Caryologia
International Journal of Cytology, Cytosystematics and Cytogenetics

Firenze University Press
www.fupress.com/caryologia

OPEN ACCESS

Citation: Masoud Sheidai, Mohammad Mohebi Anabat, Fahimeh Koohdar, Zahra Noormohammadi (2022). Identifying potential adaptive SNPs within combined DNA sequences in Genus *Crocus* L. (Iridaceae family): A multiple analytical approach. *Caryologia* 75(3): 159-167. doi: 10.36253/caryologia-1560

Received: January 31, 2022

Accepted: October 02, 2022

Published: April 5, 2023

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Identifying potential adaptive SNPs within combined DNA sequences in Genus *Crocus* L. (Iridaceae family): A multiple analytical approach

Masoud Sheidai[1,*], Mohammad Mohebi Anabat[1], Fahimeh Koohdar[1], Zahra Noormohammadi[2]

[1] *Department of Plant Sciences and Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran*
[2] *Department of Biology, School of Basic Science, Science and Research Branch, Islamic Azad University, Tehran, Iran*
*Corresponding author. E-mail: msheidai@sbu.ac.ir

Abstract. The genus *Crocus* L. of Iridaceae family contains about 160 species and is considered as a complex group of plant taxa with regard to evolutionary and phylogenetic events. Inter-specific hybridization and gene flow contribute to species genetic homogeneity in one hand and high within species genetic variability and species genetic content overlaps caused species resolution a problem. In spite of extensive molecular phylogenetic studies in this genus, nothing is known about DNA sequences or Single nucleotide polymorphisms (SNPs) which are of adaptive nature. Moreover, nothing is known about which geographical or environmental factors plays role in species local adaptation and speciation events within *Crocus* L. genus. Therefore, the present study was conducted to answer the above said questions. We used a combined molecular data set of internal transcribed spacer (*ITS*) nuclear gene and trnL-F intergenic spacer (*trnL-F*) sequences of chloroplast genome. A multiple analytical method of Canonical correlation (CCA), Redundency analysis (RDA), and Latent Factor Mixed Model (LFMM)identified a few potential adaptive SNPs. Moreover, population criterions like Tajimas' D, molecular clock test, as well as skyline-plot revealed a smooth and continuous genetic changes for most of the *Crocus* species, but the occurrence of a sudden deep nucleotide substitution for *Crocus* taxa of Iran. The impact of latitude was significantly higher on nucleotide changes compared to that of longitudinal distribution of *Crocus* species.

Keywords: *Crocus*, adaptive divergence, SNPs, speciation.

1. INTRODUCTION

The genus *Crocus* L. (family Iridaceae), has about 100 species and contains an economically important species *Crocus sativus* L., the edible saffron. The species of this genus are distributed from Western Europe and north-

western Africa to Western China. Though the Asia Minor is the center of genus diversity (Sheidai et al., 2017), many species grow in the Mediterranean region (Saxena, 2010).

Several studies were concerned with molecular phylogeny and DNA barcoding of this genus which produced valuable information on different molecular aspects of genus. Aghighiravan et al. (2019), reported that ITS barcode is the best molecular marker for phylogenetic investigation on *Crocus* L. genus. Similarly, Sheidai et al. (2017), reported a high degree of genetic variability both within and among the studied species in the genus and that ISSR molecular markers are useful in *Crocus* species delineation. Along with the species relationships, these authors also reported population fragmentation and inter-specific gene flow in these taxa.

In a recent investigation, Mohebi et al. (2021), presented both DNA barcode and chromosome number variation in the genus. These authors suggested that molecular events like horizontal gene transfer (HGT) and deep coalescence may be associated with geographical distribution and *Crocus* taxa diversification. Due to importance of this genus and also lack of knowledge on geographical association of the genetic differences in *Crocus* species, we carried out a detailed bioinformatic analyses of a combined molecular data set of ITS nuclear DNA sequences and *trnL-F* chloroplast sequences, to : 1- Identify discriminating nucleotide sequences among *Crocus* species, 2- Illustrate if these sequences are significantly associated with geographical coordinates, 3- Identify nucleotide sequences with phylogenetic importance.

For bioinformatic studies, we used different analytical approaches like discriminate analysis of principal components analysis (DAPC), which is suitable for SNP sequences, as well as both CCA (Canonical correspondence analysis), and RDA (Redundancy analysis). Moreover, some data on *Crocus* species expansion were also produced by using population genetics analysis methods of Tajimas' D value, molecular clock test, and mismatch nucleotide pair test. The findings of this research are new to *Crocus* science.

2. MATERIAL AND METHODS

In this study, ITS nuclear DNA and trnL-F sequences of 68 *Crocus* species were obtained from National Center for Bioinformatic Information (NCBI). In addition, we used two species of the genus *Romulea* as outgroup taxa because of the high similarity to *Crocus* (Goldblatt et al., 2006; Petersen et al., 2008) (Table 1).

2.1. Data analyses

Sequence alignment and curation was done by MUSCLE program implemented out in molecular evolutionary genetics analysis (MEGA) 7 program. Mismatch analysis and skyline plot was constructed in R package 4.2. These sequences were then used to construct Maximum likelihood phylogenetic tree (ML tree), by MEGA 7 program based on Kimura-Two parameters distance. The following analyses were performed to identify the SNPs which show association with geographical coordinates of *Crocus* species distribution. We should mention that these analytical approaches have different assumptions and may differ to some extent in their results. Therefore, comparing obtained results are important for drawing a solid conclusion.

2.2. Canonical correspondence analysis

In the first approach we used CCA (Canonical correspondence analysis). This method is based on regression of the SNPs and ecological features and uses an approach similar to principal components analysis (PCA), but it is utilized for discrete characteristics like SNPs (Podani, 2000; Sheidai et al., 2020). This method differs from PCA in the way that, PCA tries to maximize the variance of data in a reduced space, while CCA tries to maximize the association of data (SNPS), to ecological features studied (Podani, 2000; Sheidai et al., 2020). CCA was performed in PAST ver. 4., program.

2.3. Latent Factor Mixed Model (LFMM)

Latent factor mixed model is a method for testing associations between loci and environmental gradients using latent factor mixed models. LFMM implements an MCMC algorithm for regression analysis in which the confounding variables are modeled with unobserved (latent) factors. The program estimates correlations between environmental variables and allelic frequencies, and simultaneously infers the background levels of population structure (Frichot et al., 2013, Frichot and Francois, 2015). LFMM was performed by LFMM package in R. 4.2.

2.4. Redundency analysis (RDA)

Redundancy analysis (RDA), a form of constrained ordination which is fit for genomic data sets, where we are interested in understanding how the multivariate environment shapes patterns of genomic composition across geographical areas. RDA is based on multivariate

Table 1. The accession numbers and chromosome number of taxa in for the genus *Crocus* and outgroup representatives.

Number	Taxa	Accession number(ITS)	Accession number(trnL-F)	chromosome number	Country
1	*C. veneris*	HE801061.1	HE864222.1	2n= 16	cyprus
2	*C. etruscus*	HG518187.1	HG518216.1	2n= 8	Italy
3	*C. kosaninii*	HG518189.1	HG518206.1	2n= 14	Serbia
4	*C. baytopiorum*	LS398370.1	LT991646.1	2n= 28	Turkey
5	*C. scardicus*	HE663961.1	HE864166.1	2n= 36	Macedonia
6	*C. versicolor*	HE801142.1	HE864249.1	2n= 26	Italy
7	*C. malayi*	HE801170.1	HE864246.1	2n= 30	Croatia
8	*C. imperati*	HE801131.1	HE864231.1	2n= 26	Italy
9	*C. minimus*	HE801140.1	HE864247.1	2n= 24	Italy
10	*C. corsicus*	HE801096.1	HE864241.1	2n= 18	Italy
11	*C. cambessedesii*	HE801105.1	HE864228.1	2n= 16	Spain
12	*C. nudiflorus*	HE801146.1	HE864253.1	2n= 48	Spain
13	*C. serotinus*	HE801125.1	HE864225.1	2n= 22	Portugal
14	*C. niveus*	HE801081.1	HE864219.1	2n= 28	Greece
15	*C. goulimyi*	HE801130.1	HE864230.1	2n= 12	Greece
16	*C. ligusticus*	HE801167.1	HE864234.1	2n= 24	Italy
17	*C. kotschyanus*	HE664000.1	HE864256.1	2n= 8	Turkey
18	*C. scharojanii*	HE801135.1	HG518229.1	2n= 8	Russia
19	*C. vallicola*	HE801168.1	HE864238.1	2n= 8	Russia
20	*C. gilanicus*	HE801172.1	HE864255.1	2n= 24	Iran
21	*C. sativus*	HE801172.1	LT991682.1	2n= 24	Iran
22	*C. pallasii* sub sp. *hausknechtii*	LS398387.1	LT991663.1	2n= 14	Iran
23	*C. thomasii*	LS398411.1	LT991688.1	2n= 16	Italy
24	*C. cartwrightianus*	LS398376	LT991648.1	2n= 16	Greece
25	*C. moabiticus*	LS398392.1	LT991669.1	2n= 14	Jordan
26	*C. oreocreticus*	LS398397.1	LT991674.1	2n= 16	Greece
27	*C. asumaniae*	LS398366.1	LT991641.1	2n= 26	Turkey
28	*C. mathewii*	HE801089.1	HE864217.1	2n= 70	Turkey
29	*C. reticulatus*	LM993447.1	LM993633.1	2n= 10	Moldova
30	*C. cvijicii*	LT222444.1	HE864276.1	2n= 18,20,22	Albania
31	*C. dalmaticus*	HE801137.1	HE864242.1	2n= 24	Croatia
32	*C. sieberi subsp. sieberi*	HE663966.1	HE864171.1	2n= 22	Greece
33	*C. robertianus*	HE801134.1	HE864236.1	2n= 20	Greece
34	*C. cancellatus* subsp. *pamphylicus*	HE801128.1	HE864229.1	2n= 12	Turkey
35	*C. hermoneus*	HE801163.1	HE864268.1	2n= 12	Jordan
36	*C. abantensis*	HE664019.1	HE864239.1	2n= 8,16	Turkey
37	*C. angustifolius*	HE801136.1	LM993589.1	2n= 20	Russia
38	*C. ancyrensis*	HE663987.1	LM993597.1	2n= 10	Turkey
39	*C. gargaricus* sub sp. *gargaricus*	HE801138.1	HE864243.1	2n= 30	Turkey
40	*C. sieheanus*	HE801157.1	HE864263.1	2n= 16	Turkey
41	*C. rujanensis*	LT222441.1	HE864280.1	2n= 22	Serbia
42	*C. biflorus* sub sp. *biflorus*	HE801121.1	HE864220.1	2n= 8	Italy
43	*C. almehensis*	HE801162.1	HE864271.1	2n= 20	Iran
44	*C. danfordiae*	HE664007.1	HE864201.1	2n= 8	Turkey
45	*C. pestalozzae*	HE801141.1	HE864248.1	2n= 28	Turkey
46	*C. cyprius*	HE663962.1	HE864168.1	2n= 10	Greece
47	*C. hartmannianus*	HE801173.1	HE864264.1	2n= 20	Cyprus
48	*C. leichtlinii*	LN864711.1	HE864277.1	2n= 20	Turkey
49	*C. kerndorffiorum*	HE801159.1	HE864213.1		Turkey

Number	Taxa	Accession number(ITS)	Accession number(trnL-F)	chromosome number	Country
50	*C. nerimaniae*	HE663977.1	HE864181.1	2n= 10	Turkey
51	*C. korolkowii*	HE801139.1	HE864244.1	2n= 20	Uzbekistan
52	*C. michelsonii*	KY797650.1	HE864278.1	2n= 20	Iran
53	*C. caspius*	HE801171.1	HE864266.1	2n= 24	Iran
54	*C. alatavicus*	HE801116.1	HE864273.1	2n= 20	Uzbekistan
55	*C. naqabensis*	LS398395.1	LT997016.1	2n= 14	Jordan
56	*C. antalyensis*	HE664015.1	HE864209.1	2n= 8	Turkey
57	*C. olivieri*	HE8011	HE864216.1	2n= 6	Turkey
58	*C. candidus*	HE663981.1	HE864186.1	2n= 6	Turkey
59					
60	*C. hyemalis*	HE801060.1	HE864215.1	2n= 6	Plestin
61	*C. aleppicus*	HE801175.1	HE864267.1	2n= 16	Jordan
62	*C. veneris*	HE801062.1	HE864222.1	2n= 16	Cyprus
63	*C. carpetanus*	HE801071.1	HE864265.1	2n= 64	Turkey
64	*C. nevadensis*	HE663960.1	HE864170.1	2n= 28, 30	Spain
65	*C. fleischeri*	HE663983.1	HE864188.1	2n= 20	Turkey
66	*C. pulchellus*	HE801145.1	HE864252.1	2n= 12	Greece
67	*C. laevigatus*	HE801166.1	HE864233.1	2n= 30	Greece
68	*C. banaticus*	HE801147.1	HE864254.1	2n= 26	Romania
69	*Romulea ramiflora*	HE664012.1	HE864206.1	2n= 36	Turkey
70	*R. bulbocodium*	HE664012.1	HE864202.1	2n= 34,36,42	Turkey

regression, and models linear combinations of the environmental predictors that explain linear combinations of the SNPs. This method effectively identifies covering loci associated with the multivariate environmental features (Legendre and Legendre, 2012).

Redundency analysis is a highly flexible framework, and produce answers on: 1- What environmental conditions cause genetic divergence among the studied taxa? and 2. What is the genetic basis of local adaptation to the environment? RDA identifies linear relationships among the response and predictor matrices; if non-linear relationships are expected, other statistical frameworks may be more suitable. RDA was performed in Paleontological statistics (PAST) ver. 4, program.

Mantel test was performed with 1000 times permutations as implemented in PAST ver. 4., program to study correlation between genetic distance and geographical distance of the studied species.

Phylogenetically important SNPs was determined by character mapping of 110 SNPs obtained based on parsimony criterion as performed in Mesquite 3.6 program. We performed Tajima's D test (Tajima, 1989) to reveal if *Crocus* species DNA sequences evolved randomly ("neutrally"), or under a non-random process, including directional or balancing selection, demographic expansion or contraction. Moreover, we also carried out the molecular clock test, to show if SNP changes occurred in accordance with a uniform clock rate model of evolution during *Crocus* genus speciation events. These tests were performed by MEGA 7 program.

3. RESULTS

3.1. The species genetic difference

The preliminary analysis of combined sequences obtained after sequence alignments and curation, produced a DNA segment of 110 base pair length. The average p dis of the studied species was 0.126. Based on Kimura 2-parameters, the studied taxa differed in genetic distance from 0. 01 to 0.30. The paired mismatch plot of nucleotide difference is presented in Fig. 1. This plot shows a normal distribution in genetic difference of these species, which indicates that genetic divergence occurred in a continuous and steady mode in evolution of the genus *Crocus* L. Skyline-plot (Fig. 2), of the same species also revealed a smooth and continuous species expansion in the genus *Crocus*, with two sudden changes in population demography and sequence change which are related to the speciation events in Iran *Crocus* taxa.

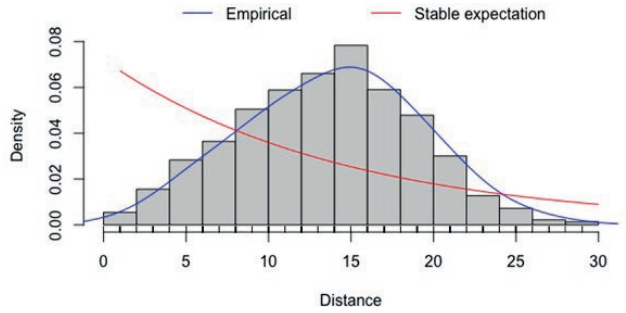

Figure 1. Mismatch plot of nucleotide difference among *Crocus* species.

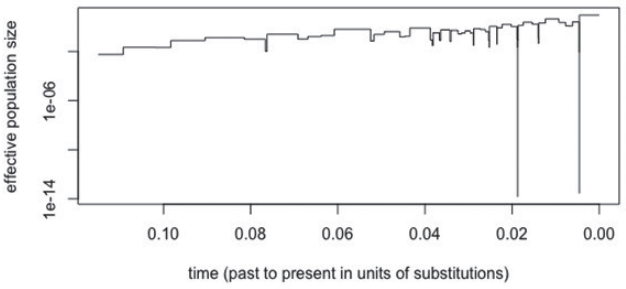

Figure 2. Skyline-plot of *Crocus* species based on combined sequence data set.

3.2. Genetic grouping of taxa

ML (Maximum likelihood) phylogenetic tree of the studied *Crocus* species based on combined molecular data set and the species geographical distribution, is presented in Fig. 3. We can place the studied species in three to four major clades. At the first glance, it is evident that species with Mediterranean distribution and those of South-West Asia (Iran, Iraq, and Afghanistan), and the neighboring regions, comprise adjacent clades, while the species growing in Europe are showing closer genetic affinity.

Genetic grouping of these species by Linear discriminating analysis (LDA), as performed in DAPC analysis is provided in Fig. 4. This plot also supports the presence of four genetic groups in the studied taa. The assignment test for the studied *Crocus* species based on DAPC analysis identified the species with genetic affinity (Fig. 5). The species n1-n68, are scattered in four major genetic groups as revealed by different cluster colors.

Linear discriminating analysis revealed that the first three discriminating analysis (DA) axes, comprise about 80% of total variation, and the first two axes have significant contribution with high Fst value (Fig. 6). DA loading obtained revealed that SNPs 74, 75 have the highest

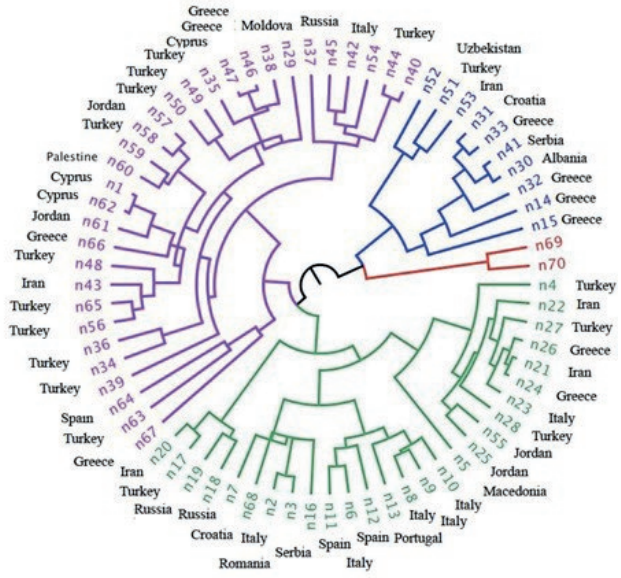

Figure 3. ML phylogenetic tree of the studied *Crocus* species and their geographical distribution. (n1-n70, as in Table 1).

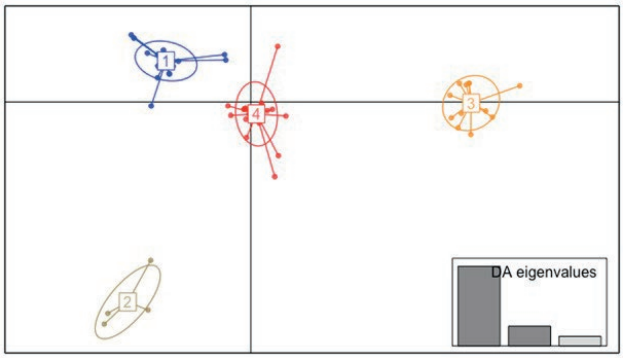

Figure 4. Genetic groups identified based on LDA analysis.

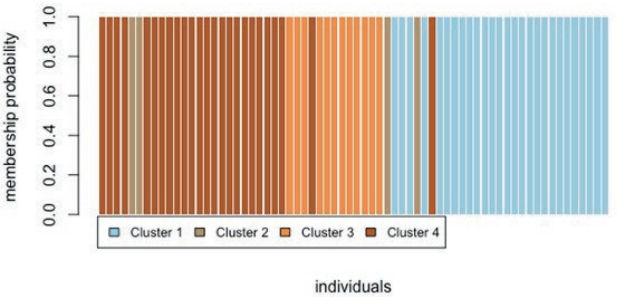

Figure 5. Assignment plot of *Crocus* species based on DAPC analysis (Individuals from left to wright are n1 to n 68 of Table 1).

discriminating power in the first LDA axis, followed by SNPs 31, and 109, in the second axis. Similarly, SNP 53 has high discriminating power in the third DA axis.

The following analyses were performed to identify the SNPs which show association with geographical coordinates of *Crocus* species distribution. We should mention that these analytical approaches have different assumptions and may differ to some extent in their results. Therefore, comparing obtained results are important for drawing a solid conclusion.

3.3. Canonical correspondence analysis

CCA plot of *Crocus* species and 110 SNPs used is provided in Fig. 7. The analysis produced two CCA axes with Eigenvalue% of 99.97and 0.028, respectively. Distribution of 110 SNPs used shows association between SNPS 31, 70, 71, 74, and 75, with latitude distribution of *Crocus* species of these three SNPS. viz. 31, 74, and 75, were identified as discriminating loci among *Crocus* taxa, by DAPC analysis. These SNPs have high association value as are placed in the first CCA axis. The SNPs 2, 93 and 94 of the second CCA axis, show a lower degree of association with longitude distribution of the studied *Crocus* taxa. From these results, we may conclude that, genetic changes of *Crocus* species towards latitude distribution was accompanied to these SNPs, which probably were associated with some important adaptive genes during *Crocus* speciation.

It becomes interesting when we plot the selected countries (geographical regions), by CCA (Fig. 8). We observe that countries like Iran, Russia, and Georgia, become separated from the other studied countries towards latitude. That means SNPs' changes occurred in these regions. The Skyline plot presented before also revealed a sudden change in nucleotide substitution and population size in Iran.

3.4. Latent Factor Mixed Model (LFMM)

Manhattan plot of LFMM analysis is presented in Fig. 9. It identified SNPs 2, 9, 63, and 79, showed a significant association with environmental features.

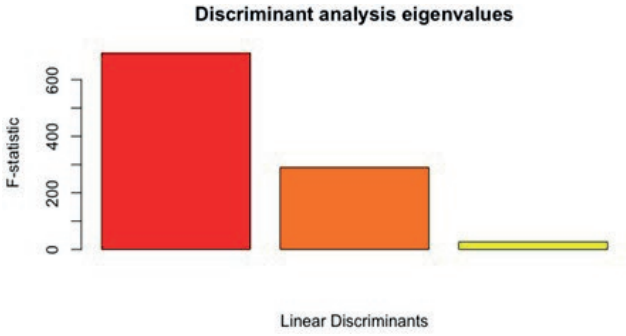

Figure 6. F-statistics of LDA analysis showing significant contribution of the first two axes in discriminating *Crocus* species.

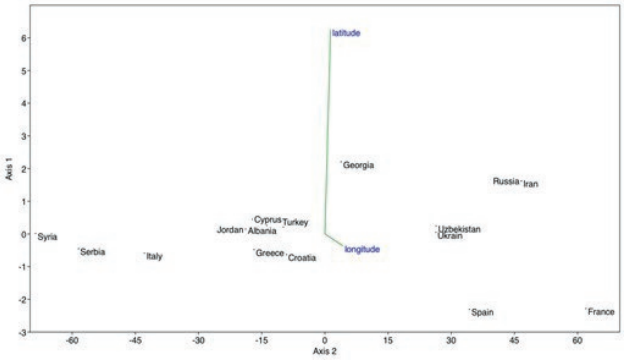

Figure 8. CCA plot of geographical regions showing separation of countries towards altitude and longitude *Crocus* species.

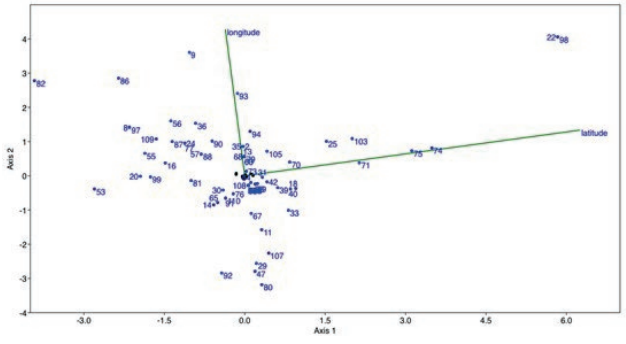

Figure 7. CCA plot of *Crocus* species showing association of few SNPs with geographical factors.

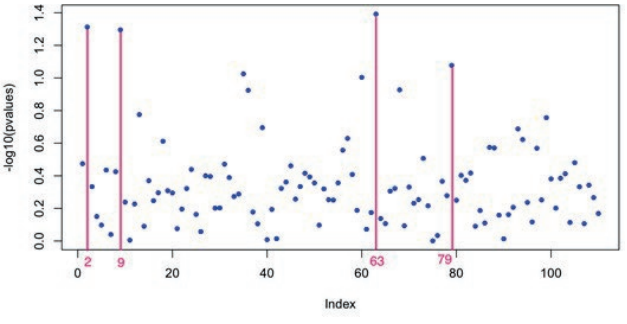

Figure 9. Manhattan plot of LFMM analysis identifying four SNPs associated with environmental features.

3.5. Redundancy analysis (RDA)

Redundancy analysis (RDA) was performed to detect the roles of geographical variables in *Crocus* species genetic subdivision, as well as the relative contribution of each variable to the population genetic structure. RDA plot is presented in Fig. 10. The SNPs 22, 71, 74, 75, 98, and 103, show association with latitude which occurs in RDA axis one with about 85% of total variance, followed by SNPs 9, 93 and 94, associated with longitude and second RDA axis with only 14% of total variance. Therefore, if we consider different association approaches utilized in this study, we can consider a few SNPs which are significantly associated with geographical factors studied. These SNPs occurred during species divergence within the genus *Crocus*.

A negative Tajima's D = -1.2, was obtained for the studied SNPs in *Crocus* species. This signifies an excess of low frequency polymorphisms relative to expectation, indicating population size expansion after a bottleneck or a selective sweep, which result in reduction in genetic diversity and formation of adaptive genotypes (species), in different geographical areas. The molecular test showed that SNP changes within the genes *Crocus* did not occurred under uniform rate of evolution and different phylogenetic clades differed in their genetic changes. This results also agree with the earlier result of skyline plot showing a deep change in SNP substitution and population size change in *Crocus* species of Iran and neighboring regions.

Mantel test performed after 1000 times permutations, produced non-significant correlation between genetic distance and geographical distance of the studied species (r = - 0.03, P = 0.7). This result indicates that nucleotide difference and change in *Crocus* taxa is not due to mere geographical distance and as indicated by different analyses reported here, genetic changes are. mainly associated with latitude distribution of these taxa.

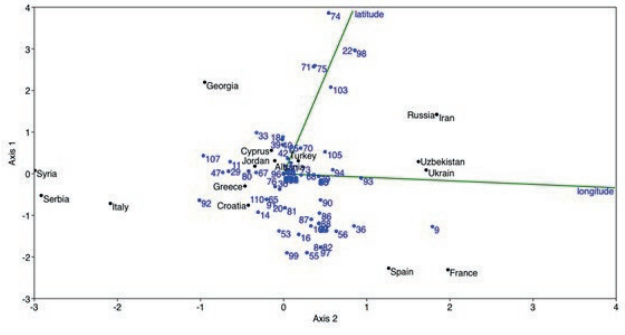

Figure 10. RDA plot of *Crocus* species showing association of few SNPs with geographical factors.

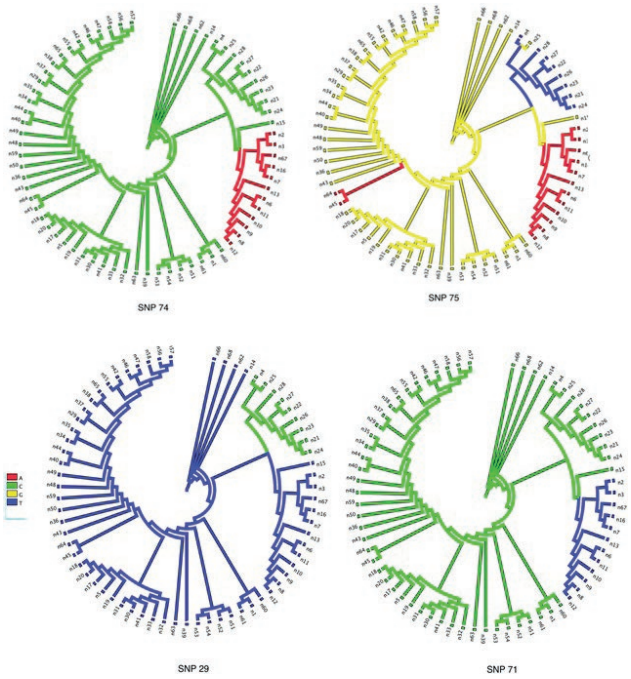

Figure 11. Character mapping of SNPs by parsimony criterion showing that these SNPs can differentiate different phylogenetic clades.

3.6. Phylogenetically important SNPs

Character mapping of the SNPs (Fig. 11), based on parsimony criterion, revealed that some of these SNPs are of phylogenetic importance as they differentiate almost a particular clade of the studied species. Interesting enough, SNPs 74 and 75. are also among these phylogenetic important SNPs. These two SNPs were identified as discriminating SNPs among *Crocus* species and also they are associated with latitude distribution of taxa, particularly *Crocus* species of Iran and neighboring areas.

4. DISCUSSION

Speciation within the genus *Crocus* is complex. A combination of diploid and polyploidisation events as well as inter-specific hybridization have been postulated for *Crocus* genus evolution (Mosolygo-L et al., 2016). Complexity at the species level has been reported by Seberg and Petersen (2009), as these authors could not delineate *Crocus* species even by utilizing different barcoding markers. However, some authors, could resolve *Crocus* species of Balkan (Mosolygo-L et al., 2016) and Iran (Sheidai et al., 2018), by using different molecular markers.

Recently, Mohebi et al. (2020), provided DNA barcode of Chloroplast DNA (trnH-psbA) region, which differentiated saffron genotypes of Iran from the other imported genotypes. Moreover, the same authors (unpublished data), provided some DNA barcode which illustrate genetic differentiation between *Crocus* taxa growing in different geographical regions and not for a particular *Crocus* species.

Nine *Crocus* L. species have been reported from Persia and some adjacent areas (Wendelbo, 1977; Matine, 1978). Taxonomy of the genus is controversial as evidenced by difficulties in *Crocus* species delineation. In spite of extensive efforts on the phylogenetic aspects of *Crocus* genus, there has been now report on ecological or geographical association of the genetic or DNA sequence changes with speciation events in this genus. The present study revealed that DNA nucleotides of both nuclear and chloroplast origin can efficiently differentiate some of the phylogenetic clades of *Crocus* taxa. Moreover, some of these sequences may be associated with geographical distribution of *Crocus* species. Some nucleotide seems to be tightly associated with latitudinal distribution of these taxa.

Tajima's test of these sequences produced a negative Tajima's D, which indicates an excess of low frequency polymorphisms relative to a selective sweep and speciation events (Tamura and Nei, 1993). We also observed almost a continuous and gradual nucleotide substitution for most of the species growing in other parts of the world, but a sudden deep change in DNA sequences of Iran *Crocus* species, which may be related to geographical adaptation as also evidenced by CCA and RDA analyses.

Different approaches used to identify the nucleotides associated with geographical variables, revealed some degree of difference. It is due the fact that CCA and RDA methods are based on linear association (regression), with different approaches, while LFMM method is a Bayesian-Model approach (Podani, 2000; Frichot and Francois, 2015). It seems therefore, using different approaches may improve understanding of associated SNPs with geographical and ecological variables. Such combined data evaluations, give insights into contemporary processes, and may explain how environmental factors influence selective and neutral genomic diversity within and among related species or different geographical populations within a single species (Segovia et al., 2020).

Presence of heterogenous environmental conditions are known to cause changes in genetic diversity of plant species and result in local adaptations even in the populations of a single species (Segovia et al., 2020). Understanding the genetic basis of local adaptation is one of the main concern of evolutionary biologists, as identifying adaptive genetic loci or SNPs improves our knowledge of the genetic mechanism of local adaptation and probably species diversification within a genus (Zhang et al., 2019).

Recent studies which are concerned with genetics of local adaptations try to answer two major questions: 1-which environmental variables play key role in the adaptive genetic divergence of a species or different species within a particular genus and shape its landscape genetic structure, and, 2-which genes or loci on the genome undergo adaptive genetic differentiation (Li et al., 2017, Zhang et al., 2019).

In general, populations' local adaptation which leads to speciation thing a genus is the act of natural selection in oppose to continuous gene flow. In plant groups such as *Crocus* genus in which species differentiation is vague due to inter-specific hybridization and a high degree of genetic affinity, local adaptation, may be expected to happen for a few genes or nucleotides, as also we demonstrated in this study. The latitude occurrence of nucleotide changes and species diversification in *Crocus* genus, may be related to a warmer and drier environmental conditions of Iran, and Afghanistan and neighboring regions in compare to those prevailing in Mediterranean countries and Europe.

In a similar study, Ingvarsson et al. (2006), characterized patterns of DNAsequence variation at the putative candidate gene phyB2 in 4 populations of European aspen (*Populous tremula*) and scored single-nucleotide polymorphisms in an additional 12 populations collected along a latitudinal gradient in Sweden. They utilized a sliding-window scan of phyB2 and identified six putative regions with enhanced population differentiation and four SNPs showed significant clinal variation. Therefore, they suggested that the clinal variation at individual SNPs is an adaptive response in phyB2 to local photoperiodic conditions. It has been suggesting that divergent selection enhances the levels of genetic differentiation not only for the sites that are the direct target of selection but also for neutral sites in the vicinity of the site(s) under selection (Charlesworth et al., 1997; Nordborg and Innan, 2003).

5. CONCLUSION

In conclusion, the present study provide data on DNA sequences changes in association with geographical variables in the genus *Crocus* and suggest that latitudinal distribution has a more profound effect on these genetic changes. Moreover, we also suggest utilizing a multiple analytical approaches for identification of both discriminating DNA nucleotides/ SNPs within a genus

and for illustrating SNPs association with geographical or ecological variables.

REFERENCES

Aghighiravan, F., Shokrpour, M., Nazeri, V., Naghavi, M.R., 2019. Phylogenetic assessment of some species of *crocus* genus using DNA barcoding. J. Genet. Resour. 5(2), 118-129.

Charlesworth, B., Nordborg, M., Charlesworth,D., 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. 70, 155–174.

Frichot, E., Schoville, S.D., Bouchard, G., Francois, O., 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol. Biol. Evol. 30 (7), 1687-1699.

Frichot, E., Francois, O., 2015. LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution: 6 (8), 925-929

Ingvarsson, P. K., Garcı´a, M.V., David Hall, D., Luquez, V., Jansson, S., 2006. Clinal Variation in phyB2, a Candidate Gene for Day-Length-Induced Growth Cessation and Bud Set, Across a Latitudinal Gradient in. European Aspen (*Populus tremula*). Genetics. 172, 1845–1853. DOI: 10.1534/genetics.105.047522

Legendre, P., Legendre, L., 2012. Numerical ecology. 3 editions. Oxford, UK.

Li, Y., Zhang, X.X., Mao, R.L., Yang, J., Miao, C.Y., Li, Z., Qiu, X.Y., 2017 Ten years of landscape genomics: challenges and opportunities. Front. Plant. Sci. 8, 2136.

Matine, F., 1978. Liste des plantes de l'herbierd'. vine (Herbarium Ministerii Iranici Agriculturae). Iridaceae, Tehran.

Mohebi Anabat, M., Riahi, H., Sheidai, M., Koohdar, F., 2020. Population genetic study and barcoding in Iran saffron (*Crocus sativus* L.). Ind. Crop. Prod. 143, 111915.

Mohebi Anabat, M., Riahi, H., Sheidai, M., Koohdar, F., 2021. A new look at the genus *Crocus* L. phylogeny and speciation: Insight from molecular data and chromosome geography. Genet. Resour. Crop. Accepted paper.

Mosolygó-L, Á., Sramkó, G., Barabás, S., Czeglédi, L., Jávor, A., Molnár, A., Surányi, G., 2016. Molecular genetic evidence for allotetraploid hybrid speciation in the genus *Crocus* L. (Iridaceae). Phytotaxa. 258 (2), 121–136.

Nordborg, M., Innan, H., 2003 The genealogy of sequences containing multiple sites subject to strong selection in a subdivided population. Genetics. 163, 1201–1213.

Podani J., 2000. Introduction to the Exploration of Multivariate Data Backhuyes, Leiden.

Petersen, G., Seberg, O., Thorsoe, S., Jorgensen, T., Mathew, B., 2008. A phylogeny of the genus *Crocus* (Iridaceae) based on sequence data from five plastid regions. Taxon. 57, 487–499.

Segovia, N. I., González-Wevar, C. A., Haye, P. A., 2020. Signatures of local adaptation in the spatial genetic structure of the ascidian Pyura chilensis along the southeast Pacific coast. Sci. Rep. 10,14098.

Sheidai, M., Tabasi, M., Mehrabian, M.R., Koohdar, F., Ghasemzadeh-Baraki, S., Noormohammadi, Z., 2017. Species delimitation and relationship in *Crocus* L. (Iridaceae). Acta. Botanica. Croat. 77(1): 10-17 .

Sheidai, M., Tabasi, M., Mehrabian, M. R., Koohdar, F., Ghasemzadeh-Baraki, S., Noormohammadi, Z., 2018. Species delimitation and relationship in *Crocus* L. (Iridaceae). Acta Bot. Croat. 77 (1), 10–17.

Tajima, F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123: 595-595.

Tamura K., Nei M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512-526.

Wendelbo, P., 1977. Tulips and Irises of Iran, and their relatives, Botanical Institute of Iran, Tehran.

Zhang, X. X., Liu, B. G., Li, Y., Liu, Y., He, Y. X., Qian, Z. H., Li, J. X., 2019. Landscape genetics reveals that adaptive genetic divergence in *Pinus bungeana* (Pinaceae) is driven by environmental variables relating to ecological habitats. BMC Evolutionary Biology. 19,160.