Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2022) 46: 214-223
© TÜBİTAK
doi:10.55730/1300-011X.2972

Inter-primer binding site (iPBS) retrotransposon markers provide insights into the genetic diversity and population structure of carrots (*Daucus*, Apiaceae)

Mehtap YILDIZ[1], **Carlos I. ARBIZU**[2,*]
[1]Department of Agricultural Biotechnology, Faculty of Agriculture, Van Yüzüncü Yıl University, Van, Turkey
[2]Directorate of Agrarian Technological Development, National Institute of Agrarian Innovation (INIA), Lima, Peru

Received: 17.10.2021 • **Accepted/Published Online:** 23.02.2022 • **Final Version:** 25.04.2022

Abstract: Studies of genetic diversity and population structure are essential as an initial step in conservation and breeding programs for modern crops. Carrot (*Daucus* spp.) is among the ten most important vegetables worldwide, however, its genetic structure and phylogenetic relationships are not totally deciphered. Here, we explored the utility of 21 inter-primer binding site (iPBS) retrotransposon markers to determine the genetic diversity and population structure of 38 accessions of *Daucus* and one accession of a related genus. The manual scoring revealed 309 bands based on their presence/absence. The dendrogram based on the UPGMA clustering algorithm and a principal coordinate analysis (PCoA) indicated the presence of four clusters. The *Daucus* species with $2n = 18$ chromosome (subclade A´) separated from the other two species *D. pusillus* and *D. muricatus,* which were positioned into two individual clusters. The other clade includes the *Daucus* from the B group. It was also noticed that few accessions were intermixed amongst clusters. Different genetic diversity parameters were estimated based on the four clusters (populations) defined by STRUCTURE software, demonstrating that clusters 3 and 4 possessed the lowest and highest diversity values, respectively. AMOVA showed variation between and within clusters of 41.85% and 58.15%, respectively. The highest population divergence (F_{st}) was observed between clusters 2 and 3 (0.579), on the other hand, clusters 1 and 4 depicted the lowest F_{st} with 0.160. Our research highlighted that iPBS markers were successful and effective to study *Daucus* genetic diversity. These results will contribute to the genetic improvement of carrots and sustainable management of its diversity.

Key words: germplasm, carrot, iPBS, *Daucus,* genetics

1. Introduction

The cultivated carrot (*Daucus caro*ta L. subsp. *sativus* Hoffm.) has high nutritional value and great economic importance. They are grown on approximately 1.5 million hectares of land worldwide per year (FAOSTAT, 2019). It is a very important source of precursors of Vitamin A and other beneficial metabolites for human health (Simon et al., 2009). Information on genetic diversity in crop species is essential for its preservation and breeding work (Comertpay et al., 2012). For multiple species, repeated controlled breeding cycles over thousands of years have resulted in limited genetic pool (Tanksley and McCouch, 1997). Different approaches at the molecular level were employed to determine the genetic diversity and relationships of carrots (*D. carota*). St. Pierre et al. (1990) failed to categorize 168 accessions of the *D. carota* from 32 countries using molecular forms of isozymes. Random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphisms (AFLP) were used by Nakajima et al. (1998) to group all accessions of *D. carota* into a major clade. Vivek and Simon (1998, 1999) employed restriction fragment length polymorphisms (RFLPs) of nuclear and organelles DNA and noticed their results were in agreement with the classification stated by Sáenz (1981). Shim and Jørgensen (2000) used AFLPs markers and demonstrated that wild and cultivated accessions of carrots grouped separately. Rong et al. (2014) found that subspecies of *D. carota* were intermingled when they reconstructed the *Daucus* phylogeny using single nucleotide polymorphisms (SNPs). Consistent with molecular studies, morphological studies did not differentiate *D. carota* at the subspecies level (Small, 1968; Arbizu et al., 2014a; Mezghani et al., 2014; Spooner et al., 2014; Tavares et al., 2014). On the other hand, successful differentiation between wild (subsp. *carota*) and cultivated (subsp. *sativus*) accessions were reported using SNP markers (Iorizzo et al., 2013). Similarly, in a larger study, Arbizu et al. (2016a) distinguished wild vs. cultivated accessions of carrots using 18,565 SNPs spanning nine chromosomes belonging to 162 accessions of *Daucus* and

* Correspondence: carbizu@inia.gob.pe

This work is licensed under a Creative Commons Attribution 4.0 International License.

two related genera. They also reported that the SNP data are applicable for phylogenetic studies of *Daucus* at the species level. Mezghani et al. (2018) conducted a study with Tunisian accessions of *D. carota* and SNPs generated by genotyping-by-sequencing (GBS). They demonstrated that *D. carota* subsp. *gummifer* has different origins from other collections of subsp. *carota* in Tunisia and some other close areas to Tunisia. In a more recent study, Martínez-Flores et al. (2020) analyzed more than two hundred accessions of subspecies of *Daucus carota* with 29,041 SNPs by GBS. They identified five independent morphotypes of *gummifer* in marine environments along the Mediterranean and near-Atlantic coasts.

DNA molecular markers have always been invaluable tools for understanding the nature of genetic diversity (Nadeem et al., 2018). During the last two decades, many molecular tools were proposed such as the retrotransposon-based DNA marker, iPBS (Kalendar et al., 2010). The iPBS is a retrotransposon marker system based on the amplification of the region covered by binding sites of the reverse transcriptase primer for two contiguous retrotransposons that are in opposite orientations (Kalendar et al., 2010). Besides being used for both, plant and animal kingdoms, iPBS is an advantageous DNA fingerprinting technique in many ways as it does not require prior sequence knowledge (Demirel et al., 2018; Karık et al., 2019). This genetic marker was successfully used in many studies including a wide range of crops such as *Cicer* (Andeden et al., 2013), *Vicia faba* (Baloch et al., 2014), *Phaseolus vulgaris* (Nemli et al., 2015), *Pisum sativum* (Baloch et al., 2015), *Myrica rubra* (Fang-Yong et al., 2014), *Solanum tuberosum* (Shah et al., 2015), *Nicotiana tabacum* (Yaldiz et al., 2018), among others.

The aim of this study was to evaluate the applicability of iPBS to characterize 38 accessions of *Daucus*, making a contribution to the conservation of its genetic diversity, and potential utilization of new carrot germplasm for breeding work.

2. Materials and methods

2.1. Study species

Thirty-eight genotypes of *Daucus* and one sample of a related outgroup genera, *Anthriscus* were examined (39 samples in total). The *Daucus* samples include the two main clades, A (contains A´) and B (Spooner et al., 2013; Arbizu et al., 2014b, 2016a, b). Twenty-six accessions labelled with "Ames" or "PI" were obtained from the United States National Plant Germplasm System maintained at the North Central Regional Plant Introduction Station (NCRPIS) in Ames, Iowa. The other 13 accessions were collected by the main author of this work (M.Y.) from different localities in Turkey (Table 1).

2.2. DNA amplification

The genomic DNA was extracted by CTAB method (Doyley & Doyle, 1987) with minor modifications incorporated by Boiteux et al. (1999) from five lyophilized leaf samples of each accession. A total of 83 IPBS primers were evaluated on four randomly selected *Daucus* accessions through PCR. As a result of this initial screening, 21 iPBS primers producing dense and polymorphic bands were selected for further analysis in our germplasm library. PCR reactions were performed according to conditions specified by Kalendar et al. (2010) in a final volume of 20 µL containing 0.2 U Taq DNA polymerase, 4 µM primer, 2 mM dNTPs,3 ng/µL template DNA, 1X PCR buffer, and 7 µL distilled water. After electrophoresis of PCR products with a 100-bp ladder (Thermo Fisher Scientific), the bands were visualized using an 'Imager Gel Doc XR+' system (Bio-Rad, USA).

2.3. IPBS marker data analysis

The IPBS band patterns were scored visually for the presence (1) or absence (0) of various molecular weight sizes. For the analysis, only polymorphic and reproducible bands were utilized. Loci with more than 10% missing data were excluded from the analysis according to the procedure given by Saldaña et al. (2021). Polymorphic information content (PIC) from dominant markers was calculated by following the equation:

$$PIC = 1 - [fi^2 + (1-fi)^2]$$

Where, fi is the frequency of amplified band (1) and (1-fi) is the frequency of absence of band (0). The genetic distances were calculated using R software v4.0.2. based on provesti's algorithm. Then, a dendrogram was constructed using the UPGMA clustering algorithm. The function *aboot* was conducted from *poppr* package v2.9.2 to apply 1000 bootstrap replicates.

The STRUCTURE program v2.3.4 was used to determine population structure (Pritchard et al., 2000). For this purpose, ten runs for each number of populations (K value) ranging from 1 to 15, using a burn-in length of 50,000 interactions followed by 150,000 Monte Carlos Markov Chain were applied. All parameters were set to default values except for the admixture model (no previous population information was considered). The most likely number of clusters was calculated by the Evanno method (Evanno et al., 2005). Membership probabilities ≥0.7 or the maximum membership probability was adopted to divide the accessions into different clusters. The R package *pophelper* v2.3.1 (Francis, 2017) was used to obtain population structure plots.

The number of populations determined by STRUCTURE software was used to conduct an analysis of molecular variance (AMOVA) using the R package *poppr*. Three genetic diversity indices were calculated using the same package: (i) Simpson's index, (ii) Shannon-

Table 1. The 38 accessions of *Daucus*, and one accession of related genera characterized in this study, improvement status, locality information and new identification.

Taxon[a]	Accessions[b]	Improvement status[c]	Location or Source[d]	Cluster assignment[e]
Anthriscus nemorosa		Wild	Van, Turkey	4
Daucus bicolor	Ames 25807	Wild	İzmir, Turkey	4
D. *carota* subsp. *carota*	1	Wild	Antalya, Turkey	1
D. *carota* subsp. *carota*	3	Wild	Van, Turkey	1
D. *carota* subsp. *carota*	2	Wild	Van, Turkey	1
D. carota subsp. *sativus*	1255020	Cultivated	Turkey	1
D. carota subsp. *sativus*	12550318	Cultivated	Turkey	1
D. carota subsp. *sativus*	349-1	Cultivated	Turkey	1
D. carota subsp. *sativus*	913-1	Cultivated	Turkey	1
D. carota subsp. *sativus*	5	Cultivated	Turkey	1
D. carota subsp. *sativus*	6	Cultivated	Turkey	1
D. carota subsp. *sativus*	7	Cultivated	Turkey	1
D. conchitae	PI 652385	Wild	Antalya, Turkey	4
D. conchitae	PI 652366	Wild	Muğla, Turkey	4
D. guttatus	Ames 25729	Wild	Qastal, Syria	4
D. guttatus	PI 652343	Wild	Syria	4
D. guttatus	PI 652339	Wild	Syria	4
D. guttatus	Ames 25724	Wild	Syria	4
D. guttatus	PI 279763	Wild	Jerusalem, Israel	4
D. muricatus	Ames 29090	Wild	Ben Arous, Tunisia	3
D. muricatus	Ames 31608	Wild	Fès-Meknès, Morocco	3
D. muricatus	Ames 31609	Wild	Fès-Meknès, Morocco	3
D. muricatus	Ames 31612	Wild	Sidi Kacem, Morocco	3
D. muricatus	Ames 31614	Wild	Chefchaouebe, Morocco	3
D. muricatus	Ames 31615	Wild	Tangier, Morocco	2
D. pusillus	PI 287113	Wild	Florida, Uruguay	2
D. pusillus	PI 341892	Wild	Uruguay	2
D. pusillus	PI 349267	Wild	Montevideo, Uruguay	2
D. pusillus	PI 661242	Wild	Oregon, United States	2
D. pusillus	Ames 29891	Wild	California, USA	2
D. setulosus	PI 652329	Wild	Peloponnese, Greece	4
D. setulosus	PI 652327	Wild	Central Greece, Greece	4
D. setulosus	PI 652331	Wild	Peloponnese, Greece	4
D. setulosus	PI 652360	Wild	Muğla, Turkey	4
D. setulosus	Ames 25596	Wild	Central Greece, Greece	4
D. setulosus	Ames 25608	Wild	Central Greece, Greece	4
Daucus aureus	PI 478858	Wild	Dijon, France	4
Daucus spp.	2	Wild	Mersin, Turkey	4
Daucus spp.	1	Wild	Mersin, Turkey	4

[a] These names correspond to those in the Germplasm Resources Information Network (GRIN) database.
[b] Plant Introduction (PI) numbers are permanent numbers assigned to germplasm accessions in the National Plant Germplasm System (NPGS). Ames numbers are assigned for carrots and other Apiaceae maintained at the North Central Regional Plant Introduction Station in Ames, Iowa, USA.
[c] It refers to whether the accessions are cultivated or wild.
[d] Location refers to where the germplasm was collected in the wild.
[e] Cluster assignment based on STRUCTURE analysis.

Wiener index, and (iii) Nei's gene diversity (expected heterozygosity). The degree of gene differentiation among clusters was estimated in terms of allele frequencies (Fst) by using the following formula:

$Fst = 1 - (H_s/H_t)$

Where, HS = average expected heterozygosity estimated from each cluster, and H_t = total gene diversity or expected heterozygosity in the total cluster as estimated from the pooled allele frequencies.

3. Results

3.1. iPBS marker analysis

The 21 iPBS primers used during the present investigation resulted in a total of 309 bands with a range of 5 to 23 bands per primer in 38 samples of *Daucus* + one related species (*Anthriscus nemorosa*), with 14.6 fragments as average (Table 2). Of the 309 bands, 306 (99.03%) were polymorphic in our germplasm collection of carrots. The PIC (polymorphic information content) values varied from 0.841 (iPBS2390) to 0.969 (iPBS2238), while 0.919 was mean PIC value (Table 3).

Table 2. The 21 iPBS primer marker sequences and annealing temperature.

Primer	Sequence (5'-3')	Annealing temperature (°C)
2229	cgacctgttctgatacca	52
2230	tctaggcgtctgatacca	53
2232	agagaggctcggatacca	55
2238	acctagctcatgatgcca	55
2239	acctaggctcggatgcca	55
2249	aaccgacctctgatacca	51
2251	gaacaggcgatgatacca	53
2253	tcgaggctctagatacca	51
2272	ggctcagatgcca	55
2277	ggcgatgatacca	52
2374	cccagcaaacca	53
2375	tcgcatcaacca	50
2383	gcatggcctcca	53
2388	ttggaagaccca	50
2390	gcaacaacccca	55
2395	tccccagcggagtcgcca	53
2400	cccctccttctagcgcca	51
2220	acctggctcatgatgcca	57
2074	gctctgatacca	50
2095	gctcggatacca	53
2228	cattggctcttgatacca	53

Table 3. Polymorphic information content (PIC) for 21 iPBS markers.

Primer	Total	Polymorphic	% Polymorphism	PIC
2229	16	15	93.75	0.931
2230	13	13	100	0.873
2232	23	23	100	0.958
2238	20	19	95	0.969
2239	16	16	100	0.940
2249	14	14	100	0.924
2251	13	13	100	0.912
2253	23	23	100	0.940
2272	19	19	100	0.948
2277	10	10	100	0.860
2374	11	11	100	0.918
2375	11	11	100	0.896
2383	17	17	100	0.926
2388	5	5	100	0.849
2390	11	10	90.91	0.841
2395	15	15	100	0.938
2400	11	11	100	0.961
2020	14	14	100	0.950
2074	15	15	100	0.931
2095	12	12	100	0.885
2228	20	20	100	0.958
Total	309	306		
Average	14.7	14.6	99.03%	0.919

3.2. Genetic diversity and population structure

A data matrix containing 39 x 309 presence/absence data was manually generated. Phylogenetic tree using Provesti's genetic distances separated *Daucus* into two clades (A and B) with bootstrap support (BS) lower than 70%. In addition, population structure analysis identified four clusters. Individuals belonging to the *D. carota* complex having 2*n* = 18 chromosomes (subclade A′) were placed with 99% BS in cluster 1. All accessions of *D. pusillus formed* cluster 2 with 100% BS; however, accession *D. muricatus* 31615 was placed within it as well. Cluster 3 comprised *D. muricatus* accessions with 100% BS. Only members of the *D. guttatus* complex except *D. aureus* 478858 + *A. nemorosa* were grouped in cluster 4 (Figure 1). The Evanno method (Evanno et al., 2005) showed that the best K value is four for our data set. Our dendrogram is in agreement with STRUCTURE analysis, which exhibited admixture for very few samples (Figure 1).

Genetic diversity and F_{st} estimation were also calculated for the four populations identified by STRUCTURE

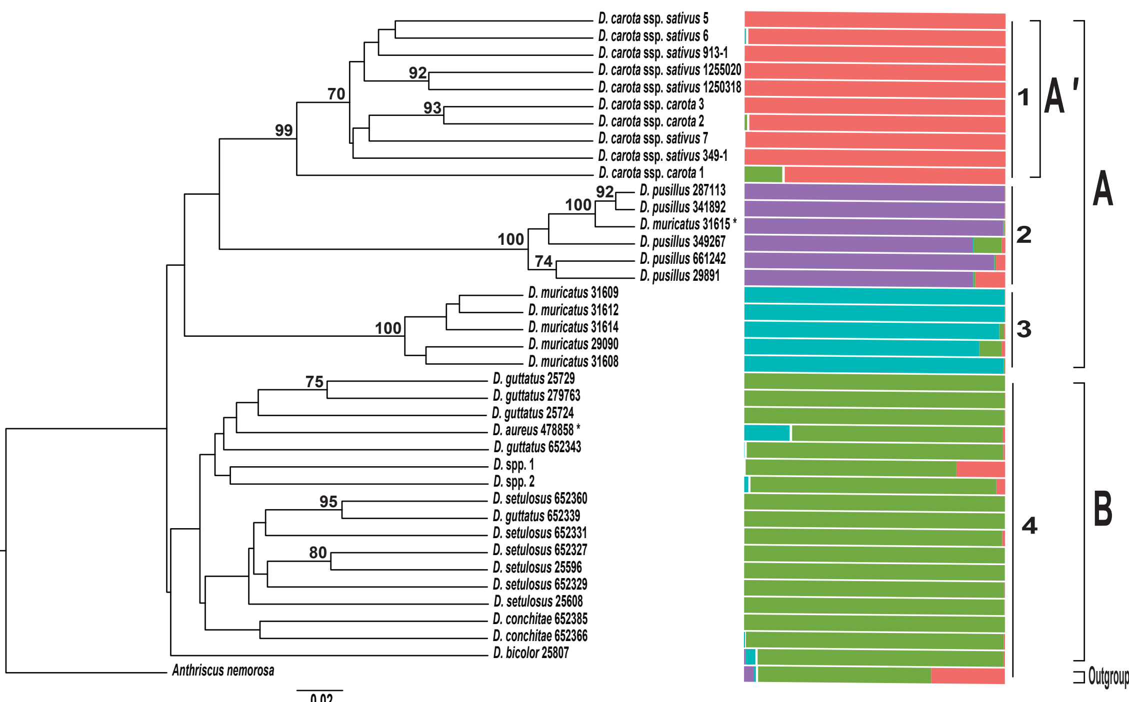

Figure 1. Dendrogram based on Provesti's genetic distance and the UPGMA clustering algorithm of 38 accessions of *Daucus* spp. and one outgroup using 21 iPBS. Each accession is represented by a horizontal bar, and each color represents a population (four in total). Numbers above branches represent bootstrap values, with only values higher than 70% shown. Clades A and B correspond to the two main groups of the *Daucus* phylogeny. Subclade A′ corresponds to *Daucus* accessions possessing 2n = 18 chromosomes. The two accessions designated by an asterisk (*) are misplaced relative to the study of Arbizu et al. (2014b).

software. The maximum and the minimum number of different alleles (N_a) ranged from 0.51 to 1.647 for clusters 3 and 4, respectively. The effective number of alleles varied slightly (1.090–1.39). Genetic diversity (Nei, 1972) ranged from 0.09 to 0.259, among the four populations sampled in this study, showing high genetic diversity for clusters 1 and 4. Shannon-Wiener index ranged from 1.61 to 2.89, and Simpson's index from 0.80–0.944, confirming high genetic diversity for all four populations of *Daucus*. The polymorphism information content (PIC) varied from 0.063 to 0.242. Also, the percentage of polymorphic loci per cluster ranged from 18.63% (cluster 2) to 83.35% (cluster 4) with an average of 44.69% (Table 4). The highest population divergence (F_{st}) was observed between clusters 2 and 3 with the highest genetic difference (0.579), on the other hand, clusters 1 and 4 gave the lowest F_{st} with the lowest genetic difference (0.160) (Table 5). The analysis of molecular variance (AMOVA) revealed that 41.85% of the total variation was found between clusters of *Daucus* while 58.15% was within clusters (Table 6).

Principal coordinate analysis (PCoA) was performed employing all 39 accessions of *Daucus* and showed that the first two PCoAs explained a total of 30.3%. The PCoA strengthened the clustering of the dendrogram and population structure by dividing the studied germplasm into four populations (Figure 2).

4. Discussion

Knowledge of the diversity in plant genetic resources help plant breeders to develop new and improved cultivars with favorable characteristics for farmers (Govindaraj et al., 2015). Molecular markers are crucial in plant breeding and are widely used today for multiple purposes. Retrotransposons are transposable elements and constitute a considerable part of the plant genome, and their repetition results in genomic variations, making them an exceptional source of molecular markers (Schulman et al. 2004). As a marker system, retrotransposons have been applied successfully in various crops for the investigation of phylogeny and genetic diversity assessment due to their user-friendliness and genotype resolution systems. Retrotransposons-based markers have been employed to explore the genetic diversity in different crop plants like grapevine species and cultivars identification (D'onofrio et al. 2010), and *Cicer* genus and its relatives (Andeden et al. 2013). These markers have been used to investigate the phylogenetic relationships in *Vitis* (Milovanov et al. 2019) and *Adonis* (Hossein-Pour et al. 2019). Moreover, iPBS-

Table 4. Genetic diversity parameters based on 21 iPBS primers among four clusters of *Daucus*.

Cluster	Number of accessions	N_a	N_e	H_e	H	lambda	PIC	PPL
1	10	1.260	1.332	0.224	2.300	0.900	0.199	58.82%
2	6	0.546	1.090	0.075	1.790	0.833	0.063	18.63%
3	5	0.510	1.112	0.090	1.610	0.800	0.072	18.95%
4	18	1.647	1.390	0.259	2.890	0.944	0.242	82.35%
Mean		0.991	1.231	0.162	2.148	0.869	0.144	44.69%

N_a: number of different alleles, N_e: number of effective alleles, H_e: expected heterozygosity, H: Shannon-Wiener index, lambda: Simpson's index, PIC: Polymorphism information content, PPL: Percentage of Polymorphic Loci.

Table 5. The F_{st} values among the four clusters inferred by STRUCTURE analysis.

Cluster	1	2	3
1			
2	0.372		
3	0.310	0.579	
4	0.160	0.253	0.214

Table 6. Analysis of molecular variance (AMOVA) using 21 iPBS primers of the genetic variation among and within four clusters of 39 accessions of *Daucus*.

Source of variation	df	SS	MS	Est. Var.	%
Between clusters	3	690.48	230.16	22.47	41.85
Within clusters	35	1092.71	31.22	31.22	58.15
Total	38	1783.19	46.93	53.69	100.00

retrotransposon markers were also used for gene mapping in barley (Manninen et al., 2000).

In this work, 309 clear and unambiguous bands were generated, of which 306 (99.02%) were found polymorphic, thus confirming a good level of polymorphism in the *Daucus* germplasm evaluated. Total polymorphic bands are higher than the value (65.1%) reported by Cavagnaro et al. (2011) with SSR markers, but in line (100%) with Baranski et al. (2012), Maksylewicz and Baranski (2013) using SSR markers, and to the value reported by Grzebelus et al. (2014) with DArT markers. Other studies on different plant species such as *Adonis* (Hossein-Pour et al., 2019) and *Capsicum* (Yildiz et al., 2020) also reported very high levels of polymorphism with iPBS markers, confirming that they are an excellent marker alternative for detecting genetic variation in plant germplasm. The mean PIC value was 0.919, which is higher than the values reported by Rong et al. (2010) and Cavagnaro et al. (2011) (0.5 and 0.67, respectively). Also, the PIC reported in this study is substantially higher than the value (0.301) reported by Grzebelus et al. (2014) using DArT markers in 94 cultivated and 65 wild carrot accessions and is also higher (0.715) than the PIC reported by Maksylewicz and Baranski (2013) employing SSR markers. Thus, the iPBS markers examined here reflect a very good level of genetic variation that may be massively used with *Daucus* germplasm for breeding work. Overall genetic diversity (i.e. expected heterozygosity, H_e) among the four clusters identified in this *Daucus* collection was 0.162. This measure of diversity is lower than the value reported by Clotault et al. (2010), Baranski et al. (2012), Iorizzo et al. (2013), and Maksylewicz and Baranski (2013) (0.73, 0.63, 0.34, 0.4, respectively). This difference may be explained due to the effect of sample size on estimating population genetic diversity, as reported by Pruett and Winker (2008).

Our dendrogram resolved that *Daucus* is placed in two main clades, A and B. Even though many of our bootstrap values are lower than 70%, our tree topology is in line with previous reports by Spooner et al. (2013) using 8 nuclear orthologs, Arbizu et al. (2014b) using 94 COS (conserved ortholog set) markers with 107 accessions of *Daucus* and related genera, and Arbizu et al. (2016a) employing 18,565 SNPs distributed along all 12 chromosomes of 150 accessions of *Daucus*. In a more recent study, Spooner et al. (2020) used 29,041 SNPs with 112 accessions of *Daucus* and identified three main clades. However, their study was focused on subspecies variation of *D. carota* "gummifer" morphotypes. The clade A in our dendrogram clustered *D. carota* complex, *D. pusillus* and *D. muricatus*, and this clustering was found in line with trees of *Daucus* reported by Spooner et al. (2013) and Arbizu et al. (2014b, 2016a). The clade B clustered genotypes belonging to *D. guttatus* complex species. When bootstrapping (100%) was performed, *D. muricatus* and *D. pusillus* were recovered as monophyletic. However, accession Ames 31615 was

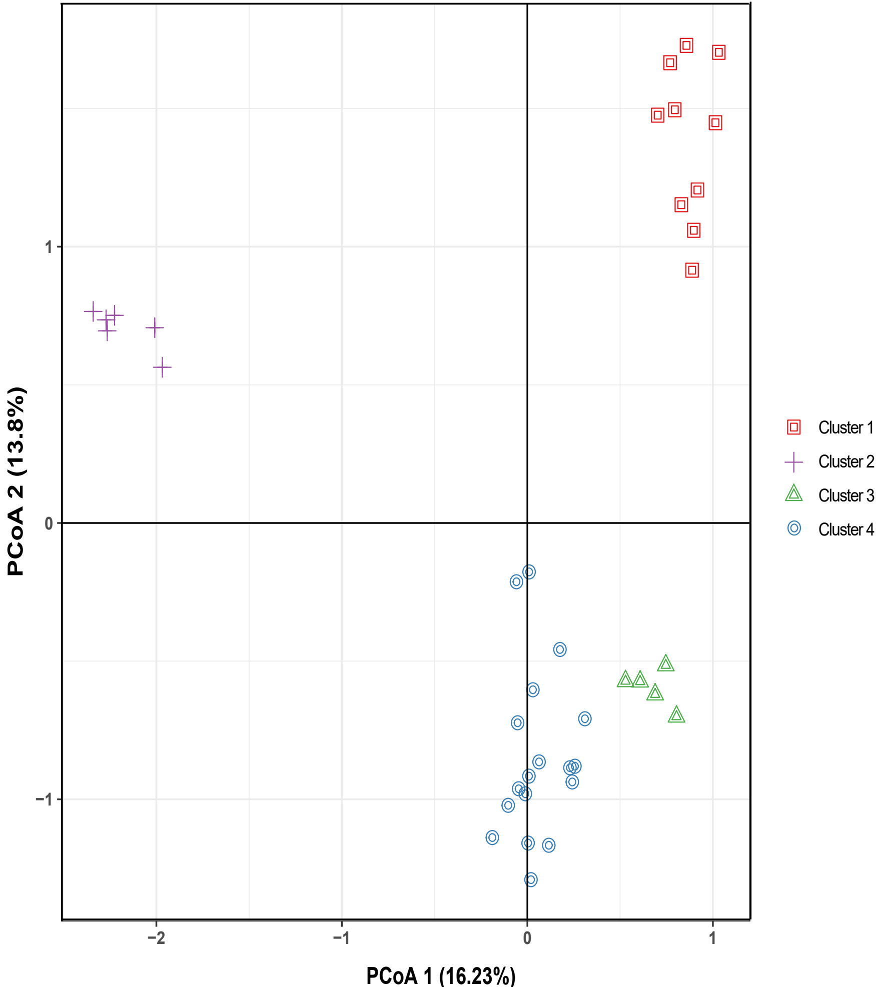

Figure 2. Principal coordinates analysis (PCoA) of 39 samples of *Daucus* and one related species based on 21 iPBS markers. Percentages of total variance explained by the first two coordinates are noted in parentheses.

intermingled within the *D. pusillus* clade. Similarly, accession PI 478858 is misplaced. These two cases may be explained due to DNA contamination or lack of discrimination of iPBS marker for species identification. Arbizu et al. (2016b) used DNA sequences (10 COS markers) and recovered four clades for the *D. guttatus* complex. However, iPBS failed to recover species of the *D. guttatus* complex as monophyletic, suggesting that iPBS markers may not work well for resolving plant species complex.

The AMOVA revealed the existence of higher variations within the *Daucus* clusters. This may be explained due to the extensive gene flow, especially within the individuals of the *D. carota* complex as reported in previous findings (Vivek and Simon 1999; Nothnagel et al. 2000; Hauser and Bjørn 2001; Hauser 2002). Selection and adaptation may also be playing a role. Two clustering methods of the *Daucus* germplasm, STRUCTURE and PCoA, revealed four major clusters. Cluster 1 consists of members of the *D. carota* complex, having ancestry coefficient values >0.7 according to the criteria suggested by Iorizzo et al. (2013). Cluster 2 contained accessions of *D. pusillus* + accession 31615 (*D. muricatus*). The only cluster containing one single species (*D. muricatus*) was number 3. Species that belong to the *D. guttatus* complex + *Anthriscus nemorosa* (with member coefficient <0.6) were placed into cluster 4. Our genetic structure results are in line with previous studies on *Daucus*. Baranski et al. (2012) identified that the population structure of *D. carota* cultivars is comprised of two clusters, Asian and Western types. Iorizzo et al. (2013) used wild and cultivated open-pollinated carrots germplasm collected from various regions of the world belonging to nine *Daucus carota* subspecies or other closely related *Daucus* species and identified

three genetically distinct clusters. A similar result was obtained by Grzebelus et al. (2014) with a set of DArT markers in a collection of 94 cultivated and 65 wild carrot accessions. We identified extra genetically distinct clusters since accessions from additional species of *Daucus* were employed in this work. In a more extended work, Arbizu et al. (2016a) used 144 accessions of *D. carota* subspecies and 18,565 SNPs, identifying eight clusters. Similar to our results, they did not recover *D. carota* subsp. s*ativus* as monophyletic. Spooner et al. (2020) reported similar findings using additional samples of *D. carota* belonging to diverse geographic origins and reported nine populations (K = 9). However, subspecies of *D. carota* were not reported as monophyletic. Other studies failed to distinguish taxa within *D. carota* using isozymes (St. Pierre et al. 1990; St. Pierre and Bayer 1991), ISSRs, and AFLPs (Bradeen et al. 2002). A plausible explanation is the recent divergence of populations of *D. carota*, and extensive fertile intercrosses among traditionally recognized subspecies. To our best knowledge, genetic clusters 2, 3, 4 reported in our study were not previously described by other investigators.

The iPBS markers should be considered as a reliable and polymorphic molecular tool since they allowed the discrimination among *Daucus* species, and determined the genetic diversity and population structure among carrot accessions. However, those markers failed to resolve species complex. The results presented here will be useful for carrot breeders as the utilization of wild crop relatives for breeding perspectives is intensifying (Khoury et al., 2013) and there is a constant demand for the development of new superior carrot cultivars (Rubatzky et al., 1999; Frese and Nothnagel, 2008). Consequently, we believe wild *Daucus* germplasm will gain more importance as they are a good source of male sterility, nutraceutical, and culinary traits, and they are also important for yield increase, among other characteristics. However, further research is needed in terms of establishing hybridization techniques with cultivated carrots under controlled conditions.

Acknowledgments

We would like to thank Prof. Dr. Philipp Simon for providing germplasm.

References

Andeden EE, Baloch FS, Derya M, Kilian B, Özkan H (2013). iPBS-retrotransposons-based genetic diversity and relationship among wild annual Cicer species. Journal of Plant Biochemistry and Biotechnology 22: 453-466. doi: 10.1007/s13562-012-0175-5

Arbizu C, Reitsma KR, Simon PW, Spooner DM (2014a). Morphometrics of *Daucus* (Apiaceae): A counterpart to a phylogenomic study. American Journal of Botany 101: 2005-2016. doi: 10.3732/ajb.1400252

Arbizu C, Ruess H, Senalik D, Simon PW, Spooner DM (2014b). Phylogenomics of the carrot genus (*Daucus*, Apiaceae). American Journal of Botany 101: 1666-1685. doi: 10.3732/ajb.1400106

Arbizu CI, Ellison SL, Senalik D, Simon PW, Spooner DM (2016a). Genotyping-by-sequencing provides the discriminating power to investigate the subspecies of *Daucus carota* (Apiaceae). BMC Evolutionary Biology 16: 1-16. doi: 10.1186/s12862-016-0806-x

Arbizu CI, Simon PW, Martínez-Flores F, Ruess H, Crespo MB et al. (2016b). Integrated molecular and morphological studies of the *Daucus guttatus* complex (Apiaceae). Systematic Botany 41: 479-492. doi: 10.1600/036364416X691948

Baloch FS, Karaköy T, Demirbaş A, Toklu F, Özkan H et al. (2014). Variation of some seed mineral contents in open pollinated faba bean (*Vicia faba* L.) landraces from Turkey. Turkish Journal of Agriculture and Forestry 38: 591-602. doi: 10.3906/tar-1311-31

Baloch FS, Alsaleh A, de Miera LE, Hatipoğlu R, Çiftçi V et al. (2015). DNA based iPBS-retrotransposon markers for investigating the population structure of pea (*Pisum sativum*) germplasm from Turkey. Biochemical Systematics and Ecology 61: 244-252. doi: 10.1016/j.bse.2015.06.017

Baranski R, Maksylewicz-Kaul A, Nothnagel T, Cavagnaro PF, Simon PW et al. (2012). Genetic diversity of carrot (*Daucus carota* L.) cultivars revealed by analysis of SSR loci. Genetic Resources and Crop Evolution 59: 163-170. doi: 10.1007/s10722-011-9777-3

Boiteux LS, Fonseca MEN, Simon PW (1999) Effects of plant tissue and DNA purification methods on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. Journal of the American Society Horticultural Science 124:32–38.

Bradeen JM, Bach IC, Briard M, le Clerc V, Grzebelus D et al. (2002). Molecular diversity analysis of cultivated carrot (*Daucus carota* L.) and wild Daucus populations reveals a genetically nonstructured composition. Journal of the American Society Horticultural Science 127 (3): 383-391. doi: 10.21273/JASHS.127.3.383

Cavagnaro PF, Chung SM, Manin S, Yildiz M, Ali A et al. (2011). Microsatellite isolation and marker development in carrot-genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. BMC Genomics 12: 1-20. doi: 10.1186/1471-2164-12-386

Clotault J, Geoffriau E, Lionneton E, Briard M, Peltier D (2010). Carotenoid biosynthesis genes provide evidence of geographical subdivision and extensive linkage disequilibrium in the carrot. Theoretical and Applied Genetics 121: 659-672. doi: 10.1007/s00122-010-1338-1

Comertpay G, Baloch FS, Kilian B, Ulger AC, Ozkan H (2012). Diversity assessment of Turkish maize landraces based on fluorescent labelled SSR markers. Plant Molecular Biology Reporter 30: 261-274. doi: 10.1007/s11105-011-0332-3

D'Onofrio C, De Lorenzis G, Giordani T, Natali L, Cavallini A et al. (2010). Retrotransposon-based molecular markers for grapevine species and cultivars identification. Tree Genetics Genomes 6: 451-466. doi: 10.1007/s11295-009-0263-4

Demirel U, Tındaş İ, Yavuz C, Baloch FS, Çalışkan ME (2018). Assessing genetic diversity of potato genotypes using inter-PBS retrotransposon marker system. Plant Genetic Resources 16: 137-145. doi: 10.1017/S1479262117000041

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19: 11-15.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x

Fang-Yong C, Ji-Hong L (2014). Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and start codon-targeted polymorphism markers. Scientia Horticulturae 170: 169-175. doi: 10.1016/j.scienta.2014.03.010

Food and Agriculture Organization of the United Nations (2019). FAOSTAT statistical database. Rome.

Francis RM (2017). Pophelper: an R package and web app to analyse and visualize population structure. Molecular Ecology Resources 17: 27-32. doi: 10.1111/1755-0998.12509

Frese L, Nothnagel (2008). *Daucus* species promise a glossy future in carrot production. in Crop wild relative vol. 6. In: Maxted N, Dulloo E (editors). Birmingham, UK: IUCN/SSC, pp. 4-6.

Govindaraj M, Vetriventhan M, Srinivasan M (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International 2015: 431487. doi: 10.1155/2015/431487

Grzebelus D, Iorizzo M, Senalik D, Ellison S, Cavagnaro P et al. (2014). Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* L.) genome, as revealed by Diversity Arrays Technology (DArT) markers. Molecular Breeding 33: 625-637. doi: 10.1007/s11032-013-9979-9

Hauser TP (2002). Frost sensitivity of hybrids between wild and cultivated carrots. Conservation Genetics 3: 75–78. doi: 10.1023/A:1014256302971

Hauser TP, GK Bjørn (2001). Hybrids between wild and cultivated carrots in Danish carrot fields. Genetic Resources and Crop Evolution 48: 499–506. doi: 10.1023/A:1012051731933

Hossein-Pour A, Karahan F, İlhan E, Ilçim A, Haliloğlu K (2019). Genetic structure and diversity of Adonis L. (Ranunculaceae) populations collected from Turkey by inter-primer binding site (iPBS) retrotransposon markers. Turkish Journal of Botany 43: 585-596. doi: 10.3906/bot-1810-1

Iorizzo M, Senalik DA, Ellison SL, Grzebelus D, Cavagnaro PF et al. (2013). Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus*) (Apiaceae). American Journal of Botany 100: 930-938. doi: 10.3732/ajb.1300055

Kalendar R, Antonius K, Smýkal P, Schulman AH (2010). iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. Theoretical and Applied Genetics 121: 1419-1430. doi: 10.1007/s00122-010-1398-2

Karık Ü, Nadeem MA, Habyarimana E, Ercişli S, Yildiz M et al. (2019). Exploring the genetic diversity and population structure of Turkish laurel germplasm by the iPBS-retrotransposon marker system. Agronomy 9: 647. doi: 10.3390/agronomy9100647

Khoury CK, Greene S, Wiersema J, Maxted N, Jarvis A et al. (2013). An inventory of crop wild relatives of the United States. Crop Science 53: 1-13. doi: 10.2135/cropsci2012.10.0585

Maksylewicz A, Baranski R (2013). Intra-population genetic diversity of cultivated carrot (*Daucus carota* L.) assessed by analysis of microsatellite markers. Acta Biochimica Polonica 60: 753-760.

Manninen O, Kalendar R, Robinson J, Schulman AH (2000). Application of BARE-1 retrotransposon markers to the mapping of a major resistance gene for net blotch in barley. Molecular Genetics and Genomics 264: 325-334. doi: 10.1007/s004380000326

Martínez-Flores F, Crespo MB, Simon PW, Ruess H, Reitsma K et al. (2020). Subspecies variation of *Daucus carota* coastal ("Gummifer") morphotypes (Apiaceae) using genotyping-by-sequencing. Systematic Botany 45: 688-702. doi: 10.1600/0363664420X15935294613527

Mezghani N, Zaouali I, Amri WB, Rouz S, Simon PW et al. (2014). Fruit morphological descriptors as a tool for discrimination of *Daucus* L. germplasm. Genetic Resources and Crop Evolution 61: 499-510. doi: 10.1007/s10722-013-0053-6

Mezghani N, Ruess H, Tarchoun N, Amor JB, Simon PW et al. (2018). Genotyping-by-sequencing reveals the origin of the Tunisian relatives of cultivated carrot (*Daucus carota*). Genetic Resources and Crop Evolution 65: 1359-1368. doi: 10.1007/s10722-018-0619-4

Milovanov A, Zvyagin A, Daniyarov A, Kalendar R, Troshin L (2019). Genetic analysis of the grapevine genotypes of the Russian *Vitis* ampelographic collection using iPBS markers. Genetica 147: 91-101. doi: 10.1007/s10709-019-00055-5

Nadeem MA, Habyarimana E, Çiftçi V, Nawaz MA, Karaköy T et al (2018). Characterization of genetic diversity in Turkish common bean gene pool using phenotypic and whole-genome DArTseq-generated silicoDArT marker information. PLoS One 13: e0205363. doi: 10.1371/journal.pone.0205363

Nakajima Y, Oeda K, Yamamoto T (1998). Characterization of genetic diversity of nuclear and mitochondrial genomes in *Daucus* varieties by RAPD and AFLP. Plant Cell Reports 17: 848-853. doi: 10.1007/s002990050496

Nei M (1972). Genetic distance between populations. The American Naturalist 106 (949): 283-292.

Nemli S, Kianoosh T, Tanyolac MB (2015). Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) accessions through retrotransposon-based interprimer binding sites (iPBSs) markers. Turkish Journal of Agriculture and Forestry 39: 940-948. doi: 10.3906/tar-1505-59

Nothnagel T, Straka P, Linke B (2000). Male sterility in populations of *Daucus* and the development of alloplasmic male-sterile lines of carrot. Plant Breeding 119: 145-152. doi: 10.1046/j.1439-0523.2000.00470.x

Pierre MS, Bayer RJ (1991). The impact of domestication on the genetic variability in the orange carrot, cultivated *Daucus carota* ssp. *sativus* and the genetic homogeneity of various cultivars. Theoretical and Applied Genetics 82: 249-253. doi: 10.1007/BF00226221

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Pruett CL, Winker K (2008). The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. Journal Avian Biology 39: 252-256.

Rong J, Lammers Y, Strasburg JL, Schidlo NS, Ariyurek Y et al. (2014). New insights into domestication of carrot from root transcriptome analyses. BMC Genomics 15: 895. doi: 10.1186/1471-2164-15-895

Rubatzky VE, Quiroz CF, Simon PW (1999). Carrots and related vegetable Umbelliferae. 1st ed. Wallingford, UK: CABI Publishing.

Sáenz Laín C (1981). Research on *Daucus* L. (Umbelliferae). Anales Jardin Botanico de Madrid 37(2): 481-533.

Saldaña CL, Cancan JD, Cruz W, Correa MY, Ramos M et al. (2021). Genetic diversity and population structure of capirona (*Calycophyllum spruceanum* Benth.) from the Peruvian Amazon revealed by RAPD markers. Forests 12: 1125. doi: 10.3390/f12081125

Schulman AH, Flavell AJ, Paux E, Ellis THN (2012). The application of LTR retrotransposons as molecular markers in plants. Methods in Molecular Biology 859: 115-153. doi: 10.1007/978-1-61779-603-6_7

Shah KR, Patel DK, Pappachan A, Prabha CR, Singh DD (2016). Characterization of a Kunitz-type serine protease inhibitor from *Solanum tuberosum* having lectin activity. International Journal of Biological Macromolecules 83: 259-269. doi: 10.1016/j.ijbiomac.2015.11.068

Shim SI, Jørgensen RB (2000). Genetic structure in cultivated and wild carrots (*Daucus carota* L.) revealed by AFLP analysis. Theoretical and Applied Genetics 101: 227-233. doi: 10.1007/s001220051473

Simon P, Pollak L, Clevidence B, Holden J, Haytowitz D (2009). Plant breeding for human nutrition. Plant Breeding Reviews 31: 325-392. doi: 10.1002/9780470593783.ch7

Small E (1978). A numerical taxonomic analysis of the *Daucus carota* complex. Canadian Journal of Botany 56: 248-276. doi: 10.1139/b78-033

Spooner D, Rojas P, Bonierbale M, Mueller LA, Srivastav M et al. (2013). Molecular phylogeny of *Daucus* (Apiaceae). Systematic Botany 38(3): 850-857. doi: 10.1600/036364413X670449

Spooner DM, Widrlechner MP, Reitsma KR, Palmquist DE, Rouz S et al (2014). Reassessment of practical subspecies identifications of the USDA *Daucus carota* L. germplasm collection: Morphological data. Crop Science 54: 706-718. doi: 10.2135/cropsci2013.04.0231

St. Pierre MD, Bayer RJ, Weis IM (1990). An isozyme-based assessment of the genetic variability within the *Daucus carota* complex (Apiaceae: Caucalideae). Canadian Journal of Botany 68: 2449-2457. doi: 10.1139/b90-311

Tanksley SD, McCouch SR (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063-1066. doi: 10.1126/science.277.5329.1063

Tavares AC, Loureiro J, Castro S, Coutinho AP, Paiva J et al. (2014). Assessment of *Daucus carota* L. (Apiaceae) subspecies by chemotaxonomic and DNA content analyses. Biochemical Systematics and Ecology 55: 222-230. doi: 10.1016/j.bse.2014.03.031

Vivek BS, Simon PW (1998). Genetic relationships and diversity in carrot and other *Daucus* taxa based on nuclear restriction fragment length polymorphisms. Journal of the American Society for Horticultural Science 123: 1053-1057. doi: 10.21273/JASHS.123.6.1053

Vivek BS, Simon PW (1999). Phylogeny and relationships in Daucus based on restriction fragment length polymorphisms (RFLPs) of the chloroplast and mitochondrial genomes. Euphytica105:183-189. doi: 10.1023/A:1003446301145

Yaldiz G, Camlica M, Nadeem MA, Nawaz MA, Baloch FS (2018). Genetic diversity assessment in *Nicotiana tabacum* L. with iPBS-retrotransposons. Turkish Journal of Agriculture and Forestry 42: 154-164. doi: 10.3906/tar-1708-32

Yildiz M, Kocak M, Nadeem MA, Cavagnaro P, Barboza K et al. (2020). Genetic diversity analysis in the Turkish pepper germplasm using iPBS retrotransposon-based markers. Turkish Journal of Agriculture and Forestry 44: 1-14. doi: 10.3906/tar-1902-10