Genetic and phytochemical diversity analysis in *Bunium persicum* populations of north-western Himalaya

Rakesh Kumar Chahota¹ · Vikas Sharma^{1,2} · Minerva Ghani¹ · Tilak Raj Sharma¹ · Jai Chand Rana³ · Shyam Kumar Sharma⁴

Received: 7 September 2016/Revised: 1 March 2017/Accepted: 6 March 2017/Published online: 20 March 2017
© Prof. H.S. Srivastava Foundation for Science and Society 2017

Abstract The present study explores morphological, genetic and phytochemical composition of *Bunium persicum* populations belonging to high altitudinal areas of Indian Himalayan region. In total, 23 morphological traits (13 quantitative and 10 qualitative traits) and 32 Random Amplified Polymorphic DNA primers were employed to infer the population structure of the species. Of the fourteen populations, five genetically diverse populations were analyzed for phytochemical diversity. Among morphological traits, inflorescence, seed and branch traits were most significant in detecting variation. At molecular level, primers TIBMBA-06 and OPR-16 were found most polymorphic with respect to Polymorphism Information Content and Marker Index values. Dendrogram grouped all populations into two major clusters while population from Shong region

out grouped separately showing its distantness from all other populations. STRUCTURE analysis was done by using Bayesian model, which characterised all populations into four clusters and some degree of admixture was also observed within individuals. Shong population showed distinct genetic makeup as also suggested by dendrogram. Phytochemical analysis showed the presence of 55 components, of which, 2-methyl-3-phenyl propanal, benzeneacetic acid, 1-phellandrene, γ -terpene, α -terpinolene, Δ 0.3-carene and sabinene were major components in its essential oils. The present study revealed high genetic and phytochemical diversity in *B. persicum* accessions from north-western Himalayan regions. Specifically, accessions from Saptal regions were having higher quantity of essential oils and can be selected for cultivation to meet the commercial demand to some extent. Further, the diversity information provided herein can be useful in management and improvement of this species through future breeding programmes.

✉ Rakesh Kumar Chahota
rkchahota@yahoo.com

Vikas Sharma
vikasam@gmail.com

Minerva Ghani
minervaghani@yahoo.in

Tilak Raj Sharma
trsharma88@gmail.com

Jai Chand Rana
ranajc2003@yahoo.com

Keywords *Bunium persicum* · Morphological traits · RAPD markers · GC–MS technique · Genetic structure · Phytochemicals

Introduction

Bunium persicum (Boiss) Fedtesch, also known as black cumin or Kala Zira, belongs to family Apiaceae. It is highly valued for its seeds, which are used as spice and medicinal constituents in many medications to cure a number of diseases. It is a small herbaceous plant that grows naturally in temperate to alpine regions (1850–3100 m, amsl) of the world and is considered to have originated from Central Asia to Northern India (Azizi et al. 2009). It is a perennial herb with a diploid chromosomal number of $2n = 14$

¹ Department of Agricultural Biotechnology, CSKHP Agricultural University, Palampur 176062, H.P, India

² Department of Botany, Punjabi University, Patiala 147002, Punjab, India

³ National Bureau of Plant Genetic Resources, New Delhi, India

⁴ CSIR-Institute of Himalayan Bioresource Technology, Palampur 176061, H.P, India

(Sheidai et al. 1996). As far as its global distribution is concerned, it is found in the high altitude regions of Iran and some parts of Afghanistan, Pakistan and Tajikistan in addition to India (Panwar 2000; Hanelt et al. 2001). In India, it is distributed in sparse patches, which are confined to few high altitude Himalayan regions of Kashmir, Himachal Pradesh, and Uttarakhand. Generally, it grows as a wild plant in the forests and grassy slopes. The specific regions having its populations comprises of Kinnaur, Lahaul-Spiti, Pangi and Bharmaur regions of Himachal Pradesh, Ladakh and Padder valley of Jammu & Kashmir and some ranges of Uttarakhand Himalayas. It is a small plant, grows up to 80 cm in height, and produces white or pink compound umbel of flowers on the terminal and lateral stems during the third year of its life (Omidbeigi 1997).

The seeds of *B. persicum* are mainly used as a carminative and spice in food industry for flavouring purposes (Pourmortazavi et al. 2005; Gincarlo et al. 2006). Seeds are reported to have stimulant, expectorant, antispasmodic and diuretic properties (Chauhan 1999; Kala 2003). Moreover, seeds are frequently used in treating diarrhea, dyspepsia, fever, flatulence, stomachic and numerous other diseases by local people (Sardari et al. 1998; Peter 2004; Panda 2004; Boskabady and Moghaddas 2004). Further, the essential oils extracted from seeds are reported to have noteworthy antioxidative, antibacterial and antifungal activities (Shahsavari et al. 2008; Moghtader et al. 2009; Talei and Mosavi 2009; Sekine et al. 2007). Besides, its high commercial utility it remained overlooked and neglected for its cultivation practices and the major portion of the seed produce is harvested extensively from its natural habitats. This existing practice of unsustainable and unscientific exploitation of this species has resulted in reduction of its natural populations and has made it a plant of special conservation concern in north-western Himalayas. Although some efforts have been done in the recent years to domesticate it in certain areas of north-western Himalayas, but due to non-availability of suitable genotypes, cultivation remained very limited. There was only one site, namely, Shong in Kinnaur district of Himachal Pradesh, where it is being cultivated by folks. However, collectively wild and cultivated resources of *B. persicum* at present are not enough to meet the requirements of its different end users. Therefore, this crop needs much more attention at present for its sustainable exploitation by the farmers and improvement through scientific methods. Thus, the current situation demands the development of scientific and sustainable approaches to be followed by growers to increase the production and to conserve the diverse resources of this plant species existing in western Himalaya.

Exploration and evaluation of genetic diversity is a prerequisite for identifying the suitable genotypes in a crop,

which can be utilized further for initiation of breeding and other improvement programmes. The breeding programmes can result in development of desired elite genetic stocks for commercial cultivation. In spite of prized herb having many therapeutic uses, the information on genetic diversity present among the different populations of *B. persicum*, especially in North-western Himalayan region is severely lacking. Similarly, information on biochemical and morphological diversity is also unavailable in this crop. The morphological characterization of many plant species is difficult because, they resemble morphologically and it is very difficult to differentiate between them (Martinez et al. 2007). This necessitates the use of molecular markers, particularly DNA markers for distinguishing accessions, removing duplicates, DNA fingerprinting of ambiguous genotypes, determining the extent of genetic diversity and signifying the path of domestication (Wunsch and Hormaza 2002). Moreover, phytochemical evaluation is needed to identify genotypes with greater quantities of specific active principles, which can be included in breeding programme for further improvements of related traits (Shahsavari et al. 2008). However, some reports on essential oil study in Iranian populations using GC/MS technique are available but no report on populations belonging to north-western Himalayan region exists. Therefore, in present study we assessed the genetic diversity of *B. persicum* populations of North-Western Himalayan region using various morphological descriptors, randomly amplified polymorphic DNA (RAPD) markers and phytochemical parameters. The results of present study can be used to select diverse lines to be used for cultivation and production as of new cultivars directly or elite genetic stocks can be developed through systematic improvement programmes in future.

Materials and methods

Plant material

Material in the form of tubers was collected during October–November from the already demarcated areas on the basis of plant density and population diversity in the month of June–July. Tubers uprooted from isolated forest areas were designated as the different wild populations (Table 1) whereas tubers of cultivated population were collected from the Shong area, which is the only *Bunium* cultivating place of the entire area under study. Tubers of equal size and weight of all the populations were planted in the month of October at Mountain Agricultural Research and Extension Center (MAREC) Sangla situated at elevation of 2670 m amsl (78°16′18.92″E 31°25′46.18″N) in the Kinnaur district of Himachal Pradesh, India. All the

Table 1 List of populations collected from different parts of North Western Himalayas

S. no.	Population name	Altitude (m)	Latitude (N)	Longitude (E)
01	Thunaar	1850	32°26'58.85"	76°27'46.20"
02	Purthi	2583	32°33'33.60"	76°58'25.83"
03	Saptal	2816	32°41'26.10"	76°41'42.45"
04	Chamrat	3103	32°49'48.36"	76°58'25.83"
05	Tindi	2905	32°45'33.01"	76°30'56.57"
06	Kukumseri	2641	32°43'28.18"	76°44'57.28"
07	Kalpa	1965	31°29'57.94"	78°10'16.73"
08	Rispa	2340	31°35'32.50"	78°10'16.73"
09	Moorang	2672	31°35'16.66"	78°10'16.73"
10	Jangi	2774	31°35'26.48"	78°10'16.73"
11	Shong	2833	31°26'44.72"	78°11'43.65"
12	Batseri	2733	31°24'26.80"	78°18'16.02"
13	Chansu	2891	31°24'54.80"	78°25'27.92"
14	Sangla	2733	31°25'46.18"	78°16'18.92"

populations were planted in a randomized block design with three replications. Forty tubers per population were planted in 2 m long plot of 2 rows in three replications with row to row distance of 40 cm and distance between two plants was 10 cm.

Morphological evaluation

Morphological data related to 23 traits were recorded at experimental fields at MAREC, Sangla under natural conditions. Of the 23 traits, 13 were quantitatively measured, while 10 were qualitative which were differentiated on the basis of a notable variation in these traits. The quantitatively measured traits were plant height, number of umbels per plant, umbelets per umbel, fertile umbels, fertile umbelets per umbel, seeds per umbelets, primary branching, secondary branching, internode distance, days to flowering, maturity group while qualitative traits were flower colour, leaf colour, leaf pigment, leaf pigment colour, stem pigment, stem pigment colour, pigment on umbel, umbel pigment colour, disease and insect attack on plant. Data for each trait was recorded on 20 randomly taken plants in three replications. Mean values of these replications were used to analyse the data. Different diversity measures such as variances, correlations and principal component analysis (PCA) were determined by using the software SYSTAT-12 and trial versions of XLSTAT and StatistiXL version 1.10.

Molecular analysis

DNA isolation and RAPD reactions

Fresh leaf samples from each plant were taken for DNA isolation. DNA was extracted using method described by

Doyle and Doyle (1990). Transparent DNA pallets were dissolved in appropriate amount of TE buffer. DNA quantification was done using 0.8% agarose gel and uncut λ DNA. All the DNA stocks were diluted to make working DNA of 13 ng/ μ l. These working dilutions were also checked on 0.8% agarose before performing PCR reactions. RAPD technique given by Williams et al. (1990) was used to study DNA polymorphisms. Initially 53 RAPD primers were screened and finally 32 polymorphic RAPD primers were selected for genetic diversity analysis. RAPD reaction mixture was made as per Sharma et al. (2015), which consisted of 1.5 μ l DNA (13 ng/ μ l), 2.5 μ l 10 $\times$ PCR buffer (10 mM Tris-pH 9.0, 50 mM KCl, 0.01% Gelatin, 1.5 mM MgCl₂), 0.2 μ l dNTPs (25 mM), 1.0 μ l MgCl₂ (25 mM), 1.0 μ l primer (2 μ M/ μ l), 0.1 μ l Taq DNA polymerase (5 U/ μ l) and deionised water to prepare total volume of 25 μ l. The PCR was run with initial cycle at 94 °C for 5 min, followed by 45 cycles of at 94 °C for 1 min, 37 °C for 1 min, and 72 °C for 2 min. The final cycle consisted of 72 °C for 7 min. Afterwards, PCR products were separated by electrophoresis using 1.5% (w/v) agarose gel. Gel electrophoresis was carried out in TBE buffer at 70 V and DNA fragments were detected using ethidium bromide. Permanent photographs were captured with the help of Gel Documentation system (Syngene, USA).

Phytochemical evaluation

Five populations, which showed high molecular diversity, were taken for essential oils analysis using Gas chromatography/mass spectrometry (GC/MS). The seeds of each sample (250 g) were ground separately and subjected to hydro-distillation for 3 h, with Clevenger apparatus (Borosil, Mumbai, India). Oils were dried with anhydrous

sodium sulfate and their percent oil content was calculated on the basis of seeds dry weight. Oils were stored at -4°C till GC/MS analysis which was done with the help of Shimadzu GC/MS QP 2010 system (Shimadzu Co., Japan), using HP-5 capillary column. Temperature of oven was kept 40°C for 4 min and raised to 220°C for 15 min at an increment of 4°C per minute. The carrier gas was Helium. Samples were run in a ratio of 1:50. Quantitative data were obtained in the form of peaks and essential oil components were identified by comparing their mass spectra with those of the spectral computer library. Confirmation of results was done by comparing retention indices to Safidkon and Rahinis' authentic compounds (Sefidkon and Rahimi 2002; Sefidkon and Jamzad 2005).

Data analysis

All prominent and polymorphic fragments detected by RAPD primers were used to prepare binary data. This data was then used for all the downstream analyses. Unweighted pair group method of arithmetic mean (UPGMA) dendrogram using Jaccard's coefficient was constructed using software DARwin (Perrier and Jacquemoud-Collet 2006). Factorial analysis was also done by DARwin software. Population genetic structure was inferred using STRUCTURE software version: 2.3.3 (Pritchard et al. 2000; Falush et al. 2007). Parameters in STRUCTURE software were set as per Sharma et al. (2015). Ten runs were performed with 3 iterations in a range of 1–14 populations for K. The length of burn-in period and number of Markov Chain Monte Carlo (MCMC) repeat after burn-in was set at 100,000. Results of STRUCTURE analysis were taken as input for STRUCTUREHARVESTER software (Evanno et al. 2005; Earl and Vonholdt 2011) which was used to calculate the best fit value of K.

Results and discussion

Morphological evaluation

A comprehensive evaluation of a species for the extent of genetic diversity present among the populations is highly beneficial for sustainable utilization of its available resources and for their management. The present study was an effort towards characterizing and correlating various morphological, molecular and phytochemical traits in *B. persicum* germplasm of western Himalayas. As morphological traits can be observed directly and easy to score, presence of phenotypic diversity increases the scope of a crop and provides opportunities to work upon. In different populations of *B. persicum* we found high variability in both qualitative and quantitative traits. Among

the ten qualitative traits recorded, leaf colour was the most variable trait followed by flower colour, leaf pigment colour, stem pigment colour, umbel pigment colour and disease and insect attack on plant. Analysis of variance for 13 quantitative traits revealed the presence of significant diversity among populations for all the traits studied. Analysis of correlation of different traits was performed to ascertain the information on association of traits with seed yield and the inter-relationship among themselves. It was observed that 1000 seed weight has positive association with seed yield ($r^2 = 0.743$, $p < 0.05$). Other traits having positive and high coefficient of correlation with seed yield were fertile umbels/umbel, primary branches, fertile umbelets/umbel, umbels/plant, umbelets/umbel and seeds/umbel. This indicates that three traits namely seed size, number of fertile umbels and number of primary branches/plant are the important traits for direct selection of the desirable plants and one or two traits from yield contributing traits may be taken into consideration. All these traits also have positive correlations among themselves (Table 2). Traits such as plant height, internodes distance, days to flowering and days to maturity were having very low correlations with seed yield and other yield contributing traits. Preliminary studies by few earlier workers also reported morphological diversity in this crop (Kapila et al. 1997; Mittal et al. 2006). The variation in these traits may be attributed to climatic condition prevailing in the isolated habitats, which are separated by extreme geographical conditions such as higher peaks and deep ravines of Greater Himalayas. In addition, adaptability to very cold climatic (-45°C), regimes by different populations can be a reason behind these variations. The principal component analysis indicated that the first principal component (PC1) was accounting for 46.2% of the variance. The important traits included in the PC1 were fertile umbel, fertile umbelets/umbel, umbel/plant, umbelets/umbel, primary branches/plant, seed weight, seeds/umbel and yield/plant (Fig. 1). While PC2 was contributing 16.2% variance and traits included in this group were days to flowering, 1000-seed weight and maturity group. Similarly, PC3 and PC4 were showing 13.3% and 7.0% variance with major traits of plant height, seeds/umbelets, secondary branching, internodes distance and days to flowering respectively (Table 3). The PCA analysis further showed that first four components were harboring 82.9% variations and major traits contributing to this variation were inflorescence related traits, primary branches/plant, seed weight, seeds/umbel, 1000-seed weight, maturity group, plant height, secondary branching and internodes distance and days to flowering. Therefore, traits directly related to seed, branching, days to flowering and maturity group can be considered very important from the improvement point of this crop. Euclidean distance

Table 2 Correlation coefficients among thirteen quantitative traits studied in *B. persicum*

	PHT	UPP	ULPU	FUPP	FULPU	SPU	PB	SB	IND	DOF	DOM	SW	SY
PHT	1.000												
UPP	0.344	1.000											
ULPU	0.589*	0.768*	1.000										
FUPP	0.375	0.988*	0.773*	1.000									
FULPU	0.631*	0.729*	0.924*	0.755*	1.000								
SPU	0.463	0.449	0.532*	0.436	0.433	1.000							
PB	0.330	0.750*	0.684*	0.768*	0.639*	0.474	1.000						
SB	0.135	0.536*	0.218	0.561*	0.405	−0.307	0.273	1.000					
IND	0.594*	0.374	0.396	0.348	0.299	0.266	0.462	0.263	1.000				
DOF	0.066	−0.051	0.072	−0.089	−0.115	0.305	0.037	−0.301	0.159	1.000			
DOM	−0.535*	0.155	−0.001	0.104	−0.095	0.088	0.035	−0.007	−0.416	0.384	1.000		
SW	0.126	0.534*	0.477	0.520*	0.406	0.560*	0.608*	0.035	0.239	0.369	0.242	1.000	
SY	0.294	0.540*	0.518*	0.552*	0.616*	0.506*	0.611*	0.292	0.207	0.187	0.145	0.743*	1.000

PHT plant height, UPP umbel per plant, ULPU umblets per umbel, FUPP fertile umbel per plant, FULPU fertile umblets per umbel, SPU seed per umbel, PB primary branches per umbel, SB secondary branches per plant, IND internode length per plant, DOF days to flowering, DOM days to maturity, SW 100-seed weight, SY seed yield per plant

Fig. 1 Principal component analysis showing biplot of different morphological variables loaded on PC1 and PC2

based cluster analysis based on morphological traits suggested two main clusters, while Sangla population out grouped (Fig. 2). Cluster I contained the nine populations belonging to Chamba, Kinnaur and Lahaul-Spiti districts. The early flowering & medium maturity, large number of secondary branches and longer internode distance characterized these populations. Cluster II was composed of four populations, all belonging to Kinnaur district and characterized by low yield and small plant size. The relationships among studied population revealed by dendrogram on the basis of quantitative traits data matrix was phenotypic in nature. The low precision of dendrogram

may be attributed to less numbers of traits recorded and continuously varying nature of these traits.

Molecular analysis

The DNA based markers provide an important tool for the precise estimation of genetic diversity with additional advantage that they are not influenced by the environmental conditions. Molecular diversity analysis using 32 polymorphic RAPD primers generated 164 bands in a size range of 100–2800 bp. Of these, 120 bands (73.1%) were found to be polymorphic across analyzed populations. The

Table 3 Eigen vector values of the first four principal components of quantitatively measured traits

Morphological trait	Component			
	1	2	3	4
Plant height	0.240	−0.291	−0.435	−0.048
No of umbels/plant	0.364	−0.022	0.239	0.010
Umblets/umb.	0.361	−0.034	−0.051	−0.237
Fertile umbles	0.367	−0.053	0.243	−0.023
Fertile umblets/Umbel	0.354	−0.141	0.038	−0.293
Seeds/umblet	0.252	0.270	−0.348	−0.312
Primary branching	0.343	0.042	0.052	0.068
Secondary branching	0.162	−0.315	0.470	0.414
Internode distance	0.211	−0.236	−0.327	0.589
Days to flowering	0.036	0.463	−0.304	0.461
Maturity group	0.000	0.523	0.379	0.023
1000-seed weight	0.274	0.365	−0.027	0.141
Yield/plant	0.302	0.201	0.047	0.047
Eigenvalue	6.014	2.112	1.736	0.916
Variability %	46.262	16.244	13.351	7.048
Cumulative %	46.262	62.506	75.857	82.906

average numbers of polymorphic bands were 3.75 among all the primers. Primer OPR-16 amplified maximum of 9 bands while maximum polymorphism information content (PIC) and marker index (MI) was detected by primers TIBMBA-06 and OPR-16, respectively. We detected mean values of PIC and MI were 0.366 and 1.51 respectively, which indicate that sufficient polymorphism is present in all populations analyzed. Other workers have employed different types of markers to study the genetic diversity of wild collections of *B. persicum* in different parts of world (Hashemi et al. 2010). In Iranian populations of *B. persicum* 86% and 75% polymorphism has been reported with RAPD and AFLP markers respectively (Azimzadeh et al.

2012) while here we recorded an overall polymorphism of 73.1%, which is comparable to these earlier reports (Table 4). The dendrogram based on RAPD polymorphism using Jaccard's dissimilarity coefficient grouped all the populations of *B. persicum* into mainly two clusters but the Shong population was separated as a different out-group (Fig. 3). Group-I was comprised of Saptal, Purthi and Thunaar populations while group-II was consisted of ten populations, namely, Kukumseri, Chamrat, Jangi, Rispa, Kalpa, Tindi, Chansu, Batseri, Sangla and Moorang. Two directional factorial analysis also revealed that various populations were distributed mainly into two groups (Fig. 4). Although, populations from same region largely grouped together, which was in agreement with earlier reports in this species from various regions of the world (Majeed and Sharma 2006; Jahansooz et al. 2007; Pezhmanmehr et al. 2007). This may be due to its long term cultivation that may have resulted into accumulation of some desirable allelic combinations in the population due to prolong selection pressure. Further, grouping of RAPD based dendrogram was in discordance to morphological grouping but showed more precision than the later. The differences in genetic diversity information generated by morphological traits and molecular markers may be accounted by numerous reasons. Generally, morphological variation is highly influenced by environmental conditions and morphological similarities or differences are not always reflected at genetic level (John et al. 1997). Moreover, adaptive value for specific traits may have accumulated in specific ecological conditions of their habitat, irrespective of the genetic origins of the populations (Steiner and los Santos 2001). Bayesian clustering grouped 14 populations into four clusters. Cluster-I was the largest having seven populations namely, Kalpa, Moorang, Rispa, Jangi, Kukumseri, Chamrat and Sangla. All the populations in this cluster were showing more than 80% membership

Fig. 2 Dendrogram based on morphological traits constructed using squared Euclidean distance and group average clustering method

Table 4 List of Random primers used and diversity indices of *B. persicum* accessions

S. no.	Primer name	Primer sequence (5'-3')	Polymorphic bands	PIC	MI
1	BP1TS2R3-12	TAGTATATGTGCTA	6	0.459	2.76
2	BP1TS1R3-10	GCAGTCATAGTCATG	5	0.245	1.22
3	TIBMBC-10	GTCGTAGCTCACGT	5	0.408	2.04
4	BP1TS2R1-2	TACTTGCATGCATGA	2	0.459	0.92
5	TIBMBA-06	CGTATCGATCGTCTT	6	0.500	3.00
6	BP1TS1F2-7	TGACCTGCATCGCTAT	5	0.133	0.66
7	TIBMBA-17	AGTCATAGCCTGCATC	1	0.245	0.24
8	TIBMBC-06	ACTGCATCAGTCATAG	4	0.469	1.88
9	TIBMBA-02	CCTCAGTCATGCATC	4	0.375	1.50
10	BP1TS2R2-4	CAGTCATCCTGCATC	6	0.444	2.67
11	BP1TS2F2-3	TTGTCCTGATCGAT	4	0.133	0.53
12	OPE-02	GGTGCGGGAA	2	0.497	0.99
13	OPR-16	CTCTGCGCGT	9	0.499	4.50
14	OPN-05	ACTGAACGCC	1	0.490	0.49
15	OPM-10	TCTGGCGCAC	4	0.448	1.79
16	OPR-02	CACAGCTGCC	5	0.500	2.50
17	OPM-04	GGCGGTTGTC	4	0.270	1.08
18	OPC-11	AAAGCTGCGG	5	0.498	2.49
19	OPS-14	AAAGGGGTCC	1	0.133	0.13
20	OPN-02	ACCAGGGGCA	1	0.133	0.13
21	OPR-20	ACGGCAAGGA	5	0.265	1.32
22	OPB-05	TGCGCCCTTC	2	0.436	0.87
23	OPS-05	TTTGGGGCCT	1	0.337	0.34
24	OPI-20	AAAGTGCGGG	8	0.490	3.92
25	OPE-12	TTATCGCCCC	5	0.496	2.48
26	OPC-07	GTCCCGACGA	2	0.293	0.59
27	OPM-03	GGGGGATGAG	2	0.245	0.49
28	OPN-03	GGTACTCCCC	4	0.448	1.79
29	OPB-03	CATCCCCCTG	7	0.499	3.49
30	OPS-16	AGGGGGTTCC	2	0.490	0.98
31	OPS-04	CACCCCCTTG	1	0.133	0.13
32	OPS-07	TCCGATGCTG	1	0.245	0.24
Mean			3.75	0.366	1.51

PIC polymorphism information content, *MI* marker index

except Sangla population. Cluster-II was having populations from Saptal, Thunaar, Purthi and Tindi. Except Tindi population, all the three populations were showing more than 95% membership. Cluster III was having the only cultivated population in Himalayan region. Cluster IV was having two populations Chansu and Batseri and were showing 100% membership in this cluster (Fig. 5). It was interesting to observe that only cultivated population from Shong formed a discrete cluster and not sharing any genetic similarity with rest of the populations. The four genetic clustering was influenced by geographic proximity and habits of the plant. Cluster-II was entirely pure and different from other three clusters (Fig. 5). It represented cultivated population of Shong, the cultivation practice

followed from historical times, selection pressure and recombination events within same population may have given rise to entirely different allelic combinations. Geographic distribution of genetic structure revealed that two gene pools (cluster-I and cluster-IV) were spread along entire region under investigation and seems to be primitive or parent gene pools, from which two others (cluster-II and cluster-III) have originated through isolation and self hybridization events. Admixture within populations (Sangla and Tindi) was also observed, which may be the result of cross breeding behaviour of different populations (Fig. 6). Thus, it can be inferred that geographic topology and habits of the plant are major factors in shaping its genetic structure.

Fig. 3 Dendrogram based on RAPD data constructed using Jaccard's similarity coefficient and UPGMA method

Fig. 4 Factorial analysis based on RAPD data. Different colors represent different geographic regions (colour figure online)

Phytochemical evaluation

The seeds of *B. persicum* are used in pharmaceutical industry for its important essential oils and as spice in food industry. Therefore, it is important to characterize these populations for various phytochemicals to the benefit of farming community and industry. Total essential oils estimation of *B. persicum* by GC/MS analysis revealed that Saptal population of Lahaul Valley yielded maximum

essential oil (6.4%) and the cultivated population from Shong possessed the lowest (3.8%) percentage of essential oils. Similar percentage of essential oil contents (5.1–6.2%) has been reported in the cultivated samples of Iranian *B. persicum* (Azizi et al. 2009). However, in Iranian *B. persicum*, essential oil content in fresh and dried seeds was reported to be 5–14 and 3–6% respectively (Azimzadeh et al. 2012). In total fifty-five compounds were identified, representing 98.59% of total essential oils of Shong

Fig. 5 STRUCTURE analysis using RAPD data, which assigned 14 populations of *B. persicum* into four genetic clusters

Fig. 6 Geographic distribution of genetic structure of *B. persicum* in north-western Himalayas, showing prevalence of two gene pools denoted by red and yellow colours (colour figure online)

population, 99.17% of Chansu region, 96.86% of Saptal region, 97.89% of Kalpa region and 99.69% of Moorang region. The main essential oil components identified in all the five populations were 2-methyl-3-phenyl propanal (11.22–30.4%), γ -terpinene (21.96–41.27%), benzenoacetic acid, α -hydroxy-ethyl ester (CAS) ethyl mandelate (15.21–21.54%), 1-phellandrene (17.44–22.48%), sabinene (1.83–2.36%) and β -myrcene (1.05–1.54%). The compound 2-methyl-3-phenyl propanal is analogous to widely reported cuminaldehyde (Table 5). Highest concentration of 2-methyl-3-phenyl propanal was observed in Kalpa population (32.62%) followed by Shong population (30.4%). Population of Moorang was found to possess highest concentration of both 1-phellandrene (21.84%) and γ -terpinene (41.27%). Considerable percentage of

acetylphenylcarbinol was detected in only two populations i.e. in Chansu (16.68%) and Kalpa (13.87%), although it was completely absent in rest of three populations. Similarly, Bicyclo (3.2.1) oct-2-ene, exo-4-(phenyl sulphonyl) (27.52%) was found to be present only in Saptal population. Further, grouping of populations based on their phytochemical components showed that, of the five populations, Chansu was most distant than other four (Saptal, Kalpa, Shong and Moorang) populations. Saptal and Kalpa populations were most similar (Fig. 7). As the prevailing climatic regimes and adaptive behaviour of plant towards these regimes can alter the genetic variation, the altered genetic components can have direct or indirect impact on the composition and biological activity of an active principle (Heywood 2002). Some earlier worker

Table 5 Various phytochemical constituents recorded in fourteen populations of *B. persicum* from western Himalayas

Components	Shong	Kalpa	Moorang	Chansu	Saptal
I. Essential oil yield/dry seed weight (%)	3.8	4.4	5.2	6	6.4
II. Essential oil components (%)					
1. α -Thujene	0.76	0.89	0.85	0.94	1.16
2. Bicyclo[3.1.1]hept-2-ene-2-ethanol,6,6-dimethyl-	0.10	–	–	–	–
3. α -Pinene	0.21	0.23	0.24	0.27	0.34
4. Bicyclo[3.2.1]oct-2-en-7-one	0.03	–	–	–	–
5. Sabinene	1.83	2.3	2.18	2.06	2.36
6. β -Pinene	0.24	0.19	0.17	0.24	0.29
7. β -Myrcene	1.05	1.21	1.21	1.44	1.54
8. Δ 3-Carene	0.11	0.12	0.13	21.72	0.16
9. α -Terpinolene	0.31	0.63	0.56	0.27	0.76
10. 1-Phellandrene	17.44	22.48	17.14	20.27	21.84
11. γ -Terpene	21.96	41.26	26.65	–	41.27
12. α -Terpinolene	0.24	0.68	0.61	0.36	0.76
13. Fenchone	0.91	–	–	0.97	–
14. L-Fenchone		0.17	0.42		
15. 3-Cyclohexen-1-ol, 5-methylene-6-(1-methylethenyl)-, acetate(CAS)	0.03	–	–	–	–
16. Linalool	0.05	–	–	–	–
17. Epoxyterpinolene	0.09	–	0.08	0.12	–
18. 3-Isopropylidene-5-methyl-hex-4-en-2-one	0.33	–	0.21	0.56	–
19. 2-Methyl-1-hepten-6-yn-3-ol	0.33	–	–	0.55	–
20. 3-Cyclohexen-1-ol, 4-methyl-1-(1-methylethyl)-CAS4-terpineol	0.25	0.26	0.33	0.43	0.3
21. 1-Cyclohexene-1-carboxaldehyde, 4-(1-methylethyl)-	0.07	1.05	1.3	0.48	0.93
22. Propanal, 2-methyl-3-phenyl-	30.4	11.22	17.00	32.62	11.91
23. 1-Cyclohexene-1-carboxaldehyde, 4-(methylethyl)	0.07	0.19			
24. Benzeneacetic acid, α -hydroxy-, ethyl ester(CAS)ethyl mandelate	21.54	–	–	–	15.21
25. 2-Caren-10-al	0.13	–	0.17	–	–
26. Phenol, 5-methyl-2-(1-methylethyl)-(CAS)thymol	0.26	–	–	0.48	–
27. (–)-Caryophyllene oxide	0.04	–	–	0.11	–
28. (+)Spathulenol	0.08	–	–	0.15	–
29. α -Bisabolol	0.04	–	–	–	–
30. 5-Methyl-3-(1-methylvinyl)-1, 4-hexadiene	–	0.07	0.08	–	–
31. Camphene	–	0.02	–	–	0.02
32. L-Fenchone	–	0.17	0.42	–	–
33. 4,8-Epoxy- <i>p</i> -menth-1-ene	–	0.05	–	–	0.06
34. 3-Cyclohexene-1-carboxaldehyde, 1,3,4-trimethyl-	–	0.1	–	–	0.07
35. 2-Methyl-1,5-heptadien-4-ol	–	0.09	0.19	–	0.08
36. 3-Cyclohexene-1-methanol, $\alpha,\alpha,4$ -trimethyl-,(S)-(CAS)p-M	–	0.1	–	–	–
37. 3-Cyclohexene-1-ol, 4-methyl-1-(methylethyl)-(CAS)4-terpineol	–	0.26	0.33		0.30
38. 1-Cyclohexene-1-carboxaldehyde,4-(1mehtylethyl)	–	1.05	1.30		0.93
39. Acetylphenylcarbinol	–	16.68	–	13.87	–
40. Bicyclo[3.1.0]hexan-3-ol, 4-methyl-1-(1-methylethyl)	–	–	0.07	–	–
41. Bicyclo[3.2.1]oct-2-ene, exo-4-(phenylsulfonyl)-	–	–	27.52	–	–
42. 1,4-Cyclohexadiene-1-methanol, 4-(1-methylethyl)-	–	–	0.52	–	–
43. 1,3,5-Cycloheptatriene, 7,7-dimethyl-(CAS) 7,7-dimethyl-1,3,5-cycloheptatriene	–	–	–	0.07	–
44. 2-Pinen-10-hydroperoxide	–	–	–	0.04	–
45. 2-Caren-10-al	–	–	0.17		
46. <i>n</i> -Hexadecanoic acid	–	–	–	0.09	–

Table 5 continued

Components	Shong	Kalpa	Moorang	Chansu	Saptal
47. 3-Menthene	–	–	–	–	0.04
48. <i>cis</i> -Sabinene hydrate acetate	–	–	–	–	0.88
49. α -Thujone	–	–	–	–	0.22
50. 5,8-Dimethylene-bicyclo[2.2.2]oct-2-ene	–	–	–	–	0.03
51. Carveol (fr.1)	–	–	–	–	0.05
52. Limonene oxide	–	–	–	–	0.03
53. Myrtenyl acetate DB5-967	–	–	–	–	0.03
54. Phellandral	–	–	–	–	0.03
55. <i>trans</i> -Caryophyllene	–	–	–	–	0.04

Fig. 7 Hierarchical clustering of five *B. persicum* populations based on quantitative data of their phytochemical constituents, constructed using squared Euclidean distance and group average method

have observed a correlation between altitudinal variations with phytochemical variations and found that higher altitude samples possess lower amount of these constituents (Vokou et al. 1993), no such correlation was found in our study. In agreement with our results, γ -terpene and cuminaldehyde are the major essential constituents in Indian black cumin collections (Shankaracharya and Shankararayanana 1988). Gamma-terpinene, *p*-cymene and β -pinene are considered to be responsible for deterioration of spice quality (Salehi et al. 2008; Thappa et al. 1991; Baser et al. 1997; Foroumadi et al. 2002) but different populations showed their variable amount in them. Therefore, we can have more choices of accession to be included in improvement programmes of this species. Further, essential oils research in *Satureja bachtiarica* species have shown significant correlation and association between morphological and phytochemical traits (Khadivi-Khuba et al. 2014) However, in this study, we did not found any statistically significant association between studied morphological, molecular and phytochemical traits of black cumin. Further, in collective examination of phytochemical and structure analysis we found that Saptal and Kalpa populations were having maximum essential oils and cuminaldehyde contents respectively; genetically these populations were having pure lines in different gene pools. Therefore inclusion of accessions from these populations

in breeding programmes can yield better results. Likewise, we can choose suitable accessions on the basis of their phytochemical contents and genetic makeup. Overall, estimates of genetic diversity of *B. persicum* germplasm using DNA markers can have several implications in crop improvement programmes.

Conclusion

Present study reported high morphological, molecular and phytochemical diversity in analyzed *B. persicum* populations from north-western Himalayas. It is a positive aspect of these populations in contemporary environment, which otherwise is more depressing towards this plant species. The divergence of *B. persicum* germplasm can be helpful in perpetuation of this species through adaptation to changing climatic conditions. Moreover, this high phytochemical diversity also indicates the higher potential of *B. persicum* germplasms in terms of its essential oil based industrial produce. Further, diverse populations pointed in this study can prove an asset for initiating breeding programmes in this crop. It is suggested that diverse lines should be conserved under in situ conditions and should be multiplied by cultivation. The areas having natural populations of *B. persicum* should be identified and developed as conservatories by reducing anthropogenic disturbances.

Submission declaration

It is submitted that work described herein has neither been published nor communicated to any other journal for publication. All authors approve and declare that this manuscript will not be published in english or in any other language, including electronically in its present format without the written permission of the copyright holder.

Acknowledgements We gratefully acknowledge Ministry of Environment and Forests (MoEF), Government of India, New Delhi for providing financial support to conduct this research work.

Compliance with ethical standards

Conflict of interest All the authors declare that there is no conflict of interest.

References

- Azimzadeh M, Amiri R, Assareh MH, Bihanta MR, Forootan M (2012) Genetic diversity of Iranian *Bunium persicum* germplasm by morphological markers and essential oil components. *J Med Plant Res* 6:1119–1129
- Azizi M, Davarenejad GBR, Woerdenbag HJ, Kayser O (2009) Essential oil content and constituent of Black Zira (*Bunium persicum* [Boiss.] B. Fedtsch.) from Iran during field cultivation (Domestication). *J Essent Oil Res* 21:78–82
- Baser KHC, Ozek T, Abduganiev BE, Abdullaev UA, Aripov KN (1997) Composition of essential oil of *Bunium persicum* (Boiss) Fedtsch from Tajikistan. *J Essent Oil Res* 9:597–598
- Boskabady MH, Moghaddas A (2004) Antihistaminic effect of *Bunium persicum* on guinea pig tracheal chains. *Iran Biomed J* 8:149–155
- Chauhan NS (1999) Medicinal and aromatic plants of Himachal Pradesh. Indus Publishing Co., New Delhi
- Doyle JJ, Doyle JE (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Earl DA, VonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. doi:10.1007/s12686-011-9548-7
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7:574–578
- Foroumadi A, Asadipour A, Arabpour F, Amanzadeh Y (2002) Composition of essential oil of *Bunium persicum* (Boiss) Fedtsch from Iran. *J Essent Oil Res* 14:161–162
- Gincarlo S, Rosa LM, Nadjafi F, Francesco M (2006) Hypoglycemic activity of two spices extracts: *Rhus Coriaria* L. and *Bunium persicum* Boiss. *Nat Prod Res* 20:882–886
- Hanelt P, Büttner R, Mansfeld R, Kilian R (2001) Mansfield's encyclopedia of agricultural and horticultural crops. Springer, Berlin
- Hashemi H, Safarnejad A, Bagheri A (2010) The use of RAPD marker for assessing the genetic diversity of *Bunium persicum* (Boiss.) *B. fedtsch* populations. *IJSN* 1:202–208
- Heywood VH (2002) The conservation of genetic and chemical diversity in medicinal and aromatic plants. In: Sener B (ed) Biodiversity: biomolecular aspects of biodiversity and innovative utilization. Kluwer Academic/Plenum Publishers, New York, pp 13–22
- Jahansooz F, Najafi AA, Sefidkon F, Ebrahimzadeh H, Mardi M (2007) Assessment of genetic diversity in ecotypes of *Bunium persicum* using AFLP markers. In: Proceeding of the 5th national biotechnology congress of Iran, Tehran
- John MA, Skroch PW, Nienhuis J (1997) Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. *Crop Sci* 37:605–613
- Kala CP (2003) Medicinal plants of Indian trans-Himalaya. Bishan Singh Mahendra Pal Singh, Dehradun
- Kapila R, Panwar K, Badiyala D (1997) Variation and association analysis in domesticated populations of Black caraway (*B. persicum*). *JMPS* 19:709–711
- Khadivi-Khuba A, Salehi-Arjmandb H, Hadianc J (2014) Morphological and phytochemical variation of *Satureja bachtiarica* populations from Iran Abdollah. *Ind Crop Prod* 54:257–265
- Majeed S, Sharma D (2006) Assessment of genetic divergence of Kalazeera of different geographical races from India. In: Proceedings of 1st international conference on biotechnological approaches for alleviating malnutrition and human health, Bangalore
- Martinez GP, Sanchez PR, Dicenta F, Howad W, Arus P, Graziel TM (2007) Genome mapping and molecular breeding in plants. Springer, Berlin, pp 229–242
- Mittal R, Chahota R, Gartan S, Katna G (2006) Genetic variability and component analysis in Kala Zeera in dry temperate area of North West Himalayas. *Crop Improv* 33:202–204
- Moghtader M, Mansori AI, Salari H, Farahmand A (2009) Chemical composition and antimicrobial activity of the essential oil of *Bunium persicum* Boiss. *Seed Iranian J Med Rom Plants* 25:20–28
- Omidbeigi R (1997) Approaches to production and processing of medicinal plants-2. Tarrahan-e-Nashr, Tehran
- Panda H (2004) Aromatic plants cultivation, processing and uses. Asia Pacific Business Press Inc, New Delhi
- Panwar KS (2000) Black caraway-Kala zira. In: Arya PS (ed) Spice crops of India. Kalyani Publishers, New Delhi, pp 172–178
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software <http://darwin.cirad.fr/darwin>
- Peter KV (2004) Handbook of herbs and spices. CRC Press, Florida
- Pezhmanmehr M, Hassani ME, Fakhre, Tabatabaie M (2007) Evaluation of genetic diversity in some ecotypes of Black Cumin (*Bunium persicum*) using RAPD markers. In: Proceeding of the 5th national biotechnology congress of Iran, Tehran
- Pourmortazavi SM, Sefidkon F, Hosseini SG (2003) Supercritical carbon dioxide extraction of essential oils from Benth. *J Agri and Food Chem* 51(18):5414–5419
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genet* 155:945–959
- Salehi P, Mohammadi F, Asghari B (2008) Seed essential oil analysis of *Bunium persicum* by hydro distillation head space solvent micro extraction. *Chem Nat Compd* 44:111–113
- Sardari S, Amin GH, Micetichm RG, Daneshlab M (1998) Antifungal activity of selected Iranian and Canadian plants. *Pharmaceutical Biol* 36:180–188
- Sefidkon F, Jamzad Z (2005) Chemical composition of the essential oil of three Iranian *Satureja* species (*S. mutica*, *S. macrantha* and *S. intermedia*). *Food Chem* 91:1–4
- Sefidkon F, Rahimi A (2002) Assessment of quantitative and qualitative changes in *Thymus kotschyanus* essence using different distillation methods. *Med Aromat Res* 15:1–22
- Sekine TM, Sugano AM, Fujii Y (2007) Antifungal effects of volatile compounds from black zira (*Bunium persicum*) and other spices and herbs. *J Chem Ecol* 33:2123–2132
- Shahsavari N, Barzegar M, Sahari MA, Naghdibadi H (2008) Antioxidant activity and chemical characterization of essential oil of *Bunium persicum*. *Plant Foods Hum Nutr* 63:183–188
- Shankaracharya NB, Shankaranarayananana ML (1988) Research notes on essential oil of *Cumin cyminum* L. and *Bunium persicum* B. Pafai J 10:33–35
- Sharma V, Sharma TR, Rana JC, Chahota RK (2015) Analysis of genetic diversity and population structure in horsegum (*Macrotyloma uniflorum*) using RAPD and ISSR markers. *Agric Res*. doi:10.1007/s40003-015-0165-7

- Sheidai M, Ahmadian P, Poorseyedy S (1996) Cytological studies in Iran zira from three genus: *Bunium*, *Carum* and *Cuminum*. *Cytologia* 61:19–25
- Steiner JJ, Los Santos GG (2001) Adaptive ecology of *Lotus corniculatus* L. genotypes: I. Plant morphology and RAPD marker characterization. *Crop Sci* 41:552–563
- Talei GR, Mosavi Z (2009) Chemical composition and antibacterial activity of *Bunium persicum* from west of Iran. *Asian J Chem* 21:4749–4754
- Thappa RK, Ghosh S, Agarwal A, Raina K, Jamwal PS (1991) Comparative studies on the major volatiles of Kalazira (*Bunium persicum*) seed of wild and cultivated sources. *Food Chem* 41:129–134
- Vokou D, Kokkini S, Bessiere JM (1993) Geographic variation of Greek Oregano (*Origanum vulgare* ssp. *hirtum*) essential oils. *Biochem Syst Ecol* 21:287–295
- Williams JG, Kubelik AR, Livak KJ, Rafaski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic marker. *Nucleic Acids Res* 18:6531–6535
- Wunsch A, Hormaza JL (2002) Cultivar identification and genetic fingerprinting of temperate fruit tree species using DNA markers. *Euphytica* 125:59–67