Australian Journal of Crop Science AJCS
AJCS 6(2):194-199 (2012) ISSN:1835-2707

Molecular diversity of Cumin (*Cuminum cyminum* L.) using RAPD markers

Alireza Bahraminejad[1, 2], Ghasem Mohammadi-Nejad[*2], Mihdzar Abdul Kadir[3], Mohd Rafii Bin Yusop[4]

[1]Department of Crop Science, Agricultural Faculty, University Putra Malaysia, P.O.B. 43400 UPM, Serdang, Selangor Darul Ehsan, Malaysia
[2]Horticultural Research Institute, Shahid Bahonar University of Kerman, Kerman, P.O.B 76169-133-Iran
[3]Department of Agriculture Technology, Agricultural Faculty, University Putra Malaysia, P.O.B. 43400 UPM, Serdang, Selangor Darul Ehsan, Malaysia
[4]Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

*** Corresponding author: Mohammadinejad@uk.ac.ir**

Abstract

Cumin (*Cuminum cyminum* L.) belonging to the family Apiaceae is one of the oldest and economically most important medicinal spices. In this study, as a comprehensive research on genetic diversity, forty nine cumin ecotypes, belong to nine Iranian regional sub-populations were assessed using RAPD markers. DNA extracted from cumin seeds through CTAB method. Twenty three RAPD markers were used for diversity assessment, in which 21 showed polymorphism. Allele frequency and polymorphism information content (PIC) of each locus were calculated by Power Marker version 3.25. Molecular variability among and within populations was assessed accordingly. Based on molecular data, Jacquard's similarity coefficient was used to detect the phylogenic relationship; subsequently dendrograms were drawn based on UPGMA using NTSYS software. Cluster analysis among the populations categorized nine populations into two groups at the similarity level of 0.43, in which class one was consisted of only Golestan population and the rest were arranged in the second group. Golestan and Northern-Khorasan populations showed the highest difference while Kerman and Esfahan populations showed the most similarity. Based on Principal Coordinate Analysis (PCOA), 79% of variation was explained by two first principal components. Populations of Semnan, Yazd and Golestan showed different reaction rather than the other populations. It is suggested that Kerman, Esfahan and Southern-Khorasan may have the same ancestors. Molecular diversity among 49 ecotypes, as the sub-populations of nine populations derived from different Iranian states, showed five different categories. Based on the result it can be concluded that there is a high potential of variability in Iranian cumin populations which are very important sources for cumin breeding objectives.

Key words: Cumin, *Cuminum cyminum* L., Genetic diversity, RAPD.

Introduction

Cumin (*Cuminum cyminum* L.) is a herbaceous, dicotyledonous annual plant, diploid (2n = 2x = 14) and allogam with hermaphrodite flowers. Recent studies have indicated its pharmaceutical and medicinal importance (Agrawal et al., 1996; Banerjee and Sarkar, 2003; Takayanagi et al., 2003; Bahraminejad et al., 2011). Because of its low water requirements, farmers are interested in cultivation of cumin in drought affected areas, where most of the other crop plants cannot be grown economically. However, production of cumin is limited due to several biotic stresses, of which wilt diseases are the most serious (Valizadeh et al., 2008; Agrawal, 1996). The economic product of this plant is seed which has pharmaceutical applications and also used as spice for thousands of years. It forms part of the food culture in the western Asia, the habitat of this plant. Thus, trade and consumption is limited to almost natural production area. Currently, in maintaining the natural structure of this plant, agricultural production technology in the west Asia has been founded (Sheidai et al., 1996). The genetic information of a landrace can allow the possible explanation of genes for traits such as disease resistance, tolerance to environmental stresses etc. The genetic variability of landraces can be conserved in the environments, where they have originated (Zeven, 1996). Genetic diversity of plants occurs over thousands of years in nature. A local population of plant is a suitable germplasm for improving plant breeding programs. Gene banks collection, identification, precise evaluation and conservation from reserves inheritance and plant populations, provide necessary information for researchers. In agriculture, food production also depends on the use of plant genotypes with high yield. Common methods to improve crops are based on the selection of desirable genotypes with genetic variation among communities; therefore knowledge of population diversity is a main prerequisite and the first step in plant breeding (Allard, 1999). Up to now, several studies have investigated the diversity and classification of plants based on morphological markers, biochemical and molecular characteristics. The RAPDs provide a useful tool for assessing genetic diversity of rare, endemic species and for resolving relationships among populations and results show that the genetic diversity of this species is high, possibly allowing it to adapt more easily to environmental variations. During the stages of phenology, yield and yield components

and morphological status of populations of cumin were investigated and some traits such as seed length, seed yield per unit area (square meter), 1000- seed weight, oil yield per plant characterized (Bonyanpour et al., 2001). In other researches, morphological markers were similarly used to evaluate genetic diversity of Parse cumin (Kapila et al., 1997; Mittal et al., 2006). The fact that the measurement of morphological traits is time and energy consuming and costly, and usually confused with environmental effects of gene expression, molecular markers are used for determining genetic differences. PCR based markers such as RAPD have widely been used in analysis of genetic distance (Bosch, 1997). RAPD markers are commonly used enabling analysis of genetic diversity of plant populations in breeding programs and germplasm collections (Moeller et al., 1999). RAPD markers are normally easier to use and cost effective compared to RFLPs (Dos Santos et al., 1994; Thormann et al., 1194). This is an indicator for determining the genetic diversity of many plants (Pezhmanmehr et al., 2009; Thormann et al., 1994). Breeding strategies in Cumin are more difficult, compared with other outcross crops, due to many unsolved problems such as difficulties in emasculation and hybridization. Evaluation of genetic diversity is the first step in breeding strategies. Comprehensive phenotypic diversity analysis of Iranian Cumin has been done to detect inter and intra variation and their phylogenic relation (Bahraminejad et al., 2011). Iran has the biggest variation of cumin in the world. Therefore, the purpose of the present study is to assess molecular diversity of forty nine cumin ecotypes, which are the sub-populations belong to nine populations of different provinces of Iran.

Results and Discussion

Totally, 252 amplified DNA fragments obtained from RAPD primers. Two primers were monomorphic among the ecotypes and the rest of 21 RAPD primers were used for genetic diversity assessment. PIC value of the Primer OPC8 showed the highest polymorphism (0.37). Marker OPU6 has highest allele frequency for major allele (0.98), while marker OPJ18 showed minimum allele frequency for major allele (0.075). Polymorphism information content (PIC) values varied from 0.075 to 0.37 with an average of 0.31, in which highest value was belonged to OPC8 (PIC = 0.37) while OPJ1 showed the lowest PIC value (PIC = 0.076) (Fig. 1) (Table 3).

Molecular variability among the populations derived from different provinces

Golestan and Northern-Khorasan populations with 8 polymorphic bands showed the highest difference while Kerman and Esfahan populations showed the lowest difference. The primer OPJ1 exhibited the most frequent allele. The lowest frequency of allele was belonged to the primer OPC13. Primers OPB14, OPB17 just showed amplified band for Yazd and Golestan populations. Primer OPJ19 in Pars, OPF5, OPQ6, and OPK14 in Semnan and OPQ6 and OPF5 in Golestan and Northern-Khorasan populations, respectively, were able to amplify polymorphic products to exhibit the diversity. OPJ19 primer showed the most segregation in Pars population, while OPI14 was the highest segregant primer in Yazd and Golestan populations. Primers OPP6, OPP7, OPQ6 present in Kerman and Southern-Khorasan, primers OPP6, OPP7 in Esfahan and primers OPK14, OPF5 in Semnan province showed the highest segregation. The mentioned populations were categorized into two clusters based on molecular data with Jaccards similarity coefficient at the level of 0.43. Class one contained only one population, Golestan, and the rest were arranged in the second group as Kerman, Esfahan, Southern-Khorasan, Khorasan-Razavi, Northern-Khorasan, Yazd and Semnan populations. The latter cluster had three subclasses as follows: the first subclass contained Yazd, Pars and Semnan, the second included Northern-Khorasan and Khorasan-Razavi while Southern-Khorasan, Esfahan and Kerman were categorized in the third subclass (Fig 2). Therefore, it can be concluded that Golestan has a different genetic back ground, while Kerman, Esfahan and Southern-Khorasan may have same ancestor (Fig 2). The genetic distances between pairs of population was the highest for Semnan-Yazd, Semnan-Pars, Semnan-Golestan and Golestan-Yazd, whereas these genetic distances was the lowest for populations Khorasan-Razavi-Kerman, Khorasan-Razavi-Southern-Khorasan and Khorasan-Razavi-Esfahan. The validity of obtained dendrogram was evaluated by cophentic coefficient (r=0.92) then dendrogram was confirmed as valuable cluster. Principal Coordinate Analysis (PCOA) was done using the first principal component from the molecular data and these data were scattered on biplot to simplify the reaction of populations according to the biplot (Fig 3). Based on the biplot analysis, populations were divided into two groups: group one showed the highest number of populations including Kerman, Esfahan, and Southern-Khorasan as classified in one group. Semnan, Yazd and Golestan showed the different reactions based on first two principle components and Golestan showed the biggest difference with the other two populations, which explained 79% of total variation in this study. It can be concluded that these markers arranged randomly on the genome, it can be confirmed that Kerman, Esfahan and Southern-Khorasan have the same origin.

Variability between the sub-populations based on RAPD markers

All of the evaluated ecotypes showed variation based on RAPD primers. The molecular assessment based on the variability among the ecotypes showed that Esfahan and Southern-Khorasan has the highest variability within all populations. Khorasan-Razavi with 18, Northern-Khorasan with 16, Yazd with 15, Golestan with 12, Semnan with 11 and Pars with 6 polymorphic loci were arranged from high to low variability within population, respectively (Table 2). The interesting point is correlation of genetic similarity with the geographic distances. It seems that ecotypes located in closer distances have higher similarities and genetic background. The dendrogram of 49 ecotypes (sub-populations) of the populations derived from different Iranian provinces showed five different categories: class1 includes ecotypes of Jat and Zarand while class 2 contains Sadoq, Khatam, Rafsanjan, Baneh and Kashmar. Class 3 includes Torbat-Jam, Birjand, and Semirom. Class 4 contains Darmian, Joopar and Ardestan. Class 5 consisted of six subclasses: the first is Gonbad, Sorkheh, Sarayan and Shahmirzad and second subclass includes Sirjan, Gonabad, Ferdows, Ravar, Sadroea, Taybad, Natanz, Bardsekan. The third subclass includes Feridan, Naien, forth subclass includes Chatrood, Mahan, Koohbanan, Bafq, fifth subclass includes Sarvestan, Bardsir, Baft, Kalateh, Estahban, Sepidan, Sivand and Bojnord sixth

Table1. List of 49 studied cumin ecotypes/sub-populations from 9 different provinces of Iran.

Population Derived from Iranian States	Subpopulations / Ecotypes from Each Province						
Pars	1-Sarvestan	2-Sepidan	3-Sivand	4-Estahban			
Yazd	5-Ardekan	6-Bafq	7-Sadoq	8-Khatam	9-Sadroea		
Golestan	10-Maraveh-Tapeh	11-Aq-Qala	12-Jat	13-Gonbad			
Kerman	14-Baft	15-Bardsir	16-Chatrood	17-Joopar	18-Kooh-banan	19-Mahan	20-Ravar
	21-Rafsanjan	22-Sirjan	23-Zarand				
Southern-Khorasan	24-Qaen	25-Nahbandan	26-Birjand	27-Sarayan	28-Darmian		
Esfahan	29-Feridan	30-Semirom	31-Ardestan	32-Naien	33-Khansar	34-Natanz	
Semnan	35-Shahmirzad	36-Sorkheh	37-Ivanaki	38-Kalateh			
Northen-Khorasan	39-Esfarayen	40-Shirvan	41-Bojnord	42-Baneh			
Khorasan-Razavi	43-Gonabad	44-Ferdows	45-Torbat-Heidarieh	46-Torbat-Jam	47-Kashmar	48-Taybad	49-Bardsekan

Table 2. Sequences of the random nucleotide primers.

OPU12	5´ TCA CCA GCC A 3´	OPQ6	5´ GAG CGC CTT G 3´
OPS17	5´ TGG GGA CCA C 3´	OPI14	5´ TGA CGG CGG T 3´
OPP7	5´ GTC CAT GCC A 3´	OPC13	5´ AAG CCT CGT C 3´
OPD7	5´ TGT CTG GGT G 3 ´	OPP6	5´ GTG GGC TGA C 3´
OPB17	5´ CGA CTG CAG T 3´	OPJ18	5´ TGG TCG CAG A 3´
OPF5	5´ CCG AAT TCC C 3´	OPC8	5´ TGG ACC GGT G 3 ´
OPJ19	5´ GGA CAC CAC T 3´	OPA9	5´ GGG TAA CGC C 3´
OPJ1	5´ CCC GGC ATA A 3´	OPJ21	5´ ACG AGG GAC T 3´
OPU20	5´ ACA GCC CCC A 3´	OPU6	5´ ACC TTT GCG G 3´
OPK14	5´ GAC GGA TCA G 3´	OPS17	5´ TGG GGA CCA C 3´
OPJ20	5´ AAG CGG CCT C 3´	OPU12	5´ TCA CCA GCC A 3´
OPE7	5´AGA TGC AGC C 3		

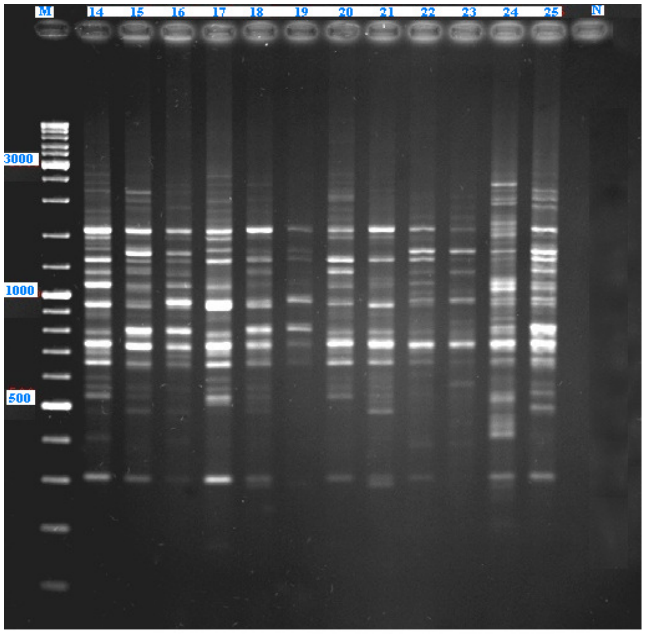

Fig 1. DNA bands amplified from 49 Cumin genotypes using RAPD OPC8 marker and electrophoresed in a 3% Agaros gel. Ladder = 1kbp bp ladder

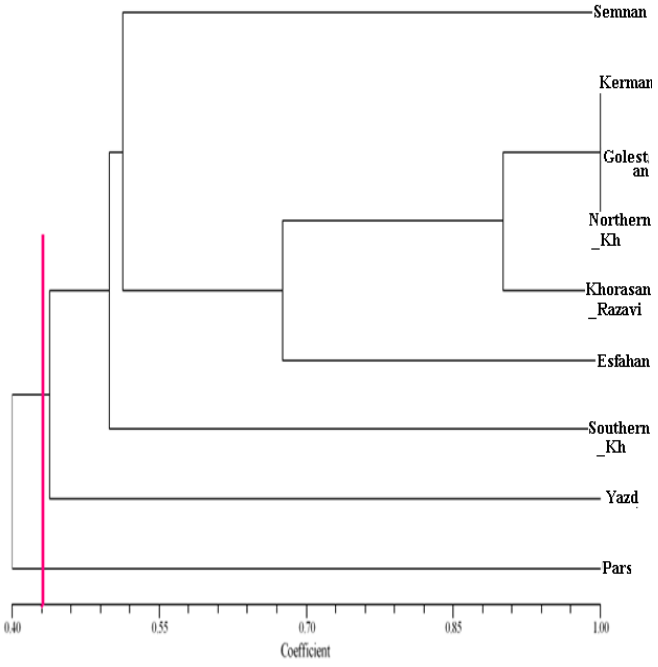

Fig 2. Dendrogram of 9 cumin populations based on RAPD markers, according to the un-weighted pair group mean algorithm (UPGMA) method based on a similarity matrix by NTYSYs 2.1 software

subclass includes Maraveh-tapeh, Esfarayen Ivanaki, Ardestan, Torbat-Heidareh, Aq-Qala, Khansar, Shirvan, Qaen and Nahbandan. More ecotypes were focused in the class 5. Disperse arrangement of sub-populations in molecular clusters can be due to low number of primers and perhaps incomplete coverage of genome by the used primers (Fig 4).

In this study wide range of similarity (0.49-0.970) was observed among Cumin ecotypes. According to the above mentioned results significant genetic diversity was observed among the cumin populations, as well as their within or among subpopulation even for those ecotypes which were located closely.

Pezhmanmehr et al. (2009) reported that there is poor correlation between the molecular similarity and the geographical distances, due to germplasm exchange in different provinces through the farmer and also, incomplete coverage of genome by markers. Considering the observed diversity among Iranian cumin ecotypes and its cross pollination behavior, these diverse genotypes can be used as a good material for breeding strategies. These ecotypes would be a promising materials for producing improved cultivars as well as exploiting the genes controlling tolerance for abiotic stresses in cumin (Allard, 1999). In general, study of molecular survey with RAPD markers showed that this marker is a suitable for germplasm management and identifying the informative markers. So it seems this molecular marker can be a useful tool for assessment of genetic diversity in different cumin ecotypes.

Materials and methods

Plant materials

Forty nine cumin ecotypes which are sub-populations belong to nine populations from different Iranian provinces was collected by authors (Table1). This collection represents the biggest variation for cumin in Iran.

DNA extraction

Instead of plant leaf material, cumin seeds were used for the DNA extraction. In addition to saving time and cost, the use of plant seeds, for DNA extraction is very useful when dealing with a large number of samples and also reduces the volume of work. But cumin seeds have a lot of secondary metabolites and essential oils that makes the DNA extraction process difficult. Several DNA extraction protocol for DNA isolation from the seeds of desired cumin were tested and finally a modified method based on CTAB (Krizman et al., 2006) was used to extract genomic DNA.

Optimization of RAPD Protocol

Amplification of genomic DNA was made on a Perkin Elmer DNA Cycler (BIOMETRA, Germany), using 23 arbitrary decamers in 25-µl reaction volumes containing 1 unit of *Taq* polymerase, 2.5µl Tris-HCl (pH 9.0), 4.2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5mM each of random primer and 50 ng of template DNA. The cycle program included an initial 5 min denaturation at 95°C, followed by 30 cycles of 1 min at 95°C, 30 sec at 35°C and 2 min at 72°C, with a final extension at 72°C for 8 min. RAPD fragments were separated on 3% agarose gels in 1X TBE buffer, stained with ethidium bromide and photographed on a UV transilluminator using a digital camera. DNA from each plant was amplified used

Table 3. Number of alleles and polymorphism information content (PIC) value of RAPD markers for 9 Cumin populations.

Marker	Major Allele requency	PIC
OPU6	0.6939	0.3346
OPU12	0.7755	0.2876
OPU20	0.8367	0.2359
OPS17	0.6122	0.3621
OPA9	0.7959	0.2721
OPJ1	0.9592	0.0752
OPJ18	0.7959	0.2721
OPJ19	0.6735	0.3431
OPJ20	0.8367	0.2359
OPJ21	0.6122	0.3621
OPP6	0.8163	0.2549
OPP7	0.8571	0.2149
OPI14	0.7959	0.2721
OPC13	0.5510	0.3724
OPC8	0.5918	0.3664
OPQ6	0.7143	0.3249
OPE7	0.9388	0.1083
OPK14	0.7143	0.3249
OPF5	0.6531	0.3505
OPB17	0.7755	0.2876
OPD7	0.8163	0.2549
OPU6	0.6939	0.3346

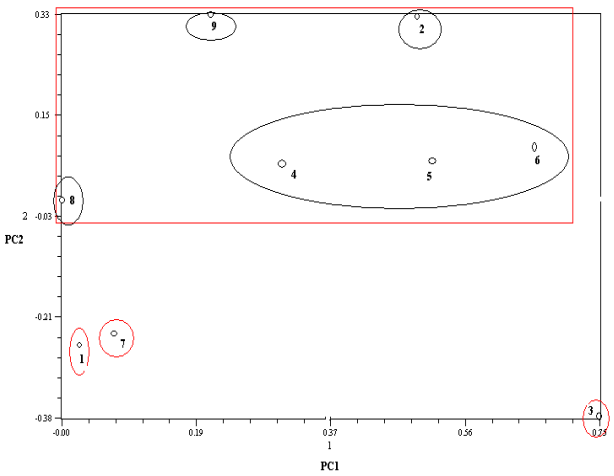

Fig 3. Biplot of two first principal components

primer and finally their banding patterns were compared (Fig 1).

Data analysis of RAPD

From the total of 23 primers, only 21 primers which showed polymorphism were selected for the analysis (Table 2). Fragment sizes were designated as amplified bands, and bands were shared as diallelic characters (present = 1, absent = 0) for each of the 21 primers, respectively. Distance matrix was performed through the Jaccard's similarity coefficient (Jaccard, 1908) using NTSYS-pc Ver. 2.1 (Rohlf, 2000).

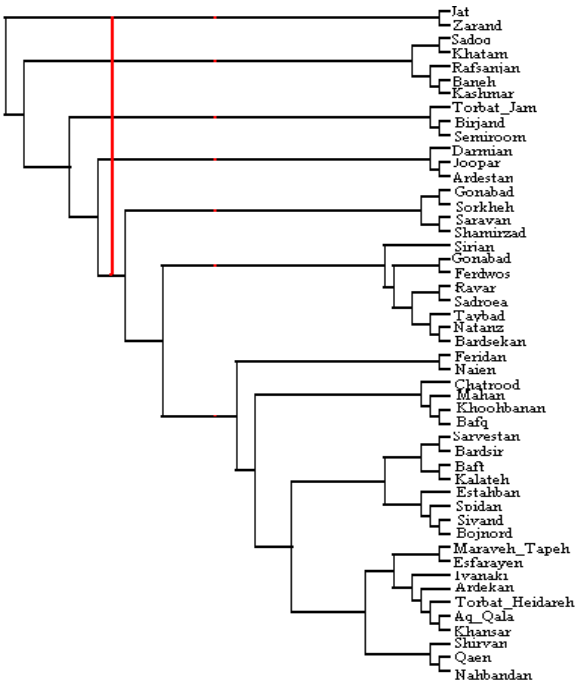

Fig 4. Dendrogram of 49 cumin ecotypes based on RAPD markers, according to the un-weighted pair group mean algorithm (UPGMA) method based on a similarity matrix by NTYSYs 2.1 software.

Principal coordinate analysis (PCOA)

Principal coordinate analysis (PCOA) was performed and consequently two-dimensional plot was drawn based on the first two main components. The program Power Marker version 3.25 (Liu et al., 2005) was used to calculate allele frequencies and alleles per locus. Polymorphism information content (PIC) which indicated the ability to distinguish between genotypes for each primer combination also was calculated.

Acknowledgments

The author would like to thank Assoc. Prof. Dr. Mohd Said Saad from Agrogene Bank, University Putra Malaysia for his guidance and Mr. Alagi Bah for his valuable comments and editing of the manuscript.

References

Agrawal S (1996) Volatile oil constituents and wilt resistance in cumin (*Cuminum cyminum* L.). Curr Sci India 71:177–178.

Allard RW (1999) Principles of Plant Breeding. John Wiley and Sons, New York.

Bahraminejad A, Mohammadi-Nejad G, Abdul Khadir M (2011) Study of genetic diversity of Cumin (*Cumin cyminum* L.) based on phenotypic characteristics. Aust J Crop Sci 5(3):301-307.

Banerjee M, Sarkar PK (2003) Microbiological quality of some retail spices in India. Food Res Int. 36:469-474.

Bonyanpour AR, Khosh-Khui M (2001) Factors influencing seed germination and seedling growth in black zira (*Bunium persicum*). J Herbs Spices Med Plants. 8:79- 86.

Bosch L, Casanas F, Sanchez E, Nuez V (1997) Variability of maize landraces from Northwest Spain. Plant Genet Resources News let. 112:90-92.

Dos Santos JB, Nienhuis J, Skroch P, Tivang J, KSlocum M (1994) Comparison of RAPD and RFLP genetic markers in determining genetic similarity among (*Brassica oleracea* L.) genotypes. Theor Appl Genet. 87:909-915.

Jaccard P (1908) Nouvelles recherchers sur la distribution florale. Bull Vaud Soc Nat. 44: 233-270

Kapila RK, Panwar KS, Badiyala D (1997) Variation and association analysis in domesticated populations of black caraway *(Bunium persicum*). J Med Aromat Plant Sci. 19: 709-711.

Kermani M, Marashi SH, Safarnejad A (2009) Investigation of genetic variation within and among two species of Cuminum spp. using AFLP markers. Iran J Chem Eng. 16(2):199–207.

Krizman M, Jakse J, Baricevic D, Javornik B, Prosek M (2006) Robust CTAB-activated charcoal protocol for plant DNA extraction. Acta Agr Slove. 87: 427-433.

Liu K, Muse SV (2005) Power Marker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 21:2128- 2129.

Mittal RK, Chahota RK, Gartan SL, Katna G (2006) Genetic variability and component analysis in kalazira *(Bunium persicum)* in dry temperate areas of north-western Himalayas. J Crop Improvement. 33: 202-204.

Moeller DA, Schaal BA (1999) Genetics relationships among native American maize accessions of the Great Plains assessed by RAPDs. Theor Appl Genet. 99:1061-106728

Pezhmanmehr M, Hassani A, Jahansooz F, Najafi A, Sefidkon F, Mardi M, Pirseiedi MA, Naghavi MR (2009) Assessment of genetic diversity in some Iranian populations of *Bunium persicum* using RAPD and AFLP markers. Iran J Bio. 7(2):93-100.

Rohlf FJ (2000) NTSYS-pc Numerical taxonomy and multivariate analysis system, Version 2.1 (New York: Exeter) Owners manue.

Sheidai M. and Ahmadian P (1996) Cytological studies in Iran Zira from three genera: *Bunium, Carum* and. *Cuminum.* Cytologia. 61:19-25.

Takayanagi T, Ishikawa T, Kitajima J (2003) Sesquiterpene lactone glucosides and alkyl glycosides from the fruit of cumin. Phytochemistry. 63:479–484.

Thormann CE, Ferreira ME, Camargo LEA, Tivang JC, Osborn TC (1994) Comparison of RFLP and RAPD markers for estimating genetic relationships within and among cruciferous species. Theor Appl Genet. 88:973-980.

Valizadeh M, Safarnejad A, Nematzadeh G, Kazemitabar S (2008) Regeneration of plantlets from embryo explants of *Bunium persicum*. J Sci Technol Agri. 11 (42):33-38.

Zeven AC (1996) Results of activities to maintain landraces and other material in some European countries in situ before 1945 and what we learn from them. Genet Res and Crop Evol. 43: 337-341.