Title: The Importance in Phylogenetic Relationships of The Regions Belonging to Nuclear and Plastid DNA among Crocus Biflorus Subspecies

Authors: Aykut YILMAZ
Received: 2020-05-26 23:15:59
Accepted: 2020-10-27 12:59:52
Article Type: Research Article
Volume: 25
Issue: 1
Month: February
Year: 2021
Pages: 55-64

How to cite
Aykut YILMAZ; (2021), The Importance in Phylogenetic Relationships of The Regions Belonging to Nuclear and Plastid DNA among Crocus Biflorus Subspecies. Sakarya University Journal of Science, 25(1), 55-64, DOI: https://doi.org/10.16984/saufenbilder.742989
Access link

New submission to SAUJS
The Importance in Phylogenetic Relationships of The Regions Belonging to Nuclear and Plastid DNA among *Crocus biflorus* Subspecies

Aykut YILMAZ*1

Abstract

The genus *Crocus* L. (Iridaceae) composed of about 200 species is taxonomically very problematic, because of introgression and backcrossing observed among closely related species. Furthermore, determination of new taxa and variable characters observed in these new taxa are other important reasons of the taxonomic problems. Recently, many molecular based studies to understand the phylogenetic relationships of the *Crocus* taxa show the presence of the problems among the taxa, especially in *C. biflorus* subspecies. As a result of this, some researchers state that the term of subspecies must be changed in the genus and the most of subspecies must be categorized as species. For these reasons, in this study, the four regions belonging to nuclear and plastid DNA (*ITS1*-5.8S rRNA-*ITS2*, *psbA-trnH IGS*, *rpoC1* and *trnL-trnF IGS*) were used to understand the identification and separation abilities of taxa studied, in addition to understanding the taxonomy of *C. biflorus* subspecies.

**Keywords:** *Crocus*, introgression, backcrossing, *C. biflorus*

---

* Corresponding Author: aykut.yilmaz@usak.edu.tr
1 Usak University, Usak, Turkey, ORCID: https://orcid.org/0000-0002-0327-8388
1. INTRODUCTION

The genus *Crocus* L. belonging to the family Iridaceae consists of about 200 recognized species distributed from western Europe and north west Africa to western China. Balkan Peninsula and Turkey are considered as the center of species diversity for the genus [1, 2, 3, 4].

The genus *Crocus* is famous with their flowers in different colours and popular ornamentals. However, *C. sativus* is the best popular species of the genus because of the Saffron known as the world’s most expensive spices [5].

The genus is very complicated and hard to understand as taxonomic and phylogenetic. In addition to intermediate or variable characters observed in closely related species caused by introgression and backcrossing [6, 7], especially the continuously determination of new taxa and not assigned to any series [8] are the main reasons of taxonomic problems and doubtful species identification within the genus. Furthermore, recent phylogenetic analyses clearly show that subspecies of *C. biflorus* occur in distinct clades even in different series on the dendrogram prepared for phylogenetic relationships [9, 10]. This situation reveal that subspecies status of the genus for *C. biflorus* is incorrect. In other words, the term “subspecies” which was brought into the genus taxonomy by Mathew (1982) [1] can not be maintained any more [2].

Molecular studies such as PCR-based and especially in last years the studies on DNA sequence informations accelerated understanding of the genus taxonomy. As a result of these studies, taxonomical situation of the genus *Crocus* changed [9, 10, 11, 12] and many taxonomic classifications belonging to the genus were revised because of incorrect phylogenetic relationships.

Furthermore, two new series named as *Isauri* and *Lyciotauri* for section *Nudiscapus* were described by Kerndorff et al. (2014; 2015) [13, 14] and some subspecies of *C. biflorus* were grouped into these series [15].

For these reasons stated, determination of phylogenetic relationships among *Crocus biflorus* subspecies and grouped according to their genetic similarity of subspecies is necessary.

In this study, different regions containing *ITS1-5.8S rRNA-ITS2*, *psbA-trnH IGS*, *rpoC1* and *trnL-trnF IGS* belonging to nuclear and plastid DNA were used to understand the taxonomy of *C. biflorus* subspecies and to contribute the solution of the still existing problems.

2. MATERIAL AND METHODS

Sequence informations of four different regions containing *ITS1-5.8S rRNA-ITS2*, *psbA-trnH IGS*, *rpoC1* and *trnL-trnF IGS* belonging to nuclear and plastid DNA were provided from National Centre of Biotechnology Information [16]. Sequence informations belonging to *Crocus biflorus* subspecies for each regions were separately examined to evaluate discrimination ability of each regions studied and to understand the taxonomy and phylogenetic relationships of *C. biflorus* subspecies. For this aim; 16 subspecies of *C. biflorus* for *ITS1-5.8S rRNA-ITS2*, 9 subspecies of *C. biflorus* for *psbA-trnH IGS*, 16 subspecies of *C. biflorus* for *rpoC1*, 17 subspecies of *C. biflorus* for *trnL-trnF IGS* and their sequence informations were performed by using Molecular Evolutionary Genetics Analysis (MEGA).

After the sequence informations for *C. biflorus* subspecies were obtained, multiple sequence alignments for each regions were seperately performed by using MEGA X [17].

After that, the alignment sequence informations for each barcoding regions studied were used to assign the variable sites, probabilities of substitution from one base to another base, transitional substitutions (%), transversional substitutions (%), transition/transversion rates for purines-pyrimidines and nucleotide frequencies (Table 5).

Moreover, the tables showing the variable sites were prepared for each regions separately (Table 1, 2, 3, 4).
Neighbour-joining dendrograms showing bootstrap values on branches were provided for each regions examined to determine the species identification abilities and phylogenetic relations among C. biflorus subspecies.

All positions containing gaps and missing data for each regions were eliminated with the complete deletion option of the program for effective analyses.

3. RESULTS AND DISCUSSIONS

3.1. Analysis results for ITS1-5.8S rRNA gene-ITS2

Sequence informations belonging to 16 subspecies of C. biflorus were provided from NCBI [8, 10]. These DNA regions containing sequence informations of ITS1-5.8S rRNA gene-ITS2 were aligned by using Molecular Evolutionary Genetics Analysis (MEGA X). The sites with missing/ambiguous data and gaps were excluded for effective analyses in the determination of alignment length and variable sites. After the exclusion of these regions, alignment length for taxa studied was established as 604 bp. Totally 62 variable sites for taxa examined were determined (Table 1).

Furthermore, the probabilities of substitutions from one base to another base were computed as transitional and transversional substitutions (Table 5). It was observed that rate of transitional substitutions with 75.72 % is higher than the transversional substitutions. In other words, it can be said that the variable sites among the C. biflorus taxa are highly caused by the substitutions between same base groups (purines; A ↔ G or pyrimidines; C ↔ T).

In addition to the rate of base substitutions, transition/transversion rate for purines ($k_1$), pyrimidines ($k_2$) and overall transition/transversion rate (R) were assigned as 3.73, 8.84 and 2.89, respectively.

The nucleotide frequencies for ITS1-5.8S rRNA gene-ITS2 of C. biflorus subspecies were analysed as 37.42 % (A+T/U) and 62.58 % (C+G) (Table 5). In other words, it can be said that the DNA region examined consists of highly G and C bases.

Finally, Neighbor-Joining (NJ) dendrogram based on the sequence informations of ITS1-5.8S rRNA gene-ITS2 for C. biflorus subspecies was drawn to evaluate the phylogenetic relationships of taxa studied (Figure 1). Furthermore, the separation ability of ITS1-5.8S rRNA gene-ITS2 was evaluated for the taxa studied. Branch lengths in dendrogram infer the evolutionary distances. The evolutionary distances in dendrogram were computed using the Maximum Composite Likelihood method [18].

Table 1

Subspecies of C.biflorus and variable sites belonging to ITS1-5.8S rRNA-ITS2 (The numbers show variable nucleotides)
The Importance in Phylogenetic Relationships of The Regions Belonging to Nuclear and Plastid DNA among C. biflorus s. l. subspecies

![Line diagram](image.png)

Figure 1 Neighbor-Joining dendrogram provided from ITS1-5.8S rRNA-ITS2

### 3.2. Analysis results for psbA-trnH IGS

Sequence informations belonging to 9 subspecies of *C. biflorus* were provided from NCBI [9]. These DNA regions containing sequence informations of *psbA-trnH IGS* were aligned by using Molecular Evolutionary Genetics Analysis (MEGA X). The sites with missing/ambiguous data and gaps were excluded for effective analyses in the determination of alignment length and variable sites.

Alignment length for taxa studied was established as 603 bp. Variable sites showing the phylogenetic relationships among the taxa examined were determined in 6 nucleotides (Table 2).

Transitional and transversional substitutions expressing the substitutions between same or different base groups were determined as 70.39 % and 29.61 %, respectively. Moreover, transition/transversion rate for purines (k₁), pyrimidines (k₂) and overall transition/transversion rate (R) were assigned as 2.30, 7.12 and 2.22, respectively (Table 5).
The nucleotide frequencies for psbA-trnH IGS of *C. biflorus subspecies* were analysed as 62.41 % (A+T/U) and 37.59 % (C+G) (Table 5).

Finally, Neighbor-Joining (NJ) dendrogram based on the sequence informations of *psbA-trnH IGS* for *C. biflorus* subspecies was drawn to evaluate the phylogenetic relationships of taxa studied and to understand the discrimination ability of *psbA-trnH IGS* for the taxa studied (Figure 2).

Table 2

Subspecies of *C. biflorus* and variable sites belonging to *psbA-trnH IGS* (The numbers show variable nucleotides)

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>Variable Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Crocus biflorus subsp. weldenii</em></td>
<td>1 1 1 1</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. leucostylosum</em></td>
<td>2 7 1 1 1 7</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. damii</em></td>
<td>9 1 1 2 3 0</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. pseudonubigena</em></td>
<td>C A T C T C</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. arvinensis</em></td>
<td>T . . . . . .</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. melantherus</em></td>
<td>. . A G A .</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. pulchricolor</em></td>
<td>. G . . . .</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. punctatus</em></td>
<td>. . . . . . A</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. striidi</em></td>
<td>T . . . . .</td>
</tr>
</tbody>
</table>

Table 3

Subspecies of *C. biflorus* and variable sites belonging to *rpoC1* (The numbers show variable nucleotides)

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>Variable Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Crocus biflorus subsp. weldenii</em></td>
<td>2 2 3 4 4 5 5</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. leucostylosum</em></td>
<td>9 7 8 1 0 3 0 6</td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. damii</em></td>
<td>2 0 7 6 7 7 6 6</td>
</tr>
</tbody>
</table>

3.3. Analysis results for *rpoC1*

The DNA sequences belonging to 16 subspecies of *C. biflorus* for *rpoC1* were provided from NCBI [9] and these sequence informations were aligned by MEGA X for further analysis.

Alignment length for 16 taxa belonging to *C. biflorus* subspecies was established as 575 bp, after the exclusion of the positions containing gaps and missing data for effective analyses. Totally 8 variable sites were determined among the taxa studied (Table 3).

The rate of transitional substitutions for *rpoC1* sequence is higher with 91.83 % than transversional substitution. In addition to transitional and transversional submissions, transition/transversion rates were assigned as 30.85 for purines (k₁) and 13.60 for pyrimidines (k₂). The overall transition/transversion rate (R) was 11.13. The nucleotide frequencies of A+T/U and G+C were determined as 58.03 % and 41.97, respectively (Table 5).

Finally, Neighbor-Joining (NJ) dendrogram based on the sequence informations of *rpoC1* for *C. biflorus* subspecies was drawn to evaluate the phylogenetic relationships of taxa studied and to understand the discrimination ability of *rpoC1* for the taxa studied (Figure 3).
The Importance in Phylogenetic Relationships of The Regions Belonging to Nuclear and Plastid DNA among Subspecies of *Crocus biflorus*

Figure 3 Neighbor-Joining dendrogram provided from *rpoC1*

3.4. Analysis results for *trnL-trnF* IGS

Sequence informations belonging to 17 subspecies of *C. biflorus* for *trnL-trnF* IGS were provided from NCBI [10]. After that, these DNA sequences for *C. biflorus* subspecies were aligned by using Molecular Evolutionary Genetics Analysis (MEGA X). The sites with missing/ambiguous data and gaps were excluded in the analysis of alignment length and variable sites.

Alignment length was established as 670 bp for taxa studied. Variable sites which were very important in species identifications and phylogenetic relationships among the taxa were determined in 9 nucleotides (Table 4).

The rates of transitional and transversional substitutions were determined as 79.77 % and 20.23 %, respectively. In other words, it can be said that the most of variable sites among the taxa were caused by the substitutions between same base groups.

Furthermore, transition/transversion rate for purines (k<sub>1</sub>), pyrimidines (k<sub>2</sub>) and overall transition/transversion rate (R) were assigned as 8.45, 7.21 and 3.65, respectively (Table 5).

The nucleotide frequencies belonging to *trnL-trnF* IGS sequences of *C. biflorus subspecies* were analysed as 65.08 % (A+T/U) and 34.92 % (C+G) (Table 5).

Finally, Neighbor-Joining (NJ) dendrogram was drawn to evaluate the phylogenetic relationships of *C. biflorus* taxa and to understand the discrimination ability of *trnL-trnF* IGS for the taxa studied (Figure 4).

Table 4

Subspecies of *C. biflorus* and variable sites belonging to *trnL-trnF* IGS (The numbers show variable nucleotides)

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>G</th>
<th>C</th>
<th>G</th>
<th>A</th>
<th>G</th>
<th>T</th>
<th>T</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Crocus biflorus subsp. stridi</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. weldenii</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. melantherus</em></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. biflorus</em></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. fibroannulatus</em></td>
<td>A</td>
<td></td>
<td>A</td>
<td>G</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. alboeornatus</em></td>
<td></td>
<td>A</td>
<td>A</td>
<td>.</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. wattiorum</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. atospermus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. arvinensis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus cf. biflorus subsp. tauri</em></td>
<td></td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. nubigena</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. caelestis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. yagatanensis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. caricus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. crewei</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. pseudonubigena</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crocus biflorus subsp. adamii</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>6</th>
<th>6</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 3 Neighbor-Joining dendrogram provided from *rpoC1*
4. CONCLUSIONS

In the comparisons of sequence informations according to species identification abilities, it can be said that the region belonging to ITS1-5.8S rRNA gene-ITS2 is more efficient among the DNA regions examined for Crocus taxa. This region phylogenetically separated the taxa in two groups, besides it identified all taxa studied. Furthermore, variable sites that was important in species identification and phylogenetic relationships among taxa was observed in highest rate on ITS1-5.8S rRNA gene-ITS2.

It can be stated that although other DNA regions examined (psbA-trnH IGS, rpoC1 and trnL-trnF IGS) clearly separated some taxa from each other, these regions were insufficient in the separation and identification of all taxa studied. Variable sites expressing the substitutions based on the sequence informations are the most important characters in the evaluation of taxa phylogenetically. However, these regions showing the sequence changes among the taxa were observed on several nucleotides. In other words, it can be stated that the sequence informations belonging to psbA-trnH IGS, rpoC1 and trnL-trnF IGS were highly preserved for C. biflorus taxa.

Recent molecular studies show that taxa belonging to C. biflorus subspecies were not grouped together, even occur in distinct clades [9, 10]. In other words, it is observed that subspecies status of the genus for C. biflorus is incorrect and can not be maintained any more [2].

Harpke et al. (2016) [2] updated the Mathew’s study (1982) [1] which present the nineteen subspecies of Crocus biflorus in Turkey and they stated as a result of this study that all subspecies of C. biflorus ranged from Balkan Peninsula to Caucasus and Iran represent the independent lineages and should be treated at species level. Similarly, Addam et al. (2019) [19] states as a result of the studies based on morphological and molecular genetic in the Crocus genus that the most of the subspecies must be categorized as species because of their genetic divergences. This opinion is still controversial among scientists and...
it has not completely resolved as in the number of taxa.

All of them make necessary the phylogenetically evaluation of *C. biflorus* taxa and the examination of their subspecies status in more detail.

For these reasons, in this study, **ITS1-5.8S rRNA-ITS2** belonging to nuclear DNA and three regions (**psbA-trnH IGS, rpoC1** and **trnL-trnF IGS**) belonging to plastid DNA were used to understand the taxonomy of *C. biflorus subspecies* and to contribute the solution of the still existing problems. Moreover, the discrimination abilities of each DNA regions examined and nucleotide substitutions were analysed for *C. biflorus* taxa. Although the different DNA regions and their combinations are very important for effective phylogenetic analysis, some barcoding regions are not enough for species identification and discrimination on the dendrogram prepared to show phlogenetic relationships. This study results could provide important data for further studies in the selection of usefull regions, in addition to understanding of taxonomic relationships of *C. biflorus subspecies*.

**Appendix**

**ITS1-5.8S rRNA-ITS2:**

HE663991, HE663973, HE664018, HE663980, HE663975, HE664004, HE663976, HE664016, HE663972, HE664014, HE663978, HE664003, HE663969, LN864707, HE664017, HE664013

**psbA-trnH IGS:**

EU110184, EU110183, EU110150, EU110202, EU110195, EU110140, EU110134, EU110185, EU110129

**rpoC1:**

EU110612, EU110523, EU110533, EU110539, EU110544, EU110550, EU110605, EU110595, EU110637, EU110527, EU110530, EU110593, EU110522, EU110594, EU110524, EU110560

**trnL-trnF IGS:**

HE864180, HE864177, HE864182, HE864185, HE864208, HE864210, HE864257, HE864275, HE864227, HE864220, HE864183, HE864211, HE864207, HE864174, HE864165, HE864212, HE864198

**Acknowledgements**

The author would like to thank to NCBI for sequence informations used in this study and the authors sharing the sequence informations in NCBI.

**Funding**

The author has no received any financial support for the research, authorship or publication of this study.

**The Declaration of Conflict of Interest/ Common Interest**

No conflict of interest or common interest has been declared by the authors.

**The Declaration of Ethics Committee Approval**

This study does not require ethics committee permission or any special permission.

**The Declaration of Research and Publication Ethics**

The authors of the paper declare that they comply with the scientific, ethical and quotation rules of SAUJS in all processes of the paper and that they do not make any falsification on the data collected. In addition, they declare that Sakarya University Journal of Science and its editorial board have no responsibility for any ethical violations that may be encountered, and that this study has not been evaluated in any academic publication environment other than Sakarya University Journal of Science.
REFERENCES


[18] K. Tamura, M. Nei, and S. Kumar, “Prospects for inferring very large phylogenies by using the neighbor-joining method,” Proceedings of the National Academy