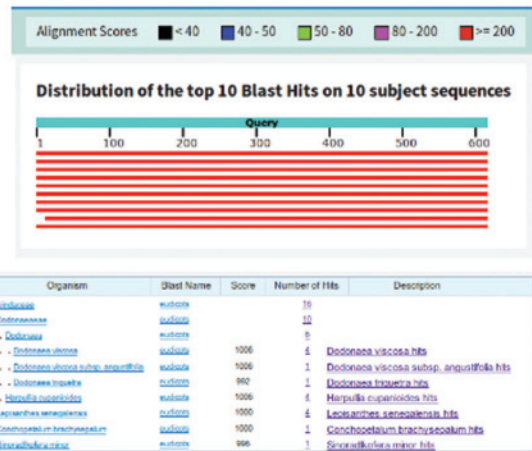


Figure 1: The NCBI blast results for retrieved *rbcL* sequences

(A) *Foeniculum vulgare*

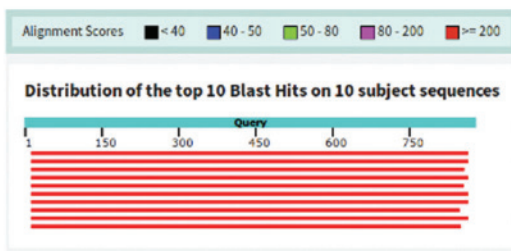
Description	Max Score	Total Score	Query Cover	E value	Per. Ident.	Accession
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare maturase K (matK) gene, partial cds, chloroplast</a>	1634	1634	96%	0.0	100.00%	MG946954.1
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare chloroplast, complete genome</a>	1634	1634	96%	0.0	100.00%	KR011054.1
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare isolate NMW4090 maturase K (matK) gene, partial cds, chloroplast</a>	1621	1621	96%	0.0	100.00%	JN894477.1
<input checked="" type="checkbox"/> <a href="#">Anethum graveolens chloroplast, complete genome</a>	1617	1617	96%	0.0	99.65%	KR011055.1
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare voucher MoSCL-269 maturase K (matK) gene, partial cds, chloroplast</a>	1617	1617	95%	0.0	100.00%	KP900773.1
<input checked="" type="checkbox"/> <a href="#">Anethum graveolens maturase K (matK) gene, partial cds, chloroplast</a>	1617	1617	96%	0.0	99.65%	EU018225.1
<input checked="" type="checkbox"/> <a href="#">Cuminum cyminum maturase K (matK) gene, partial cds, chloroplast</a>	1608	1608	96%	0.0	99.41%	MG946952.1
<input checked="" type="checkbox"/> <a href="#">Foeniculum vulgare voucher MoSCL-269 maturase K (matK) gene, partial cds, chloroplast</a>	1606	1606	95%	0.0	100.00%	KU499092.1
<input checked="" type="checkbox"/> <a href="#">Anethum graveolens maturase K (matK) gene, partial cds, chloroplast</a>	1602	1602	96%	0.0	99.18%	MG947047.1
<input checked="" type="checkbox"/> <a href="#">Anethum graveolens voucher A10 maturase K (matK) gene, partial cds, chloroplast</a>	1592	1592	95%	0.0	99.64%	KU499093.1

**Foeniculum vulgare maturase K (matK) gene, partial cds; chloroplast**

Sequence ID: [MG946954.1](#) Length: 885 Number of Matches: 1

Range 1: 36 to 885 [GenBank](#) [Graphics](#) [Next Match](#) [A View](#)

Score	Expect	Identities	Gaps	Strand
1634 bits(850)	0.0	850/850(100%)	0/850(0%)	Plus/Plus
Query 13	GTGTTAGAGATACTAATACCTACCCAGCCCATCTGGAAATATGGTTCAAACCTCTCGC	72		
Sbjct 36	GTGTTAGAGATACTAATACCTACCCAGCCCATCTGGAAATATGGTTCAAACCTCTCGC	95		
Query 73	TATTGGGTAAGAGCGCTTCTCTTTACATTTAATAGATCTTCTCCACGAGTATCGT	132		
Sbjct 96	TATTGGGTAAGAGCGCTTCTCTTTACATTTAATAGATCTTCTCCACGAGTATCGT	155		
Query 133	AGTTGGAACTCCAAATAAAGCCAGTCTGTTTTTCAAAAAGAAATCAAAGGTTTTTC	192		
Sbjct 156	AGTTGGAACTCCAAATAAAGCCAGTCTGTTTTTCAAAAAGAAATCAAAGGTTTTTC	215		
Query 193	TTCTGCTATATAATCTCATCTATGTGAATACGAATCCATCTCTGCTTTTTTCGTAAAC	252		
Sbjct 216	TTCTGCTATATAATCTCATCTATGTGAATACGAATCCATCTCTGCTTTTTTCGTAAAC	275		
Query 253	CAATCTTCTCATTATGCTCAAGCTCTCTGGAACCTCTTGAACGAATCTTTTCTAT	312		
Sbjct 276	CAATCTTCTCATTATGCTCAAGCTCTCTGGAACCTCTTGAACGAATCTTTTCTAT	335		
Query 313	GGAAAAATAGAACATCTTGGACTGTGAAGCTTTTGATAGGCCCTTCAGGACAATCTA	372		
Sbjct 336	GGAAAAATAGAACATCTTGGACTGTGAAGCTTTTGATAGGCCCTTCAGGACAATCTA	395		
Query 373	TGGTTGTTAAGGACCTTTTATGCATTATATAGGTATCAAGGAAATCAATCTCGCT	432		
Sbjct 396	TGGTTGTTAAGGACCTTTTATGCATTATATAGGTATCAAGGAAATCAATCTCGCT	455		
Query 433	TCAAAAAGGACGCCCCCTTTGATGAAAAATGGACATATTTTGTCAATTTATGGAAA	492		
Sbjct 456	TCAAAAAGGACGCCCCCTTTGATGAAAAATGGACATATTTTGTCAATTTATGGAAA	515		
Query 493	TGTCATTTTTACCTATGGCTCAGCCGGGACGGATCTGTATAAACCAATATATAATAAT	552		
Sbjct 516	TGTCATTTTTACCTATGGCTCAGCCGGGACGGATCTGTATAAACCAATATATAATAAT	575		
Query 553	TCCCTAGCTCTTCTGGGCTATCTATCAAGTGGCGACTAATCCTTCAATGGTACGCGAT	612		
Sbjct 576	TCCCTAGCTCTTCTGGGCTATCTATCAAGTGGCGACTAATCCTTCAATGGTACGCGAT	635		
Query 613	CAAAATGCTAGAAAATGCATTTAATAATGATAATCTTAAATAAGTTCGATACCTTGT	672		
Sbjct 636	CAAAATGCTAGAAAATGCATTTAATAATGATAATCTTAAATAAGTTCGATACCTTGT	695		
Query 673	CCAATTGTTCTCTGATTGGATCATTGGCTAAGGCGAGATTTGTAACTGATTTGGGAC	732		
Sbjct 696	CCAATTGTTCTCTGATTGGATCATTGGCTAAGGCGAGATTTGTAACTGATTTGGGAC	755		
Query 733	CCTATTAGTAAGGCGTTTGGACTGATTATCAGATTCTGATATTGTTCCGATTATAGG	792		
Sbjct 756	CCTATTAGTAAGGCGTTTGGACTGATTATCAGATTCTGATATTGTTCCGATTATAGG	815		
Query 793	AGTATCTGCAGAAATCTTCTCATTATTATAGTGGATCTCACAAAAAGGTTTGTAT	852		
Sbjct 816	AGTATCTGCAGAAATCTTCTCATTATTATAGTGGATCTCACAAAAAGGTTTGTAT	875		
Query 853	CGAATAAAGT 862			
Sbjct 876	CGAATAAAGT 885			



Organism	Blast Name	Score	Number of Hits	Description
<a href="#">Asteraceae</a>	<a href="#">matK</a>		10	
- <a href="#">Asteraceae</a>	<a href="#">matK</a>		9	
- <a href="#">Foeniculum vulgare</a>	<a href="#">matK</a>	1634	9	<a href="#">Foeniculum vulgare hits</a>
- <a href="#">Anethum graveolens</a>	<a href="#">matK</a>	1617	4	<a href="#">Anethum graveolens hits</a>
- <a href="#">Cuminum cyminum</a>	<a href="#">matK</a>	1608	1	<a href="#">Cuminum cyminum hits</a>

Figure 2: (continued)

(B) *Nitraria retusa*

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/> <a href="#">Nitraria tangutorum chloroplast, complete genome</a>	1488	1488	100%	0.0	98.17%	MK347423.1
<input checked="" type="checkbox"/> <a href="#">Nitraria sibirica chloroplast, complete genome</a>	1488	1488	100%	0.0	98.17%	MK347422.1
<input checked="" type="checkbox"/> <a href="#">Nitraria roborowskii chloroplast, complete genome</a>	1488	1488	100%	0.0	98.17%	MK347421.1
<input checked="" type="checkbox"/> <a href="#">Peganum harmala chloroplast, complete genome</a>	1373	1373	100%	0.0	95.73%	MK347420.1
<input checked="" type="checkbox"/> <a href="#">Peganum harmala trnK gene, intron, and maturase K (matK) gene, complete cds; chloroplast genes for chloroplast products</a>	1350	1350	100%	0.0	95.24%	AY177887.1
<input checked="" type="checkbox"/> <a href="#">Nitraria roborowskii voucher D453 maturase K (matK) gene, partial cds; chloroplast</a>	1344	1344	89%	0.0	98.37%	JF954729.1
<input checked="" type="checkbox"/> <a href="#">Nitraria sibirica voucher D1164 maturase K (matK) gene, partial cds; chloroplast</a>	1333	1333	89%	0.0	98.22%	JF954733.1
<input checked="" type="checkbox"/> <a href="#">Nitraria sphaerocarpa voucher D449 maturase K (matK) gene, partial cds; chloroplast</a>	1325	1325	89%	0.0	98.08%	JF954738.1
<input checked="" type="checkbox"/> <a href="#">Nitraria roborowskii voucher D1575 maturase K (matK) gene, partial cds; chloroplast</a>	1317	1317	88%	0.0	98.34%	JF954730.1
<input checked="" type="checkbox"/> <a href="#">Spondias tuberosa chloroplast, complete genome</a>	1173	1173	96%	0.0	91.65%	KU790502.1

**Nitraria tangutorum chloroplast, complete genome**

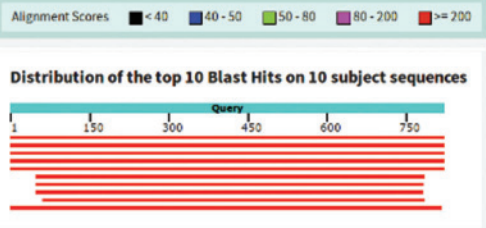
Sequence ID: [MK347423.1](#) Length: 159383 Number of Matches: 1

Range 1: 2485 to 3303 [GenBank](#) [Graphics](#) [Next Match](#) [Previous](#)

Score	Expect	Identifies	Gaps	Strand
1488 bits(774)	0.0	804/819(98%)	0/819(0%)	Plus/Plus

```

Query 1  ATATATTTTTCGATACAAACTCTTTTTTGGAGTCCACTGTGATAATGAGAAGA 60
Sbjct 2485  ATATATTTTTCGATACAAACTCTTTTTTGGAGTCCACTGTGATAATGAGAAGA 2544
Query 61  TTTCTGCATATATGCACAAATCGATCGATAATGAGAATCGGAGGAATCGCCCGAGTTC 120
Sbjct 2345  TTTCTGCATATATGCAGAAATCGATCGATAATGAGATTCGAGGAATCGCCCGAGTTC 2604
Query 121  GGTTCATTAATGGGATGACCTAATGTGTACAAAACCCGGCTTAGTCAATGATCCAATC 180
Sbjct 2605  GGTTCATTAATGGGATGACCTAATGTGTACAAAACCCGGCTTAGTCAATGATCCAATC 2664
Query 181  AGAGGAATAATGGGAATGTTGTATCGAACTTCCATGGCATTATCTATTAGAAATGAA 240
Sbjct 2665  AGAGGAATAATGGGAATGTTGTATCGAACTTCCATGGCATTATCTATTAGAAATGAA 2724
Query 241  TTTTCTAACATTTGAGTCCGTACCACCAAGGATTTACTCCACATTAGAAAGATAGCCT 300
Sbjct 2725  TTTTCTAACATTTGAGTCCGTACCACCAAGGATTTAGTCCACATTAGAAAGATAGCCT 2784
Query 301  AGAAGTTGATAGAATGTTGAATAGTGGTTTATATGGACCTTCCGGTTGAGACCAC 360
Sbjct 2785  AGAAGTTGATAGAATGTTGAATAGTGGTTTATATGGACCTTCCGGTTGAGACCAC 2844
Query 361  ACATGAAAATGCGATTCGCATAAATGACAAAGTAAATTTCCATTTATTCTAGAAAGA 420
Sbjct 2845  ACATGAAAATGCGATTCGCATAAATGACAAAGTAAATTTCCATTTATTCTAGAAAGA 2904
Query 421  GGGGTATCTTTGAAAGCGAATAGATTTTCTTGATATCAACATAATGCAGGAAGGA 480
Sbjct 2905  GGGGTATCTTTGAAAGCGAATAGATTTTCTTGATATCAACATAATGCAGGAAGGA 2964
Query 481  TCCTTGAAACAAACATAAGATGCTCGAAAATCATTAGCAACACTTCGGCAAGATATTCT 540
Sbjct 2965  TCCTTGAAACAAACATAAGATGCTCGAAAATCATTAGCAACACTTCGGCAAGATATTCT 3024
Query 541  ACTTTTCCATAGAAATATATTGCTCAAGAAAGACTCCAGAAGATGTTGATCGTAAATGA 600
Sbjct 3025  ACTTTTCCATAGAAATATATTGCTCAAGAAAGACTCCAGAAGATGTTGATCGTAAATGA 3084
Query 601  GACGATTGGTTACAGATAAAAAAGAAAGATAGATTATATTCACATAAATGAGAATTAT 660
Sbjct 3085  GACGATTGGTTACAGATAAAAAAGAAAGATAGATTATATTCACATAAATGAGAATTAT 3144
Query 661  AGGAAGAAATAATCTTTGATTACTTTTTGAAAAATCGAAATCGATTTCTTTGGAGTA 720
Sbjct 3145  AGTAAGAAATAATCTTTGATTACTTTTTGAAAAATCGAAATCGATTTCTTTGGAGTA 3204
Query 721  ATAAAACATTTCCAATTAATAACTCATGGAGAAAGAACCGTAATAAATGTAAGAAAG 780
Sbjct 3205  ATAAGACTATTTCCAATTAATAACTCATGGAGAAAGAACCGTAATAAATGTAAGAAAG 3264
Query 781  GCSTCTTTTACCGAGTAGCGAAGGGTTTGAACCAAGATT 819
Sbjct 3265  GCSTCTTTTACCGAGTAGCGAAGGGTTTGAACCAAGATT 3303
    
```



Organism	Blast Name	Score	Number of Hits	Description
Sapindales	eufiledb		14	
- Nitrariaceae	eufiledb		13	
- Nitraria	eufiledb		11	
- - Nitraria tangutorum	eufiledb	1488	1	<a href="#">Nitraria tangutorum hits</a>
- - Nitraria sibirica	eufiledb	1488	4	<a href="#">Nitraria sibirica hits</a>
- - Nitraria roborowskii	eufiledb	1488	3	<a href="#">Nitraria roborowskii hits</a>
- - Nitraria sphaerocarpa	eufiledb	1344	3	<a href="#">Nitraria sphaerocarpa hits</a>
- Peganum harmala	eufiledb	1373	2	<a href="#">Peganum harmala hits</a>
- Spondias tuberosa	eufiledb	1173	1	<a href="#">Spondias tuberosa hits</a>

Figure 2: (continued)

(C) *Dodonaea viscosa*

Description	Max Score	Total Score	Query Cover	E value	Per Ident	Accession
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa chloroplast, complete genome</a>	1506	1506	98%	0.0	98.89%	<a href="#">MF155892.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa subsp. angustifolia maturase K (matK) gene, partial cds: chloroplast</a>	1496	1496	96%	0.0	99.12%	<a href="#">KP110036.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa voucher Dod71 maturase K (matK) gene, partial cds: chloroplast</a>	1471	1471	94%	0.0	99.36%	<a href="#">MK125127.1</a>
<input checked="" type="checkbox"/> <a href="#">Trochocarpa sp. A CPL-2012 maturase K (matK) gene, partial cds: chloroplast</a>	1444	1444	93%	0.0	99.22%	<a href="#">JQ667215.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa subsp. angustifolia voucher OM2918 maturase K (matK) gene, partial cds: chloroplast</a>	1442	1442	92%	0.0	99.35%	<a href="#">JQ24958.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa voucher Abbott9229 maturase K (matK) gene, partial cds: chloroplast</a>	1436	1436	92%	0.0	99.22%	<a href="#">JX517869.1</a>
<input checked="" type="checkbox"/> <a href="#">Majidea fosteri maturase K (matK) gene, partial cds: chloroplast</a>	1435	1435	96%	0.0	97.98%	<a href="#">JN191124.1</a>
<input checked="" type="checkbox"/> <a href="#">Harpullia ramiflora chloroplast maturase K-like gene, complete sequence</a>	1435	1435	98%	0.0	97.78%	<a href="#">AF314605.1</a>
<input checked="" type="checkbox"/> <a href="#">Dodonaea viscosa voucher Hosam00025 maturase K (matK) gene, partial cds: chloroplast</a>	1429	1429	91%	0.0	99.47%	<a href="#">JX495703.1</a>
<input checked="" type="checkbox"/> <a href="#">Harpullia longipetala voucher WP200497 maturase K (matK) gene, partial cds: chloroplast</a>	1410	1410	93%	0.0	98.31%	<a href="#">GQ248131.1</a>

**Dodonaea viscosa chloroplast, complete genome**

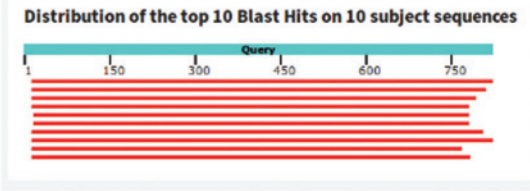
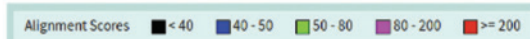
Sequence ID: [MF155892.1](#) Length: 159375 Number of Matches: 1

Range 1: 2344 to 3153 [GenBank](#) [Graphics](#) [View/Map](#) [History](#)

Score	Expect	Identities	Gaps	Strand
1506 bits(783)	0.0	801/810(99%)	0/810(0%)	Plus/Plus

```

Query 15   TCGATCAAAACTCTTTTTTGAAGATCCACTGTGATAATGAGAAAGATTTCTGCATAT   74
Sbjct 2344   TCGATCAAAACTCTTTTTTGAAGATCCACTGTGATAATGAGAAAGATTTCTGCATAT   2483
Query 75   ACGGCGAAACCGGTCGATAATATGAGAATCTGAGGAATCGGCCAAGTCGGCTTACTAAT   134
Sbjct 2484   ACGGCGAAACCGGTCGATAATATGAGAATCTGAGGAATCGGCCAAGTCGGCTTACTAAT   2463
Query 135  GGGATGCGCCTAATGGGTTACAAAACCGGCCCTTAGTCAATGATCCAATCAGAGGAATAAT   194
Sbjct 2464   GGGATGCGCCTAATGGGTTACAAAACCGGCCCTTAGTCAATGATCCAATCAGAGGAATAAT   2523
Query 195  AGGAATGGTTGCTCGAACCTTTCATAGCATTATCTATTAGAAATGAAATTTCTAGCAT   254
Sbjct 2524   AGGAATGGTTGCTCGAACCTTTCATAGCATTATCTATTAGAAATGAAATTTCTAGCAT   2583
Query 255  TTGACTCCGTAACCCAAAGTTTTAGTGGCCACTGGAAAAATAGCCAGAAAGTGGAT   314
Sbjct 2584   TTGACTCCGTAACCCAAAGTTTTAGTGGCCACTGGAAAAATAGCCAGAAAGTGGAT   2643
Query 315  AGAATCTTTGTATAAGTAGTTTATATGAATCCTCCGAGTTGAGACCACAGCTGAAAAAT   374
Sbjct 2644   AGAATCTTTGTATAAGTAGTTTATATGAATCCTCCGAGTTGAGACCACAGCTGAAAAAT   2703
Query 375  CCATTGCCATAAATGACAAGGTAATATTTCCATTATTTCATCAGAAGGGGTGTTTTTTT   434
Sbjct 2784   CCATTGCCATAAATGACAAGGTAATATTTCCATTATTTCATCAGAAGGGGTGTTTTTTT   2763
Query 435  TGAAGCCAGAATGGATTTCTTGGATATCAAGATAATGCATGAAAGGATCCTTGAACAA   494
Sbjct 2764   TGAAGCCAGAATGGATTTCTTGGATATCAAGATAATGCATGAAAGGATCCTTGAACAA   2823
Query 495  GCATAAGATGCTCGAAAAATCATTAGCAAAAGACTCGCAAGATGTTCTACTTTCCATA   554
Sbjct 2824   GCATAAGATGCTCGAAAAATCATTAGCAAAAGACTCGCAAGATGTTCTACTTTCCATA   2883
Query 555  AAAATGGATTGCTCAAGAAAGGACTCCAGAAAGATGTTGATCGTAAATGAGAAGATTGGTT   614
Sbjct 2884   AAAATGGATTGCTCAAGAAAGGACTCCAGAAAGATGTTGATCGTAAATGAGAAGATTGGTT   2943
Query 615  ACGGAGAAAAAAGAGATGGATTCATATCATATACATGAGAATTATATAGGAAAAAGAA   674
Sbjct 2944   ACGGAGAAAAAAGAGATGGATTCATATCATATACATGAGAATTATATAGGAAAAAGAA   3003
Query 675  AAATCTTGGATTACTTGTAAAAAATAGAAATGGATTTCTTTGGAGTAATAAGCAATT   734
Sbjct 3084   AAATCTTGGATTACTTGTAAAAAATAGAAATGGATTTCTTTGGAGTAATAAGCAATT   3063
Query 735  CCAATTATAATATTCGTAAGAAAGAAATCACAATAGATGTAAGAAAGAGGCATCTTTTAC   794
Sbjct 3064   CCAATTATAATATTCGTAAGAAAGAAATCACAATAGATGTAAGAAAGAGGCATCTTTTAC   3123
Query 795  TCAGTAGCGAAAGGTGTGAGCCAGATTTTC   824
Sbjct 3124   CCAGTAGCGAAAGGTGTGAGCCAGATTTTC   3153
    
```



Organism	Blast Name	Score	Number of Hits	Description
<i>Persea indica</i>	<a href="#">outdb</a>		11	
<i>Dodonaea</i>	<a href="#">outdb</a>		10	
- <i>Dodonaea</i>	<a href="#">outdb</a>		2	
- - <i>Dodonaea viscosa</i>	<a href="#">outdb</a>	1506	1	<a href="#">Dodonaea viscosa hits</a>
- - - <i>Dodonaea viscosa subsp. angustifolia</i>	<a href="#">outdb</a>	1496	1	<a href="#">Dodonaea viscosa subsp. angustifolia hits</a>
- - <i>Majidea fosteri</i>	<a href="#">outdb</a>	1435	1	<a href="#">Majidea fosteri hits</a>
- <i>Harpullia ramiflora</i>	<a href="#">outdb</a>	1435	1	<a href="#">Harpullia ramiflora hits</a>
- <i>Harpullia longipetala</i>	<a href="#">outdb</a>	1410	1	<a href="#">Harpullia longipetala hits</a>
- <i>Trochocarpa sp. A CPL-2012</i>	<a href="#">outdb</a>	1444	1	<a href="#">Trochocarpa sp. A CPL-2012 hits</a>

Figure 2: (continued)

(D) *Rumex nervosus*

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/> <a href="#">Commnicarpus scandens chloroplast partial trnK gene intron and matK gene for maturase K</a>	1508	1508	98%	0.0	99.28%	<a href="#">FN868306.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus sinuatus voucher Richard W. Spellenberg 7144 (NMC) maturase K (matK) gene, partial cds, chloroplast</a>	1513	1513	94%	0.0	99.51%	<a href="#">KY952383.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus peduncululosus voucher Mats Thulin 1301 (UPS) maturase K (matK) gene, partial cds, chloroplast</a>	1513	1513	94%	0.0	99.51%	<a href="#">KY952383.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus reniformis voucher Mats Thulin 4200 (UPS) maturase K (matK) gene, partial cds, chloroplast</a>	1011	1011	94%	0.0	99.50%	<a href="#">KY952388.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus greenwayi voucher Mats Thulin 606 (UPS) maturase K (matK) gene, partial cds, chloroplast</a>	1511	1511	94%	0.0	99.50%	<a href="#">KY952377.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus decipiens voucher Erin Tripp et al. 4127 (NMC) maturase K (matK) gene, partial cds, chloroplast</a>	1511	1511	94%	0.0	99.50%	<a href="#">KY952375.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus arabicus voucher Richard W. Spellenberg 7217 (NMC) maturase K (matK) gene, partial cds, chloroplast</a>	1511	1511	94%	0.0	99.50%	<a href="#">KY952368.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus ambiguus voucher Mats Thulin 11015 (UPS) maturase K (matK) gene, partial cds, chloroplast</a>	1511	1511	94%	0.0	99.50%	<a href="#">KY952366.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus squamosus var. squamosus voucher Erin Tripp et al. 4049 (NMC) maturase K (matK) gene, partial cds, chloroplast</a>	1509	1509	94%	0.0	99.50%	<a href="#">KY952365.1</a>
<input checked="" type="checkbox"/> <a href="#">Commnicarpus reniformis voucher Mats Thulin et al. 8337 (UPS) maturase K (matK) gene, partial cds, chloroplast</a>	1506	1506	94%	0.0	99.38%	<a href="#">KY952389.1</a>

**Commnicarpus scandens chloroplast partial trnK gene intron and matK gene for maturase K**  
 Sequence ID: [FN868306.1](#) Length: 2300 Number of Matches: 1

Range 1: 1065 to 1900 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1558 bits(810)	0.0	832/838(99%)	2/838(0%)	Plus/Minus

Query 17 AGTCGATGATATACCTTTATTGATACAAACTCTTTTCGTGAAGATGCCACTAGTAAT 76  
 Sbjct 1900 AGTCGAGATATATACCTTTATTGATACAAACTCTTTTCGTGAAGA-GCCACTA-TAAT 1843

Query 77 AATGAGAAAGATTTCTGCGTATACGCCCAAATCTATCAATAATCGGAATCGATAAAT 136  
 Sbjct 1842 AATGAGATAGATTTCTGCGTATACGCCCAAATCTATCAATAATCGGAATCGATAAAT 1783

Query 137 CGGTCCAAACTGACTTACTAATGGGATGCCCTAATACATTACAAAATTCGCTTTAGCCA 196  
 Sbjct 1782 CGGTCCAAACTGACTTACTAATGGGATGCCCTAATACATTACAAAATTCGCTTTAGCCA 1723

Query 197 ACGATCCAAC CAGAGGAATAATTTGGAACATGGTATCGAACTTTTAAATGGTATTTTCTA 256  
 Sbjct 1722 ACGATCCAAC CAGAGGAATAATTTGGAACATAGATATCGAACTTTTAAATGGTATTTCTA 1663

Query 257 TTAGAAATGGATTTTCTAACATTTGACTCCGACCACCTGAAGAAATGGGTCGCACACTTG 316  
 Sbjct 1662 TTAGAAATGGATTTTCTAACATTTGACTCCGACCACCTGAAGAAATGGGTCGCACACTTG 1603

Query 317 AAAGAAAACC CAGAAAGTCGAGGGAATGCTTTGATAATCGATTGATATATATTTTCTTG 376  
 Sbjct 1602 AAAGAAAACC CAGAAAGTCGAGGGAATGCTTTGATAATCGATTGATATATATTTTCTTG 1543

Query 377 ATTGAGACCACACATAAAAATAAGATTGCCAAAATTTGATAAGGTAATATTTCCATTAT 436  
 Sbjct 1542 ATTGAGACCACACATAAAAATAAGATTGCCAAAATTTGATAAGGTAATATTTCCATTAT 1483

Query 437 GCATCAGAAAAGATGTACCTTTTGAAGAAAGAAATGATTTTCTTGATATCGAATAAAT 496  
 Sbjct 1482 GCATCAGAAAAGATGTACCTTTTGAAGAAAGAAATGATTTTCTTGATATCGAATAAAT 1423

Query 497 CGGTAAATGGATCTTTGAAAAGCCATAAGATAATCGGACAAATTTAGTTAAAACCTTGA 556  
 Sbjct 1422 CGGTAAATGGATCTTTGAAAAGCCATAAGATAATCGGACAAATTTAGTTAAAACCTTGA 1363

Query 557 CTAGATATTCATCTTTCCCGGAAATAAATTCGTTCAAGAGGGCTTCAAAAAGATGTTG 616  
 Sbjct 1362 CTAGATATTCATCTTTCCCGGAAATAAATTCGTTCAAGAGGGCTTCAAAAAGATGTTG 1303

Query 617 ATCGTAAATAAGAGGATCGGTTGCGGAGAAAACAAAATGGATTGATTTATATACAT 676  
 Sbjct 1302 ATCGTAAATAAGAGGATCGGTTGCGGAGAAAACAAAATGGATTGATTTATATACAT 1243

Query 677 AGAAATATATAAGAACAAGAAATAATCTTTGAGTCTTTTGCAAAATCGAAATGGAAAT 736  
 Sbjct 1242 AGAAATATATAAGAACAAGAAATAATCTTTGAGTCTTTTGCAAAATCGAAATGGAAAT 1183

Query 737 TTTTTGAAGTAATAAGACTATTCGATTACGATTACGATACTATAAGAAAATTCGTA 796  
 Sbjct 1182 TTTTTGAAGTAATAAGACTATTCGATTACGATTACGATACTATAAGAAAATTCGTA 1123

Query 797 ATAAATGCGAAGAAGAGGCATCTTCCACCAGTAGCGAAGATTTGAACCAAGATTC 854  
 Sbjct 1122 ATAAATGCGAAGAAGAGGCATCTTCCACCAGTAGCGAAGATTTGAACCAAGATTC 1065

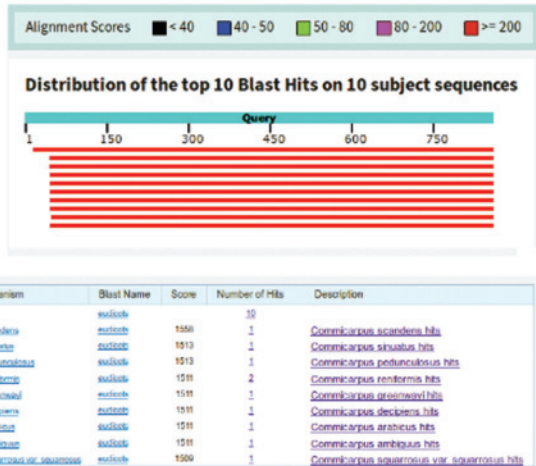
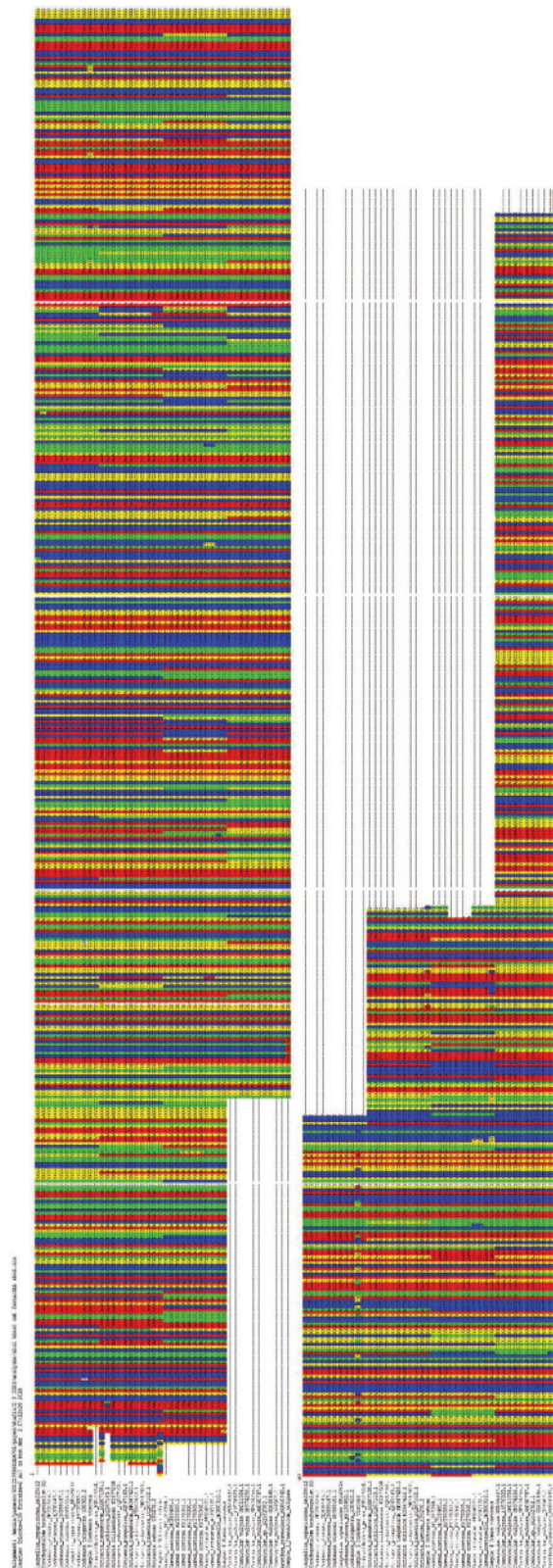
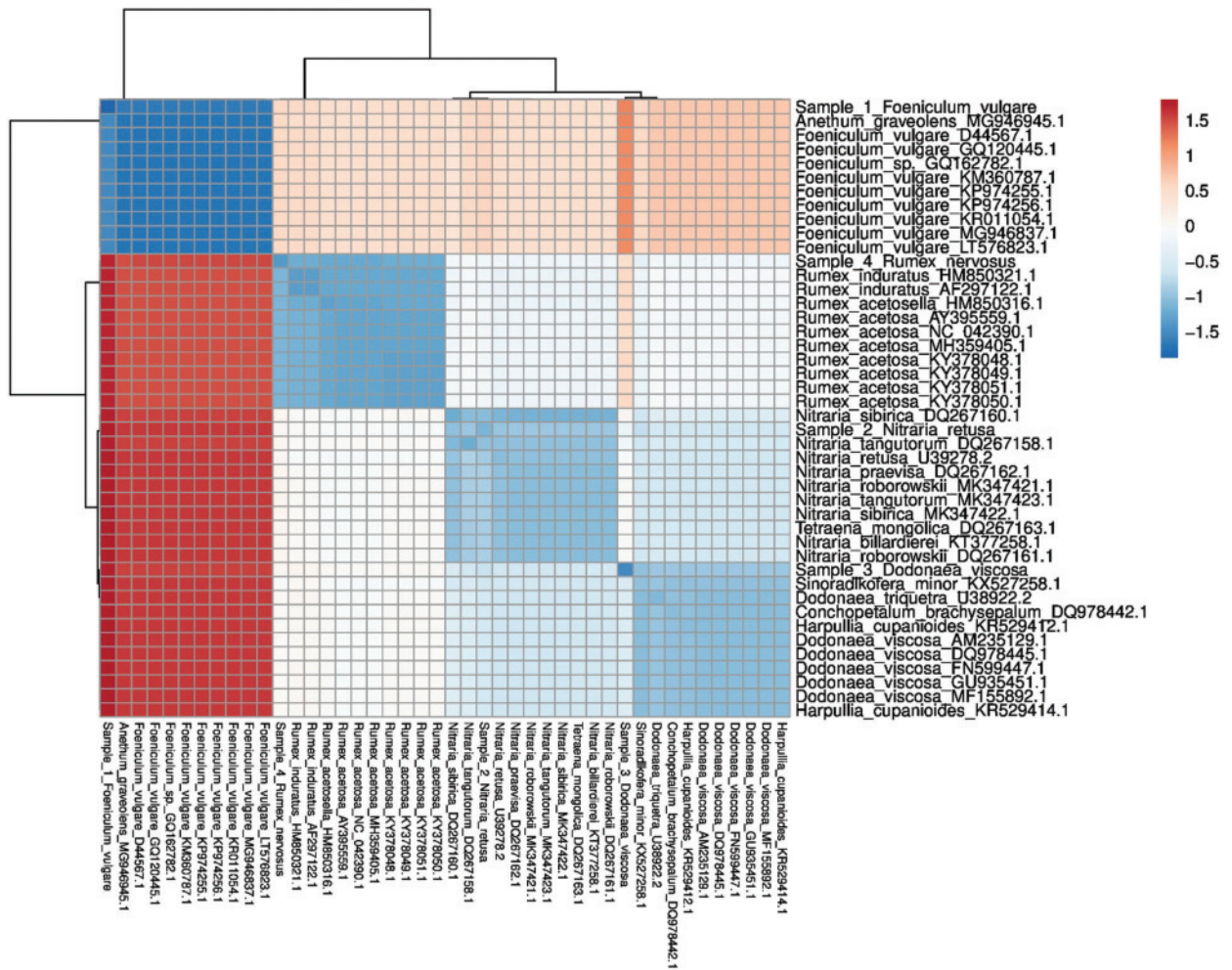


Figure 2: The NCBI blast results for retrieved *matK* sequences



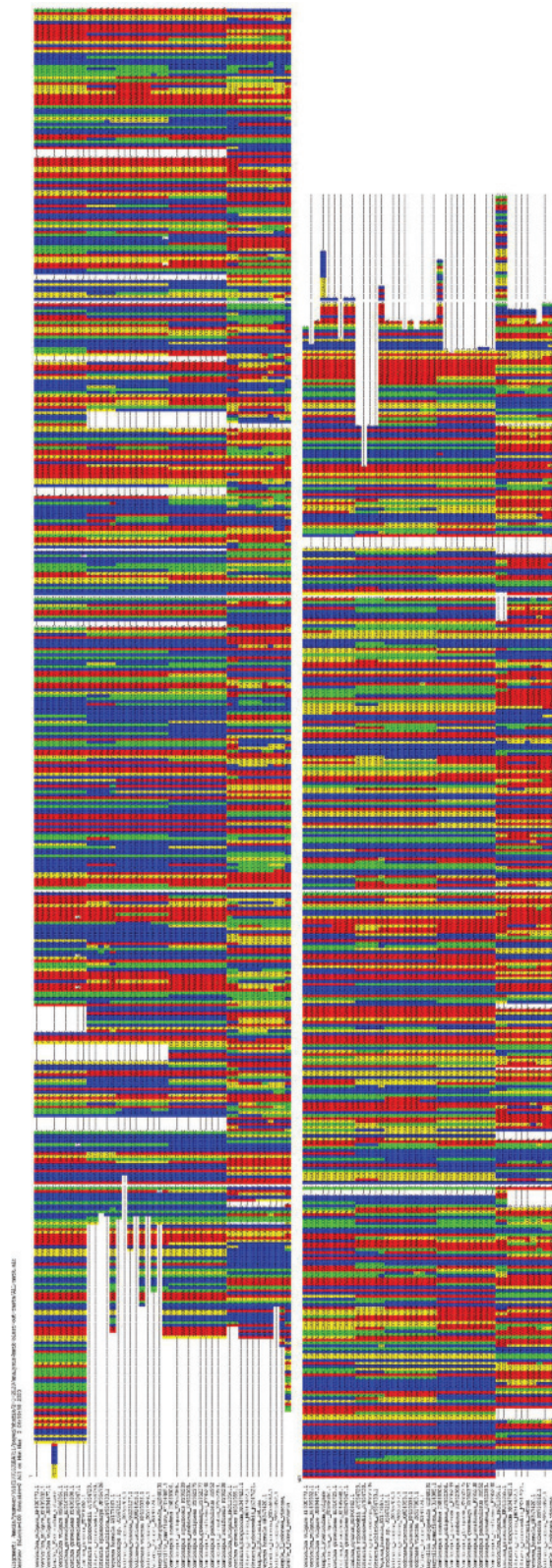
**Figure 3:** The sequence alignment of retrieved *rbcL* sequences in addition to similar plant sequences retrieved from NCBI



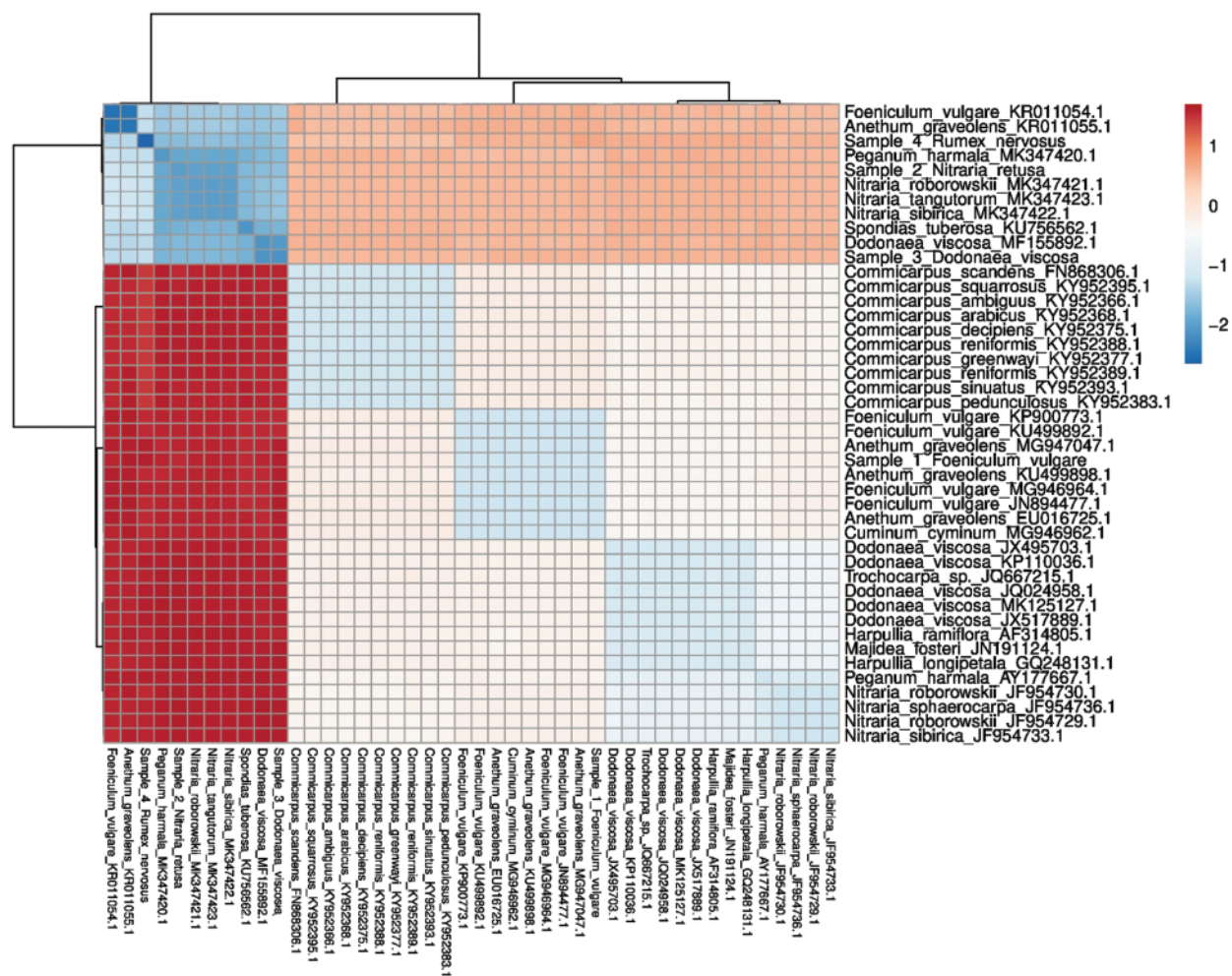
**Figure 4:** The heatmap analysis of *rbcL* sequences depending on sequence similarities, where red color indicates high sequence difference and blue color indicates high sequence similarities







**Figure 6:** The sequence alignment of retrieved *matK* sequences in addition to similar plant sequences retrieved from NCBI



**Figure 7:** The heatmap analysis of *matK* sequences depending on sequence similarities, where red color indicates high sequence difference and blue color indicates high sequence similarities



technique was not reliable for plant identification, where *matK* and *rbcL* must be used in dual DNA-barcoding procedures.

**Acknowledgement:** The author is thankful to University of Tabuk, Tabuk, Kingdom of Saudi Arabia for supporting my research.

**Funding Statement:** The author(s) received no specific funding for this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Abdel-Sattar, E., Abou-Hussein, D., Petereit, F. (2015). Chemical constituents from the leaves of *Euphorbia ammak* growing in Saudi Arabia. *Pharmacognosy Research*, 7(1), 14–17. DOI 10.4103/0974-8490.147136.
2. Alsamman, A. M. (2019). The art of bioinformatics learning in our Arabic world. *Highlights in BioScience*, 2, 1–10.
3. Alshehri, M. A., Aziz, A. T., Alzahrani, O., Alasmari, A., Ibrahim, S. et al. (2019). DNA-barcoding and species identification for some Saudi Arabia seaweeds using *rbcL* gene. *Journal of Pure and Applied Microbiology*, 13(4), 2035–2044. DOI 10.22207/JPAM.13.4.15.
4. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z. et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389–3402. DOI 10.1093/nar/25.17.3389.
5. Awadh Ali, N. A., Al Sokari, S. S., Gushash, A., Anwar, S., Al-Karani, K. et al. (2017). Ethnopharmacological survey of medicinal plants in Albaha region, Saudi Arabia. *Pharmacognosy Research*, 9(4), 401–407. DOI 10.4103/pr.pr\_11\_17.
6. Barthet, M. (2006). *Expression and function of the Chloroplast-Encoded Gene Matk, United States*. The faculty of the Virginia Polytechnic Institute and State. Virginia Polytechnic Institute and State University.
7. Chester, K., Zahiruddin, S., Ahmad, A., Khan, W., Paliwal, S. et al. (2017). Bioautography-based identification of antioxidant metabolites of *Solanum nigrum* L. and exploration its hepatoprotective potential against D-galactosamine-induced hepatic fibrosis in rats. *Pharmacognosy Magazine*, 13(50), 179–188. DOI 10.4103/pm.pm\_432\_16.
8. Cutarelli, G. G., Capuano, F. C. F. (2018). Species identification by means of mitochondrial cytochrome b DNA sequencing in processed anchovy, sardine and tuna products. *Food and Nutrition Sciences*, 9(4), 369–375. DOI 10.4236/fns.2018.94029.
9. CBOL Plant Working Group, Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S. et al. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), 12794–12797. DOI 10.1073/pnas.0905845106.
10. Hebert, P. D. N., Cywinska, A., Ball, S. L., DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313–321. DOI 10.1098/rspb.2002.2218.
11. Hollingsworth, P. M., Graham, S. W., Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS One*, 6(5), e19254. DOI 10.1371/journal.pone.0019254.
12. Kazi, M. A., Reddy, C. R. K., Jha, B. (2013). Molecular phylogeny and barcoding of caulerpa (bryopsidales) based on the *tufA*, *rbcL*, 18S rDNA and ITS rDNA genes. *PLoS One*, 8(12), e82438. DOI 10.1371/journal.pone.0082438.
13. Liu, Y., Wang, K., Liu, Z., Luo, K., Chen, S. et al. (2013). Identification of medical plants of 24 *ardisia* species from China using the *matK* genetic marker. *Pharmacognosy Magazine*, 9(36), 331–337. DOI 10.4103/0973-1296.117829.
14. Lucas, C., Thangaradjou, T., Papenbrock, J. (2012). Development of a DNA barcoding system for seagrasses: successful but not simple. *PLoS One*, 7(1), e29987. DOI 10.1371/journal.pone.0029987.
15. Metsalu, T., Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research*, 43(W1), W566–W570. DOI 10.1093/nar/gkv468.

16. Parson, W., Pegoraro, K., Niederstätter, H., Föger, M., Steinlechner, M. (2000). Species identification by means of the cytochrome b gene. *International Journal of Legal Medicine*, 114(1–2), 23–28. DOI 10.1007/s004140000134.
17. Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U. M. et al. (2010). Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. *PLoS One*, 5(10), e13674. DOI 10.1371/journal.pone.0013674.
18. Shafik, D. I., Alsamman, M. A. (2019). A rapid method for DNA extraction of cotton mature fiber suitable for PCR fingerprinting. *Cold Spring Harbor Laboratory*, 1, 22–57.
19. South, A. (2011). rworldmap : a new R package for mapping global data. *R Journal*, 3(1), 35–43. DOI 10.32614/RJ-2011-006.
20. Steven, G. N., Subramanyam, R. (2009). Testing plant barcoding in a sister species complex of pantropical Acacia (Mimosoideae, Fabaceae). *Molecular Ecology Resources*, 9(7), 172–180. DOI 10.1111/j.1755-0998.2009.02642.x.
21. Sugita, M., Shinozaki, K., Sugiura, M. (1985). Tobacco chloroplast tRNALys (UUU) gene contains a 2.5-kilobase-pair intron: an open reading frame and a conserved boundary sequence in the intron. *Proceedings of the National Academy of Sciences of the United States of America*, 82(11), 3557–3561. DOI 10.1073/pnas.82.11.3557.
22. Tallei, T. E., Kolondam, B. J. (2015). DNA barcoding of sangihe nutmeg (*myristica fragrans*) using matK gene. *HAYATI Journal of Biosciences*, 22(1), 41–47. DOI 10.4308/hjb.22.1.41.
23. Tan, P., Lim, P., Lin, S., Phang, S., Draisma, S. A. et al. (2015). Foliose Halymenia species (Halymeniaceae, Rhodophyta) from Southeast Asia, including a new species. *Halymenia Malaysiana Sp. Nov., Botanica Marina*, 58(3), 203–217.
24. Tanaka, S., Ito, M. (2020). DNA barcoding for identification of agarwood source species using TrnL-TrnF and matK DNA sequences. *Journal of Natural Medicines*, 74(1), 42–50. DOI 10.1007/s11418-019-01338-z.
25. Thakur, V. V., Tiwari, S., Tripathi, N., Sapre, S. (2016). DNA barcoding and phylogenetic analyses of mentha species using. *Annals of Phytomedicine*, 5(1), 59–62.
26. Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673–4680. DOI 10.1093/nar/22.22.4673.

### Supplementary

#### Sample\_1\_matK-1\_Foeniculum\_vulgare

```
GGGGGTTTTTTAGTGTAGAGATACTAATACCCTACCCAGCCCATCTGGAAATATTGGTTCA
AACTCTTCGCTATTGGGTAAGACGCTTCTTTACATTTATTAAGATTCTTCTCCACGAGT
ATCGTAGTTGGAATACTCCAAATAAGCCAGTTCTTGTTTTTCAAAAAGAAATCAAAGGTTTTT
CTTCGTCCTATATAATTCTCATCTATGTGAATACGAATCCATCTTCGTCTTTTTTCGTAACCAATCT
TCTCATTTATGCTCAACGCTTCTGGAACCCTTCTGAACGAATCTTTTTCTATGAAAAATAGA
ACATCTTGACTTGTAGAAGCTTTTGATAAGGCCTTTCAGGACAATCTATGGTTGTTTAAGGAC
CCTTTCATGCATTATATTAGGTATCAAGGAAAATCAATTCTCGCTTCAAAGGGACGCCCTTTT
GATGAAAAAATGGACATATTTTTTTGTCAATTTATGGAAATGTCAATTTTTACCTATGGTCTCAGCC
GGGACGGATCTGTATAAACCAATTATATAATAATTCCCTAGCTCTTCTGGGCTATCTATCAAGTGC
GCGACTAAATCCTTCAATGGTACGCAGTCAAATGCTAGAAAATGCATTTATAATTGATAATCCTA
TTAATAAGTTTCGATACTCTTGTTCCAATTGTTCCCTCTGATTGGATCATTGGCTAAGGCGAGATTTT
GTAACGTATTGGGGCACCCATTAGTAAGGCGGTTTGGACTGATTTATCAGATTCTGATATTGTTG
TCCGATTTAGGAGTATCTGCAGAAATCTTCTCATTATTATAGTGGATCCTCACAAAAAAGAGT
TTGTATCGAATAAAGTGGGGGGGTTTTTTTTT
```

#### Sample\_2\_matK-2\_Nitraria\_retusa

```
ATATATTTTATTCGATACAAACTCTTTTTTTTTGAGGATCCACTGTGATAATGAGAAAGATTT
CTGCATATATGCACAAATCGATCGATAATATGAGAATCGGAGGAATCGGCCAGGTCGGTTTAC
TAATGGGATGACCTAATGTGTTACAAAACCCCGCCTTAGTCAATGATCCAATCAGAGGAATAAT
```

GGGAATGTTTGTATCGAACTTCCTCATGGCATTATCTATTAGAAATGAATTTTCTAACATTTGAGT  
 CCGTACCACCAAAGGATTTACTCCCACATTAGAAAGATAGCCTAGAAAGTTGATAGAATGTTTG  
 AATAGGTGGTTTATATGGACCCTTCCTGGTTGAGACCACACATGAAAATGCGATTGCCATAAATT  
 GACAAAGTAATATTTCCATTTATTCATCAGAAGAGGCGTATCTTTTGAAGCGAGAATAGATTTTCC  
 TTGATATCTAACATAATGCAGGAAAGGATCCTTGAACAAACATAAGATGTCCTGAAAATCATTAG  
 CAACAACTTCGGCAAGATATTCTACTTTTCCATAGAAATATATTCGCTCAAGAAAGACTCCAGAA  
 GATGTTGATCGTAAATGAGACGATTGGTTACAGATAAAAAAGAAGATAGATTCATATTCACATA  
 AATGAGAATTATATAGGAAGAAGAATAATCTTTGATTACTTTTTGAAAAAATCGAAATCGATTTT  
 TTTGGAGTAATAAACTATTCCAATTAATACTCATGGAGAAAGAACCCTAATAAATGTAAAG  
 AAGAGGCGTCTTTTACCGAGTAGCGAAGGGTTTGAACCAAGATT

Sample\_3\_matK-1\_Dodonaea\_viscose

TAGATATTTATACATCGATACAACTCTTTTTTTTTGAGGATCCACTGTGATAATGAGAAAG  
 ATTTCTGCATATACGCGCAAACCGGTCGATAATATGAGAATCTGAGGAATCGGCCCAAGTCGGC  
 TTAATAATGGGATGCCCTAATGGGTTACAAAAACGCGCCTTAGTCAATGATCCAATCAGAGGAA  
 TAATAGGAATGGTTGTCTCGAACTTCTTCATAGCATTATCTATTAGAAATGAATTTTCTAGCATT  
 GACTCCGTACCACCAAAGTTTTAGTTGCCACTGGAAAAATAGCCAGAAAGTGGATAGAATC  
 TTTGTATAAGTAGTTTATATGAATCCTTCCGAGTTGAGACCACACGTGAAAATTCCATTGCCATA  
 AATTGACAAGGTAATATTTCCATTTATTCATCAGAAGGGGTGTTTTTTTTGAAGCCAGAATGGAT  
 TTTCTTGATATCTAAGATAATGCATGAAAGGATCCTTGAACAAGCATAAGATGTCCTGAAAATC  
 ATTAGCAAAGACTCCGACAAGATGTTCTACTTTTCCATAAAAATGGATTTCGCTCAAGAAGGACT  
 CCAGAAGATGTTGATCGTAAATGAGAAGATTGGTTACGGAGAAAAAGAAGATGGATTCATAT  
 TCATATACATGAGAATTATATAGGAAAAAGAAAAATCTTGGATTACTTGTTAAAAAATAGAAAT  
 GGATTTCTTTGGAGTAATAAGACAATTCCAATTATAATTCGTAAAGAAAGAATCACAAATAGAT  
 GTAAAGAAGAGGCATCTTTTACTCAGTAGCGAAAGGTGTGAGCCACG ATTC

Sample\_4\_matK-5\_Rumex\_nervosus

CCAGTTCGAGCCATGGAGTCGATGTATATACTTTATTTCGATACAACTCTTTTTTCGTGAAG  
 ATGCCACTAGTAATAATGAGAAAGATTTCTGCGTATACGCCCAAATCTATCAATAATATCGGAATC  
 TGATAAATCGGTCCAAACTGACTTACTAATGGGATGCCCTAATACATTACAAAATTTTCGCTTTAG  
 CCAACGATCCAACCAGAGGAATAATTGGAAGTATGGTATCGAACTTTTTAATGGTATTTTCTATTA  
 GAAATGGATTTTCTAACATTTGACTCCGCACCCTGAAGAATTGAGTCGCACACTTGAAAGAA  
 AACCCAGAAAGTCGAGGGAATGCTTTGATAATCGATTGATATATATTTTTCTTGATTGAGACCAC  
 ACATAAAAATAAGATTGCCAAAAATTGATAAGGTAATATTTCCATTTATGCATCAGAAAAGATGT  
 ACCTTTTGAAGAAAGAATTGATTTTCTTTCGATATCGAATATAATGCGTAAATGGATCTTTGAAAA  
 GCCATAAGATAATCCGACAATTTTAGTTAAAACCTTTGACTAGATATTCTATCTTTCCGCGGAAA  
 TAAATTCGTTCAAGAAGGGCTTCAAAGATGTTGATCGTAAATAAGAGGATCGGTTGCGGAGA  
 AAAACAAAATGGATTCGTATTCATATACATAGAAATTATATAAGAACAAGAATAATCTTTGAGT  
 CCTTTTTGCAAAAATCGAAATGGAATTTTTTGAAGTAATAAGACTATTCCGATTACGATTACGAT  
 ACTCATAAAGAAAAATCGTAATAAATGCGAAGAAGAGGCATCTTTC  
 ACCCAGTAGCGAAGAGTTTGAACCAAGATTTT

Sample\_1\_rbcL-1\_Foeniculum\_vulgare

CGGTAGCTGCCGAAACAACAACCTGGTACATGGACCACTGTGTGGACCGATGGACTTACCA  
 GTCTTGATCGTTACAAAGGGCGCTGCTACGGAATCGAGCCCCTTGCTGGAGAAGAAAATCAATA  
 TATCGCTTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTC  
 CATTGTAGGTAATGTATTTGGGTTCAAAGCCCTGCGCGCTCTACGTCTGGAAGATCTGCGAATC  
 CCCGTTGCTTATGTTAAAACCTTTCCAAGGACCGCCTCATGGCATCCAAGTTGAGAGAGATAAAT  
 TGAACAAGTATGGTCGTCCCCTGTTGGATGTACTATTAAACCTAAATTGGGGTTATCCGCTAAAA

ACTACGGTAGAGCGGTTTATGAATGTCTCCGCGGTGGACTTGATTTTACCAAAGACGATGAGAA  
 TGTGAACTCCCAACCATTTATGCGTTGGAGAGATCGTTTCTTATTTTGTGCCGAAGCAATTTATA  
 AAGCACAGGCTGAAACTGGTGAAATCAAAGGGCATTACTTGAATGCTACTGCGGGTACATGTG  
 AAGAAATGATGAAAAGGGCTATATTTGCCAGAGAATTGGGAGTTCCTATCGTAATGCATGATTAC  
 TTAACAGGGGGATTCACTGCAAATACTAGCTTAGCCCATTTATGCCGAGATAATGGCCTACTTCT  
 TCACATCCACCGTGCAATGCATGCAGTTATTGATAGACAGTTGAAACATGGAATACACAACCGT  
 GTACTAGCTAAAGCGTTACGTATGTCTGGTCCACAAC

Sample\_2\_rbcL-2\_Nitraria\_retusa

GATTCAGCCGGTGTTAAGATTATAAATTGACTTATTATACTCCTGAATATGAAACCAAAGATA  
 CTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCACCCGAAGAAGCAGGGG  
 CTGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACGGTATGGACCGATGGGCTTACCAG  
 CCTTGATCGTTACAAAGGACGATGCTACAACATCGAGCCTGTTGCTGGAGAAGAAAATCAATAT  
 ATATGTTATGTAGCTTATCCTTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCC  
 ATTTGGGTAATGTATTTGGGTTCAAAGCCCTGCGTGCTCTACGTCTAGAGGATCTACGAATCCC  
 TCCTGCGTATATTA AAACTTTCCAAGGCCCGCCTCACGGTATCCAGGTTGAGAGAGATAAATTG  
 AACAAGTATGGCCGTCCCCTATTGGGATGTACTATTAACCTAAATTGGGGTTATCCGCTAAGAA  
 CTACGGTAGAGCAGTTTATGAATGTCTACGTGGTGGACTTGACTTTACCAAAGATGATGAGAAC  
 GTGAACTCGCAACCATTTATGCGTTGGAGAGACCGTTTCTTATTTTGC GCGGAAGCTATTTATAA  
 AGTGCAGGCTGAAACAAGTGAATCAAAGTCACTACTTAAAT GCTACTGT

Sample\_3\_rbcL-4\_Dodonaea\_viscosa

AGCCGGTGTTAAGATTATAAATTGACTTATTATACTCCTGACTATGTAACCAAAGATACTGAT  
 ATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGGGTTCCGCCTGAGGAAGCAGGGGGCCGCG  
 GTAGCTGCGGAATCTTCTACTGGTACATGGACAACCTGTGTGGACCGATGGGCTTACCAGCCTTG  
 ATCGTTATAAAGGACGATGCTACAACATTGAGCCTGTTGCTGGAGAAGAAAATCAATATATATG  
 TTATGTAGCTTATCCTTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGT  
 GGGTAATGTATTTGGGTTTAAAGCCCTGCGCGCTCTACGTCTAGAGGATCTACGAATCCCCTCCCG  
 CGTATTCGAAAACCTTTCCAAGGCCCGCCTCACGGTATCCAAGTTGAGAGAGATAAATTGGACAA  
 GTATGGACGTCCCCTATTGGGATGGGCTATTAACCTAAATTGGGATTATTCTCTAAGAACTACTG  
 TAGAACATTTTATGAATGTCGACGTCGTGGACTTCACGTGGCCAACCATGGTGAGGAGGTTTAC  
 TCCAACCGCTTATACGTTTGAAGAAACCGTTTTTTTGT

Sample\_4\_rbcL-5\_Rumex\_nervosus

GTTAGAGAATACAAATTGACTTATTATACTCCTGACTATGAACCCCATGACCATGATATCTTG  
 GCAGCATTTTCGAGTAACTCCTCAACCTGGAGTTCACCAGAAGAAGCAGGGGGCCGCGGTAGCT  
 GCCGAATCTTCTACTGGTACATGGACAACCTGTGTGGACCGATGGGCTTACCAGCCTTGATCGTT  
 ACAAAGGACGATGCTACCACATCGAGCCTGTTTCTGGAGAAGAAAGTCAGTTTATTGCTTATGT  
 AGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTGGGTA  
 ATGTATTTGGGTTCAAAGCCCTGCGTGCTCTACGTTTGGAGGATTTGCGAATCCCCTCCTGCTTAT  
 ACGAAAACCTTTCCAAGGCCCGCCTCATGGTATCCAAGTTGAGAGAGATAAATTGAACAAATATG  
 GACGTCCCCTATTGGGATGTACTATTAACCGAAATTGGGGTTGTCCGCTAAGAACTACGGCCG  
 AGCAGTTTATGAATGTCTTCGTGGCGGACTTGATTTTACCAAAGATGATGAAAACGTGAACTCC  
 CAACCATTTATGCGTTGGAGAGACCGGGTCTTATTCTGTGGGGAAGCTATTTTAAATCACAGTC  
 TGAAACACGTGATTTAAATGACATTACTTGAATGCTACTGC