Pak. J. Bot., 45(S1): 333-338, January 2013.

DEVELOPMENT AND CHARACTERIZATION OF BARLEY CORE COLLECTION: A STRATEGY FOR GERMPLASM MANAGEMENT

REHAN NAEEM[1,2] AND BUSHRA MIRZA[1*]

[1]*Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan*
[2]*Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology, Pakistan*
[*]*Corresponding author's e-mail: bushramirza@qau.edu.pk; phone: +92(51)90643007 fax: +92(51)9219888*

Abstract

Redundant materials in germplasm collections cause hurdles in gene bank management. Present study was conducted to analyze redundancies using Simple Sequence Repeats (SSRs) and to evaluate genetic diversity of barley germplasm comprising of 404 *Hordeum vulgare* landraces belonging to 8 different regions (Pakistan. India, Iran, Nepal, Iraq, Turkmenistan, Uzbekistan Kazakhstan). In total 50 alleles were detected. Number of alleles observed was highest for Bmag0023 (11) followed by HVLOX (9), HVM54 (8), Bmag0382 (7), Bmag0500 (5), Bmag0490 (5) and HVID (5) locus. Allele size range was largest for HVLOX while it was lowest for Bmag0500. Allele frequencies were found highest for allele HVM54-120 (0.39) followed by HVID-80 (0.38) and Bmag0382-66 (0.35). Core collection for the barley germplasm comprised of 42 patterns including 14 unique and 28 frequent patterns. The number of patterns observed in landraces of different regions was in the order of Pakistan> India> Nepal> Turkmenistan> Iraq> Iran> Uzbekistan> Kazakhstan. These patterns were clearly clustered into two groups using Jaccard's cluster analysis (group A and group B). Principal Component Analysis (PCA) also distributes these patterns into three groups. The overall results revealed that differentiation of patterns using Jaccard's cluster analysis was in accordance with PCA which revealed that these patterns were distributed with respect to their geographic regions. Furthermore, it was concluded that development of core collection is the one strategy besides other strategies to reduce redundancies in gene bank management and make the handling of germplasm easier.

Introduction

Barley (*Hordeum vulgare* ssp. *vulgare*) is one of the oldest crops cultivated in the world. At present, the collection of cultivated barley maintained in all gene banks have raised two fold in the world germplasm centers (Valkoun & Konopka, 2004). Similarly, the collection of cultivated barley maintained at the Plant Genetic Resource Institute, NARC, Islamabad is found to be around 1274 accessions. Conservation for such huge germplasm is not only labor-intensive, but also high-priced and time-consuming. To alleviate these management difficulties, use of core collections has been suggested. It can be fulfilled by the use of sub-sets of the whole collection, known as core collections (Frankel & Brown, 1984). Various criteria have been proposed to analyze such type of core collections, which include morphological, agronomic, eco-geographical, biochemical and molecular traits. Morphological and agronomic traits both have certain limitations for the discrimination of barley. Consequently biochemical studies like isozymes, Hordein (Naeem *et al.,* 2006, Pervaiz *et al*., 2010) and molecular markers have received much more attention in recent years. Microsatellites, also known as Simple Sequence Repeats (SSRs) are found to be popular among various laboratories for conduction genetic diversity studies (Rabbani *et al*., 2010). SSRs are an excellent molecular marker system for many types of genetic analyses, including germplasm surveys and phylogenetic studies (Liu *et al*., 1996, Turi *et al*., 2012). New techniques based on SSR variation have been used regularly for evaluating and characterizing core collections and to determine evolutionary relationships within and between populations of the genus *Hordeum* (Saghai-Maroof *et al*., 1994). These are short (1-6 bp long) tandemly repeated DNA sequences and are highly polymorphic as a result of frequent variation in the number of times the core sequence is repeated (Saghai-maroof *et al*., 1994).

The present study was analyzed on the hypothesis that germplasm collections have redundant materials which add to the cost of gene bank management that can be reduced by the development of core collection. The objectives of the study are the management of barley core collection using SSRs and their characterization.

Materials and Methods

In this study, 404 landraces of *H. vulgare* L. from 8 different geographic regions including, Pakistan, India, Iran, Nepal, Iraq, Turkmenistan, Uzbekistan and Kazakhstan were collected through the help of gene bank at Plant Genetic Resource program (PGRP), NARC, Islamabad. Seeds were germinated at 25ºC in a green house. Genomic DNA was extracted using the modified rapid cetyltrimethylammonium bromide (CTAB) method. The quality of extracted DNA was analyzed using agarose-gel electrophoresis and by spectrophotometer at 260/280 nm. Finally DNA samples were diluted at a concentration of 25 ng/µL using deionized water and stored at -25ºC.

Genetic variability for these landraces was evaluated by mapped SSR markers (Liu *et al*., 1996). Those microsatellite markers were selected for this study that have locus exclusively on sigle chromosome. Polymerase chain reaction (PCR) was performed in a reaction volume of 25 µL containing 25 ng/ µL genomic DNA, 10X PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 1 units of Taq polymerase (Fermentas, USA) and 0.25 µM forward and reverse primer. Depending on the primer pair used, DNA amplifications were performed in a thermocycler using regular and touchdown PCR protocol. After PCR amplification 2 µL of 6X electrophoresis loading buffer was added to each reaction mixture and separated on 8% polyacrylamide gel. A 20 to 300 bp DNA ladder (Fermentas, USA) was used as a size standard. The amplified DNA fragments of each SSR locus

were scored for the presence (1) or absence (0) of clear bands. Allele frequencies were calculated and unbiased genetic distances of SSR alleles were determined using Jaccard's coefficient. SSR patterns were analyzed based on Sj = n11/n11+n01+n10 where n is the number of character code in both OTU (Operational taxonomic unit), n01 is the number of character coded 0 for first OTU and 1 for the other and n10 is the number of character coded 1 for the first OTU and 0 for the other. The distance matrix was used to reveal associations between landraces based on the Un-weighed Pair Group Method with Arithmetic Averages (UPGMA). Principal Component Analysis (PCA) was performed on the basis of SSR patterns using the statistical program "Statistica". Two principal components (PCA1, PCA2) were further included as independent variables in order to test grouping of populations.

Results

Number of alleles, allele size range and allele frequencies for the barley germplasm are shown in Fig. 1. All the SSRs markers studied in the present study were polymorphic and in total 50 alleles have been detected for 404 landraces of barley germplasm. The number of alleles for these loci were found maximum for Bamag0023 (11) while they were found lowest for Bamag0500 (5), Bmag490 (5) and HVID (5) as shown in Fig. 1. Similarly allele size range was found maximum for HVLOX (51 bp-120 bp yielding 69 bp difference) while it was found minimum for Bmag0500 (62 bp-88 bp) as depicted in Fig. 1. Allele frequencies were found highest for allele HVM54-120 (0.39) followed by HVID-80 (0.38) and Bmag0382-66 (0.35) as shown in Fig. 1.

Average pair wise differences between alleles of SSRs markers were compared by Jaccard's coefficient. The landraces with alleles having same similarity index value were grouped to be represented by a single pattern. In this way, 42 patterns across 7 SSR loci for 404 landraces of barley germplasm were identified which represent core collection of the barley germplasm. Out of 42 patterns identified, 14 patterns were considered as unique i.e. restricted to only one landraces, while remaining 28 were considered as frequent patterns, observed in more than one landrace. Out of 14 unique patterns, five were observed in Pakistan landraces, three each in India and Nepal landraces, one in Iraq landraces, while two unique patterns were observed in Turkmenistan landraces as shown in Table 1.

Out of 28 frequent patterns, 17 patterns were observed specifically in landraces of one region only while 11 patterns were observed as shared patterns found in landraces of more than one region. Out of 17 region specific patterns, 11 were observed in Pakistan landraces while remaining 6 patterns were appeared as 2 each in India, Iran and Iraq landraces. Whereas, 11 shared patterns comprised of Pakistan, India, Iran, Nepal, Iraq, Turkmenistan, Uzbekistan and Kazakhstan landraces. The numbers of patterns observed in different regions are shown in Table 2. It observed that highest numbers of patterns were observed for Pakistan (20) followed by India (12), Nepal (11), Iran (6), Iraq (6), Turkmenistan (6), Uzbekistan (2) and Kazakhstan (1) landraces.

In order to reveal genetic relationships among 42 patterns, similarity coefficients were determined and a dendrogram was constructed which depicted the relationships among the patterns. At a genetic similarity level of about 3.5 the patterns were clearly clustered into 2 groups (Group A and Group B). Group B comprised of landraces from India, Iran and Pakistan while group A comprised of landraces of Pakistan, Iran, Nepal, Iraq, Turkmenistan, Uzbekistan and Kazakhstan as shown in Fig. 2. At a genetic similarity level of about 2.7 the patterns were clustered into 10 clusters. Group A consisted of 7 clusters having most of the landraces from Pakistan (119), India (52), Iran (22), Nepal (51), Turkmenistan (17), Uzbekistan (5) and Kazakhstan (1). Similarly group B comprised of 3 clusters having landraces of only from Pakistan (41) in cluster VIII, India (19) and Iran (1) in cluster IX while India (5) and Iran (33) in cluster X. PCA was performed to analyze the relationships within the patterns. Molecular variation of about 38% at these SSR loci was observed for first two principal components as shown in Fig. 3.

It was found that Pakistan, India, Iran, Nepal, Iraq, Turkmenistan, Uzbekistan and Kazakhstan patterns were clumped into three groups. Group A comprised of majority of population shared patterns of Pakistan, India, Iran, Nepal, Iraq, Uzbekistan, Turkmenistan and Kazakhstan while Group B comprised of majority of population specific patterns of Pakistan, India, Iran, Nepal and Iraq. Group C comprised of all population specific patterns except pattern which was shared by 5 landraces of India and 14 landraces of Iran. PCA indicated that landraces belonging to Pakistan, India, Iran, Nepal and Iraq were clumped in one group while landraces belonging to Turkmenistan, Uzbekistan and Kazakhstan were clustered in separate group. The overall results revealed that differentiation of patterns with PCA was in accordance with the geographic regions as previously shown by the cluster analysis.

Discussion

In the present study, it was observed that number of alleles increased with increase in allele size range as maximum of 11 alleles was found for Bmag0023 having 54 bp allele size range difference (66 bp-120 bp), while the lowest number of alleles (5) were observed for Bmag0500 having allele size range 26 bp difference (62 bp-88 bp). Difference in allele number, allele size range and allele frequencies indicates the difference in association of the germplasm under different environments which depicts that the germplasm has the potential for genetic resource conservation. Allele size range also depicts the differentiation among barley groups under different environments and these are due to variations in the repeat length of the microsatellite region. In previous studies, Matus & Hayes (2002) found 11 alleles in barley germplasm for HVM54 and 5 for Bmag0382, while in another study, total of 173 alleles were detected across 18 loci and noticed 17 alleles for HVM68 and only 2 alleles for HVDHN (Turpeinen *et al*., 2001). In another report, Baek *et al*., (2003) observed the largest difference in allele size range for the locus HVM5, with fragments ranging from 162 to 224 bp, followed by 56bp difference for loci BMS64 and HVM68. In the present study, in certain cases allele size range observed was not in accordance with expected size range reported in the literature for cultivated barley. This may be attributed to insertion/deletion events in the flanking sequences of microsatellites as suggested previously (Huang *et al*., 1998).

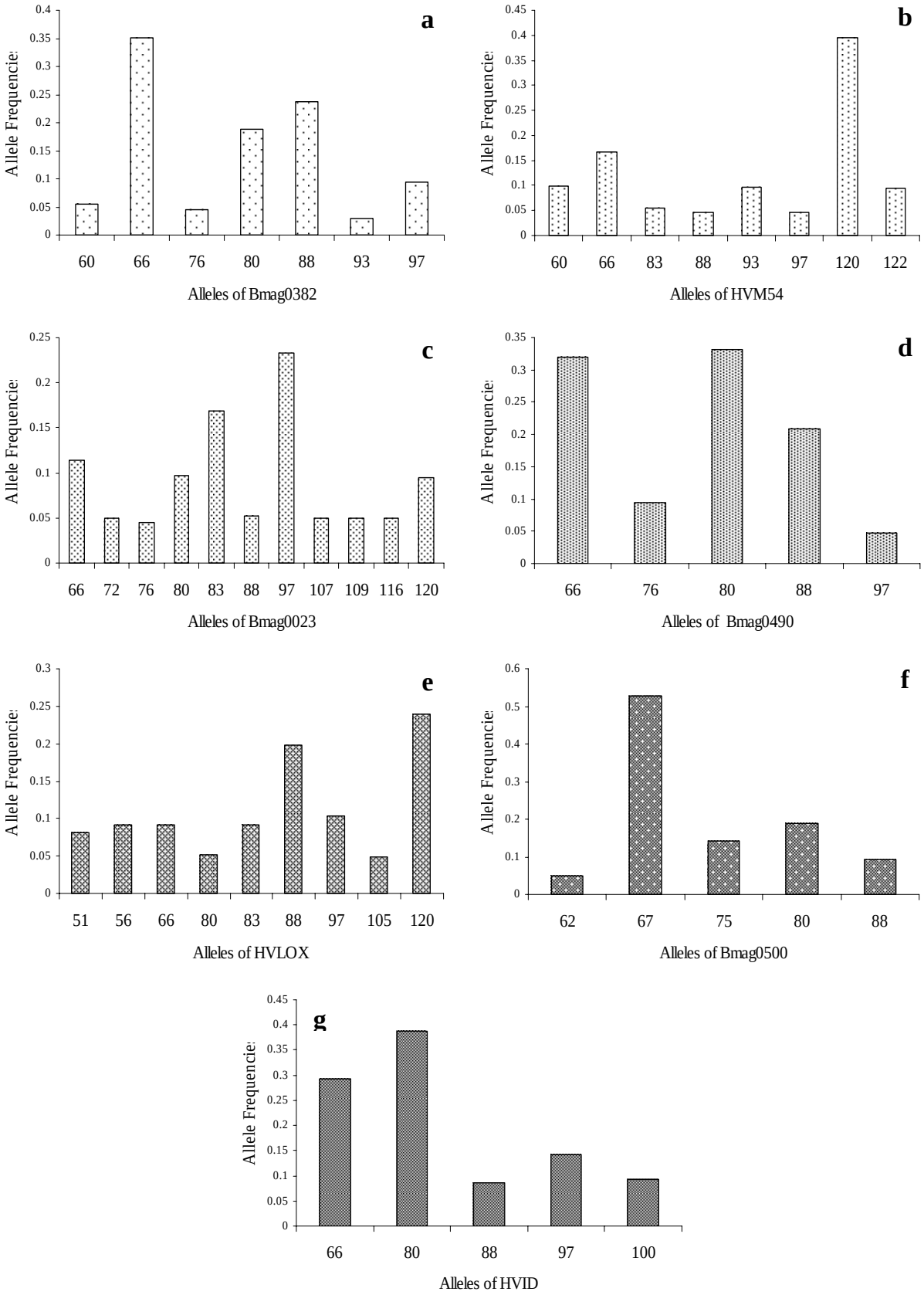

Fig. 1. Allele frequencies of *Hordeum vulgare* L. germplasm by SSR markers
a. Bmag0382; b. HVM54; c. Bmag0023; d. Bmag0490; e. HVLOX; f. Bmag0500; g. HVID

Table 1. Unique and frequent patterns of *Hordeum vulgare* L. germplasm by SSR markers.

Unique patterns		Frequent patterns			
		Population specific		Population shared	
Patterns	Frequency of landraces	Patterns	Frequency of landraces	Patterns	Frequency of landraces
P-6, P-10, P-13, P-19, P-20	5(Pak)	P-1, P-2, P-3, P-4, P-7, P-8, P-9, P-11, P-12, P-14, P-18	136(Pak)	P-5, P-15, P-16	5(Pak),1(N) 3(Pak),1(N) 2(Pak),2(N)
P-23, P-25, P-30	3(Ind)	P-27, P-28	27(Ind)	P-17	9(Pak),10(T)
P-35, P-36, P-37	3(N)	P-33, P-34	37(Iran)	P-21	7(Ind), 1(Iran), 3(N), 2(T), 4(U), 1(K)
P-41, P-42	2(T)	P-38, P-40	32(Iraq)	P-22	16(Ind), 1(N),1(Iraq), 1(T), 1(U)
P-39	1(Iraq)	-	-	P-24	3(I), 3(Iran), 9(N), 2(Iraq), 2(T)
-	-	-	-	P-26, P-29	5(Ind), 14(Iran) 12(Ind), 1(Iran)
-	-	-	-	P-31	2(Ind),16(N)
-	-	-	-	P-32	1(Ind), 15(N), 2(Iraq)
14	14	17	232	11	158

*Pak-Pakistan, Ind-India, Iran, N-Nepal, Iraq, T-Turkmenistan, U-Uzbekistan, K-Kazakhstan P-Patterns

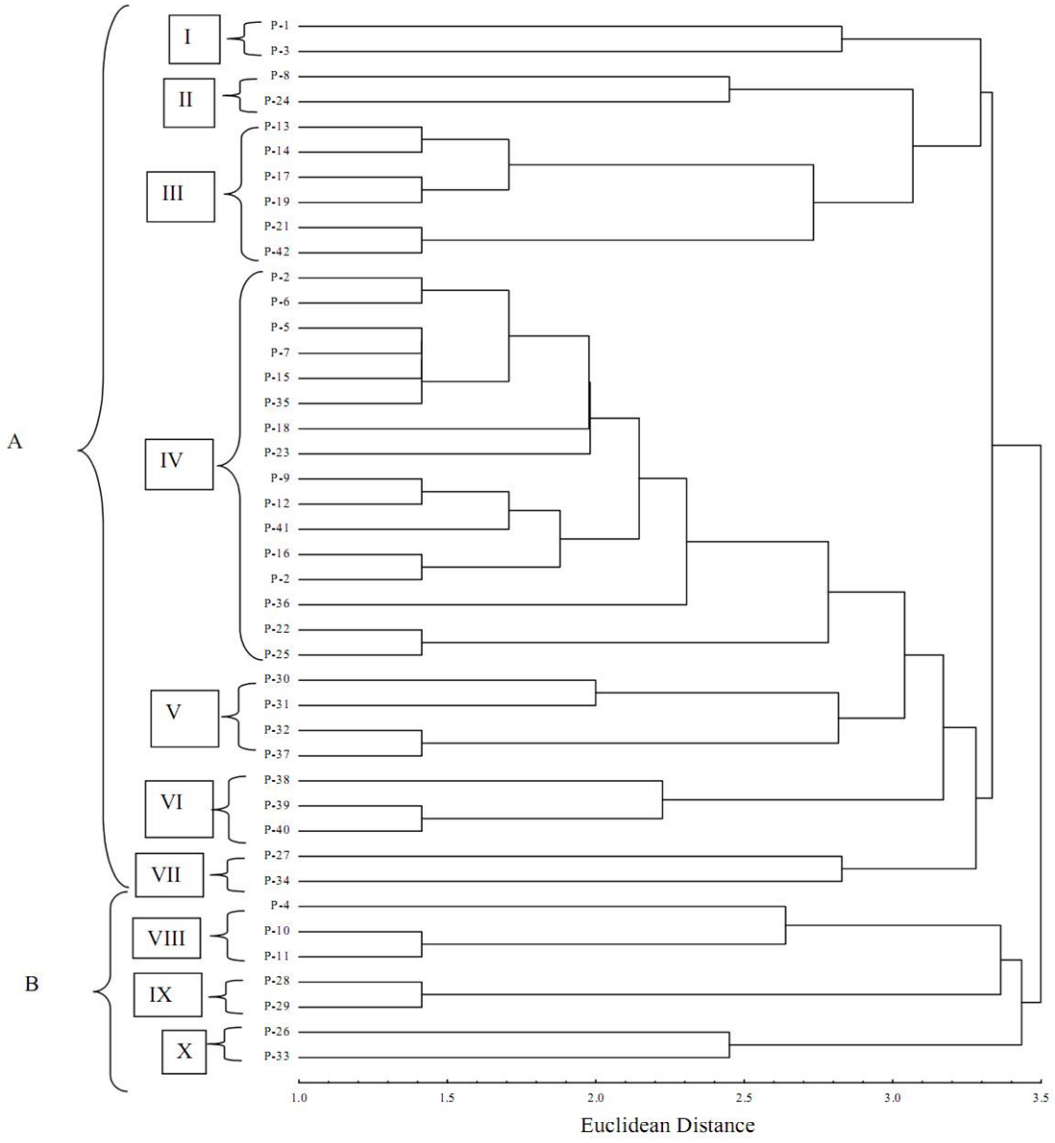

Fig. 2. Cladogram of barley Core Collection developed using SSRs by Unweighed Pair Group Average (UPGMA).

Table 2. Frequency of landraces within patterns and their regions for *Hordeum vulgare* L. germplasm by SSR markers.

Regions	Landraces	Patterns	Frequency of landraces within patterns
Pakistan	160	20	P-1(19), P-2(4), P-3(20), P-4(19), P-5(5), P-6(1), P-7(3), P-8(19), P-9(2), P-10(1), P-11(21), P-12(3), P-13(1), P-14(24), P-15(3), P-16(2), P-17(9), P-18(2), P-19(1), P-20(1)
India	76	12	P-21(7), P-22(16), P-23(1), P-24(3), P-25(1), P-26(5), P-27(20), P-28(7), P-29(12), P-30(1), P-31(2), P-32(1)
Iran	56	6	P-21(1), P-24(3), P-26(14), P-29(1), P-33(19), P-34(18)
Nepal	51	11	P-5(1), P-15(1), P-16(2), P-21(3), P-22(1), P-24(9), P-31(16), P-32(15), P-35(1), P-36(1), P-37(1)
Iraq	38	6	P-22(1), P-24(2), P-32(2), P-38(20), P-39(1), P-40(12)
Turkmenistan	17	6	P-17(10), P-21(2), P-22(1), P-24(2), P-41(1), P-42(1)
Uzbekistan	5	2	P-21(4), P-22(1)
Kazakhstan	1	1	P-21(1)

*P-Patterns

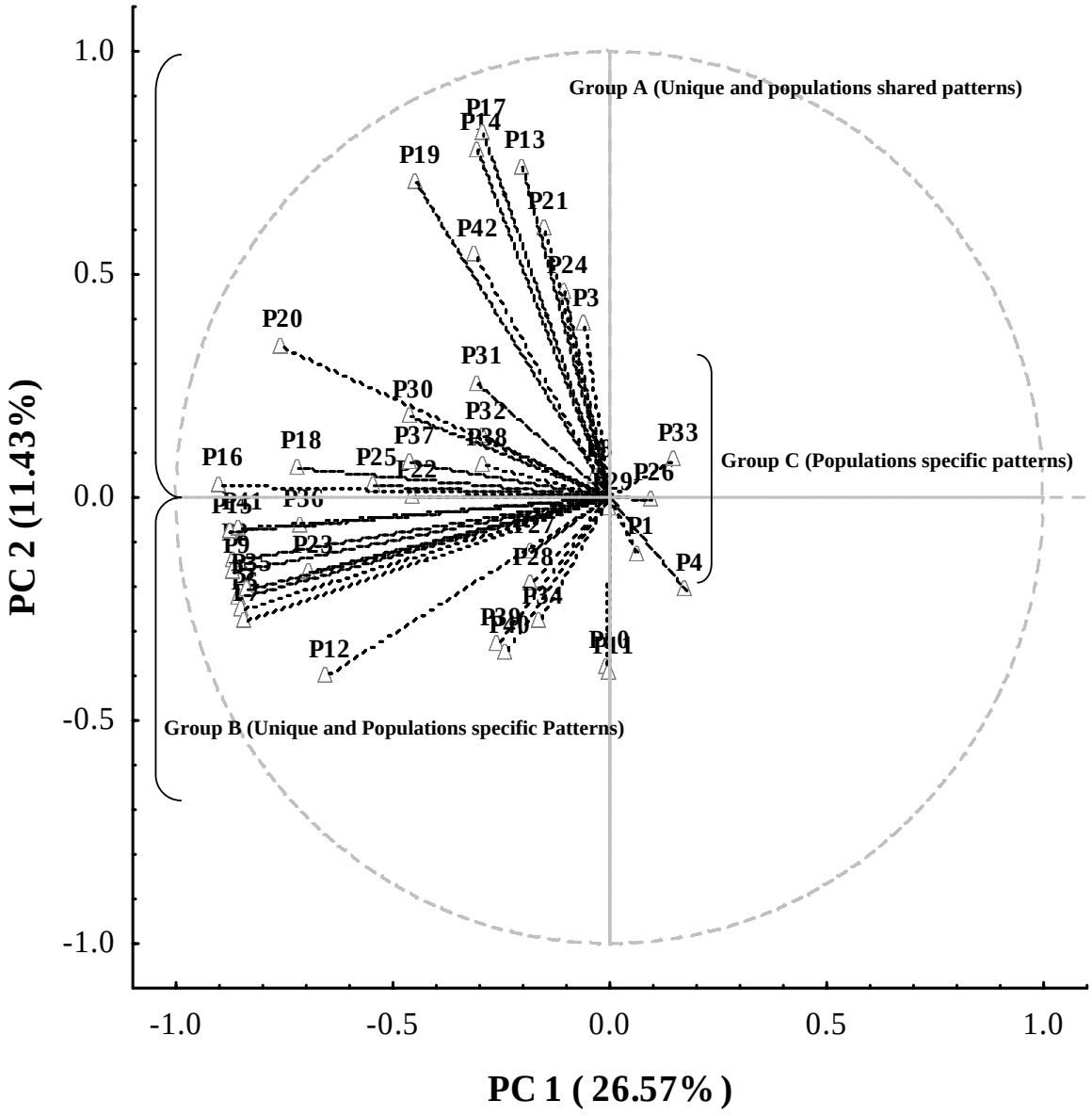

Fig. 3. Ordination of Core Collection for *Hordeum vulgare* L. germplasm by SSRs.

In this study, 42 SSR patterns (fingerprints) were identified, including unique, specific and shared ones which represented the core collection of the barley germplasm. It was observed that the similarities in banding patterns can be used to determine genetic relatedness among landraces. Greater genetic similarity was observed among populations from the same environment than among populations from different environments. Volis *et al*., (2003) has suggested the same phenomenon and stressed that natural selection is environmentally induced and genetic similarity among population were reflected by their similarities in their habitat and requirements which are independent of their geographic distance. In the present study, the total number of patterns depicting the genetic diversity for cultivated barley was in the order of Pakistan> India> Nepal> Turkmenistan> Iraq> Iran> Uzbekistan> Kazakhstan. A similar study was carried by Nevo (1992), who compared the wild barley landraces and proposed an order of Iran< Turkey < Israel for these three countries. It was believed that landraces collected from different areas could be identified on the basis of genetic patterns. Genetic relationships of 42 patterns were studied by UPGMA cluster analysis, and 10 clusters were identified. Each cluster comprised of patterns, which have landraces coming from same region. Similarly, Hamza *et al*., (2004) clearly differentiated the barley genotypes by UPGMA according to their geographic regions, row number and end use. They also associated the barley cultivars with their adaptive traits. In our study, the results of cluster analysis were compared with PCA which indicated that clustering of patterns was in accordance to the geographic regions as previously shown by the cladogram. In a recent report, it was stressed that clustering patterns of the subsp. *spontaneum* were in accordance to their geographical origin as accessions from East Asia, Tibet, Afghanistan, Pakistan, Turkmenistan, Azerbaijan, Iraq, and Iran were grouped in the same cluster while accessions originated from Israel were grouped in other cluster (Matus & Hayes, 2003).

The overall results of this study can be summarized as the fact that, core collection developed for the germplasm comprised of 42 SSR patterns (404 accessions), including 14 unique and 28 frequent patterns. It was observed that accessions of South Asia belonging to Pakistan, India, Iran, Nepal, Iraq were more related as compared to landraces of Central Asia belonging to Turkmenistan, Uzbekistan and Kazakhstan. Pakistan accessions showed highest genetic variability as compared to landraces of other regions. Therefore, it is concluded that these germplasm collections have redundant materials which cause problems in gene bank management. The markers used in this study are highly informative and useful for the development of core collection to reduce redundancies and gene bank costs leading to efficient gene bank managements.

Acknowledgements

This research was funded by Pakistan Science Foundation (PSF) Islamabad, Pakistan. Plant Molecular Biology Laboratory, Department of Biochemistry, Quaid-i-Azam University, Islamabad, Pakistan also acknowledges the contribution of Plant Genetic Resource Program (PGRP), National Agricultural Research Center (NARC) for providing barley landraces.

References

Baek, H.J., A. Beharav and E. Nevo. 2003. Ecological genomic diversity of microsatellites in wild barley, *Hordeum spontaneum*, populations in Jordon. *Theor. Appl. Genet.*, 106: 397-410.

Frankel, O.H. and A.H.D. Brown. 1984. Plant genetic resources today: a critical appraisal. In: *Crop genetic resources: conservation and evaluation*. (Eds.): J.H.W. Holden and J.T. Williams, George Allen and Unwin: Winchester, UK. pp. 249-268.

Hamza, S., W.B. Hamida, A. Rebai and M. Harrabi. 2004. SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. *Euphytica*, 135: 107-118.

Huang, W.G., G. Cipriani, M. Morgante and R. Testolin. 1998. Microsatellite DNA in *Actinidia chenensis*: isolation, characterization and homology in related species. *Theor. Appl. Genet.*, 97: 1269-1278.

Liu, W., R.M. Biyashev and M.A. Saghai-Maroof. 1996. Development of simple sequence repeat DNA markers and their integration in to a barley linkage map. *Theor. Appl. Genet.*, 93: 869-876.

Matus, I.A. and P.M. Hayes. 2003. Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. *Genome*, 45: 1095-1106.

Naeem, R., I. Ahmed, R. Asgher and B. Mirza. 2006. Characterization of *Hordeum vulgare* accessions of Pakistan by Hordein SDS-PAGE. *Cereal Res. Commun.*, 34: 1247-1254.

Nevo, E. 1992. Origin, evolution, population genetics and resources for breeding of wild barley, *Hordeum spontaneum*, in the Fertile Crescent. In: *Barley, Genetics, Biochemistry, Molecular Biology and Biotechnology*. (Eds.): P.R. Shewry, CAB International Wallindford, Oxon, UK. pp. 19-43.

Pervaiz, Z.K., M.A. Rabbani, Z.K. Shinwari, M.S. Masood and S.A. Malik. 2010. Assessment of genetic variability in rice (*Oryza sativa* l.) germplasm from pakistan using RAPD markers. *Pak. J. Bot.*, 42(5): 3369-3376.

Rabbani, M.A., M.S. Masood, Z.K. Shinwari and K.Y. Shinozaki. 2010. Genetic analysis of basmati and non-basmati pakistani rice (*Oryza sativa* L.) cultivars using microsatellite markers. *Pak. J. Bot.*, 42(4): 2551-2564.

Saghai-Maroof, M.A., R.M. Biyashev, G.P. Yang, Q. Zhang and R.W. Allard. 1994. Extraordinarily polymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. *Proc. Nat. Acad. Sci.*, 91: 5466-5470.

Turi, N.A., Farhatullah, M.A. Rabbani and Z.K. Shinwari. 2012. Genetic diversity in the locally collected brassica species of Pakistan based on microsatellite markers. *Pak. J. Bot.*, 44(3): 1029-1035.

Turpeinen, T., T. Tenhola, O. Manninen, E. Nevo and E. Nissila. 2001. Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. *Mol. Ecol.*, 10: 1577-1591.

Valkoun, J. and J. Konopka. 2004. *Global inventory of barley genetic resources* (Proceedings of the 9th International Barley Genetics Symposium), Brno, Czech Republic, pp. 31-38.

Volis, S., I. Shulgina, D. Ward and S. Mendlinger. 2003. Regional subdivision in wild barley allozyme variation: adaptive or neutral. *J. Hered.*, 94: 341-351.

(Received for publication 1 September 2012)