PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 10(2): 302–309
https://doi.org/10.14719/pst.2153

HORIZON
e-Publishing Group

RESEARCH ARTICLE

Molecular phylogenetic analysis of *Tulipa* (*Liliaceae*) from Aksu-Zhabagly Nature Reserve

Arailym Sarsen[1], Malika Saginova[1], Zhiger Akishev[1,2], Saniya Aktayeva[1,2], Shuga Manabayeva[1] & Bekbolat Khassenov[1,2*]

[1]National Center for Biotechnology, Astana, 010000, Kazakhstan
[2]S. Seifullin Kazakh Agro Technical University, Astana 010001, Kazakhstan

*Email: khassenov@biocenter.kz

OPEN ACCESS

ARTICLE HISTORY

Received: 27 September 2022
Accepted: 02 February 2023

Available online
Version 1.0 : 19 March 2023
Version 2.0 : 01 April 2023

Check for updates

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS,UGC Care etc. See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE

Sarsen A, Saginova M, Akishev Z, Aktayeva S, Manabayeva S, Khassenov B. Molecular phylogenetic analysis of *Tulipa* (Liliaceae) from Aksu-Zhabagly Nature Reserve. Plant Science Today. 2023; 10(2): 302–309. https://doi.org/10.14719/pst.2153

Abstract

Barcodes are conserved sequences of genomic, plastid and mitochondrial DNA that can be utilized to uniquely identify an unidentified specimen to its species when conventional identification methods are inapplicable. Among prokaryotic and eukaryotic species, nuclear ribosomal internal transcribed spacer (ITS) sections are one of the most often utilized DNA markers in DNA barcoding and phylogenetic research. In addition to the ribosomal genes, the plastid genes are the most suitable for identifying plant species. The Aksu-Zhabagly Nature Reserve is the oldest nature reserve in Central Asia and is home to 1,312 vascular plant species, 44 of which are categorized as threatened or endangered in Kazakhstan's red data book. In this study, a collection of specimens of uncommon tulip species was compiled, along with their morphological identification and DNA barcoding. The ITS region and parts of the *matK* and *ycf1b* genes of tulip plastid DNA were sequenced. The evolutionary link between species of tulips was investigated. Phylogenetic study predicted 2 *Tulipa* subclades. *Tulipa* species have substantially preserved *MatK* genes. Tulips' *ycf1b* gene has evolved more slowly than other Liliaceae family members. Nuclear and plastid DNA sequences investigated *Tulipa* species evolutionary relationships. The findings about the ITS region of nuclear DNA were more definite. Overall, our work shows that genetic data will be important in determining species concepts in this genus, however, even with a molecular perspective pulling apart closely related taxa can be extremely challenging.

Keywords

Tulipa, phylogeny, ITS, *matK*, *ycf1*

Introduction

The correct identification of species and the assessment of their phylogenetic relationships have been challenges throughout the development of biological research. Tulips, which belong to the genus *Tulipa*, are perennial herbaceous bulbiferous geophytes that bloom in the spring. The blossoms are typically red, pink, yellow or white and they are enormous, elaborate and highly colored. Over time, the number of species in the genus *Tulipa* fluctuated between 81 species (1) and 100 species (2). Despite the fact that the study of *Tulipa*'s taxonomic characteristics began in the 18th century, the classification of tulips has been amended multiple times. The vast majority of studies concentrated on morphological or cytogenetic characteristics. The taxonomy of *Tulipa* was determined through multivariate analysis of 34 morphological characteristics (2). To get around

the "conventional" identification problems, scientists have begun to use molecular character comparisons as an alternative to morphological identification (3, 4). In the past few decades, DNA profiling techniques based on the PCR method have emerged as a powerful tool in plant systematics and have become a substantial, cost-effective and reliable technique for phylogenetic research (5). Inter simple sequence repeat (ISSR) profiling was performed to study *Tulipa* from Iran (6), while the Amplified fragment length polymorphism (AFLP) technique was utilized to examine *Tulipa orphanidea* L. from Turkey (7). DNA profiling techniques are especially effective for intraspecific research on the molecular genetic diversity of species. Since there is no way to compare DNA profiles between species due to the lack of common amplicons, these methods are not employed for species identification. On the other hand, in families of closely related species, it is feasible to detect shared amplicons that can be used to distinguish between members of the same family. Therefore, one common drawback of DNA profiling methods is that they aren't ubiquitous. Barcodes are regions of DNA in chromosomes, plastids, or mitochondria that are conserved from one individual to the next (8). DNA barcoding technique allows for the assessment of evolutionary relationships between species from a very small sample as well as the identification of species at various developmental stages. DNA barcoding, which employs molecular markers that may be used repeatedly for DNA-based identification, has emerged as a useful technique, particularly for those who lack distinguishing physical characteristics (10, 11). Research into developing universal DNA barcoding markers for land plants is ongoing. DNA barcoding research has used markers for over 20 loci/genes (reviewed in CBOL Plant Working Group, 2009) (12). DNA barcoding loci should be chosen based on several factors, including primer universality and sequence variation. Commonly proposed fern areas include matK, rbcL, trnH-psbA, and trnL-F (13-16). It is difficult to find a suitable DNA locus for DNA barcoding in closely related species. Extensive use of the *trnL* intron and *trnL-trnF* intergenic spacer in chloroplasts has led to their widespread application in the study of intraspecific and interspecific evolutionary relationships (17). Among the most popular regions for phylogenetic analysis across closely related genera and subgenera, the nuclear ribosomal internal transcribed spacer (ITS) is present in a wide variety of organisms (18, 19). Turktas et al. examined the evolutionary relationships of *Tulipa* in Turkey using DNA sequences from the *trnL-trnF* and ITS regions (20). Central Asia's first nature reserve can be found in the Aksu-Zhabagly region. It is situated in the southern part of the Republic of Kazakhstan and has an area of 57,774 hectares. In 2015, UNESCO added Aksu-Zhabagly Nature Reserve (AZNR) to the list of World Heritage Sites. (https://en.unesco.org/biosphere/aspac/aksu-zhabagly). The Aksu-Zhabagly Nature Reserve is home to seventeen tulip species that are critically endangered elsewhere but are protected by law in the Republic of Kazakhstan according to their incorporation in the red data book.

The present study was undertaken to determine the genetic diversity and to establish relationships between different accessions of rare tulip species found in Aksu-Zhabagly Nature Reserve. Barcode identification was accomplished by sequencing the ITS region of nuclear DNA and the *matK* and *ycf1b* genes of plastid DNA that were isolated from tulip samples. Barcodes are used on the tulips, and the molecular identity of each specimen is checked by an expert. Phylogenetic relationships among tulip species were investigated by analyzing plastid and nuclear DNA sequences. The aim of this research was to enhance our knowledge of species concepts among all of Kazakhstan's wild-growing *Tulipa* species in order to better inform tulip conservation efforts, our knowledge of tulip evolution and the broader taxonomic placement of Kazakh tulip species.

Materials and Methods

Plant materials

In July 2021, tulip specimens were collected from the areas of their natural growth in the Aksu-Zhabagly Nature Reserve: *T. greigii*, *T. gréigii* var. Red-Yellow, *T. kaufmanniana*, Hybrid of *T. kaufmanniana*/*T. greigii*, *T. kaufmanniana*, *T. turkestanica*, *T. bifloriformis*, the coordinates of which are presented in Table 1. The coordinates (latitude and longitude) and absolute altitude of the site of each sample were determined. Fig. 1 illustrates the locations of the tulip samples. The leaves of tulip specimens were dehydrated in silica gel and then frozen at -80 degrees Celsius. The primary specimens for

Table 1. Information about tulip species collected in Aksu-Zhabagly Nature Reserve, analyzed by ITS, *matK*, *ycf1b* sequencing with GenBank accession numbers

Plant	Coordinates	Altitude, m	ITS accession number	*matK* accession number	*Ycf1b* accession number
Tulipa bifloriformis	42°23’34’’N70°37’13’’E	1960	ON870406	ON982483	ON885953
Tulipa greigii	42°2’39’’N70°25’29’’E	1830	ON870407	ON982484	ON885954
Tulipa turkestanica	42°26’50’’N70°22’47’’E	1340	ON870408	ON982485	ON885955
Tulipa greigii var. Red-Yellow	42°20’39’’N70°26’42’’E	1830	ON870409	ON982486	ON885956
Tulipa kaufmanniana	42°20’51’’N70°29’9’’E	1980	ON870410	ON982487	ON885957
Tulipa kaufmanniana	42°20’51’’N70°28’12’’E	2060	ON870411	ON982488	ON885958
Hybrid of Tulipa kaufmanniana/Tulipa greigii	42°20’49’’N70°28’6’’E	2010	ON870412	ON982489	ON885959

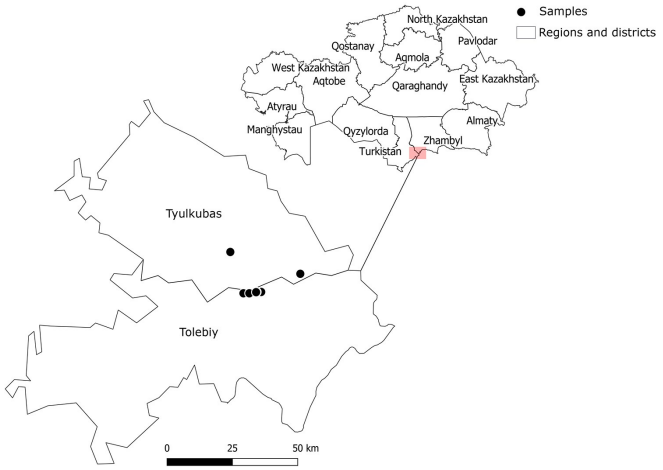

Fig. 1. Locations of sample collection of tulip specimen locations in the Aksu-Zhabagly Nature Reserve.

each species of tulip gathered are listed in Table 1.

DNA isolation, PCR and Sanger sequencing

Total DNA was extracted from frozen leaves of tulip specimens by the CTAB method (21). The tissue sample (150 mg) was collected in a glass ball-containing 2 mL Eppendorf Safe-Lock microcentrifuge tube. The samples were placed in the TissueLyser II adapters and powdered for 5 minutes by shaking at 30 Hz. To each tube, add the pre-warmed (65 °C) extraction buffer (1.4 M NaCl, 25 mM EDTA, 2% CTAB 100 mM Tris-HCl, pH 5.0), fill to 1 mL, vortex to thoroughly mix, and incubate the samples for 1 hr at 65 °C. An equal volume of isoamyl alcohol/chloroform mixture was added and vortexed to thoroughly mix before centrifuging the tubes at maximum speed for 5 min. Transferred the upper aqueous layer to a new 2 mL microcentrifuge tube containing 600 µl 2-propanol, vortexed well and centrifuged the tubes at maximum speed for 2 min in a microcentrifuge. Supernatant was removed and the pellet washed with 1.8 mL 70% EtOH, by gently inversions. Centrifuge at high speed for 2 min, then discard the ethanol. At 55°C for 10-30 min, the DNA pellet was dissolved immediately in 300 µl of 1xTE (1 mM EDTA, 10 mM Tris-HCl, pH 8.0) with RNAse A. The DNA quality was determined using agarose electrophoresis. The concentration of DNA was determined using a NanoVue plus spectrophotometer (General Electric).

The list of the primer pairs used in this study is shown in Table 2. The pair of primers ITS_U1/ITS_U4 was used for amplification of the nuclear ITS region. The pair of primers ITS_U1/ITS_U2 and ITS_U3/ITS_U4 were used only for sequencing of the originally amplified ITS region, respectively. Primers MatK472F/MatK1248R and Ycf1bF/Ycf1bR were used for amplification and sequencing of the maturase K gene and Ycf1 gene regions from plastid DNA respectively.

PCR amplification was carried out in a Bio-Rad Thermal Cycler T100 under the following conditions: initial denaturation step at 95 °C for 5 min, followed by 30 amplifications at 95 °C for 45 s, at 55 °C for 30 s and 72 °C for 30 s, followed by a final extension of 72 °C for 3 min.

For the reaction volume at 25 µL contain: 100 ng DNA template, 1x Taq-reaction buffer (with 1.5 mM MgCl2), 0.2 mM of each dNTP, 0.2 µM of each primer and 1 U of Taq DNA Polymerase. PCR products were analyzed by the gel electrophoresis with 1x TAE buffer and cleaned by QIAquick PCR Purification Kit (Qiagen). Purified PCR products were sequenced in both directions using the primers employed for amplification.

Sequencing reactions were carried out using a Big Dye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems, USA). DNA sequencing was performed with Applied Biosystems ABI 3730xl 96-capillary DNA analyzer (Applied Biosystems, USA). Generated sequences were submitted to the National Center for Biotechnology Information (NCBI). GeneBank accession numbers are provided in Table 1.

Tulip species identification and phylogenetic analysis

Chromatograms after sequencing were analyzed and edited with SnapGene Viewer 6.1 software. Edited sequences were analyzed by BLAST searches (22) for preliminary analysis. Multiple sequence alignment (Fig. 1) was carried out by the Mega 11 program (23). Sequences for ITS region, *matK* and *ysf1* genes for *T. lehmanniana, T. clusiana, T. chrysantha, T. turkestanica, T. tarda, T. sprengeri, T. cretica, T. uniflora, T. iliensis, T. silvestris, T. humilis, Lilium sulphureum, L. pumilum* were retrieved from GenBank.

Results and Discussion

Sequencing and analysis

The sequences of internal transcribed spacers (ITS) of ribosomal genes are the best-known sequences for phylogenetic studies of eukaryotes and prokaryotes. The ITS gene is found between the structure genes of ribosomal RNA in eukaryotes: 18S, 5.8S, and 26S. The ribosomal genes are a single group of nuclear genes that are organized in tandem repeats. The ribosomal gene cluster is made up of a transcribed region (18S, 5.8S, and 26S rRNA), internal transcribed spacers (ITS1 and ITS2), and external transcribed spacers (ETS1 and ETS2). The variable portion of the ITS region is species-specific and differs even in closely related organisms; hence, it is utilized in phylogenetic analyses (19, 25). The ITS region is commonly employed as a phylogenetic identifier to classify plants by genus, species and subspecies (26, 27). ITS sequences are located in the ribosomal cluster of the nuclear genome and are present in all living species. The restriction of the ITS region to conserved areas (18S and 26S rRNA) permits the use of universal primers (25). The ITS region possesses the desired diversity, making it simple to differentiate between closely related organisms. This is because the nuclear genome accumulates synonymous substitutions at nearly the same rate, whereas the mitochondrial and chloroplast genomes accumulate synonymous substitutions at different rates (28). The ITS region has a high copy number, up to 30000 copies per nuclear (29) and is adequate for rapid PCR analysis, cloning and sequencing (500-700 bp for flowering

plants) (25). Due to the biparental inheritance of ITS, hybrids can be distinguished through the localization of the ITS region in the nuclear genome (19). ITS_U1 and ITS_U4 primers (Table 2) were chosen for amplifying ITS region of DNA *Tulipa* samples and the 2 pairs ITS_U1/ ITS_U2 and ITS_U3/ITS_U4 (Table 2) were used for sequencing and identification ITS1 and ITS2 respectively.

Fig. 2A, 3A illustrate the multiple sequence alignment of a segment of the ITS region obtained from Tulipa species in the Aksu-Zhabagly Nature Reserve. The lengths of the ITS sequences ranged from 592 bp to 745 bp, with the alignments measuring 583 bp. The G+C concentration of the ITS sections exhibited less variance between samples and was, on average, 59%. Approximately 91% of the eukaryotes in the ITS could be identified. In the ITS region, tulip specimens demonstrate an average of 99 % conservation. One of the significant disadvantages of the ITS marker is that the ITS region is characterized by a high amount of homoplasy, which in some circumstances decreases the discriminating capacity of ITS sequences for phylogenetic investigations (30). To remedy this deficiency, plastid DNA genes with superior discriminatory ability were employed (31). The *matK* gene, which codes for maturase K on plastid DNA, is found even in non-chlorophyll plants (32). The chloroplast maturase K gene (*matK*) is one of the most variable angiosperm coding genes and has been proposed as a "barcode" for terrestrial plants. The *matK* gene has a length of around 1570 base pairs and encodes a maturase protein. The coding region of *matK* is normally found within an intron of the chloroplast *trnK* gene, with the exception of certain ferns in which it encodes tRNALys (UUU) (33). Being a coding area, the extremely rapid evolution of *matK* has made it useful in phylogenetic reconstructions at high taxonomic levels, such as Order or Family, and occasionally at low taxonomic levels, such as Genus or Species (34-37). Although considerable divergence in the *matK* sequence for higher taxonomic categories results in unclear locations and relationships for some evolutionary clades, *matK* is extremely beneficial for the identification of plant families using DNA barcoding (38-40). In systematics, primers designed for the *trnK* region are used to amplify and sequence the complete *matK* sequence (41, 42). However, for DNA barcoding, a segment of 600-800 bp is usually sufficient for species identification. Jing YU et al. proposed restricting themselves to a 600-800 bp region of the maturase K gene, for which universal primers can be chosen. MatK472F and MatK1248R primers (Table 2) were chosen to identify *Tulipa* plants and the 800 bp fragment of maturase K was amplified and sequenced.

Fig. 2B,3B depicts the multiple sequence alignment of the *matK* gene for *Tulipa* species. *Tulipa* samples varied in the average number of alignment characters, had *matK* sequence lengths ranging from 486 bp to 743 bp, with an average length of 356 bp. *matK* sequences included an average of 32% G+C. The identity of *matK* was approximately 99.7 % certain, and the conservation rate was one hundred percent.

The sequence length of *ycf1b* varied from 611 to 796 base pairs. The average number of alignment characters for *ycf1b* in *Tulipa* samples was 622 base pairs. The average G+C content of *ycf1b* was 29%. The identification of *ycf1b* was approximately 96% certain, and its conservation was 100%.

The multiple sequence alignment of the *ycf1b* gene is depicted (Fig. 2C, 3C). The plastid gene *ycf1* encodes a protein consisting of roughly 1800 amino acids. Recent investigations demonstrated that *ycf1* is necessary for plant viability and encodes Tic214, an important component of the *Arabidopsis* TIC complex (44). *ycf1* spans the short single copy (SSC) and inverted repeat (IR) portions of the plastid genome. The IR region of *ycf1* is short (less than one kilobase in length) and preserved. In contrast, the sequence variability of the *ycf1* SSC region is substantial in seed plants. This portion of the *ycf1* gene is more variable than *matK* in the majority of taxa studied to date and has been utilized in molecular systematics at low taxonomic levels (47-49). Two areas inside *ycf1*, *ycf1a* and *ycf1b* are anticipated to exhibit the highest species-level nucleotide diversity among angiosperm plastid genomes (50). Because *ycf1* is too long (5709 bp for *Nicotiana tabacum*) and too variable to permit the design of universal primers, it has not received much attention for DNA barcoding or molecular systematic purposes at low taxonomic levels; however, the high variability of *ycf1b* indicates its potential value in DNA barcoding of land plants (51). To identify tulip specimens, Ycf1bF and Ycf1bR

Table 2. The sequence of primers used for PCR amplification and sequencing of nuclear ribosomal internal transcribed spacer (ITS) and plastid *matK*, *ycf1b* regions

Target	Primer ID	Sequence (5'-3')	Tm, ° C*	Reference
Nuclear region	ITS_U1	GGAAGTAGAAGTCGTAACAAGG	58.5	(56)
	ITS_U2	GCGTTCAAAGATTCGATGATTC	58.2	(56)
	ITS_U3	CATCGATGAAGAACGCAGC	59.7	(56)
	ITS_U4	GGTTTCTTTTCCTCCGCTTA	58.1	(56)
Plastid region	MatK472F	CCCRTYCATCTGGAAATCTTGGTTC	63.9	(43)
	MatK1248R	GCTRTRATAATGAGAAAGATTTCTGC	58.2	(43)
	Ycf1bF	TCTCGACGAAAATCAGATTGTTGTGAAT	63.2	(51)
	Ycf1bR	ATACATGTCAAAGTGATGGAAAA	56.1	(51)

Fig. 2. Multiple sequence alignment of a fragment of ITS region (**A**), *matK* (**B**), and *ycf1b* (**C**) for *Tulipa* species were collected in the Aksu-Zhabagly Nature Reserve.

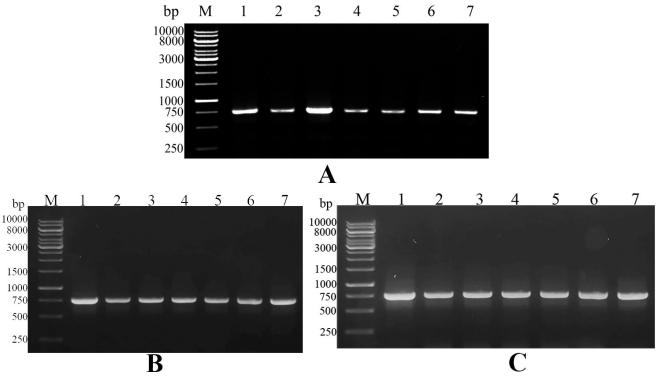

Fig. 3. PCR amplification of fragment of ITS region (A), matK (B) and ycf1b (C) for *Tulipa* samples.

primers (Table 2) were used, and the 800 bp *ycf1b* fragment was amplified and sequenced.

Phylogenetic analysis

The ITS region of tulip species was compared to that of *Lilium sulphureum* (obtained from GenBank) to serve as the outgroup for the Liliaceae family. All of the phylogenetic trees (individual ITS, *matK* and *ycf1b*) were constructed using Neighbor-joining (NJ) and Maximum-parsimony (MP) approaches. Phylogenetic analysis anticipated that the clade of studied *Tulipa* would split in 2 and the results are shown in Fig. 4a (NJ) and 4b (MP). The first clade consists of the three subclades of *Tulipa* species. *T. greigii* is in the first group, followed by *T. kaufmanniana*, and finally a hybrid (*T. kaufmanniana* x *T. greigii*) and *T. greigii* var. Red-Yellow in the third. *T. lehtmanniana*, *T. clussiana* and *T. chrysantha* from the NCBI GenBank are most closely related to the Kazakh plants *T. kaufmanniana*, *T. kaufmanniana*/*T. greigii* and *T. greigii* var. Red-Yellow, as determined by the ITS sequence. However, the specimen of *T. greigii* is excluded from this group of related plants.

Additionally, 3 sequences of *Tulipa* species and

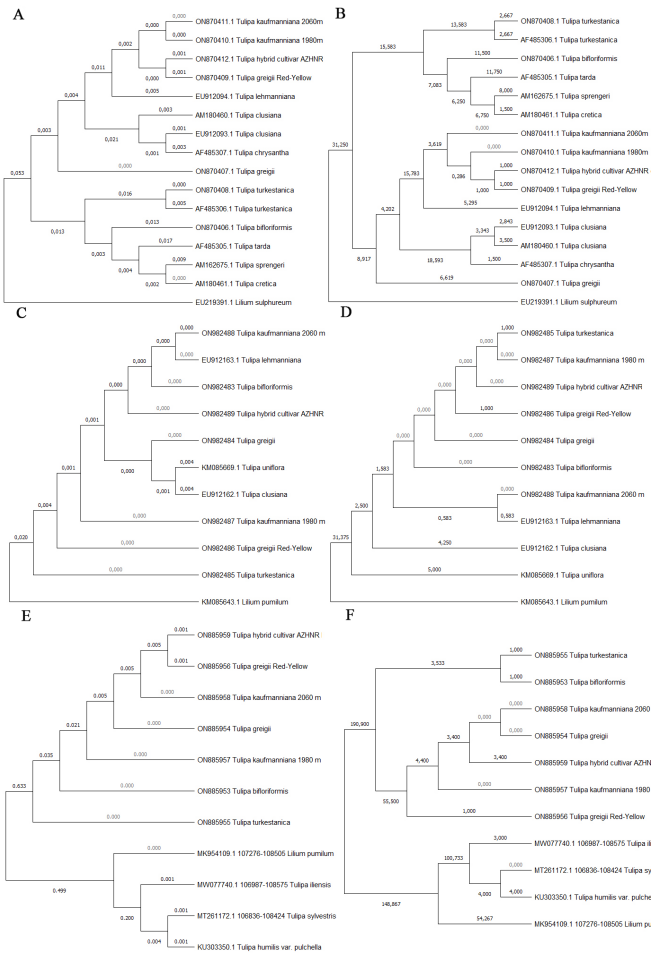

Fig. 4. Majority-rule consensus tree of ITS (A,B), matK (C,D) and ycf1b (E,F) data was obtained from the Neighbor-joining (A,C,E) and Maximum-parsimony (B,D,F) method. Gene distance values are shown above each branch.

Lilium pumilum, which was chosen as a representative of the outgroup of the Liliaceae family, were obtained from GenBank and analyzed for the *matK* gene. Inconclusive results were found when utilizing NJ and MP methods to construct a phylogenetic tree for the *matK* region of plastid DNA. The entirety of the relationship is marked by a lack of consistency (Fig. 4c, 4d). An analysis of the sequence of the *matK* gene in *Tulipa* specimens revealed that the *matK* sequence is highly conserved across all of the different species of *Tulipa*. There is only one known genetic substitution that is related to another species in the *T. turkestanica* species.

GenBank sequences of *Tulipa* species and *Lilium pumilum,* chosen as the outgroup of the *Liliaceae* family, were used to investigate the *ycf1b* gene in tulip samples. All the tulip samples were found to be grouped together on a single branch, as illustrated in Fig. 4e, 4f. These results indicate that the *ycf1b* gene has evolved more slowly in the tulip species than it has in other Liliaceae family members, which is likely due to a lack of gene recombination as a result of its self-incompatibility.

The evolutionary ties among the Tulipa species were examined using nuclear and plastid DNA sequences. Studying information from each sequencing area separately could result in incorrect conclusions about phylogenetic relationships (52). Following this observation, we looked into the phylogenetic relationship

between *Tulipa* species using the ITS, *matK* and *ycf1* sequences. The *Tulipa* family's *ycf1b* locus is poorly represented in GenBank. Unfortunately, we are unable to draw any firm conclusions about the phylogenetic relationship of the species under study because of the *matK* region. This issue stems from GenBank's depletion of plastid DNA sequences for *Tulipa* species. The study's findings about the ITS region of nuclear DNA were more definitive. Despite the limited information provided by GenBank, the study was able to make meaningful contributions to the understanding of *Tulipa* species phylogeny. The data collected agree with the findings of the *Tulipa* study (20) and other closely related species (53). The genus *Tulipa* is split into 4 different subgenera: *Tulipa*, *Eriostemones*, *Clusianae* and *Orithyia* (54). *Tulipa* is a genus with 56 species, the majority of which are located in Central Asia (4). Four species, *T. kaufmanniana*, *T. greigii*, *T. greigii* var. *red-yellow* and a hybrid of *T. kaufmanniana* and *T. greigii*, make up the subgenus *Tulipa* (4). Previous studies using morphology and cytology to classify organisms have placed these taxa in the same family as our findings suggest they do here (55). About 20 species belong to the subgenus Eriostemones, which is widespread across Eurasia and Siberia (3). As part of this study, previous research has established that both *T. turkestanica* and *T. bifloriformis* are in the genus *Tulipa* (3, 20). This new study expands upon the previous ones by using molecular techniques to delve deeper into the phylogenetic connections between the taxa.

Conclusion

Tulipa species from the Aksu-Zhabagly Nature Reserve had their nuclear ITS and plastid *matK* and *ycf1b* sequences reported for DNA barcoding and phylogenetic study. Phylogenetic trees for the ITS region were generated using neighbor-joining and maximum-parsimony approaches and produced comparable findings. Phylogenetic analysis anticipated that the clade of *Tulipa* would split in 2 subclades. *MatK* gene is highly conserved across all of the different species of *Tulipa*. There is only one known genetic substitution that is related to another species in the *T. turkestanica* species. The *ycf1b* gene has evolved more slowly in the tulip species than it has in other *Liliaceae* family members. This is likely due to a lack of gene recombination as a result of its self-incompatibility with other plant species. The evolutionary ties among the *Tulipa* species were examined using nuclear and plastid DNA sequences. The findings about the ITS region of nuclear DNA were more definitive. In conclusion, the study was able to use both nuclear and plastid DNA sequencing to draw firm conclusions about the evolutionary ties among the *Tulipa* species. The evolutionary relationships between species of *Tulipa* will be examined in greater detail, including the use of new barcode markers.

Acknowledgements

The authors wish to thank Prof. Ruslan Kalendar (Nazarbayev University, Kazakhstan) for editing, proofreading, and critical reading of the manuscript.

This research was funded by the Ministry of Agriculture of the Republic of Kazakhstan, grant number BR10764998.

Authors contributions

Conceptualization, BK.; methodology ZA., investigation AS., MS., SA., ZA. AT., SM.; writing – original draft preparation, BK.; writing – review and editing, BK., ZA.; supervision, BK.; funding acquisition, BK. All authors have read and agreed to the published version of the manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None.

References

1. Smith WW. The Genus *Tulipa*. Nature. 1940; 1940/09/01;146 (3699): 379-80. https://doi.org/10.1038/146379a0.
2. Christenhusz MJM, Govaerts R, David JC, Hall T, Borland K, Roberts PS, et al. Tiptoe through the tulips - cultural history, molecular phylogenetics and classification of *Tulipa* (Liliaceae). Botanical Journal of the Linnean Society. 2013; 07/01;172(3): 280-328. https://doi.org/10.1111/boj.12061.
3. Raamsdonk LWDv, Vries Td. Biosystematic studies in *Tulipa* sect.*Eriostemones* (Liliaceae). Plant Systematics and Evolution. 2004;179:27-41. https://doi.org/10.1007/BF00938017.
4. Raamsdonk LWD, Vries Td. Species relationships and taxonomy in *Tulipa* subg. *Tulipa* (*Liliaceae*). Plant Systematics and Evolution. 2004;195:13-44. https://doi.org/10.1007/BF00982313.
5. Zarrei M, Wilkin P, Ingrouille M, Leitch I, Buerki S, Fay M et al. Speciation and evolution in the *Gagea reticulata* species complex (Tulipeae; Liliaceae). Molecular phylogenetics and evolution. 2011; 11/17;62 624-39. https://doi.org/10.1016/j.ympev.2011.11.003.
6. Kiani M, Memariani F, Zarghami H. Molecular analysis of species of *Tulipa* L. from Iran based on ISSR markers. Plant Systematics and Evolution. 2012; 10/01;298 https://doi.org/10.1007/s00606-012-0654-0.
7. Morgil H, Şik L, Erol O. Genetic diversity by AFLP analysis within *Tulipa orphanidea* L. (Liliaceae) populations in Manisa. Celal Bayar University Journal of Science. 2017; 11/21;13: 913-17. https://doi.org/10.18466/cbayarfbe.334831.
8. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proceedings Biological sciences. 2003 Feb 7;270(1512): pp. 313-21. PubMed PMID: 12614582. Pubmed Central PMCID: PMC1691236. https://doi.org/10.1098/rspb.2002.2218.
9. de Vere N, Rich T, Trinder S, Long C. DNA barcoding for plants. Methods in Molecular Biology (Clifton, NJ). 2015; 01/01;1245: 101-18. https://doi.org/10.1007/978-1-4939-1966-6_8.
10. Ahrens D, Monaghan MT, Vogler AP. DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae). Mol Phylogenet Evol. 2007; Jul;44(1),: 436-49. PubMed PMID: 17420144. https://doi.org/10.1016/j.ympev.2007.02.024.

11. Kesanakurti PR, Fazekas AJ, Burgess KS, Percy DM, Newmaster SG, Graham SW et al. Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. Molecular ecology. 2011 Mar;20(6),: 1289-302. PubMed PMID: 21255172. https://doi.org/10.1111/j.1365-294X.2010.04989.x.

12. Hollingsworth PM, Graham SW, Little DP. Choosing and Using a Plant DNA Barcode. PloS one. 2011;6(5),: https://doi.org/10.1371/journal.pone.0019254.

13. de Groot GA, During HJ, Maas JW, Schneider H, Vogel JC, Erkens RHJ. Use of rbcL and trnL-F as a Two-Locus DNA Barcode for Identification of NW-European Ferns: An Ecological Perspective. PloS one. 2011;6(1),: e16371. https://doi.org/10.1371/journal.pone.0016371.

14. Ebihara A, Nitta JH, Ito M. Molecular Species Identification with Rich Floristic Sampling: DNA barcoding the Pteridophyte Flora of Japan. PloS one. 2010;5(12): p e15136,: https://doi.org/10.1371/journal.pone.0015136.

15. Li F-W, Kuo L-Y, Rothfels CJ, Ebihara A, Chiou W-L, Windham MD et al. *rbcL* and *matK* Earn Two Thumbs Up as the Core DNA Barcode for Ferns. PloS one. 2011;6(10),: pe26597. https://doi.org/10.1371/journal.pone.0026597.

16. Nitta JH. Exploring the utility of three plastid loci for biocoding the filmy ferns (*Hymenophyllaceae*) of Moorea. TAXON. 2008;57(3):725-36. https://doi.org/10.1002/tax.573006.

17. Mes THM, Kuperus P, Kirschner J, Stepanek J, Oosterveld P, Storchová H, et al. Hairpins involving both inverted and direct repeats are associated with homoplasious indels in non-coding chloroplast DNA of *Taraxacum* (Lactuceae: Asteraceae). Genome / National Research Council Canada = Génome / Conseil national de recherches Canada. 2000; 09/01;43,: 634-41. https://doi.org/10.1139/g99-135.

18. Navarro F, Suárez-Santiago VN, Blanca G. A new species of *Haplophyllum* A. Juss. (Rutaceae) from the Iberian Peninsula: evidence from morphological, karyological and molecular analyses. Annals of Botany. 2004;94 4,: 571-82. https://doi.org/10.1093/aob/mch176.

19. Alvarez I, Wendel JF. Ribosomal ITS sequences and plant phylogenetic inference. Mol Phylogenet Evol. 2003 Dec;29(3): pp. 417-34. PubMed PMID: 14615184. https://doi.org/10.1016/s1055-7903(03)00208-2.

20. Turktas M, Metin ÖK, Baştuğ B, Ertuğrul F, Saraç YI, Kaya E. Molecular phylogenetic analysis of *Tulipa* (Liliaceae) based on noncoding plastid and nuclear DNA sequences with an emphasis on Turkey. Botanical Journal of the Linnean Society. 2013; 2013/04/25;172(3):270-79. https://doi.org/10.1111/boj.12040.

21. Kalendar R, Boronnikova S, Seppänen M. Isolation and Purification of DNA from Complicated Biological Samples. In: Besse P, editor. Molecular Plant Taxonomy: Methods and Protocols. New York, NY: Springer US; 2021; 57-67. https://doi.org/10.1007/978-1-0716-0997-2_3

22. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990 Oct 5;215(3): pp. 403-10. PubMed PMID: 2231712. https://doi.org/10.1016/s0022-2836(05)80360-2.

23. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution. 2021;38(7):3022-27. https://doi.org/10.1093/molbev/msab120.

24. Lewin B. Chromatin and gene expression: constant questions, but changing answers. Cell. 1994 Nov 4;79(3): pp. 397-406. PubMed PMID: 7954807. https://doi.org/10.1016/0092-8674(94)90249-6.

25. Baldwin B, Sanderson M, Porter J, Wojciechowski M, Campbell C, Donoghue M. The its Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. Annals of the Missouri Botanical Garden. 1995; 01/01;82pp. 247. https://doi.org/10.2307/2399880.

26. Nieto Feliner G, Rosselló JA. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Mol Phylogenet Evol. 2007 Aug;44(2): 911-19. PubMed PMID: 17383902. https://doi.org/10.1016/j.ympev.2007.01.013.

27. Schultz J, Müller T, Achtziger M, Seibel PN, Dandekar T, Wolf M. The internal transcribed spacer 2 database--a web server for (not only) low level phylogenetic analyses. Nucleic Acids Research. 2006 Jul 1;34(Web Server issue), pp. W704-07. PubMed PMID: 16845103. Pubmed Central PMCID: PMC1538906. https://doi.org/10.1093/nar/gkl129.

28. Drouin G, Daoud H, Xia J. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol Phylogenet Evol. 2008 Dec;49 (3): pp. 827-31. PubMed PMID: 18838124. https://doi.org/10.1016/j.ympev.2008.09.009.

29. Dubouzet JG, Shinoda K. Phylogenetic analysis of the internal transcribed spacer region of Japanese *Lilium* species. Theoretical and Applied Genetics. 1999;98pp. 954-60. https://doi.org/10.1007/s001220051155.

30. Cronn RC, Small RL, Haselkorn TS, Wendel JF. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. American Journal of Botany. 2002;89 4pp. 707-25. https://doi.org/10.3732/ajb.89.4.707.

31. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC et al. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PloS one. 2008 Jul 30;3(7): pp. e2802. PubMed PMID: 18665273. Pubmed Central PMCID: PMC2475660. https://doi.org/10.1371/journal.pone.0002802.

32. Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, Joseph J et al. Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). Ann Bot. 2003 Jul;92(1): pp. 107-27. PubMed PMID: 12824072. Pubmed Central PMCID: PMC4243627. https://doi.org/10.1093/aob/mcg087.

33. Neuhaus H, Link G. The chloroplast tRNALys(UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. Current Genetics. 1987;11(4): 251-57. PubMed PMID: 2834093. https://doi.org/10.1007/bf00355398.

34. Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V et al. Angiosperm phylogeny based on *matK* sequence information. Am J Bot. 2003 Dec;90(12): pp. 1758-76. PubMed PMID: 21653353. https://doi.org/10.3732/ajb.90.12.1758.

35. Müller KF, Borsch T, Hilu KW. Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: contrasting *matK*, *trn*T-F and *rbc*L in basal angiosperms. Mol Phylogenet Evol. 2006 Oct;41(1): pp. 99-117. PubMed PMID: 16904914. https://doi.org/10.1016/j.ympev.2006.06.017.

36. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G et al. A proposal for a standardised protocol to barcode all land plants. Taxon. 2007 May;56(2): 295-99. PubMed PMID: WOS:000247420000004. https://doi.org/10.1002/tax.562004.

37. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. In: Proceedings of the National Academy of Sciences of the United States of America. 2008 Feb 26;105(8):

2923-28. PubMed PMID: 18258745. Pubmed Central PMCID: PMC2268561. https://doi.org/10.1073/pnas.0709936105.

38. Qiu YL, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M et al. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature. 1999 Nov 25;402(6760):404-47. PubMed PMID: 10586879. https://doi.org/10.1038/46536.

39. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics (Oxford, England). 2009 Jun 1;25(11): 1451-52. PubMed PMID: 19346325. https://doi.org/10.1093/bioinformatics/btp187.

40. Hausner G, Olson R, Simon D, Johnson I, Sanders ER, Karol KG et al. Origin and evolution of the chloroplast *trnK* (*matK*) intron: a model for evolution of group II intron RNA structures. Mol Biol Evol. 2006 Feb;23(2): pp. 380-91. PubMed PMID: 16267141. https://doi.org/10.1093/molbev/msj047.

41. Xiao-Xian L, Zhe-Kun Z. The higher-level phylogeny of monocots based on *matK*, *rbcL* and 18S rDNA sequences. Journal of Systematics and Evolution. 2007;45(2): 113-33. https://doi.org/10.1360/aps06148.

42. Ya-Ling W, Yong L, Shou-Zhou Z, Xing-Sheng Y. The utility of *matK* gene in the phylogenetic analysis of the genus *Magnolia*. Journal of Systematics and Evolution. 2006;44(2): 135-47. https://doi.org/10.1360/aps040013.

43. YU J, XUE J-H, ZHOU S-L. New universal *matK* primers for DNA barcoding angiosperms. Journal of Systematics and Evolution. 2011;49(3): 176-81. https://doi.org/10.1111/j.1759-6831.2011.00134.x.

44. Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M et al. Uncovering the protein translocon at the chloroplast inner envelope membrane. Science. 2013;339(6119): 571-74. https://doi.org/10.1126/science.1229262.

45. Oliver MJ, Murdock AG, Mishler BD, Kuehl JV, Boore JL, Mandoli DF et al. Chloroplast genome sequence of the moss *Tortula ruralis*: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. BMC genomics. 2010;11(1)1-8. https://doi.org/10.1186/1471-2164-11-143.

46. Wolf PG, Der JP, Duffy AM, Davidson JB, Grusz AL, Pryer KM. The evolution of chloroplast genes and genomes in ferns. Plant Molecular Biology. 2011;76(3): 251-61. https://doi.org/10.1007/s11103-010-9706-4.

47. Neubig KM, Whitten WM, Carlsward BS, Blanco MA, Endara L, Williams NH et al. Phylogenetic utility of *ycf1* in orchids: a plastid gene more variable than *matK*. Plant Systematics and Evolution. 2009;277(1): 75-84. https://doi.org/10.1007/s00606-008-0105-0.

48. Gernandt DS, Hernández-León S, Salgado-Hernández E, Pérez de La Rosa JA. Phylogenetic relationships of *Pinus* subsection Ponderosae inferred from rapidly evolving cpDNA regions. Systematic Botany. 2009;34(3): 481-91. https://doi.org/10.1600/036364409789271290.

49. Drew BT, Sytsma KJ. The South American radiation of *Lepechinia* (Lamiaceae): phylogenetics, divergence times and evolution of dioecy. Botanical Journal of the Linnean Society. 2013;171(1): 171-90. https://doi.org/10.1111/j.1095-8339.2012.01325.x.

50. Dong W, Liu J, Yu J, Wang L, Zhou S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PloS one. 2012;7(4): pp. e35071. https://doi.org/10.1371/journal.pone.0035071.

51. Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S et al. ycf1, the most promising plastid DNA barcode of land plants. Scientific Reports. 2015 2015/02/12;5(1): 8348. https://doi.org/10.1038/srep08348.

52. Doyle JJ. Gene Trees and Species Trees: Molecular Systematics as One-Character Taxonomy. Systematic Botany. 1992;17pp. 144-63. https://doi.org/10.2307/2419070.

53. Potter D, Luby JJ, Harrison RE. Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Systematic Botany. 2000;25(2): 337-48, 12. https://doi.org/10.2307/2666646

54. Everett D, Fay MF, Christenhusz MJM, Wilford R, Royal Botanic Gardens K. The Genus *Tulipa*: Tulips of the World: Kew Publishing; 2013.

55. Raamsdonk LWD, Eikelboom W, Vries Td, Straathof TP editors. The systematics of the genus *Tulipa* 11997.

56. Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S. Barcoding the Kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. Molecular Ecology Resources. 2015; 07/03pp. https://doi.org/10.1111/1755-0998.12438.

57. Kalendar R. A guide to using FASTPCR software for PCR, *in silico* PCR and oligonucleotide analysis. In: Basu C editor. PCR Primer Design. New York, NY: Springer US; 2022; p. 223-43. https://doi.org/10.1007/978-1-0716-1799-1_16