RESEARCH

Open Access

Geographical and genetic clines in *Dracocephalum kotschyi* X *Dracocephalum oligadenium* hybrids: landscape genetics and genocline analyses

Masoud Sheidai[1*], Fahimeh Koohdar[1*] and Javad Mazinani[1]

Abstract

Conservation and management of medicinally important plants are among the necessary tasks all over the world. The genus *Dracocephalum* (Lamiaceae) contains about 186 perennials, or annual herb species that have been used for their medicinal values in different parts of the world as an antihyperlipidemic, analgesic, antimicrobial, antioxidant, as well as anticancer medicine. Producing detailed data on the genetic structure of these species and their response against climate change and human landscape manipulation can be very important for conservation purposes. Therefore, the present study was performed on six geographical populations of two species in the *Dracocephalum* genus, namely, *Dracocephalum kotschyi*, and *Dracocephalum oligadenium*, as well as their inter-specific hybrid population. We carried out, population genetic study, landscape genetics, species modeling, and genetic cline analyses on these plants. We present here, new findings on the genetic structure of these populations, and provide data on both geographical and genetic clines, as well as morphological clines. We also identified genetic loci that are potentially adaptive to the geographical spatial features and genocide conditions. Different species distribution modeling (SDM) methods, used in this work revealed that bioclimatic variables related to the temperature and moisture, play an important role in *Dracocephalum* population's geographical distribution within IRAN and that due to the presence of some potentially adaptive genetic loci in the studied plants, they can survive well enough by the year 2050 and under climate change. The findings can be used for the protection of these medicinally important plant.

Keywords Computational methods, *Dracocephalum* taxa, Genetic structure, Hybrid zone, Selection

*Correspondence:
Masoud Sheidai
msheidai@sbu.ac.ir
Fahimeh Koohdar
f_koohdar@yahoo.com
[1] Department of Plant Sciences and Biotechnology, Faculty of Life
Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran

Introduction

Dracocephalum L., contains about 186 perennials, or annual herb species that mostly grow in alpine and semi-dry regions of temperate Asia and Europe with only one species occurring in North America [1]. These species have been used for their medicinal values in different parts of the world as an anti-hyperlipidemic, analgesic, antimicrobial, antioxidant, as well as anticancer medicine [2–5].

Dracocephalum genus is paraphyletic with *Hyssopus* L. and *Lallemantia* Fisch. & C.A.Mey., and this genus originated in Central and West Asia and Southern Siberia, and dispersed from Central and West Asia into the QTP and adjacent areas twice independently during the Pliocene [6].

Eight to ten *Dracocephalum* species have been reported in Flora Iranica [7], with five species endemic to Iran. Koohdar et al. [8], reported the occurrence of *D. oligadenium* Bornm. & Gauba in Iran. Before, Jamzad [9] considered *D. oligadenium* and *Dracocephalum kotschyi* Boiss. as synonyms and reported only the occurrence of *D. kotschyi* [9].

The occurrence of inter-specific hybrid plants is important as it brings about two genomes of divergent but related species together. Koohdar et al. [8] also reported hybridization between *D. oligadenium* and *D. kotschyi* by two morphological traits such as habitat form and margin of stem leaves and molecular and anatomical studies in a locality with intermediate position between the two parental species habitat.

Hybridization events can produce new genetic and phenotypic traits that can help the species' ecological adaptation and improve the fitness of the plants and their genotype-environment interaction [10]. Natural inter-specific hybridization occurs with high frequency in the Lamiaceae family see e.g. [11–13],

The gene flow and natural selection shape geographic patterns of genotypic and phenotypic variation in plants, and therefore to understand the impact of these evolutionary forces on specific closely related species groups we have to carry out several complementary lines of investigation like, the gene flow and connectivity, to produce a fine-scale genetic structure for the studied species and populations as well as at a landscape scale; and study the spatial patterns of genetic variation and interacting impacts of geography and climate on gene flow and selection. Moreover, identifying the genes or genetic regions (or SNPs), which are potentially adapted to the local environment through light on their role in plant response to the environmental conditions and climate [14].

The field of landscape genetics which combines population genetics and spatial landscape analyses can produce a better understanding of the above said phenomenon and processes by using different computational approaches like multiple regression, redundancy analysis (RDA), canonical analysis (CCA), latent factor mixed model (LFMM), etc. Landscape genetics tries to identify the present-time adaptive genetic loci and investigate the adaptive potential of populations in response to future landscape and climatic changes [15]. Moreover, it can suggest conservation and management strategies based on the genetic response of the plant populations [16].

Similarly, species distribution modeling (SDM) methods, can produce insight into the present-time species pattern of geographical distribution and the species' response to future climate changes. Global warming and climate change are great threats of the 21st century that may result in disaster in global biodiversity, with serious adverse ecological consequences, such as droughts, wildfires, and invasive pest outbreaks. These changes can result in the loss of plant species and lowered productivity, food crop scarcity, etc. Climate change may negatively affect the plant populations' genetic diversity and genetic connectivity, which in turn may bring about a more genetic homogenization of plant populations and result in lowering the adaptive potentials of plant species in the coming years [17].

The hybrid zones are geographic regions where genetically divergent taxa meet and hybridize and are thought to be stable over evolutionary time and can be maintained by either the environmental gradients or by a balance between selection against hybrids and the dispersal of parentals into the zone. Within the hybrid zone, the genetic exchange among two or more divergent species may occur differently in different regions of their genome through recombination and introgression [18].

Gompert and Buerkle [19] proposed the cline theory concept and the genomic clines method for mapping components of reproductive isolation in a hybrid zone. This method produces data related to the balance between selection and gene flow in the maintenance of hybrid zones. The estimates obtained in the genocline method are based on e cline width and linkage disequilibrium considering the change in frequency of marker genotypes along a genome-wide admixture gradient. These estimates are referred to as genomic clines and differ from geographical clines, which estimate changes in the population frequency of characters (alleles or phenotypes) along geographical gradients [19].

Based on the cline theory concept, the patterns of introgression could be measured by using cline analysis of the genetic and morphological features across the hybrid zone [18]. In these analyses, the shape of a cline is modeled for both genetic loci and morphological traits by combining three equations that describe a sigmoid shape at the center of a cline and two exponential decay curves on either side of the central cline. The results obtained then identify the traits under the influence of cline [18].

The genocline method multinomial regression analysis is based on a null model of neutral introgression and identify molecular markers with patterns of introgression inconsistent with neutral expectations which are thought to be linked to the genes that contribute to reproductive isolation (markers with reduced introgression or a deficit of heterozygotes), and the genes related to the hybrid vigor (markers with increased introgression or an excess of heterozygotes) [19].

The aims of the present study were: 1- To investigate gene flow and genetic admixture in two species of *Dracocephalum kotschyi* and *D. oligadenium*, 2- Investigate and identify geographical and genetic clines in these two species and their hybrid plants based on spatial analyses of landscape genetics and genocline analyses, and 3- Perform species distribution modeling and study the effects of climate change on these taxa.

Materials and methods
Plant material
We used information from Koohdar et al. [8], in this study plant materials (70 plant specimens) were collected from 7 geographical populations and used for morphological, anatomical, and ISSR molecular studies (Table 1).

Data analyses
Genetic diversity and admixture analysis
The ISSR bands obtained [8] were used for further analyses. Total Band Patterns for Binary (Diploid) Data by Populations studied was performed. The genetic diversity parameters and AMOVA test were determined by GenAlex program 6.4. Grouping of the cultivars was performed by principal coordinate analysis (PCoA), and discriminant analysis of principal components (DAPC). Gen flow and admixture analysis were performed by DAPC and admixture plot by Adgentet, LEA, and LFMM packages in R 4.2. Gene flow versus genetic differentiation for ISSR loci was computed for each locus using POPGENE 3.1. The assignment test and proportion of genetic admixture were determined by STRUCTURE software 2.3 [20].

Landscape genetic analysis
Association studies
Association between genetic data and geographical variables was determined by redundancy analysis (RDA), after 999 permutations, LFMM (Latent factor mixed model), as well as pcadapt analyses in related packages of RDA, LDA, LFMM, and pcadapt in R 4.2 [21, 22].

Spatial PCA (sPCA)
We used the spatial PCA method (sPCA) to study the spatial pattern of *Dracocephalum* genetic variability and structure [23] in the Adegenet package in R 4.2.

Genocline analyses
We used three R packages namely AFLPism, HZAR, and HIest for genetic cline analyses. These packages carry out the analyses for dominant molecular markers like ISSR. The package AFLPsim was used for genetic-cline estimation of dominant marker profiles in hybridizing populations [24]. The genome scan method used in AFLPism is the "bal&gar-ca method" which is based on the departure of the theoretically expected frequencies for each band in each hybrid category under neutral introgression. HZAR package was used to study the cline model. HZAR provides functions for fitting molecular genetics and morphological data from hybrid regions to classical genetic models using the Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm [18, 24].

We used the HIest package which provides approaches to find maximum likelihood estimates of S (The maximum likelihood estimate of the ancestry index), and H (The maximum likelihood estimate of the interclass heterozygosity), based on the likelihood functions as described by Fitzpatrick [25]. With respect to two ancestral species or parental populations (P1 and P2), the index of ancestry (S) is the proportion of alleles of an individual that descend from the alleles of population P1, and interclass heterozygosity (H) is the proportion of loci of an individual that have one allele from each ancestral population.

Species distribution modeling (SDM)
We used a combination of methods to perform SDM, such as Maxent (a maximum entropy modeling method

Table 1 Populations studied their locality and ecological features

NO	Province	Locality	Altitude (m)	Longitude	Latitude	Humidity	Average rainfall	wind speed	Average temperature
1	Qazvin	Evan	1796	3629	5026	50	24	10.5	11
2	Mazandran	Namarestagh	2370	3603	5203	35	11	8	25
3	Qazvin	Niroogah	1318	3617	5001	47	23	11	10
4	Gilan	Roodbar	1473	3648	4922	60	36	8.5	14
5	Mazandran	Noor	2043	3612	5148	59	53	1.3	25
6	Tehran	Fasham	2217	3557	2133	35.5	22	7.5	11
7	Mazandran	Rineh	2026	3552	5210	30	12	9	24

tailored for presence-only species [26], Dismo analyses, GLM (General linear model), and Random Forest model. These analyses were performed by MAXENT software and under Dismo package in R. 4.2.

In species distribution modeling we used projected climate data layers for the current (~1950–2000) and the year 2050 period (average for years 2050–2061) based on 19 bioclimatic variables at a resolution of 5 minutes' spatial resolution in WorldClim database. (of a longitude/latitude degree, this is about 9 km at the equator).

We used bioclimatic variables derived from the monthly temperature and rainfall values. These bioclimatic variables are coded according to the information of https://www.worldclim.org/data/bioclim.html.

Morphological cline analyses

Morphological data (X1. Plant height (cm), X 2. Length of basal leaf (mm), X3. Width of basal leaf (mm), X4. Length of petiole in basal leaf (mm), X5. Length of stem leaf (mm), X6. Width of stem leaf (mm), X7. Length of petiole in stem leaf (mm) 8 Length of inflorescence leaf (mm), X9. Width of inflorescence leaf (mm), X10. Length of petiole in inflorescence leaf (mm), X11. Size of inflorescence leaf arista (mm), X12. Length of bracteole, X13. Width of bracteole, X14. Size of bract arista, X15. Length of calyx (mm), X16. Length of corolla(mm), X17. The number of vine, X18. Width of calyx (mm), X19. Size of tooth in calyx(mm), X20. Length of style, X21. Length of areole arm, X22. Length of stamen, X23. Habitat, X24. Margin of stem leaves) were standardized (Mean=0, variance=1), and used for CCA (Canonical correspondence analysis), and the Mantel test between morphological distance and geographical distance. For both the CCA analysis and Mantel test, geographical variables longitude, altitude, and latitude of the species and populations studied were used. These analyses were performed in PAST software, ver. 4. Similarly, the HZAR package in R 4.2, was used for cline analysis of morphological data. We also used the phylogenetic PCA methods of both Revell [27, 28], and Jombart et al. [29], for studying the quantitative morphological structuring in response to environmental conditions.

The Revell method of phylogenetic PCA is carried out with data transformations and data reduction procedures (such as a size correction and principal components analysis, PCA), after correcting for nonindependence among the observations for the studied species and populations.

The phylogenetic Principal Component Analysis (pPCA) [29] is derived from the spatial Principal Component Analysis (sPCA) [23], implemented in the adegenet package.

pPCA is designed to investigate phylogenetic patterns a set of quantitative traits. The analysis returns principal components maximizing the product of variance of the scores and their phylogenetic autocorrelation (Moran's I), therefore reflecting life histories that are phylogenetically structured. Large positive and large negative eigenvalues correspond to global and local structures.

We used three samples of the studied species each (parental species and the hybrid), for these analyses which were performed in related packages of phytools, and adephylo in R 4.2.

Results

Genetic diversity and gene flow analyses

The results of analyses presented in the following paragraphs are based on 73 ISSR loci obtained in 42 plant samples of the studied *Dracocephalum* populations. Details of ISSR bands/ or loci are provided in Table 2. The studied populations contained 0–4 private bands.

The estimated genetic diversity parameters like the percentage of genetic polymorphism and expected gene diversity in three studied *Dracocephalum* populations are provided in Table 3. These data indicate that a low to moderate (24–52% genetic polymorphism), is present in these populations. AMOVA produced a significant genetic difference ($p=0.01$), (Table 4).

The paired AMOVA indicated that significant genetic differences exist among all the paired-sampled populations (Table 5). Moreover, Bayesian MCMC analysis of STRUCTURE, produced a moderate to high Fst value for most of the studied populations, particularly for the hybrid population (0.6), which indicates this hybrid population differs in its genetic content from the parental populations (see below data):

Mean value of Fst-1=0.1346.
Mean value of Fst-2=0.4465.
Mean value of Fst-3=0.0027.
Mean value of Fst-4=0.3468.
Mean value of Fst-5=0.4595.
Mean value of Fst-6=0.4356.
Mean value of Fst-7=0.6807.

Table 2 Total Band patterns for binary (diploid) data by populations studied

Population	Pop1	Pop2	Pop3	Pop4	Pop5	Pop6	Pop7
No. Bands	38	18	41	25	28	37	37
No. Bands Freq. >= 5%	38	18	41	25	28	37	37
No. Private Bands	1	0	4	0	1	1	1
No. LComm Bands (<=25%)	0	0	0	0	0	0	0
No. LComm Bands (<=50%)	14	6	12	7	10	14	14

Abbreviations: Pop1-7 are: Populations 1–3=*Dracocephalum kotschyi*, 4–6=*D. oligadenium*, and 7=The hybrid population

Table 3 Genetic diversity parameters determined in *Dracocephalum* populations studied

Pop	N	Na	Ne	I	He	uHe	%P
Pop1	6.000	1.041	1.260	0.245	0.159	0.173	52.05%
Pop2	6.000	0.493	1.102	0.106	0.066	0.072	24.66%
Pop3	6.000	1.068	1.206	0.212	0.132	0.144	50.68%
Pop4	6.000	0.589	1.156	0.137	0.092	0.100	24.66%
Pop5	6.000	0.671	1.145	0.138	0.089	0.098	28.77%
Pop6	6.000	0.973	1.297	0.256	0.172	0.188	46.58%
Pop7	6.000	1.000	1.248	0.239	0.154	0.169	49.32%

Pop1-7 are: Populations 1–3=*Dracocephalum kotschyi*, 4–6=*D. oligadenium*, and 7=The hybrid population

N=The number of plant samples studied, Na=No. of different alleles, Ne=No. of effective alleles, I=Shannon's Information Index, He=Expected heterozygosity, uHe=Unbiased expected heterozygosity

Table 4 AMOVA table showing significant genetic differences among *Dracocephalum* populations studied

Source	df	SS	MS	Est. Var.	%
Among Pops	6	164.571	27.429	3.459	34%
Within Pops	35	233.667	6.676	6.676	66%
Total	41	398.238		10.135	100%
Stat	Value	P (rand >=data)			
PhiPT	0.341	0.001			

The preliminary results on genetic differentiation (Gst), versus gene flow (Nm), are provided in Table 6. The results are for the ISSR loci with a higher value of Gst and Nm. In general, data obtained revealed that a few loci have a higher differentiating power (Gst>0.70), among the studied *Dracocephalum* populations, while some other loci have a higher gene flow value (Nm>3.00), contributing to genetic similarity and homogenization of the studied populations.

Admixture and assignment plots, as well as PCoA ordination (Fig. 1, A and B), supported revealed the occurrence of gene flow and genetic admixtures among *Dracocephalum* populations and also identified the plants and populations involved in gene flow (Fig. 1).

Assignment results of the STRUCTURE analysis also indicate a genetic admixture of the studied plants and populations, but also revealed that plants in the hybrid population (Pop 7), are genetically somewhat different from the other populations, and that have received relatively more genetic contribution from plants of the populations 4 (0.369), from *D. oligadenium* (Table 7).

Table 6 Loci with a higher value for Gst and Nm in *Dracocephalum* populations studied

Locus	Sample size	Ht	Hs	Gst	Nm
Locus3	42	0.1190	0.1110	0.0676	6.8940
Locus4	42	0.3836	0.3397	0.1144	3.8694
Locus14	42	0.0246	0.0227	0.0756	6.1117
Locus18	42	0.0743	0.0655	0.1184	3.7233
Locus20	42	0.1190	0.1110	0.0676	6.8940
Locus21	42	0.0485	0.0455	0.0638	7.3341
Locus26	42	0.2530	0.2267	0.1039	4.3111
Locus27	42	0.2375	0.2008	0.1545	2.7359
Locus29	42	0.3991	0.3539	0.1132	3.9181
Locus30	42	0.0485	0.0455	0.0638	7.3341
Locus31	42	0.3083	0.2859	0.0728	6.3706
Locus41	42	0.3383	0.0924	0.7268	0.1880
Locus42	42	0.2955	0.0682	0.7693	0.1500
Locus43	42	0.4632	0.1284	0.7227	0.1918
Locus58	42	0.1856	0.1739	0.0634	7.3923
Locus59	42	0.0970	0.0883	0.0901	5.0521
Locus65	42	0.1907	0.1675	0.1214	3.6190
Locus66	42	0.2305	0.2103	0.0876	5.2061
Locus67	42	0.0246	0.0227	0.0756	6.1117
Locus71	42	0.0743	0.0655	0.1184	3.7233
Locus72	42	0.0743	0.0655	0.1184	3.7233
Locus73	42	0.0246	0.0227	0.0756	6.1117

Abbreviations: Ht=Total heterozygosity, Hs=Heterozygosity in populations, Gst=Differentiating power, and Nm=gene flow

The genetic grouping

Despite gene flow among the studied populations, we have three distinct genetic groups as revealed by

Table 5 Pairwise population PhiPT values in *Dracocephalum* populations studied. (PhiPT values below diagonal. Probability, P (rand >=data) based on 999 permutations)

Pop1	Pop2	Pop3	Pop4	Pop5	Pop6	Pop7	
0.000	0.011	0.005	0.002	0.001	0.005	0.028	Pop1
0.271	0.000	0.002	0.005	0.001	0.001	0.004	Pop2
0.258	0.359	0.000	0.001	0.001	0.001	0.006	Pop3
0.378	0.356	0.413	0.000	0.001	0.002	0.004	Pop4
0.346	0.466	0.223	0.542	0.000	0.005	0.003	Pop5
0.222	0.417	0.385	0.415	0.502	0.000	0.297	Pop6
0.167	0.263	0.262	0.362	0.401	0.020	0.000	Pop7

Pop1-7 are: Populations 1–3=*Dracocephalum kotschyi*, 4–6=*D. oligadenium*, and 7=The hybrid population
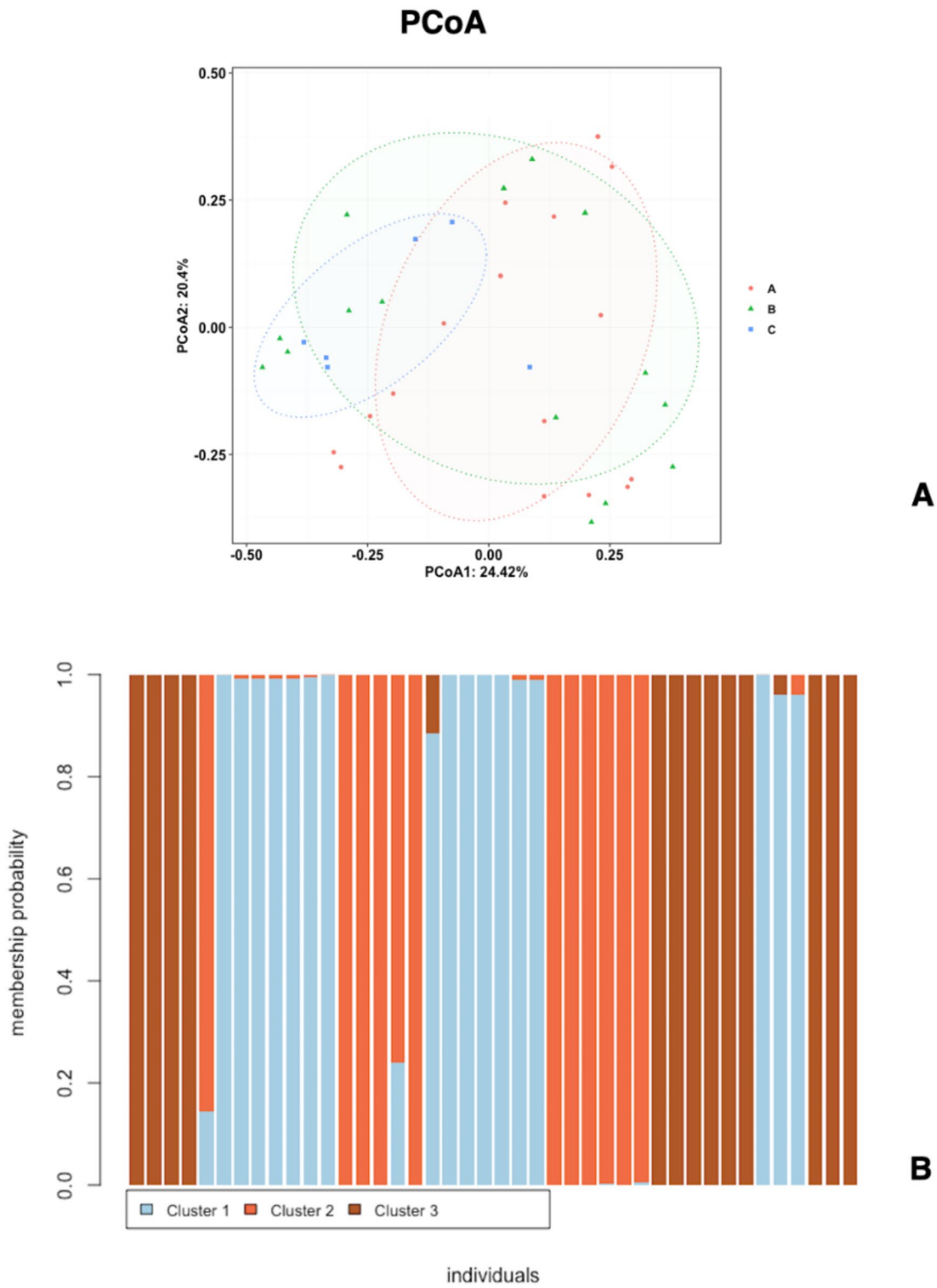
Fig. 1 PCoA plot (**A**), and DAPC plot (**B**), showing genetic admixture among *Dracocephalum* populations. A-C in Fig, A, are: A = *Dracocephalum kotschyi*, B = *D. oligadenium*, and C = The hybrid population

Table 7 The results of the assignment test in STRUCTURE, showing inferred population membership of *Dracocephalum* plants studied

Given Pop	Inferred Clusters							Number of Individuals
	1	2	3	4	5	6	7	
1:	0.033	0.180	0.029	0.056	0.135	0.014	0.553	6
2:	0.004	0.942	0.012	0.012	0.021	0.005	0.004	6
3:	0.159	0.083	0.136	0.010	0.591	0.012	0.009	6
4:	0.012	0.934	0.004	0.032	0.012	0.004	0.002	6
5:	0.006	0.011	0.006	0.013	0.939	0.003	0.021	6
6:	0.059	0.014	0.038	0.354	0.005	0.500	0.030	6
7:	0.008	0.021	0.018	0.369	0.099	0.473	0.011	6

Pop1-7 are: Populations 1–3=*Dracocephalum kotschyi*, 4–6=*D. oligadenium*, and 7=The hybrid population
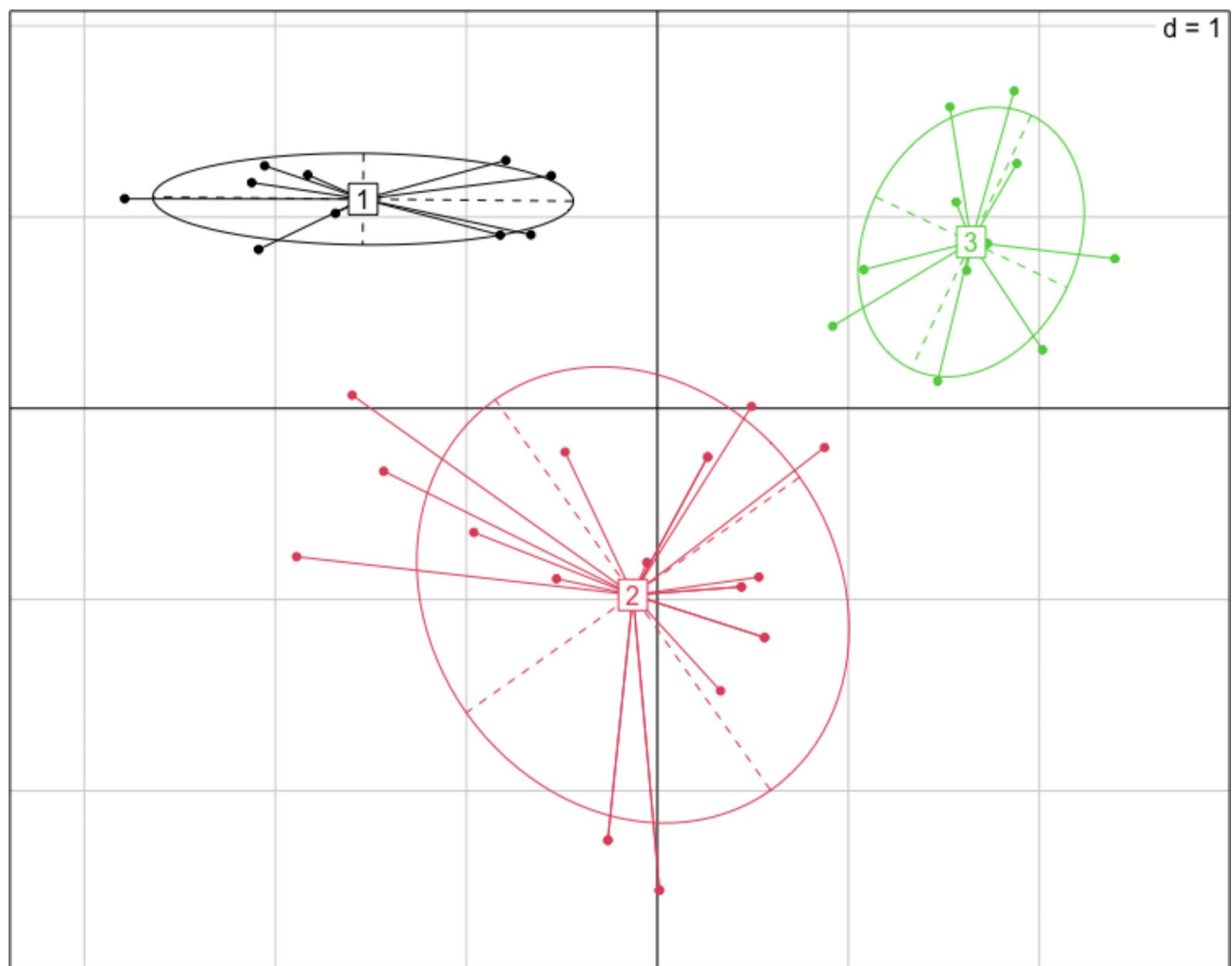
Fig. 2 DAPC genetic grouping of *Dracocephalum* plants studied with three separate populations predefined analysis

grouping of the studied plants by DAPC analysis (Fig. 2). This result indicates that we have plants that differed significantly in their genetic content, as also evidenced by AMOVA analysis (Significant phi value with $p<0.01$). When we performed the same analysis by considering only two genetic groups, DAPC produced two broad genetic groups with some degree of overlaps in between, which indicate the occurrence of inter-mixed individuals

and support the presence of hybrid plants (Fig. 3). Therefore, these results indicate that the hybrid population is genetically distinct from the parental populations.

Landscape genetic results
sPCA results
Spatial PCA (sPCA) analyses, revealed that the studied populations' genetic structure is influenced by
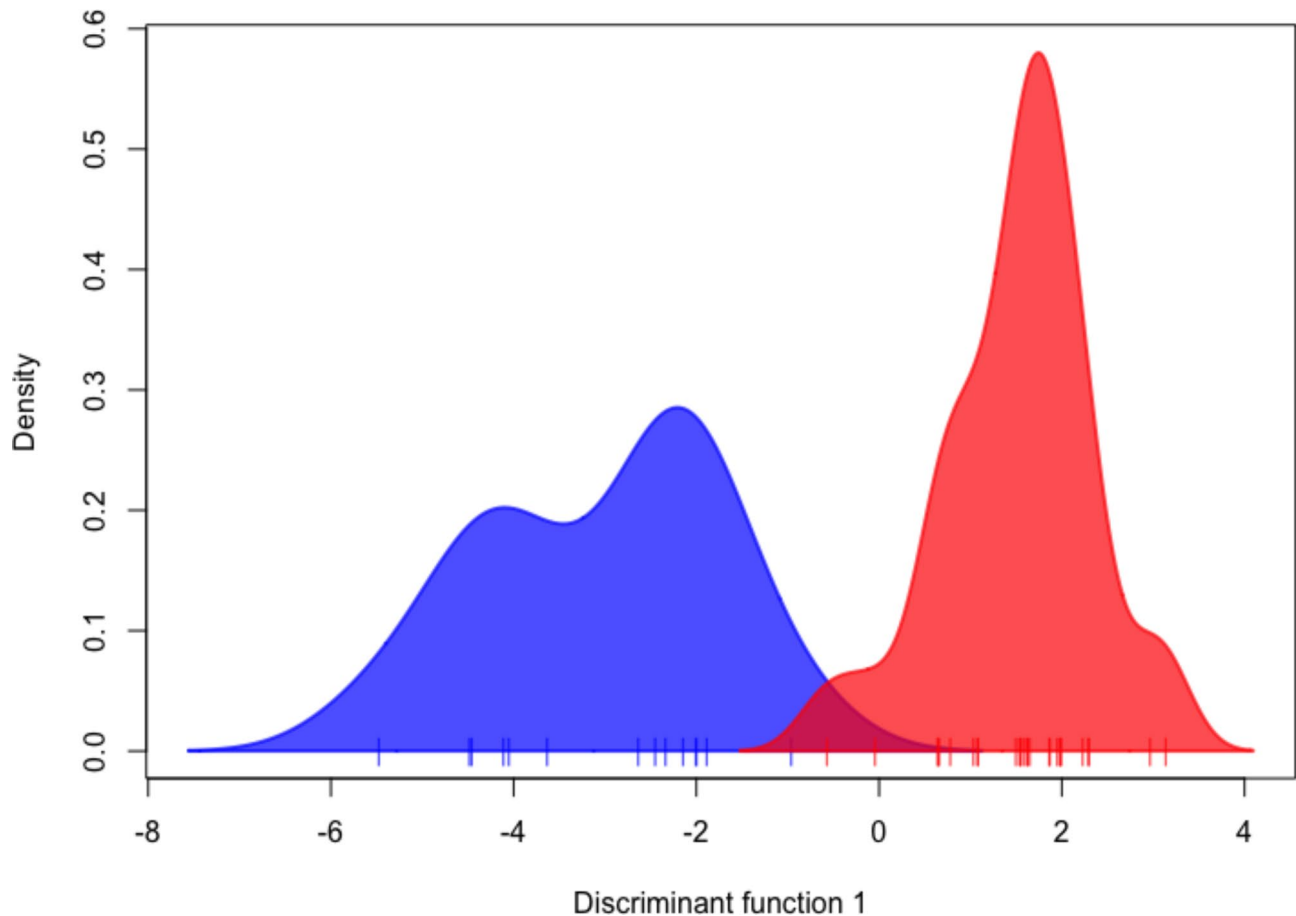
Fig. 3 DAPC genetic grouping of *Dracocephalum* plants with two populations predefined analysis, showing the overlap individuals

geographical variables (longitude and latitude), and both global and local environmental conditions affect their genetic content. These results are evidenced by the presence of both positive and negative Eigenvalues (Fig. 4, C), as well as a significant global test ($p=0.001$, Fig. 4, D), and local test ($p=0.002$), which were performed after 999 permutations.

The connection network (Fig. 4, A), and potential genetic/geographical clines (Fig. 4, B) were identified by sPCA analysis. These analyses also identified those ISSR genetic loci which are potentially adapted and selected in a geographical cline context (Fig. 4, E). These loci are ISSR loci number 35, 43, 44 and 61. The Mantel test also produced a significant correlation between genetic distance and geographical distance ($p=0.01$), among the studied populations. Therefore, closely placed plants are genetically more similar due to gene flow.

Association studies
We used different computational approaches to identify ISSR genetic loci which are potentially adaptive to the geographical variables. It is done as each of these methods utilizes a different statistical procedure for this

purpose, and it is always better to conclude the results based on different methods of analysis.

Redundancy analysis (RDA) which is based on the multiple regression method produced a significant association between ISSR genetic loci and geographical variables i.e. longitude, latitude, and altitude after 999 permutations ($F=2.74$, $p=0.001$).

The ANOVA analysis of the RDA method performed for the importance of these geographical variables on the genetic structure of *Dracocephalum* species and populations also produced a significant effect for all three variables ($p=0.001$). RDA analysis identified ISSR loci 10, 25 and 6 as the outliers and candidate genetic loci that are significantly associated with the geographical variables studied.

A similar analysis performed by the pcadapt package to detect genetic markers involved in biological adaptation is based on Principal Component Analysis (PCA). The test statistics histogram of the studied genetic loci is presented in Fig. 5, A. It reveals that although most of the p values obtained show a uniform distribution, some loci with small p values are considered to be the outlier loci. This is also supported by the Q-Q plot of the
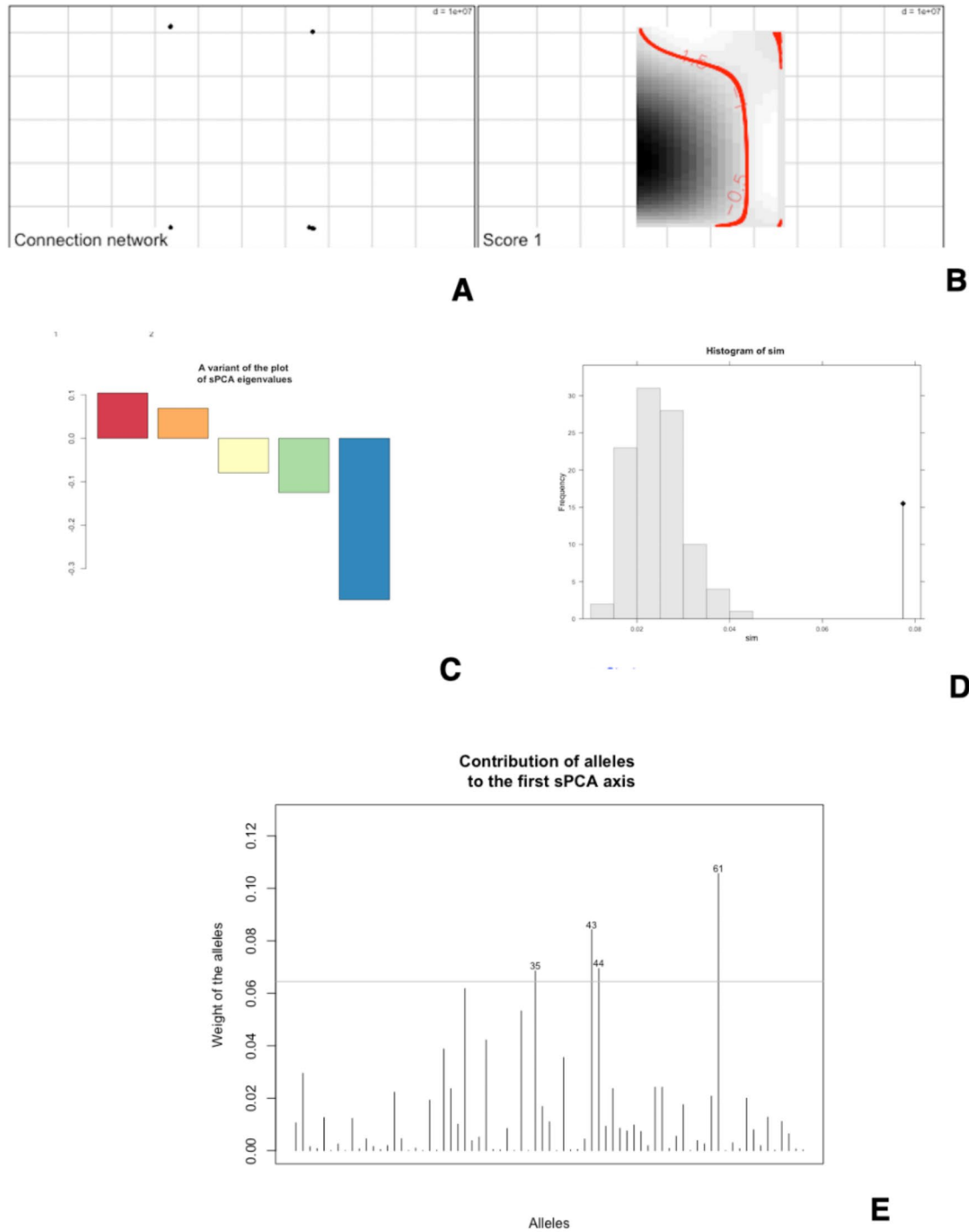
Fig. 4 sPCA results related to landscape genetic study, showing connection network (**A**), the presence of geographical genetic cline (**B**), Eigen values (**C**), Gtest significant results (**D**), as well as contributing ISSR loci in spatial genetic adaptation of *Dracocephalum* plants studied, and contribution of SCoT loci adapted and selected in a geographical cline context (**E**)
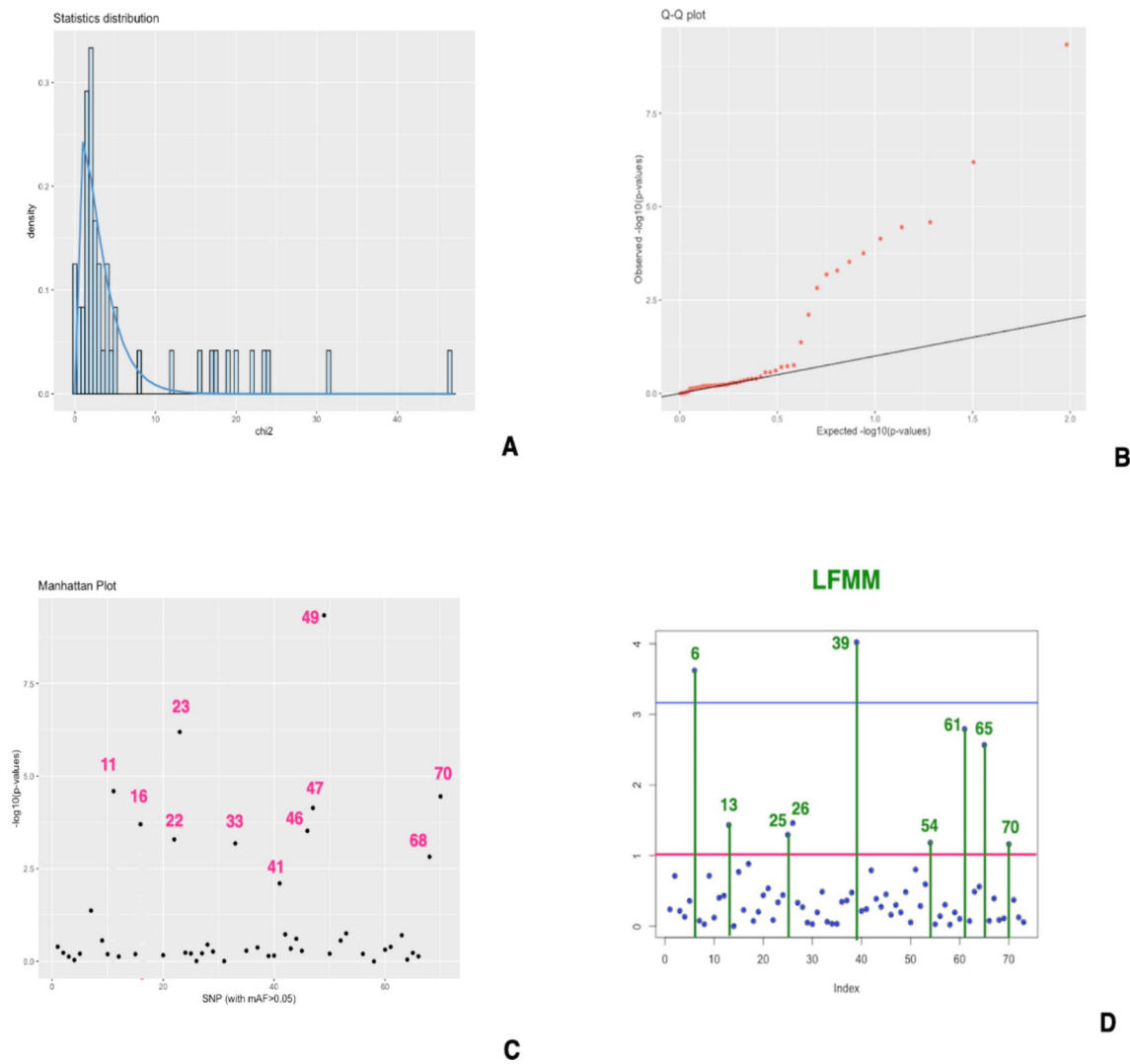
Fig. 5 Plots of association study, **A** and **B** = Alleles frequency plot, and Q-Q plot of the pcadapt package showing that a few genetic markers were involved in biological adaptation. **C**, and **D** = The Manhattan plot of pcadapt analysis and LFMM analysis respectively, identifying the potentially adaptive loci in *Dracocephalum* populations studied

pcadapt analysis (Fig. 5, B). Finally, the Manhattan plot of pcadapt analysis (Fig. 5, C), identified these outlier loci. The results of the Bayesian-MCMC method of LFMM analysis also identified the locally adapted ISSR loci in the studied *Dracocephalum* populations (Fig. 5, D). In total, a few loci from 3 to 10 genetic regions with adaptive potentials by these methods.

Genetic cline analyses

The result of AFLPism and identification of the genetic loci contributing to the hybrid zone is presented in Fig. 6. In total 13 ISSR loci were significantly associated with the hybrid zone adaptation. These loci are from both parental species and the hybrid plants formed, therefore not only the combination of parental alleles/genetic loci from parents through introgression has played adaptation to the hybrid zone, but some new genetic loci from the hybrid plants possibly formed due to recombination or new genetic arrangement of loci with hybrids, also play role in the survival of the hybrids within the clinal selection.

HZAR genetic clines with a 95% confidence interval related to the ISSR loci which are associated with the hybrid zone cline are presented in Fig. 7.

The estimated mean center and the width of these loci' genetic clines are for the ISSR locus 1, were, the mean center = 160 Km, and mean width = 350 Km; for locus
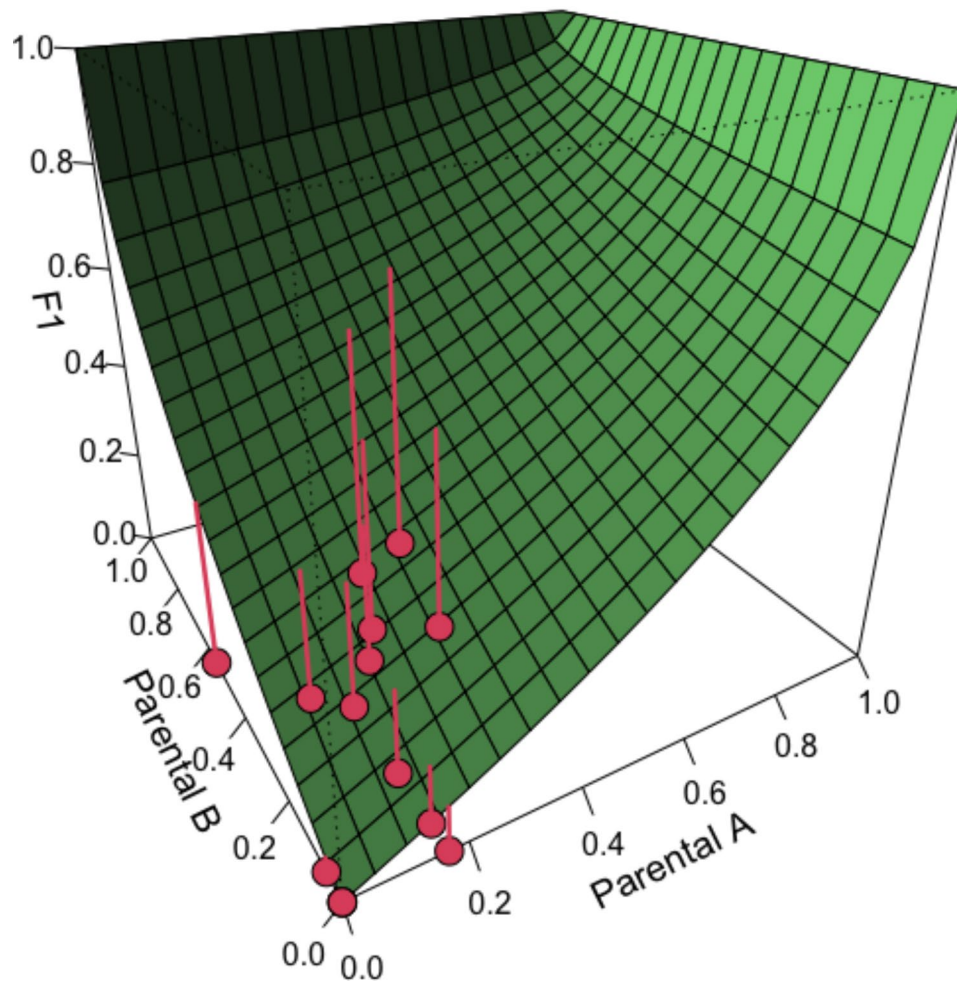
Fig. 6 Genocline analysis of AFLPism shows that a few ISSR loci are potentially adaptive in both parental species as well as the hybrid plants (F1), in *Dracocephalum* plants studied

3=310, and 350 Km; locus 4=140, and 320 Km; locus 6=115, and 325 Km, locus 11=98, and 135 Km, locus 16=370, and 370 Km, locus 23=350, 400 Km, locus 49=100, and 300 Km, and for locus 16=160, and 350 Km, respectively.

Similarly, the estimates of the S (The maximum likelihood estimate of the ancestry index), H (The maximum likelihood estimate of the interclass heterozygosity), and logLik (The maximum log-likelihood), for some of the selected ISSR loci that were identified by different methods as an outlier or adaptive loci revealed that the condition of S=0.5, and H=1, only occurred for ISSR loci 33, 46, 60 and 61. Therefore, these results indicate that a few loci are adaptive due to heterozygosity of the parental genotypes and are in agreement with the AFLPism results presented in earlier paragraphs.

Species distribution modeling (SDM)
The four different modeling of species distribution used for *Dracocephalum* species and populations produced similar results for both the present time as well as the year 2050 and in response to climate change (Fig. 8, A-D). The AUC values obtained for all models were above 0.8, with good accurate results.

The predicted distribution due to climate change is not much different from the present-time distribution of these populations and these medicinally valuable plants survive in response to climate change.

The importance of bioclim variables determined by random forest and Maxent models are provided in Fig. 9, A-D. Maxent model jackknife analysis revealed that bioclim variables 18=Precipitation of Warmest Quarter, 17=BIO17=Precipitation of Driest Quarter, and
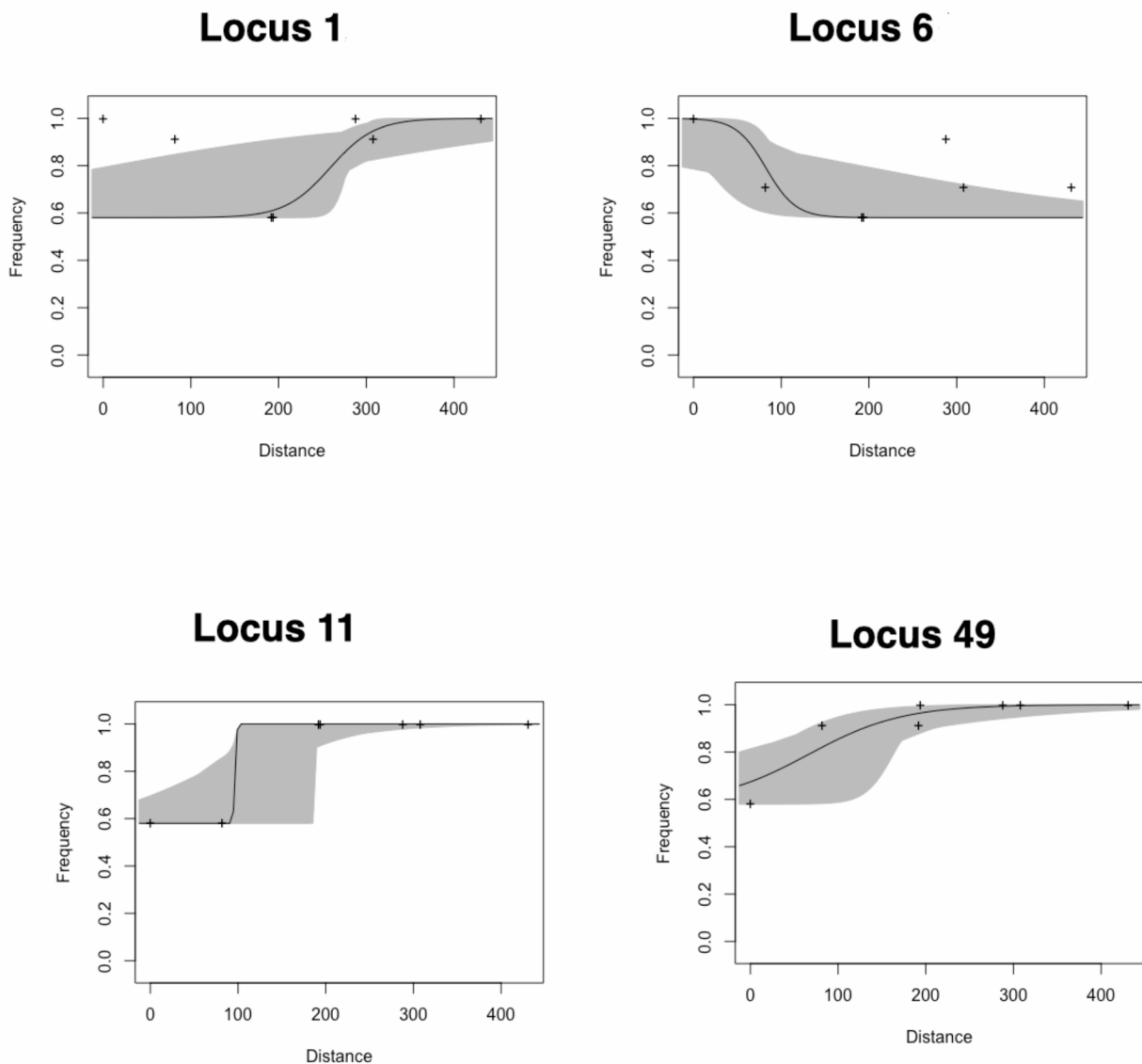
Fig. 7 Representative Hzar genetic clines with 95% confidence interval

14 = Precipitation of Driest Month, are the most important bioclim variables for the species *D. kotschyi,* while variables the same variables along with 13 = Precipitation of Wettest Month, and 8 = Mean Temperature of Wettest Quarter are important for the species *D. oligadenium,* and the hybrid population.

The variables which were identified as important for the present time distribution of the studied populations by random forest analysis based on the lowest Gini value were BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp)), BIO11 = Mean Temperature of

Coldest Quarter, BIO8 = Mean Temperature of Wettest Quarter, and BIO12 = Annual Precipitation.

Similarly, the important variables for plants' distribution in the year 2050 and due to climate change were BIO6 = Min Temperature of Coldest Month, BIO12 = Annual Precipitation, BIO5 = Max Temperature of Warmest Month, and BIO8 = Mean Temperature of Wettest Quarter. Therefore, a combination of variables related to precipitation and temperature determines *Dracocephalum*'s geographical distribution in the country.

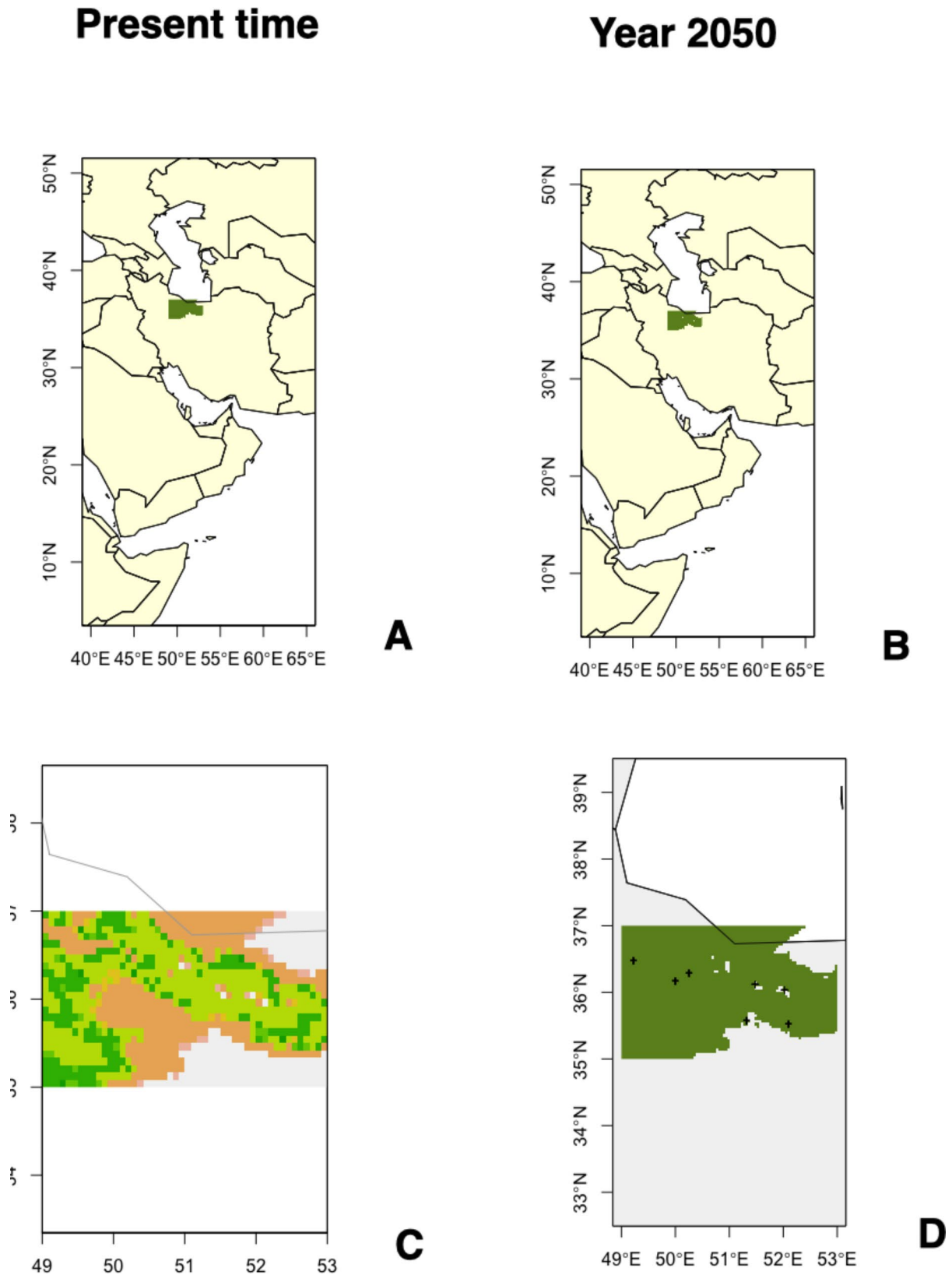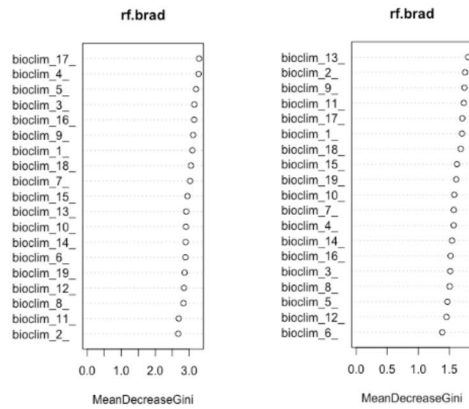Latent factor mixed model (LFMM) analysis of the genetic data in response to the current temperature and

Present time

Year 2050

Fig. 8 A-D = Representative figures of the four different modeling of predicted species distribution
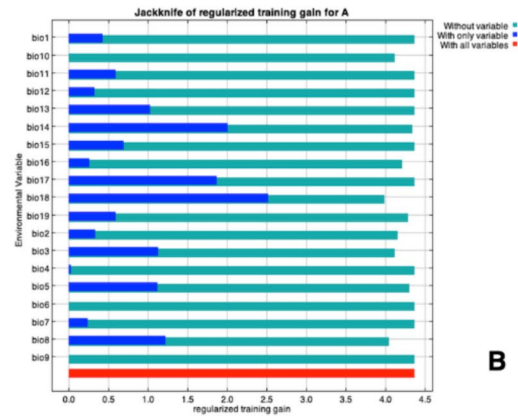
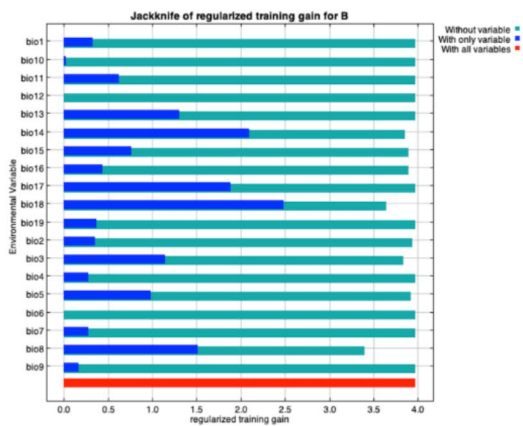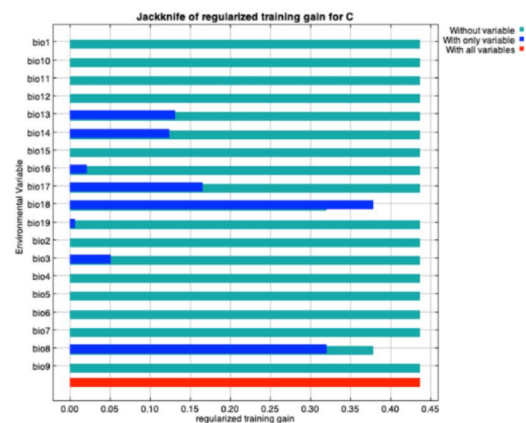**Fig. 9** The importance of bioclim variables determined by random forest (**A**), and Maxent models (**A**-**C**). Species A-C are A = *Dracocephalum kotschyi*, B = *D. oligadenium*, and C = The hybrid population

humidity and by year 2050, under influence of climate change, identified some of the ISSR genetic loci that are significantly associated with these variables and are of potential adaptive value (Fig. 10). Therefore, these genetic regions could be selected in *Dracocephalum* plants and utilized for conservation purpose.

Morphological cline analyses

Canonical correspondence analysis (CCA) and Mantel test produced significant results ($p < 0.001$). CCA plot identified morphological characters associated with geographical variables (Fig. 11). Length of the petiole in the basal leaf (X4), Length of the petiole in stem leaf (X7), and Size of the inflorescence leaf arista (X17), are associated with longitude. Similarly, the number of vines (X17), and The margin of stem leaves (X24), show association with the latitude, while, the length of petiole in inflorescence leaf (X10), and The width of bracteole (X13), are associated with the latitude.
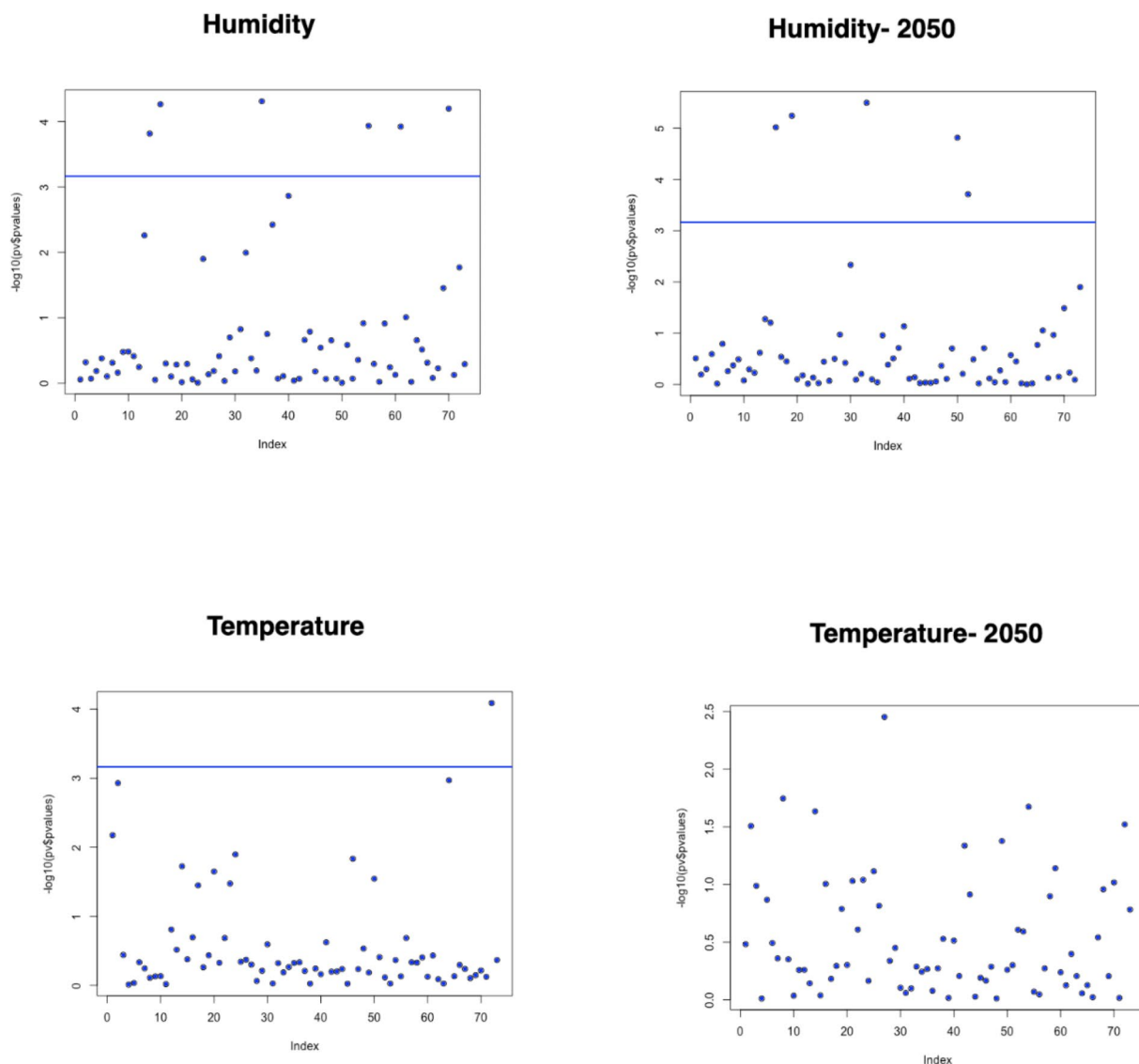
Fig. 10 The Manhattan plots of LFMM analysis identifying ISSR loci which are potentially adaptive to the current temperature and humidity as well as the year 2050

Mantel test after 9999 permutations revealed a significant association between morphological changes and the studied geographical variables. The representative plots of the maximum-likelihood cline and observed morphological data over the associated fuzzy cline region (95% credible cline region), are provided in Fig. 12, while the cline center and width of important characters are presented in Table 8.

In general, a smooth sigmoid cline curve was obtained for most of the morphological characters, which shows the selection acts on these quantitative traits. However, a few characteristics like the length of the inflorescence leaf, the margin of stem leaves, and the length of stamen, have a stepped cline shape, which indicates that either linkage disequilibrium combined with the dispersal of genes creates a stepped cline., and that most of the change in these traits' expression occurs in a narrow range in the middle of a cline.

Phylogenetic PCA results

Both pPCA and phyloPCA methods produced similar results (Fig. 13). In general, pPCA produced significant eigenvalues, and both global and local structuring of morphological features have occurred in the studied species and populations. From these analyses, it seems that the global structuring of the morphological characters has occurred in hybrid plants (C1-C3 plants in Fig. 13, B).
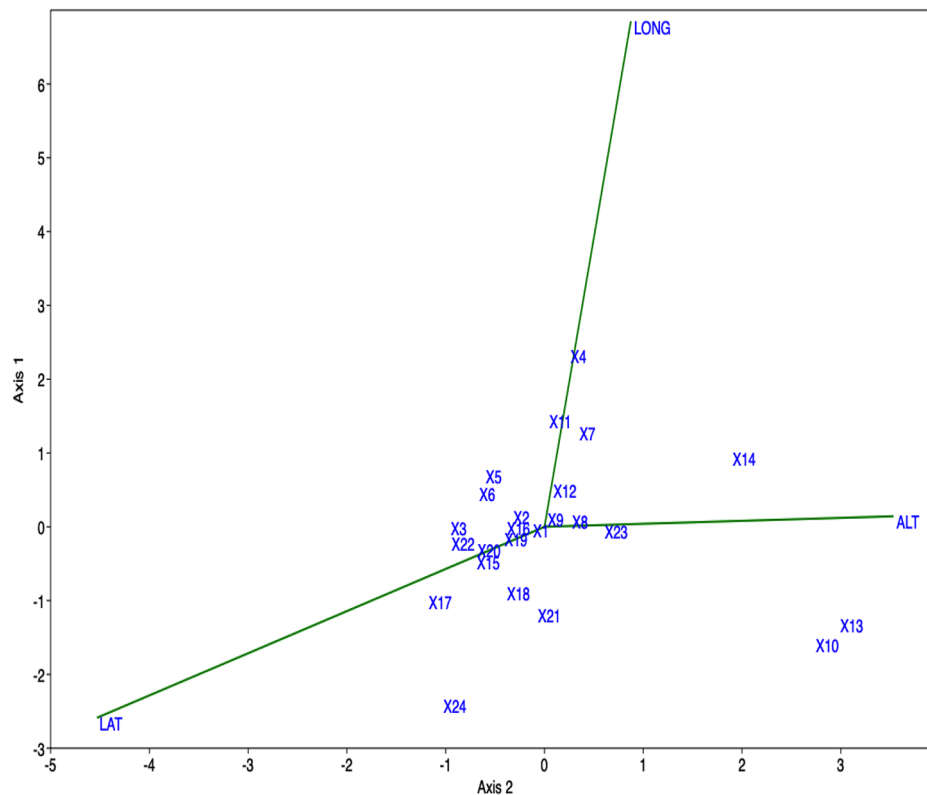
Fig. 11 CCA plot showing morphological characters that are significantly associated with geographical variables. (X1-X24 are morphological characters of the Table 7)

These hybrid plants occupy almost an intermediate position in the pPCA tree.

The most structured morphological features in the studied species are namely, the length of basal leaf (X2), length of stem leaf (X5), width of inflorescence leaf (X9), the number of vines (X17), and width of calyx (X23), respectively.

The Moran' I index estimated produced significant results ($p<0.0$), for morphological characters X2, X5 and X9 which indicates changes in these characters occurred similarly, in neighboring localities with similar environmental variables.

Discussion
Genetic diversity and gene flow
The present study revealed a low to moderate polymorphism in seven groups of *Dracocephalum* which were from two species namely *D. kotschyi* and, *D. oligadenium*, and their hybrid population.

A similar study performed on some other populations of *D. kotschyi*, within Iran using ISSR markers also reported within-population genetic variability with Gst and Nm mean indices of 0.27 and 1.31, as well as within and among the populations genetic variance was 74% and 26% respectively [30].

It seems the same scenario exists for the other species of the genus *Dracocephalum* reported from our country and elsewhere. For example, Koohdar et al. [31], studied the genetic diversity within populations of the medicinally important *Dracocephalum thymiflorum* L., which has also confined geographical populations in Iran, by ISSR molecular markers and reported that the studied populations contained a high within-population genetic variability and have a strong genetic differentiation by AMOVA test ($P=0.01$). Structure analysis and K-Means clustering revealed the occurrence of limited gene flow among these populations. Similarly, the genetic diversity investigation in another endangered related species of *Dracocephalum ruyschiana* L. in Norway by using SNPs data showed 4 genetic groups [32].

Finally, Dosta´lek et al. [33], by examining the genetic diversity of 12 populations of *Dracocephalum austriacum* L. in order to investigate its extinction risks in Europe, concluded that genetic diversity plays a vital role in seed production in this species.

The genetic variability of the geographical populations within a single species is important within the speciation events context, and as suggested by Knaus [34], If we take the species to be the unit of distinction, the infrataxa (the subspecies, the variety, and the ecotype) are
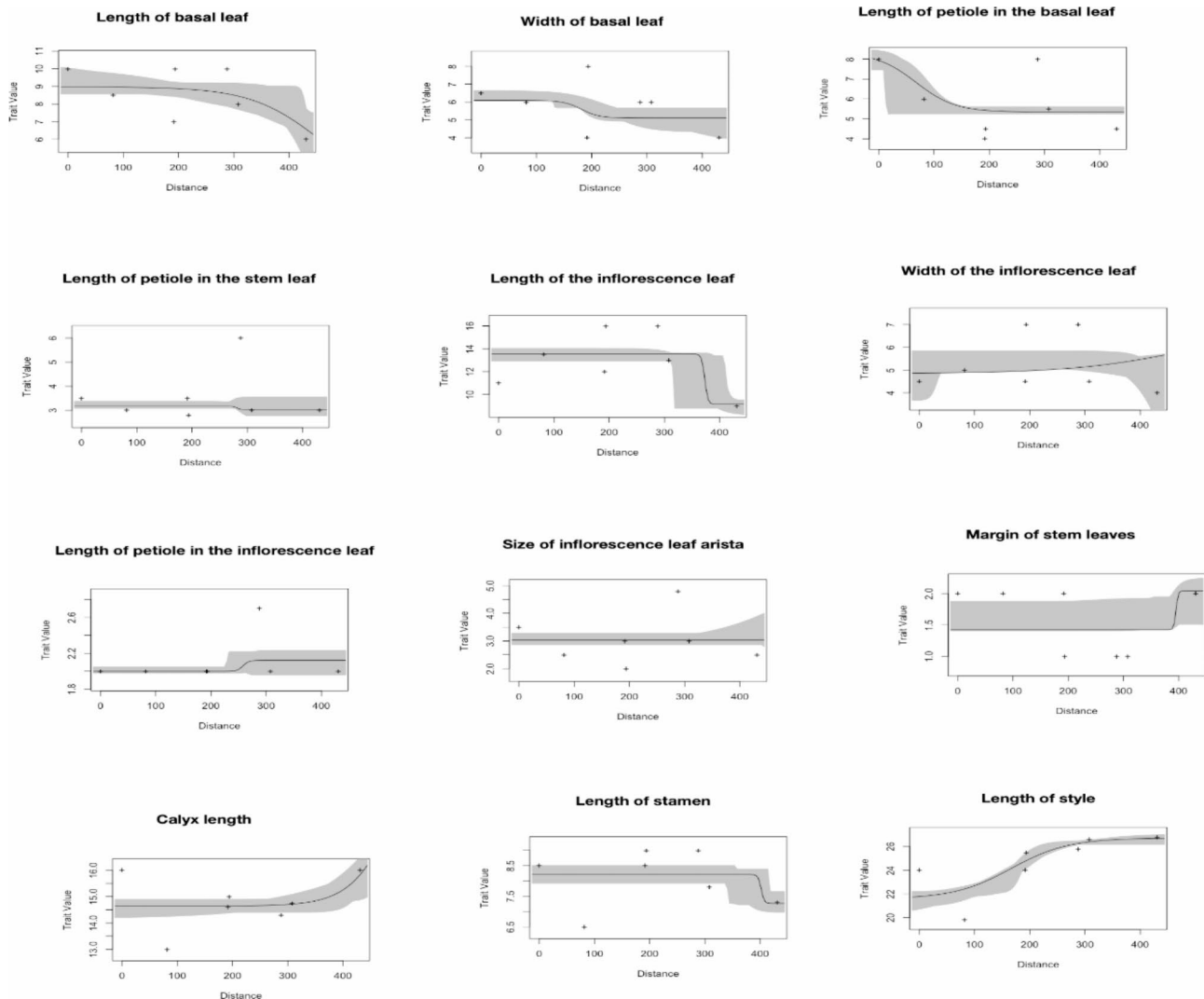
Fig. 12 Representative plots of the maximum-likelihood cline and observed morphological data over the associated fuzzy cline region (95% credible cline region)

Table 8 The cline center and width for important morphological characters

Characters	Mean center	Mean width
Length of basal leaf	443.8	363.1
Width of basal leaf	175.5	66.5
Length of petiole in the basal leaf	57.8	72
Length of petiole in the stem leaf	285.5	6.5
Length of the inflorescence leaf	362	39
Width of the inflorescence leaf	502.5	319
Length of petiole in the inflorescence leaf	262	20
Size of the inflorescence leaf arista	540	251
Margin of the stem leaves	367.5	25
Length of the stamen	378	40
Calyx length	523.5	198
Corolla length	176.5	157.5

consequently non-distinct. The process in which a group of organisms diverges from being one cohesive group to becoming two or more distinct groups is the process of speciation. Stebbins [35] also contributed the idea that species are systems of populations, which resemble each other, yet contain genetically different ecotypes that could be arranged in a continuous series. These allopatric infra-specific categories are usually recognized as infra-taxa. Such populations, and genetic diversity, may also be correlated with some important morphological features in some of the taxonomic species. For example, Ockendon [36], reported morphological diversity in several geographical populations of the *Linum perenne* group in Europe, which showed continuously changed features among populations with no sharp differences in some characters, but with a significant difference in some of the quantitative characters. PCA plot showed that the
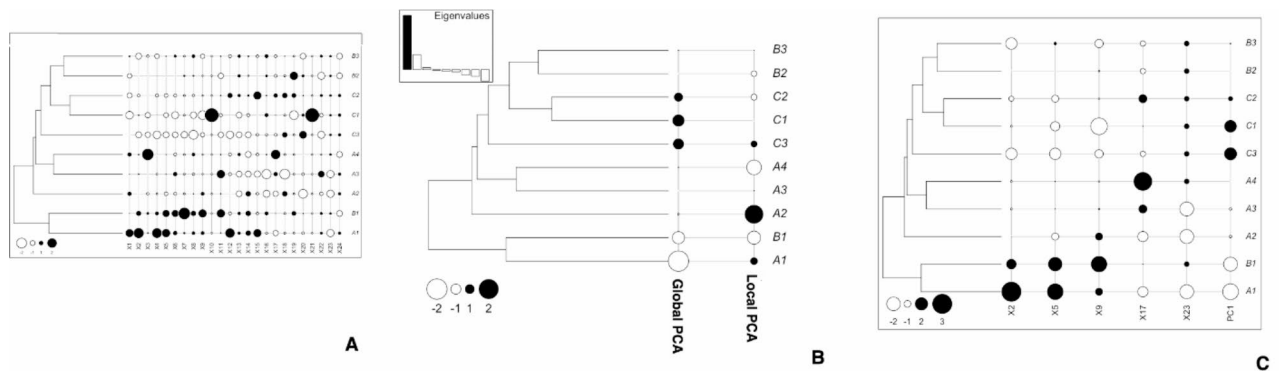
Fig. 13 Phylogenetic PCA results of morphological data. **A** = pPCA tree of all morphological characters and their pPCA loadings. **B** = pPCA showing global and local structuring of morphological characters. **C** = pPCA tree of most structured morphological characters among *Dracocephalum* species studied. Abbreviations: A1-A3, B1-B3, and C1-C3 are samples of the parental species and hybrid, respectively

different populations were separated from each other, and therefore, different geographical populations were considered to be ecotypes within each subspecies.

Landscape genetics and genetic cline analyses
In the present study sPCA analysis of geographical spatial features and their effect on the genetic structuring of diversity in *Dracocephalum*, populations revealed that both global and local spatial features affect such structure, and also identified the presence of spatial, geographical genetic clines within the studied populations. Moreover, different gene scan analysis methods used identified some of the genetic loci which are potentially adaptive to geographical variables. The genocline analyses also identified some genetic loci which contribute to species identification within *Dracocephalum* parental species, and also for establishing the hybrid population studied. The STRUCTURE assignment test also identified a higher genetic contribution of one population to the genetic structure of the hybrid plants.

In a similar study, Walsh et al. [37], carried out both geographic and genomic cline analyses in avian species and reported variable patterns of introgression among marker types. Most of the markers studied exhibited gradual clines, while, steeper clines and therefore, a strong selection were observed for the loci associated with traits related to tidal marsh adaptations. However, narrow clines occurred at mitochondrial and sex-linked markers. Patterns of asymmetrical introgression were noticed, which may be driven by differences in mating strategy or differences in population density between the two species.

Hybridization and introgression are important forces that can shape the evolution of a particular species, and obtaining knowledge on this aspect of evolution is achieved by studying taxa that maintain a genetic distinction with ongoing gene flow [37]. Similarly, studying the genetic architecture of hybrid zones provides knowledge

to understand the sequence of genetic changes that accompany or facilitate speciation and produce data on the interplay between gene flow as homogenizing force facilitating the spread of alleles and maintaining species as cohesive units, and natural selection promoting population divergence [38].

The present study produced the introgression value as revealed by the H index for only a few markers, while fixation of parental alleles ($S=1$), was obtained for most of the markers studied.

Diagnostic markers (i.e. markers that are fixed or highly differentiated between two parental species) are predicted to be under divergent selection, exhibiting reduced introgression, and are likely to be associated with genomic regions under selection [37].

HZAR genetic clines analysis of the hybrid zone cline produced a variable cline mean center for the genetic loci studied ranging from 100 Km for locus 49, to 370 Km, for locus 16. However, the values obtained for the cline width in the studied genetic loci, were relatively more similar for some of the loci, while ranging from 135 Km to 400 Km, for other genetic loci studied. As stated by Walsh et al. [37], A wider cline width, may indicate extensive introgression and recombination within and well outside of the overlap zone, but the clines varying in width but constrained to the same center are indicative of differential introgression across the hybrid zone.

The width of a cline in gene frequency is approximately proportional to gene flow (sigma) divided by the square root of per-locus selection (square root of s). Gene flow also causes gametic correlations (linkage disequilibria) between genes that differ across hybrid zones. Correlations are stronger when the hybrid zone is narrow, and rise to a maximum roughly equal to s. Thus, cline width and gametic correlations combine to give estimates of gene flow and selection [39].

We found differential introgression values for the genetic loci studied. This variable introgression across

markers suggests that while most loci exhibit uninhibited movement and sharing possibly through gene exchange among the nearby plants (as also revealed by significant Mantel test and IBD (isolation by distance)), certain loci do not freely cross the species boundaries and therefore may be important in reproductive isolation. These loci are probably those which contribute to the species and populations differentiation as revealed by Gst value and the loci with a higher contribution to sPCA geographical spatial genetic structuring and by AFLPism gene scan presented before.

We observed almost a wide cline with a moderate magnitude of gene flow and genetic polymorphism in the *Dracocephalum* populations studied.

It is a well-known fact that the balance between selection on hybrids and dispersal of genes determines cline width. The clines are expected to be narrow when the species are facing strong selection against the hybrids, or, if gene flow is restricted only among a few individuals, or if there are steep environmental gradients. Similarly, the clines will be wider if the selection is weak, gene flow is more extensive, or environmental gradients are gradual [39].

The cline shapes we obtained was a smooth sigmoid curve. These cline shapes are expected to be present when the selection force acts on single genes or quantitative traits [39]. In contrast, linkage disequilibrium combined with the dispersal of genes can distort the smooth shape of the cline, creating a stepped cline in which most of the change in allele frequency or trait expression occurs in a narrow range in the middle of a cline.

The two species of *Dracocephalum* were studied and the hybrid population was growing in distinct, separate geographical areas. It is expected that in these situations mosaic hybrid zones occur which can be as widely distributed as the distribution of habitats and parental species, and therefore, the hybrid genotypes are either not strongly associated with specific habitats or occupy intermediate or novel habitats. This may result in a wider geographical distribution of the genetic variability, within closely related species [39]. A similar situation has been reported for two species of sunflower namely, *Helianthus annuus* L. and *H. petiolaris* Nutt. which have overlapping ranges and form local hybrid swarms in the western United States, and for the hybrids between *Salix sericea* Marsh. and *S. eriocephala* Michx., with the hybrids being found throughout the sympatric range of these species in eastern North America [39].

The present study identified very few ISSR loci with S=0.5, and H=1, indicative of heterozygosity in the hybrid zone plants. Almost all other loci had ancestry value (S)=1.

It is suggested that maximizing the log-likelihood for dominant markers like ISSR gives an unbiased estimate of S and H in the studies performed on a sufficiently large number of markers. Though these dominant markers are less informative about heterozygosity, work well as long as there is a mixture of loci for which the dominant allele is more common in ancestral species 1 and other loci for which the dominant allele is more common in ancestral species 2 [25].

Different species distribution modeling (SDM) methods, used in this work revealed that bioclimatic variables related to the temperature and moisture, play an important role in the *Dracocephalum* population's geographical distribution within the country and that due to the presence of some potentially adaptive genetic loci in the studied plants, they can survive well enough by the year 2050 and under climate change. This is a piece of good news for the conservation of these medicinally important plants and based on genetic data obtained a plan for conservation and management of them may be suggested to the organizations involved in the future.

Morphological cline
The present study revealed that the hybrid zone cline affects morphological characters in *Dracocephalum* species and populations studied. We obtained both a smooth sigmoid cline curve as well as stepped cline shape curves for morphological characters. Therefore, selection acts on most of the morphological characters, but for a few characters, most of the change in these traits' expression occurs in a narrow range in the middle of a cline.

The hybrid zones' investigations are mostly based on the neutral markers to measure cline width and linkage disequilibrium, which are combined with variation at phenotypic traits and in this way identify the traits that are involved in reproductive isolation or estimating levels of selection [40].

Comparing the genetic and trat clines that are under selection indicates the extent to which the overall genome is under indirect selection. It is expected that the unlinked neutral alleles may pass across the clines (neutral diffusion) unless many loci are under strong disruptive selection. In this case, the genome can not introgress (see for example [41],

HZAR morphological cline curves were obtained after estimating the variance and the covariance between quantitative traits. These in turn yield information about the genetic basis of quantitative trait variation, specifically the effective number of loci contributing to the difference between species, and provide an estimate of dispersal that is directly comparable to estimates of dispersal obtained using linkage disequilibrium between molecular markers, respectively [40].

In conclusion, the present study reports new findings on the genetic structure of a hybrid population between two *Dracocephalum* species and provides data on both

the geographical and genetic clines of these populations. We also identified genetic loci that are potentially adaptive to the geographical spatial features and genocline condition. The present findings may be used to conserve these medicinally important plant taxa.

Abbreviations

AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
CCA	Canonical correspondence analysis
DAPC	Discriminant analysis of principal components
GLM	General linear model
Gst	Genetic differentiation
Gtest	Global test
HSBU	Herbarium of Shahid Beheshti University
IBD	Isolation by distance
ISSR	Inter-Simple Sequence Repeat
LDA	Linear discriminant analysis
LEA	Landscape and ecological association studies
LFMM	Latent factor mixed model
Ltest	Local test
MCMC	Markov chain Montacarlo
Nm	The number of migrants (Gene flow)
P%	Polymorphism percentage
Pas	Pseudo-absences
PCA	Principal component analysis
Pcadapt	Performs principal component analysis
PCoA	Principal coordinate analysis
pPCA	The phylogenetic Principal Component Analysis
RCP	Representative Concentration Pathways
RDA	Redundancy analysis
SDM	Species distribution modeling
SDMs	Species distribution modeling
sPCA	Spatial PCA method

Data availability
All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate
FK undertook the formal identification of the plant material used in our study. The authors have complied with all relevant institutional and national guidelines and legislation in experimental research and field studies on plants, including the collection of plant materials for this study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran.

References

1. Brach AR, Song H, eFloras. New directions for online floras exemplified by the flora of China project. Taxon. 2006;55(1):188–92.
2. Sajjadi SE, MovahedianV, Attar, Yektaeian A. Antihyperlipidemic effect of hydroalcoholic extract, and polyphenolic fraction from *Dracocephalum Kotschyi* Boiss. Pharm Acta Helv. 1998;73:167–70.
3. Jahaniani F, Ebrahimi SA, Rahbar-Roshandel N, Mahmoudian M. Xanthomicrol is the main cytotoxic component of *Dracocephalum Kotschyi* and a potential anticancer agent. Phytochemistry. 2005;66:1581–92.
4. Naderifar M, Sonboli A, Gholipou A. Pollen morphology of Iranian *Dracocephalum* L. (Lamiaceae) and its taxonomic significance. Bangladesh J Plant Taxon. 2015;22:99–110.
5. Zeng Q, Jin HZ, Qin JJ, Fu JJ, Hu HJ, Li JH, Yan L, Chen M, Zhang DW. Chemical constituents of plants from the genus *Dracocephalum*. Chem Biodivers. 2010;7:1911–29.
6. Chen Ya P, Turdimatovich Tu O, Nuraliev MS, Lazarević P, Drew BT, Xiang Ch L. Phylogeny and biogeography of the northern temperate genus *Dracocephalum* s.l. (Lamiaceae). Cladistics. 2022;38(4):429–51.
7. Rechinger KH. Dracocephalum. In: Rechinger KH, editor. Flora Iranica. Akademische Druck-U: Verlagsanstalt, Graz, Austria.; 1982. pp. 218–30.
8. Koohdar F, Attar F, Talebi SM, Sheidai M. Contemporary interspecific hybridization between *Dracocephalum Kotschyi* and *Dracocephalum oligadenium* (Lamiaceae): evidence from morphological, anatomical and molecular data. Acta Biol Szeged. 2018;62(2):123–9.
9. Jamzad Z, Dracocephalum L. Flora of Iran. Iran: Research Institute of Forests and Rangelands; 2012.
10. Freeland JR, Kirk H, Peterson SD. Molecular Ecology. UK: Wiley-Blackwell; 2011.
11. Gobert V, Moja S, Colson M, Taberlet P. Hybridization in the section *Mentha* (Lamiaceae) inferred from AFLP markers. Am J Bot. 2002;89:2017–23.
12. Idowu JA, Oziegbe M. Mitotic and meiotic studies on two species of *Ocimum* (Lamiaceae) and their F1 hybrids. Bot Lith. 2017;23(1):59–67.
13. Mamadalieva NZ, Sharopov F, Satyal P, Azimova SS, Wink M. Composition of the essential oils of three Uzbek *Scutellaria* species (Lamiaceae) and their antioxidant activities. Nat Prod Res. 2017;31:1172–6.
14. Sork VL. Gene flow and natural selection shape spatial patterns of genes in tree populations: implications for evolutionary processes and applications. Evol Appl. 2016;9:291–310.
15. Genetics KAWL. Landscape effects on gene flow and population genetic structure. Essentials of Landscape Ecology. (Oxford; online edn, Oxford Academic, 22 Aug. 2019), accessed 21 Apr. 2023.
16. Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP. Putting the 'landscape' in landscape genetics. Heredity. 2007;98(3):128–42.
17. Guan B, Gao J, Chen W, Gong X, Ge G. The effects of climate change on landscape connectivity and genetic clusters in a small subtropical and warm-temperate tree. Front Plant Sci. 2021;10:671336.
18. Derryberry EP, Derryberry GE, Maley JM, Brumfield R. HZAR: hybrid zone analysis using an R software package. Mol Ecol Resour. 2013;14(3):652–63.
19. Gompert Z, Buerkle CA. INTROGRESS: a software package for mapping components of isolation in hybrids. Mol Ecol Resour. 2010;10:378–84.
20. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–59.
21. Legendre P, Legendre L. Numerical ecology. Elsevier; 2012.
22. Frichot E, Schoville SD, Bouchard G, François O. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol. 2013;30(7):1687–99.
23. Jombart T, Devillard S, Dufour AB, Pontier D. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity. 2008;101(1):92–103.
24. Balao F, García-Castaño JL, AFLPsim. An R package to simulate and detect dominant markers under selection in hybridizing populations. Plant Methods. 2014;10:40.
25. Fitzpatrick BM. Estimating ancestry and heterozygosity of hybrids using molecular markers. BMC Evol Biol. 2012;12:131.
26. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190(3–4):231–59.
27. Revell LJ. Size-correction and principal components for interspecific comparative studies. Evolution. 2009;63:3258–68.

28. Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3:217–23.

29. Jombart T, Pavoine S, Devillard S, Pontier D. Putting phylogeny into the analysis of biological traits: a methodological approach. J Theor Biol. 2010;264:693–701.

30. Maskani Sereshkeh F, Azizi A, Noroozisharaf A. Structure of genetic diversity among and within populations of the endemic Iranian plant *Dracocephalum kotschyi*. Hortic Environ Biotechnol. 2019;60:767–777.

31. Koohdar F, Sheidai M, Attar F, Talebi Seyed-Mehdi. Population genetic structure and genetic diversity in *Dracocephalum thymiflorum* L. (Lamiaceae) populations in Iran. Mol Plant Breed. 2015;6(19):1–7.

32. Kyrkjeeide MO, Westergaard KB, Kleven O, Evju M, Endrestøl M, Brandrud MK, Stabbetorp O. Conserving on the edge: genetic variation and structure in northern populations of the endangered plant *Dracocephalum ruyschiana* L. Conserv Genet. 2020;21:707–18. Lamiaceae.

33. Dostálek T, Münzbergová Z, Plačková I. Genetic diversity and its effect on fitness in an endangered plant species, *Dracocephalum austriacum* L. Conserv Genet. 2010;11:773–83.

34. Knaus BJ. A fistful of astragalus: phenotypic and genotypic basis of the most taxon-rich species in the North American flora. Ph.D. USA;, Thesis. Oregon State University; 2008.

35. Stebbins GL. Concepts of species and genera. In: Editorial C, editor. Flora of North America. Volume 1. New York: Oxford University Press; 1993. pp. 229–46. 11.

36. Ockendon DJ. Taxonomy of the Linum perenne group in Europe. Watsonia. 1971;8:205–35.

37. Walsh J, Shriver WG, Olsen BJ, Kovach AI. Differential introgression and the maintenance of species boundaries in an advanced generation avian hybrid zone. BMC Evol Biol. 2016;22:16:65.

38. De La Torre A, Ingvarsson P, Aitken S. Genetic architecture and genomic patterns of gene flow between hybridizing species of *Picea*. Heredity. 2015;115:153–64.

39. Mallet J, Barton N, Lamas G, Santisteban J, Muedas M, Eeley H. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in heliconius hybrid zones. Genetics. 1990;124(4):921–36.

40. Gay L, Crochet PA, Bell DA, Lenormand Th. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. Evolution. 2008;62:2789–806.

41. Grahame JW, Wilding CS, Butlin RK. Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. Evolution. 2006;60:268–78.

Publisher's Note