

Article

# Molecular Identification and Phylogenetic Placement of *Rosa arabica* Crép. (Rosaceae), a Critically Endangered Plant Species

Ahmed EL-Banhawy <sup>1,\*</sup>, Carmen Acedo <sup>2</sup> , Sameer Qari <sup>3,\*</sup>  and Ahmed Elkordy <sup>2,4</sup> 

<sup>1</sup> Botany Department, Faculty of Science, Suez Canal University, Ismailia 41522, Egypt

<sup>2</sup> Biodiversity and Environment Management Department, Faculty of Biological and Environmental Sciences, University of León, 24071 León, Spain; c.acedo@unileon.es (C.A.); aelkordy@science.sohag.edu.eg (A.E.)

<sup>3</sup> Biology Department, Aljumum University College, Umm Al-Qura University, Makkah 21421, Saudi Arabia

<sup>4</sup> Botany and Microbiology Department, Faculty of Science, Sohag University, Sohag 82524, Egypt

\* Correspondence: ahmedbanhaw@science.suez.edu.eg (A.E.-B.); shqari@uqu.edu.sa (S.Q.)

Received: 29 October 2020; Accepted: 7 December 2020; Published: 9 December 2020



**Abstract:** The Egyptian narrowly endemic and critically endangered plant species *Rosa arabica* Crép. was studied employing a taxonomic and molecular approach. Morphological investigations, distance analysis, and phylogenetic reconstruction revealed that *R. arabica* is a distinct species with great affinity to *R. canina* and differentiated from *R. rubiginosa*. Molecular identification based on the sequences of multiple markers single or in combination ITS, *matK*, *rbcL*, and *trnL-F* succeeded in identifying *R. arabica* at genus and species levels. We evaluated the potential of each marker and a combination of the nuclear ITS -Internal Transcribed Spacer- with one of the plastid markers, *matK*, *rbcL*, or *trnL-F*, to accurately identify *Rosa* species. All of them were successful in identifying *R. arabica*. Classification based on DNA sequences shows that *R. arabica* is placed within section *Caninae* in a clade comprising *R. canina* and *R. rubiginosa*. Moreover, *R. arabica* is closely related to other European *Rosa* species. In conclusion, our results indicate that the four DNA markers can provide species resolution in the context of the genus *Rosa* and relatives, aiming to characterize morphology and genetic diversity in the ecological and economically important genus *Rosa*.

**Keywords:** biosystematics; *Caninae*; *Rosa arabica*; phylogeny; endemic; Saint Catherine; Sinai; Egypt

## 1. Introduction

The family Rosaceae is a large sub-cosmopolitan family located mainly in the temperate and warm areas of the Northern Hemisphere [1]. Heywood [2] stated that the family consists of 122 genera and 3370 species. Christenhusz et al. [3], reported 54–75 (90) genera and 2950 species. The genus *Rosa* L. comprises approximately 200 species [4]. Recently, ref. [5] updated those numbers as 5819 species in 109 genera and 310 species in the Genus *Rosa*.

Conventional taxonomy by Wissemann [6] divides the genus into four subgenera, three of which are monotypic or contain two species (*Hulthemia* Dumort. [7], *Platyrhodon* (Hurst) Rehder [8], and *Hesperhodos* Cockerell [9]) and subgenus *Rosa* L. with 11 sections [10]. In 2005, ref. [11] stated that all the subgenera of *Rosa* are “best treated as sections, and not as separate genera”. They formally transferred all subgenera to sections. This was later supported by [12–14].

Hybridization is common within the genus *Rosa*. Hybrids contribute to the diversity of the genus, increasing the difficulty in reconstructing the species relationships based on morphology. Numerous researches investigated the phylogeny of the genus *Rosa*, most of which suggested that the divisions of most subgenera and sections based on morphology were artificial [12,15–17].

The Rosaceae family is represented in Egypt by seven taxa belonging to six genera, namely, *Crataegus azarolus* L., *Crataegus × sinaica* Boiss., *Cotoneaster orbicularis* Schltdl., *Potentilla supina* L., *Rosa arabica* Crép., *Rubus sanctus* Schreb., and *Sanguisorba minor* Scop. Of these taxa, only *Potentilla supina* is ubiquitous, while the others are either of rare or of sporadic occurrence [18–20].

The genus *Rosa* is represented in Egypt by only one species, i.e., *R. arabica* Crép. Its vernacular name is Saint Catherine wild rose, or “Ward Barre”. This species is endemic and is on the red list of the world’s most threatened species; it has a narrow distribution restricted to Mount Catherine, South Sinai, Egypt [21].

*Rosa rubiginosa* is a native rose species to Europe and Northern Asia [22]. *R. arabica* was proposed by Crépín, who classified *R. arabica* as an Asian variety of the European *R. rubiginosa* L. [23]. *R. arabica* and *R. rubiginosa* were treated as closely related on the basis of their morphology, as well as the many floras, monographs, and electronic sites treated *R. arabica* among the synonyms of *R. rubiginosa* [6,8,23,24]. According to The Plant List database (TPL) [25], the name *R. arabica* is reported as an unresolved name, i.e., not established as either accepted or a synonym. The Weeds of Australia website [26] reports both *R. arabica* and *R. eglanteria* as synonyms of *R. rubiginosa*.

Prior to this study, no phylogenetic study of *R. arabica* was undertaken. The phylogenetic analyses of the genus *Rosa* did not include samples of this species [12,15–17].

The origin of endemism is an important evolutionary and taxonomic question. Is *R. arabica* a distinct species? Is it an infraspecific taxon of *R. rubiginosa*? What precisely are its origin, phylogenetic position, and route of evolution? The current study uses classical and cutting-edge taxonomic tools to answer these questions.

## 2. Materials and Methods

### 2.1. Plant Collection

The present study was based on fresh and dried herbarium materials. The fresh materials were collected from Mount Catherine (2629 m), Saint Catherine, South Sinai, Egypt. Herbarium materials were obtained from Egyptian herbaria: Suez Canal University (SCUI, Ismailia, Egypt), Sohag University (SHG, Sohag, Egypt), and Cairo University herbarium (CAI, Cairo, Egypt). The herbarium materials for non-Egyptian species were obtained from the Florida Museum of Natural History, University of Florida Herbarium (FLAS), Gainesville, FL, USA. Herbarium acronyms follow The Index Herbariorum [27].

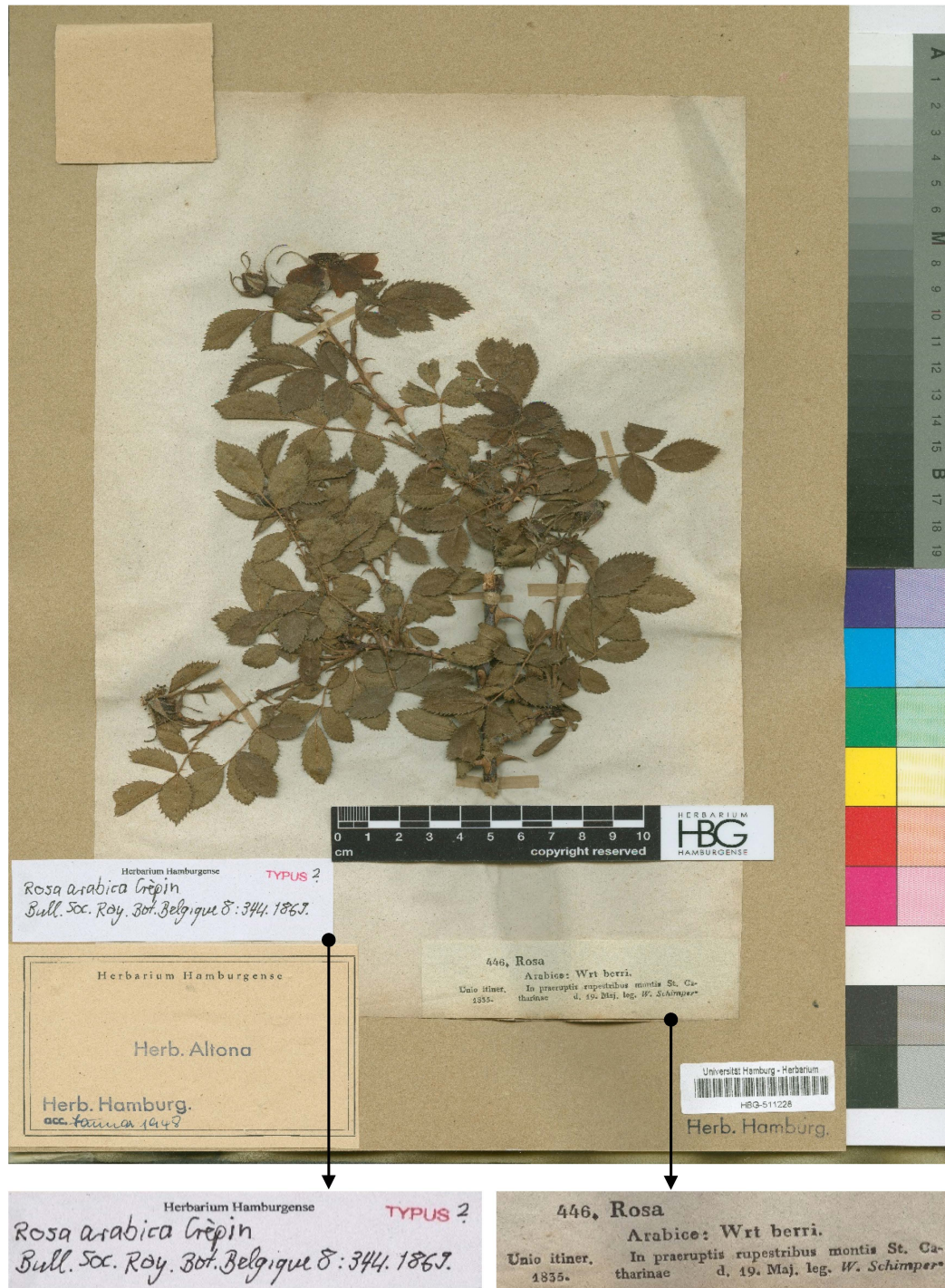
### 2.2. Selected Specimens Examined

The following specimens were examined: *Rosa arabica*, Egypt, Southern Sinai, Wadi El-Arbaeen, Kahf El-Ghola: 28°32′49″ north (N), 33°56′93″ east (E), 1 May 2016, Ahmed EL-Banhawy & Ahmed Elkordy AEB#301, (SCUI s.n.); Wadi El-Arbaeen, Kahf El-Ghola, 18 June 2005, K. Abdelkhalik, (SHG s.n.); Wadi El-Arbaeen, (Encl. 3): 28°32′56.9″ N, 33°56′56.2″ E, Alt. 1650 m, 2 April 2004, K. Shaltout et al., (Herbarium of Elsalam Botanical Garden, Sharm Elsheikh, s.n.); Saint Catherine, April 1940, M. Hassib (CAI); *Rosa eglanteria* L. USA, Pennsylvania: 1.2 mi. E–NE of Pricetown, Berks County, dry soil in the field, W.C. Brumbach, 4283, 15 June 1950 (FLAS); New York: old field on South Hill near Ithaca, Tompkins Co11, October 1940, E.S. Ford #638 (FLAS); North Carolina: Orange County, University of North Carolina campus, cultivated, Chapel Hill, 9 July 1957, Max Newbery 236 (FALS); *Rosa rubiginosa* L. USA, Oregon: Clackamas County, one mile south of Canby, in open pasture, 12 June 1943 R. H. Belton, 1943. (FLAS); Pennsylvania: Berks County, 1.2 mi. S.W. of Weavertown, dry soil along the highway, W. C. Brumbach 3745, 19 June 1944 (FLAS).

### 2.3. Systematic Treatment

#### Identification and Nomenclature

Examined specimens were identified according to the latest available literature, [18–20,22]. An image of several original specimens collected by W. Schimper in 1835 from Saint Catharine, South Sinai, Egypt, was downloaded from the internet repository of the Herbarium Hamburgense (Figure 1).



**Figure 1.** Type of *R. arabica*, collected by W. Schimper from Saint Catharine, South Sinai, Egypt, 19 May 1835, from the original collection, conserved in the Herbarium Hamburgense (HBG 511228). Available at Herbarium Hamburgense. Reproduced with permission.

To assess the traits of the original specimen *R. arabica*, we studied the specimens collected in the *loco classic*. There is no certainty about the type specimen of this species due to the presence of several W. Schimper specimens, and nobody has assigned the type.

A survey of international floras, as well as 10 online databases, was conducted to evaluate the status of *R. arabica* nomenclature (Table 1).

#### 2.4. DNA Extraction and PCR Amplification

Fresh leaf materials used for molecular analyses were collected and preserved in silica gel. DNA was extracted using the cetyltrimethylammonium bromide (CTAB) protocol with some modifications [28]. The PCR amplification was performed in 15 µL volume for ITS -Internal Transcribed Spacer-, *matK*, *rbcL*, and *trnL-F* containing 5 U/µL Taq DNA polymerase with 25 µM MgCl<sub>2</sub>, 10 µM dNTPs, and 10 µM of each primer. Amplifications were conducted using an Applied Biosystems®-Veriti™ 96-well thermal cycler (Thermo Fisher Scientific-Fisher Scientific AS-Postboks 114, Smestad-0309 Oslo–Norway). The thermal cycling program for amplification of the ITS region was as follows: 95 °C for 2 min, 34 cycles of 95 °C for 45 s, 58 °C for 45 s, and 72 °C for 90 s, and a final extension at 72 °C for 5 min; that for the *matK* region was as follows: 95 °C for 3 min, 40 cycles of 94 °C for 30 s, 49 °C for 1 min, and 72 °C for 1 min, and a final extension at 72 °C for 10 min; that for the *rbcL* region was as follows: 95 °C for 6 min, 30 cycles of 95 °C for 45 s, 48 °C for 45 s, and 72 °C for 90 s, and a final extension at 72 °C for 6 min; that for the *trnL-F* region was as follows: 95 °C for 5 min, 15 cycles of 95 °C for 45 s and 60 °C for 1 min, with an extension at 72 °C for 2 min, followed by 20 cycles of 95 °C for 45 s and 54 °C for 1 min, and a final extension at 72 °C for 2 min. The primers used in this study are shown in Table 2.

**Table 1.** Nomenclature status and resolution of *Rosa arabica* according to modern biological databases.

Query	Retrieved Name	Name Resolution	Synonym	Taxonomic Status	Database /Flora
	<i>R. arabica</i> Crép.	Accepted	None	Species	IDigBio
	<i>R. arabica</i> Crép.	Accepted	<i>R. rubiginosa</i>	Species	Open Tree of Life
	<i>R. arabica</i> Crép.	Accepted	None	Species	IPNI
	<i>R. arabica</i> (Crép. ex Boiss.) Déségl.	Accepted	None	Species	PoWo
<i>R. arabica</i>	<i>R. arabica</i> Crép.	Accepted	<i>R. rubiginosa</i> var. <i>arabica</i> (Crépin) Boiss.	Species	Catalog of Life
	<i>R. arabica</i>	Synonym	<i>R. arvensis</i> Huds.	Species	GBIF
	<i>R. rubiginosa</i> L.	Synonym	<i>R. arabica</i> Crép. <i>R. eglantheria</i> L.	Species	Weeds of Australia
	<i>R. arabica</i> (Crép. ex Boiss.) Déségl.	Not recognized	Not recognized	Not recognized	Tropicos
	<i>R. arabica</i> Crép.	Unresolved	None	Unresolved	TPL
	<i>R. arabica</i> (Crép. ex Boiss.) Déségl.	Unresolved	None	Unresolved	
	<i>R. arabica</i> (Crép.)	Ambiguous	None	Ambiguous	WFO

TPL: The Plant List [29]; Tropicos: the online, nomenclatural database of the Missouri Botanical Garden [30]; GBIF: Global Biodiversity Information Facility [31]; IDigBio: Integrated Digitized Biocollections [32]; IPNI: International Plant Name Index [33]; PoWo: Plant of the World Online [34]; Open Tree of Life [35]; Catalog of Life [36]; Weeds of Australia [26]; WFO: World Flora Online [37].

#### 2.5. DNA Sequencing

PCR products were purified with ExoSAP-IT (USB Corporation, Cleveland, OH, USA) according to manufacturer recommendations. PCR products were sent to Macrogen Spain for direct sequencing in both directions with an ABI 3730XL Genetic Analyzer (Life Technologies Corporation, Carlsbad, CA, USA).

These novel DNA sequences of *R. arabica* were deposited in the GenBank under the following accession numbers: ITS, MT358870; *matK*, MT416573; *rbcL*, MT415957; *trnL-F*, MT427590.

**Table 2.** DNA primer sequences used in molecular analysis. F, forward; R, reverse.

Region	Primer F/R	Primer Sequence (5'–3')	Reference
ITS	KRC	GCACGCGCGCTACACTGA	[38]
	AB102	TAGAATTCCCCGGTTCGCTCGCCGTTAC	[39]
<i>matK</i>	K1R	ACCCAGTCCATCTGGAAATCTTGGTTC	[40]
	K3F	CGTACAGTACTTTTGTGTTTACGAG	
<i>rbcL</i>	<i>rbcL</i> a-F	ATGTCACCACAAACAGAGACTAAAGC	[41]
	<i>rbcL</i> a-R	GTAAAATCAAGTCCACCRCG	
	<i>trnL</i> <sup>(UAG)</sup>	CTGCTTCCTAAGAGCAGCGT	
<i>trnT-trnL</i>	<i>trn-b</i>	TCTACCGATTTTCGCCATATC	[42]
	<i>trn-a</i>	CATTACAAATGCGATGCTCT	
<i>trnL</i> (intron)	<i>trn-d</i>	GGGGATAGAGGGACTTGAAC	
	<i>trn-c</i>	CGAAATCGGTAGACGCTACG	

## 2.6. Molecular Data Analysis

### 2.6.1. Molecular Identification

There is no general agreement for a single method that supports species discrimination using DNA sequence data. During this study, molecular identification and phylogenetic analysis were implemented using multiple approaches [43,44].

### 2.6.2. BLAST (Basic Local Alignment Search Tool) and Reference Datasets

A total of four markers (one nuclear and three plastid) were selected for this study; 39 ITS, 26 *matK*, 24 *rbcL*, and 16 *trnL*-F sequences of *Rosa* taxa were used in our study. Verified representative sequences of each taxon were tentatively identified using the BLASTN algorithm available on the NCBI –National Center for Biotechnology Information website. Additional sequences were obtained and included in the datasets. GenBank accession numbers and similarity matching percentage are presented in the Supplementary Materials (Table S1).

A reference sequence dataset was constructed, which consisted of sequences matching 92–99% in sequence similarity [45]. For the newly generated sequences, forward and reverse reads were assembled and edited into contigs in GENEIOUS® v.R9 (Biomatters Ltd., Berkeley, CA, USA, 94709-1405, <https://www.geneious.com>) using a personal license (C.A.). Four data matrices were constructed: ITS, *matK*, *rbcL*, and *trnL*-F. The ingroup was selected to cover most of the major sections in the genus *Rosa*. Datasets of each marker were initially aligned using ClustalW [46] or MAFFT algorithms [47], implemented in Geneious, using default alignment parameters.

### 2.6.3. Tree-Based Analysis

Analyses were run on the CIPRES portal [48]. The aligned DNA sequences for three chloroplast DNA (cpDNA) and one nuclear DNA (nrDNA) were used to construct four single markers and three combined datasets. The optimal nucleotide substitution model was estimated using MrModeltest [49] and executed in MrBayes blocks. Monte Carlo Markov chain (MCMC) was conducted using MrBayes 3.0b4 [50]. Four heated MCMC chains were run over 10 million generations, using general time reversible (GTR) plus gamma distribution substitution rates, random seed trees, and the default starting value for the nucleotide substitution model. Trees were sampled every 1000 generations, resulting in 20,001 trees. The first 25% “burn-in” trees were deleted from the analysis. A 50% majority rule consensus tree was constructed to get the posterior probabilities (PP). For each analysis, two independent runs were executed using initial parameters. Posteriori probabilities >0.5 at a given branch were considered strong support for the existence of that branch [44,51,52].

### 3. Results

#### 3.1. Taxonomic Identity of *Rosa arabica*

Shrubs and shoots were 0.5–1.5 m tall, erect or scrambling, terete, sparsely prickly, usually curved or hooked, with or without a stout base. Branchlets were dark brownish red, 5.0–10 mm in diameter, pubescent or glabrate. Prickles were dark brownish, stout, falcate up to 1.5 cm. Leaves were alternate, odd-pinnately compound, including petioles 0.5–1.2 × 0.2–0.3 cm. Stipules were 4–10 × 2–4 mm, apex acute, margin doubly serrate. Leaflets were 5–7 in number, broadly elliptic to elliptic obovate, 10–30 × 6–18 mm, abaxially green, sparsely pubescent with glandular trichome, adaxially green, glabrous, sparsely pubescent with glandular trichome, or slightly setose along midrib, apex acute, margin doubly glandular–serrate, base rounded. Flowers were solitary, 3.5–4.5 cm in diameter, pedicel ca. 10 mm long, setose–glandular trichome, as long as or longer than the fruit, bracts lanceolate, 8.0–12 × 2.0–5.0 mm, margins stipitate–glandular, glabrous, and eglandular surfaces. The hypanthium was globose, bright red, setose to sparsely glandular trichome. Sepals were five in number, reddish green, broadly lanceolate, 9–21 × 2.5–4.0 mm, abaxially setose–glandular trichome, adaxially densely tomentose, attenuate apex, margin lobed, pinnatifid, recurved, and persistent after anthesis. Petals were five in number, free, pink to deep rose, fragrant, obovate, 9.0–20 × 6.0–10 mm, entire margin, emarginate apex, and cuneate base. Stamens were numerous filaments, 3.0–5.0 mm. Styles were filiform, 4.0–6.0 mm long, exerted, slightly longer than stamens. Fruits were hip, usually ripe with bright red color, globose, 10–21 × 8.0–19 mm, sparsely setose glandular trichome, shiny. Achenes were ovoid to sub elliptic, trigonous, 3.5–4.0 × 2.5–3.0 mm, ventral side roof like with a longitudinal suture, acute apex, and mostly truncate base. Main differences with *R. rubiginosa* are listed in Table 3.

**Table 3.** List of morphological characters for *R. arabica* and *R. rubiginosa*.

Characters	Taxa	
	<i>R. arabica</i> Crép.	<i>R. rubiginosa</i> L.
Habitus	Erect or scrambling	Erect
Life form	Shrub	Shrub
Plant height	0.5–1.5 m	1.5–2 m
Prickles	Stout, falcate up to 1.5 cm	Falcate or curved, sometimes mixed with acicles and glandular setae
Leaflets shape	Broadly elliptic obovate	Suborbicular to broadly ovate to obovate
Leaflet No.	5–7	5–7
Leaflets length	10–30 mm	10–25 mm
Leaflets width	6–18 mm	8–15 mm
Leaflets margin	Doubly serrate	Compound–serrate
Leaflet base	Rounded	Rounded
Leaflet upper surface	Glabrous, sparsely glandular, or slightly setose along midtrip	Glabrous or pubescent
Leaflet lower surface	Sparsely glandular	Pubescent or densely glandular–viscid
Pedicels length	10 mm	10–15 mm
Pedicels surface	Setose–glandular	Densely stipitate or setose–glandular
Sepals nature	Recurved, persistent after anthesis	Erect and persistent after anthesis
Sepals dorsal surface	Setose–glandular	Glandular
Sepals ventral surface	Densely tomentose	Glandular
Hypanthium Shape	Globose	Sub globose, ovoid or ellipsoid
Hypanthium surface	Setose–sparsely glandular	Glabrous or glandular–hispid
Hypanthium color	Bright red	Bright red

Flowering occurred from May to June, while fruiting occurred from June to July.

Its distribution is restricted to Egypt, Southern Sinai, Endemic.

Type: *Rosa arabica* Crépin. Herbarium Hamburgense. Bull. Soc. Roy. Bot. Belgique 8(2):344. 1869. Egypt, South Sinai, Saint Catharine, 1835, W. Schimper s.n. (HBG 511228) [53] (Figure 1).

To determine the current state of the name of *R. arabica*, we surveyed 10 electronic databases (Table 1). The results showed that there is some ambiguity in nomenclature and taxonomic status. The Plant List Database (TPL) retrieved the name *R. arabica* without its related taxonomic or nomenclature status. The World Flora Online (WFO) database considered *R. arabica* as an ambiguous name, while the Catalog of Life database showed *R. arabica* as an accepted name. Moreover, the Weeds of Australia retrieved *R. arabica* as a synonym for *R. rubiginosa*, while the name *R. arabica* is not recognized by the Tropicos database. The International Plant Name Index (IPNI), Plant of the World Online (PoWo), Global Biodiversity Information Facility (GBIF), Integrated Digitized Biocollections (IDigBio), and Open Tree of Life databases accepted the name *R. arabica* as a distinct species.

### 3.2. Molecular Identification Approach

For the first time, DNA sequences of ITS, *matK*, *rbcL*, and *trnL-F* markers were generated and used for molecular identification of *R. arabica*. Generally, the single markers ITS, *matK*, *rbcL*, and *trnL-F* succeeded in identifying the query sequence at the genus and species level and provided support to discriminate the species *R. arabica* (Table S2). The ITS and *matK* markers displayed the highest-level discriminatory power.

### 3.3. Phylogenetic Relationship

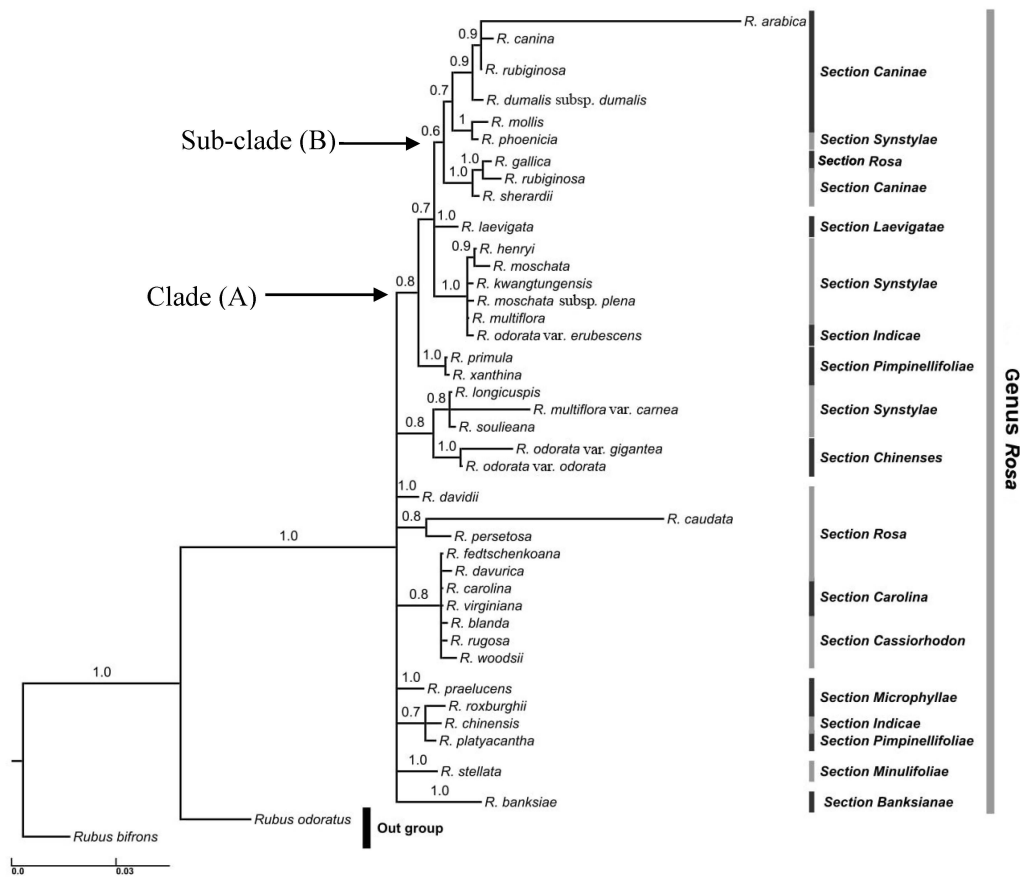
Bayesian inference (BI) of the four single markers ITS, *matK*, *rbcL*, and *trnL-F*, combined datasets ITS + *matK* and ITS + *rbcL*, and a concatenated dataset of all four markers was conducted.

The dataset of the single marker ITS (Figure 2), and the combined ITS + *matK* (Figure 3), retrieved two congruent phylogenetic trees. The ITS phylogenetic tree comprised 39 *Rosa* species, as well as two outgroup taxa: *Rubus bifrons* and *Rubus odoratus*. Despite some polytomies, a moderately resolved clade A consisted of 18 taxa. In clade A, *R. primula* and *R. xanthine* were sister groups. The subclade B “posterior probabilities (PP) = 0.6” comprised nine *Rosa* species including *R. arabica* (Figure 2).

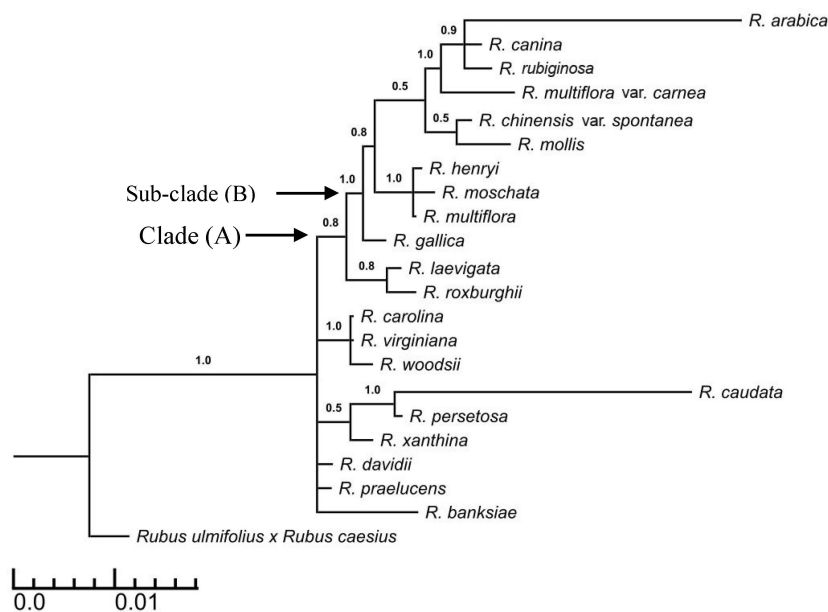
All taxa under investigation belonged to the genus *Rosa*. The current study shows genus *Rosa* subdivided into 12 sections. Clade A includes six sections: section *Caninae* “eight taxa” section *Synstaylae* “six taxa” section *Pimpinellifoliae* “two taxa” and two more sections with a single taxon, section *Rosa*, and section *Laevigatae* (Figure 2).

The phylogenetic tree obtained from the combined dataset comprised 21 *Rosa* species and *Rubus ulmifolius* × *R. caesius* as an outgroup. Clade A was composed of 12 *Rosa* species. *R. laevigata* and *R. roxburghii* were sisters to subclade B. Subclade B (PP = 1) comprised 10 *Rosa* species including *R. arabica*. Most of the main clades were moderately to highly supported (0.7–1) (Figure 3).

Nei’s [54] distance analysis (Table S2) applied to the ITS marker between *R. arabica* and *R. rubiginosa* and *R. canina* was 96.25 and 81.93. In *matK*, the distance between *R. arabica* and *R. canina* and *R. rubiginosa* was 99.70 and 99.79. In *rbcL*, the distance between *R. arabica* and *R. canina* and *R. rubiginosa* was (97.45 and 97.97). In *trnL-F*, the divergence between *R. arabica* and *R. canina* and *R. rubiginosa* was 48.78 and 48.84.



**Figure 2.** Bayesian phylogenetic tree of ITS marker showing the relationships between *R. arabica* and representatives of most *Rosa* sections. Numbers above branches are posterior probability (PP) values >0.6. Accessions numbers of sequences retrieved from GenBank are detailed in Table S1 (Supplementary Materials). Clade A and Sub-Clade B defined in the text.



**Figure 3.** Bayesian phylogenetic tree of combined datasets of ITS + *matk* markers, showing the relationships between *R. arabica* and representatives of most *Rosa* sections and outgroups. Numbers above branches are PP values >0.5. Accessions numbers of sequences retrieved from GenBank are detailed in Table S1 (Supplementary Materials). Clade A and Sub-Clade B defined in the text.



#### 4. Discussion

According to the IUCN – International Union for Conservation of Nature—criteria of plant conservation, the Saint Catherine wild rose *R. arabica* is considered to be one of the most critically endangered plant species in Egypt [21], which justified our interest in its phylogenetic placement.

Tomljenović and Pejić [55] reviewed the taxonomy of the genus *Rosa*. They discussed the efforts to classify and systematize roses from the 16th century until recently. They determined that species discrimination based on morphological characterization was very challenging over the last three centuries due to (1) few described species, (2) a low number of apparent morphological differences between closely related species, and (3) extensive hybridization and polyploidy. Tomljenović and Pejić [55] also suggested that modern tools of classification, such as molecular markers and phylogenetic analyses, as well as traditional morphological methods, would be of great help in clarifying the phylogenetic relationships within the genus *Rosa*.

The inability to retrieve the correct information for a particular plant species nomenclature status using taxonomic or biological databases might arise from the accumulation of outdated information (e.g., TPL has been static since 2013 but was used as the starting point for the Taxonomic Backbone of the WFO). We recommend that plant databases should be curated regularly with the findings of recent taxonomic and ecological studies incorporated. This is especially necessary where the plants concerned are rare and threatened, as is the case with *R. arabica*.

Although there is no certainty about the type specimen of the name *R. arabica* Crép., it was possible to study several specimens cited by [23] that are part of the original material; however, a more detailed and specific study is needed to confirm typification of the name.

*Rosa arabica* has never been the subject of extensive taxonomic or phylogenetic investigation. Even though the identification and nomenclature of *R. arabica* seem to be straight forward, they have represented a challenge even for some experts. Although *R. arabica* and *R. rubiginosa* have been treated as closely related based on their morphology, and many floras and databases treated *R. arabica* as a synonym of *R. rubiginosa* [6–8,23,24], several critical distinctive morphological traits differentiate *R. arabica* and *R. rubiginosa*. The most important diagnostic features for *R. arabica* are stout prickles, falcate up to 1.5 cm (falcate or curved in *R. rubiginosa*), leaflets shaped broadly elliptic to obovate (suborbicular in *R. rubiginosa*), the texture of the leaflet upper surface being glabrous, sparsely glandular, or slightly setose along midrib (glabrous or pubescent in *R. rubiginosa*), and the hypanthium texture being setose to sparsely glandular in *R. arabica* (glabrous or glandular–hispid in *R. rubiginosa*), Table 3.

By applying tree-based analysis, the query sequence was assigned to a species if clustered with sequences from their correct taxon with strong support value [44]. The current study recommends the use of ITS, *matK*, *rbcL*, and *trnL-F* as molecular marker candidates for identification of *R. arabica*.

Sectional classification of the genus *Rosa* was partially retrieved using analyses of a single dataset of DNA sequences of the nuclear marker ITS, as well as a combined dataset of DNA sequences of the chloroplast marker (ITS + *matK*). The genus *Rosa* is composed of 12 sections: *Caninae*, *Synstaylae*, *Rosa*, *Chinensis*, *Laevigatae* *Pimpinellifoliae*, *Carolina*, *Cassiorhodon*, *Microphyllae*, *Indica*, *Minulifoliae*, and *Banksianae* (Figures 2 and 3). Although the sectional classification of the genus *Rosa* is beyond the scope of the current study, the sectional classification is congruent with previous studies [16]. The present study aimed to place *R. arabica* within its related section within the genus *Rosa* (*s.l.*). According to our Bayesian analysis of the DNA sequences of ITS and ITS + *matK*, *R. arabica* was placed within section *Caninae* (Figures 2 and 3). Morphologically, section *Caninae* is characterized by pink or white flowers. *R. arabica* exhibits a range of intermediate morphological characters between *R. canina* and *R. rubiginosa*, especially in terms of the petal color; *R. arabica* shows pale to deep rose petals [56].

Distance analysis of DNA sequences represents another excellent tool to compare closely related species. Ussery [57] recommends using multi-locus data distance analysis instead of a single locus. In the current study, we carried out distance analysis using the ITS, *matK*, *rbcL*, and *trnL-F* data sequences to compare *R. arabica* and its closely related species within the section *Caninae*. The distance analyses

indicated that *R. arabica* is much more closely related to those two species (Table S2). All members of section *Caninae* are distributed in Europe, North America, Asia, and North Africa [6].

According to [58], the Egyptian flora is a mix between South Mediterranean and East Asia. The Sinai Peninsula is the contact point between those two geographical routes. Thus, *R. arabica* might be a hybrid and, after having been isolated in its minimal distribution range, became a new species.

## 5. Conclusions

The result of our taxonomical and molecular study confirmed the identity of *R. arabica*. A better taxonomy and molecular distance comparison of this *Rosa* species group support the distinct identity of the Red-Listed *R. arabica*. The phylogenetic results allow redefining its position in the genus *Rosa* and its close relatives, as well as its accurate identification using the identity of molecular markers. The results show that *R. arabica* is a sister of *R. canina* and a relative of *R. rubiginosa* belonging to section *Caninae*. The use of two molecular markers (ITS and *matK*) can help provide insights into *Rosa* species-level taxonomy, and this may be essential in identifying the correct taxonomy or new species. This will be useful especially where morphology-based identification is difficult. The use of a molecular identification approach for *Rosa* revealed more accurate identification than classical morphologically based taxonomy. A better understanding of the taxonomy and the phylogenetic relationships in *Rosa* is fundamental toward the proposal of conservation strategies. We show that the use of the selected molecular markers represents a powerful tool in cases where correct identification is essential, e.g., for the recognition of critically endangered species that must be protected.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2075-1729/10/12/335/s1>, Table S1. BLAST results for each marker (single locus) of tested samples. Database search match for similarities and phylogenetic relationships of ITS, *matK*, *rbcL*, and *trnL-F* sequences; Table S2. Matrix of pairwise divergences of ITS, *matK*, *rbcL*, and *trnL-F* for all samples studied.

**Author Contributions:** Conceptualization, A.E.-B.; methodology, A.E.-B. and C.A.; software, A.E.-B. and C.A.; formal analysis, A.E.-B., C.A., and A.E.; data curation, C.A. and A.E.-B.; investigation, A.E.-B., S.Q., and C.A.; resources, A.E.-B., S.Q., A.E., and C.A.; writing—original draft preparation, A.E.-B.; writing—review and editing, A.E.-B., C.A., A.E., and S.Q.; funding acquisition, A.E., C.A., and S.Q. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Spanish Ministerio de Educación under the program “Campus de Excelencia Internacional CEI Triangular E<sup>3</sup> (Grants 2017)”, which provided important financial support for the postdoctoral stay of Ahmed Elkordy at the University of León (Spain). The Research Team of the University of León 395 TaCoBi (Taxonomy and Biodiversity Conservation) partially supported the experimental study.

**Acknowledgments:** The authors would like to express thanks to Pamela Soltis and Douglas Soltis, Florida Museum of Natural History and Herbarium (FLAS), University of Florida, Gainesville, FL, USA. The authors would also like to thank the staff and institutions of the Herbaria consulted for helping us with the original materials, particularly Matthias Schultz (HBG), who provided permission to use the picture. They also extend the most profound gratitude to Stephanie Bird, Royal Horticulture Society (RHS), Wisely, UK, John Gaskin, Pest Management Research Unit, US Department of Agriculture, Sidney, Montana, and Felix Llamas, University of León for reading the manuscript and making valuable suggestions.

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

## References

1. Mabberley, D.J. *The Plant-Book: A Portable Dictionary of the Vascular Plants*; Cambridge University Press: Cambridge, UK, 1997.
2. Heywood, V.H.D.; Moore, I.; Richardson, W.T.S. *Flowering Plants of the World*; Oxford University Press: Oxford, UK, 1993.
3. Christenhusz, M.J.; Fay, M.F.; Chase, M.W. *Plants of the World: An Illustrated Encyclopedia of Vascular Plants*; University of Chicago Press: Chicago, IL, USA, 2017.
4. Ku, T.; Robertson, K. *Rosa* (Rosaceae). *Flora China* **2003**, *9*, 339–381.
5. Bisby, F.A.; Roskov, Y.; Orrell, T.M.; Nicolson, D.; Paglinawan, L.E.; Bailly, N.; Kirk, P.M.; Bourgoin, T.; Baillargeon, G.; Ouvrard, D. *Species 2000 and ITIS Catalogue of Life*; 2020-09-01 Beta; Species 2000: Leiden, The Netherlands, 2020.

6. Wissemann, V. Conventional taxonomy of wild roses. S. 111–117. In *Encyclopedia of Rose Science*; Roberts, A., Debener, T., Gudin, S., Eds.; Academic Press: London, UK, 2003.
7. Focke, W. Rosaceae en Engler u. Prantl. *Die Nat. Pflanz.* **1888**, *3*, 53.
8. Rehder, A.; Dudley, T.R. *Manual of Cultivated Trees and Shrubs Hardy in North America: Exclusive of the Subtropical and Warmer Temperate Regions*; Macmillan: New York, NY, USA, 1940.
9. Cockerell, T. *Rosa stellata*. *Nature* **1913**, *90*, 571. [[CrossRef](#)]
10. Bean, W.J. *Trees and Shrubs Hardy in the British Isles*; John Murray: London, UK, 1973; Volume 1.
11. Wissemann, V.; Ritz, C.M. The genus *Rosa* (Rosoideae, Rosaceae) revisited: Molecular analysis of nrITS-1 and atp B-rbc L intergenic spacer (IGS) versus conventional taxonomy. *Bot. J. Linn. Soc.* **2005**, *147*, 275–290. [[CrossRef](#)]
12. Fougère-Danezan, M.; Joly, S.; Bruneau, A.; Gao, X.-F.; Zhang, L. Phylogeny and biogeography of wild roses with specific attention to polyploids. *Ann. Bot.* **2014**, *115*, 275–291. [[CrossRef](#)]
13. Zhu, Z.-M.; Gao, X.-F.; Fougère-Danezan, M. Phylogeny of *Rosa* sections *Chinenses* and *Synstylae* (Rosaceae) based on chloroplast and nuclear markers. *Mol. Phylogen. Evol.* **2015**, *87*, 50–64. [[CrossRef](#)]
14. Liu, C.; Wang, G.; Wang, H.; Xia, T.; Zhang, S.; Wang, Q.; Fang, Y. Phylogenetic relationships in the genus *Rosa* revisited based on rpl16, trnL-F, and atpB-rbcL sequences. *Hortscience* **2015**, *50*, 1618–1624. [[CrossRef](#)]
15. Erlanson, E.W. Phylogeny and polyploidy in *Rosa*. *New Phytol.* **1938**, *37*, 72–81. [[CrossRef](#)]
16. Bruneau, A.; Starr, J.R.; Joly, S. Phylogenetic relationships in the Genus *Rosa*: New evidence from chloroplast DNA sequences and an appraisal of current knowledge. *Syst. Bot.* **2007**, *32*, 366–378. [[CrossRef](#)]
17. Koopman, W.J.M.; Wissemann, V.; De Cock, K.; Van Huylbroeck, J.; De Riek, J.; Sabatino, G.J.H.; Visser, D.; Vosman, B.; Ritz, C.M.; Maes, B.; et al. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). *Am. J. Bot.* **2008**, *95*, 353–366. [[CrossRef](#)]
18. Tackholm, V.; Drar, M. *Students' Flora of Egypt*; Cairo University: Cairo, Egypt, 1956.
19. El-Hadidi, M. Materials from excursion flora of Egypt (EFE). *Taeckholmia* **1995**, *15*, 74–75.
20. Boulos, L. *Flora of Egypt*; Al Hadara Publishing: Cairo, Egypt, 1999; Volume 1.
21. Omar, K. *Rosa Arabica*; The IUCN Red List of Threatened Species: London, UK, 2017. [[CrossRef](#)]
22. Klååtersky, I. *Rosa, L.* In *Flora Europaea*; Tutin, T., Heywood, V., Burges, N., Moore, D., Valentine, D., Walters, S., Eds.; Cambridge University Press: Cambridge, UK, 1981; pp. 25–32.
23. Crépin, F. *Primitiae Monographiae Rosarum: Matériaux Pour Servir à L'histoire des Roses*; Imprimerie C. Annoot-Braeckman: Ghent, Belgium, 1869.
24. Qiu, X.; Zhang, H.; Wang, Q.; Jian, H.; Yan, H.; Zhang, T.; Wang, J.; Tang, K. Phylogenetic relationships of wild roses in China based on nrDNA and matK data. *Sci. Hortic.* **2012**, *140*, 45–51. [[CrossRef](#)]
25. Kalwij, J.M. Review of 'The Plant List, a working list of all plant species'. *J. Veg. Sci.* **2012**, *23*, 998–1002. [[CrossRef](#)]
26. De la Cruz, R. The usefulness of weed diversity in slash/mulch bean production: Difficulties in herbicide use. In *Tapado Slash/Mulch: How Farmers Use It and What Researchers Know about It*; Cornell International Institute for Food, Agriculture and Development: Brighton, UK, 1994; pp. 233–237.
27. Thiers, B. Continuously Updated: Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff. New York Botanical Garden's Virtual Herbarium. Available online: <http://sweetgum.nybg.org/science/ih/> (accessed on 7 May 2020).
28. Doyle, J. DNA protocols for plants. In *Molecular Techniques in Taxonomy*; Springer: Berlin, Germany, 1991; pp. 283–293.
29. The Plant List. 2013. Available online: <http://www.theplantlist.org/> (accessed on 9 April 2020).
30. The Online, Nomenclatural Database of the Missouri Botanical Garden. 2018. Available online: <http://www.tropicos.org> (accessed on 12 April 2020).
31. Abderabbi, K.A.; Adda, H.; Benhassaini, O.M. Leaf morphological and anatomical traits variation of *Artemisia herba-alba* in a steppe zone of Algeria. *Bulg. J. Agric. Sci.* **2018**, *24*, 631–637.
32. Integrated Digitized Biocollections. Available online: <https://www.idigbio.org/portal/records/3f406bb6-d54d-443b-92ce-016166dace53#citation> (accessed on 12 February 2020).
33. Khan, A.S.; Ali, S.; Khan, I.A. Morphological and molecular characterization and evaluation of mango germplasm: An overview. *Sci. Hortic.* **2015**, *194*, 353–366. [[CrossRef](#)]

34. Plant of the World Online. Available online: <http://powo.science.kew.org/taxon/731621-1> (accessed on 16 February 2020).
35. De Azeredo, R.M.A.; Mendes, M.A.; Joko, C.Y.; Delgado, M.N. Effect of expansion time and sunlight radiation on the functional and anatomical traits of mango tree leaves. *Rev. Agrogeambiental* **2018**, *9*, 4. [[CrossRef](#)]
36. Catalogue of Life. Available online: <http://www.catalogueoflife.org/col/search/all/key/Rosa+arabica/fossil/1/match/1> (accessed on 16 February 2020).
37. WFO. World Flora Online. 2020. Available online: <http://www.worldfloraonline.org/> (accessed on 8 December 2020).
38. Torrecilla, P.; Catalán, P. Phylogeny of broad-leaved and fine-leaved *Festuca* lineages (Poaceae) based on nuclear ITS sequences. *Syst. Bot.* **2002**, *27*, 241–251.
39. Douzery, E.J.P.; Pridgeon, A.M.; Kores, P.; Linder, H.P.; Kurzweil, H.; Chase, M.W. Molecular phylogenetics of Desea (Orchidaceae): A contribution from nuclear ribosomal ITS sequences. *Am. J. Bot.* **1999**, *86*, 887–899. [[CrossRef](#)] [[PubMed](#)]
40. Lee, H.L.; Yi, D.K.; Kim, J.S.; Kim, K.J. Development of plant DNA barcoding markers from the variable noncoding regions of chloroplast genome. In Proceedings of the International Barcode of Life Conference, Taipei, Taiwan, 16–21 September 2007.
41. Wong, K.-L.; But, P.P.-H.; Shaw, P.-C. Evaluation of seven DNA barcodes for differentiating closely related medicinal *Gentiana* species and their adulterants. *Chin. Med.* **2013**, *8*, 16. [[CrossRef](#)] [[PubMed](#)]
42. Taberlet, P.; Gielly, L.; Pautou, G.; Bouvet, J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **1991**, *17*, 1105–1109. [[CrossRef](#)]
43. Gong, W.; Liu, Y.; Chen, J.; Hong, Y.; Kong, H. DNA barcodes identify Chinese medicinal plants and detect geographical patterns of *Sinosenecio* (Asteraceae). *J. Syst. Evol.* **2015**, *54*, 83–91. [[CrossRef](#)]
44. Mao, Y.-R.; Zhang, Y.-H.; Nakamura, K.; Guan, B.-C.; Qiu, Y.-X. Developing DNA barcodes for species identification in Podophylloideae (Berberidaceae). *J. Syst. Evol.* **2014**, *52*, 487–499. [[CrossRef](#)]
45. Larranaga, N.; Hormaza, J.I. DNA barcoding of perennial fruit tree species of agronomic interest in the genus *Annona* (Annonaceae). *Front. Plant Sci.* **2015**, *6*, 589. [[CrossRef](#)] [[PubMed](#)]
46. Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Multiple sequence alignment using ClustalW and ClustalX. *Curr. Prot. Bioinform.* **2003**, *1*, 1–22. [[CrossRef](#)]
47. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
48. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES science gateway for inference of large phylogenetic trees. In Proceedings of the 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010.
49. Posada, D.; Crandall, K.A. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818. [[CrossRef](#)]
50. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [[CrossRef](#)]
51. Elias, M.; Joron, M.; Willmott, K.; Kaiser, V.; Silva-Brandão, K.; Freitas, A.; Mejia, C.A.; Pineres, L.; Brower, A.; Jiggins, C. Phylogenetic hypothesis, pattern of speciation and evolution of wing pattern in neotropical *Napeogenes* butterflies (Lepidoptera: Nymphalidae). *J. Insect Sci.* **2007**, *7*, 13–14.
52. El-Banhawy, A.; Al-Juhani, W. DNA barcoding and phylogeny of *Phlomis aurea* (Lamiaceae) endemic to Sinai Peninsula, Egypt. *Pak. J. Bot.* **2019**, *51*, 1263–1271. [[CrossRef](#)]
53. Herbarium Hamburgense. Available online: [http://www.herbariumhamburgense.de/herbarsheets/disk\\_batch04/medium/HBG-511228.jpg](http://www.herbariumhamburgense.de/herbarsheets/disk_batch04/medium/HBG-511228.jpg) (accessed on 10 May 2020).
54. Nei, M.; Tajima, F.; Tateno, Y. Accuracy of estimated phylogenetic trees from molecular data. *J. Mol. Evol.* **1983**, *19*, 153–170. [[CrossRef](#)]
55. Tomljenovic, N.; Pejić, I. Taxonomic review of the genus *Rosa*. *Agric. Consp. Sci.* **2018**, *83*, 139–147.
56. Shams, E.; Sadek, A.; Hosni, H. Morphological and anatomical characteristics of endemic *Rosa arabica* (Rosaceae) from Sinai, Egypt. *Taekholmia* **2019**, *39*, 34–43. [[CrossRef](#)]
57. Ussery, D. *Encyclopedia of Genetics*; Academic Press: New York, NY, USA, 2001.

58. Faried, A.; El-Banhawy, A.; Elqahtani, M. Taxonomic, DNA barcoding and phylogenetic reassessment of the Egyptian *Ephedra*, L. (Ephedraceae). *Int. J. Environ. Sci.* **2018**, *17*, 1–13.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).