The Bioscan 11(4):3035-3041,2016(Supplement on Genetics and Plant Breeding) www.thebioscan.com
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES

GENETIC DIVERSITY ANALYSIS AMONG ELITE GENE POOL OF INDIAN MUSTARD USING SSR MARKERS AND PHENOTYPIC VARIATIONS

RAM AVTAR*, BABITA RANI, MINAKSHI JATTAN, MANMOHAN, NISHA KUMARI AND ASHA RANI[1]
Department of Genetics and Plant Breeding,
[1]Department of Molecular Biology, Biotechnology and Bioinformatics,
CCS Haryana Agricultural University, Hisar -125 004, INDIA
e-mail: ramavtarola@yahoo.in

KEYWORDS
Genetic diversity
Indian mustard
SSR markers
Phenotypic
Polymorphism

Received on :
13.10.2016

Accepted on :
07.12.2016

***Corresponding author**

ABSTRACT
Genetic diversity analysis was performed among 96 germplasm lines of *Brassica juncea* using 83 SSR primers and 20 phenotypic variables. Sixteen primers were found to be polymorphic and gave a total of 47 alleles which varied from 2 to 5 with an average of 2.9 alleles per primer. The mean PIC value from all the polymorphic primers was high *i.e.* 0.529. A wide range (0.38 to 0.96) of Jaccard's similarity coefficients was generated between these lines. Total number of germplasm lines in each cluster ranged from 2-31 and 2-26 in the dendrograms generated based on molecular data and phenotypic data, respectively. Comparison between both types of dendrograms showed that some of the germplasm lines were common in four clusters *i.e.* CI, CIII, CIV and CVI. Principal Component Analysis explained 74.70% of the total variation based on molecular data and 70.41% based on phenotypic data. The differences obtained in the results based on SSR marker data and phenotypic data might be due to the impact of environment on phenotypic characters. Therefore; genetic diversity analyzed based on SSR markers is more reliable in the present study and will be helpful in selecting diverse genotypes which can be used as parents in hybridization programme.

INTRODUCTION

Brassica juncea (L.) Czern & Coss., commonly known as 'Indian mustard'; is one of the important sources of edible oil in India and it contributes a major share in mustard production globally. India is the second largest rapeseed-mustard growing country after China, occupying 20.23% area and contributing 11.7% share to the global production. It is utilized worldwide as an oilseed, a condiment, vegetable, green manure, forage and fodder and cultivated primarily in tropical and sub-tropical countries (Gangapur *et al.*, 2010). In India *per capita* oil consumption has increased tremendously due to the increasing population and improving life standards. To meet out the present oil requirements, there is an urgent need to increase the yield potential of *B. juncea*. The development of hybrids with high seed yield and oil content has an important role to fill the gap between oil production and human population. Heterosis can only occur when parental cultivars used for F_1 production differ in gene frequencies (Falconer and Mackay, 1997). Selection of parents based on diversity estimates is vital criterion to initiate a hybrid development programme. Assessment of genetic diversity in *B. juncea* using phenotypic characters has previously been done by many researchers (Vaishnava *et al.*, 2006, Alie *et al.*, 2009 and Singh *et al.*, 2010). Isozyme loci have been used as markers in a number of genetic studies including genetic diversity in *B. juncea* (Kumar and Gupta, 1985 and Arunachalam *et al.*, 1996). However, these parameters are influenced by environmental factors and the developmental stage of the plant. Molecular marker assisted breeding approach seems to be more suitable to hasten the selection and development of new improved genotypes, as it is independent of environmental factors and developmental stage of the plant. There is an increasing number of reports where molecular markers like Restriction Fragment Length Polymorphism (RFLP), Song *et al.*, 1988 and Diers and Osborn, 1994; Random Amplified Polymorphic DNAs (RAPDs), Yildirim *et al.*, 2010 and Khan *et al.*, 2011; Amplified Fragment Length Polymorphism (AFLP), Sun *et al.*, 2001 and microsatellites or Simple Sequence Repeats (SSRs), Wang *et al.*, 2009; have been used to study genome organization, varietal differences and diversity analysis in Brassicas. During the last few years, tremendous emphasis has been particularly laid on the development of gene-based SSR markers since they are reproducible, co-dominant in nature, abundant, widely distributed throughout the genome, easily scorable, highly polymorphic, show multi-allelic variation and have greater potential for linkage to loci associated with agronomic phenotypes. Simple sequence repeat markers have been used for genetic diversity analysis in a number of crops including Indian bread wheat (Mir *et al.*, 2011), rice (Rahman *et al.*, 2012), and maize (Sivaranjini *et al.*, 2014). A large number of SSR markers have been developed and extensively characterized in *B. rapa* (Suwabe *et al.*, 2002 and Ramchiary *et al.*, 2011) and *B. napus* (Wang *et al.*, 2012 and Li *et al.*, 2013). However, only few SSR markers are reported in *B.*

juncea (Hopkins *et al*., 2007, Vinu *et al*., 2013 and Pratap *et al*., 2015). Hence, the present study was planned to see the relative efficiency of SSR markers and phenotypic variables to differentiate 96 elite germplasm lines of Indian mustard in to diverse groups/lines. These diverse lines can be further utilized as parents in hybridization program.

MATERIALS AND METHODS

Plant material

Ninety-six germplasm lines (Table. 1) of Indian mustard (*Brassica juncea*) were grown in the pot house as well as in paired rows of 5 m length each at a spacing of 30 x 10-15 cm row to row and plant to plant at Oilseeds Research Area, Department of Genetics and Plant Breeding, CCS HAU, Hisar during *rabi*, 2015-16.

DNA extraction and SSR marker analysis

Genomic DNA of all the germplasm lines was extracted from the leaves of four weeks old plants grown in the pot house using 2% CTAB (cetyl trimethyl ammonium bromide) extraction method (Saghai-Maroof *et al*., 1984). Quality and quantity of DNA was checked by agarose gel electrophoresis (0.8% agarose gel) and UV spectrophotometer. The concentration of DNA was checked by comparing with lambda DNA (50 ng/ μL). The PCR amplifications were performed using a total of eighty three SSR primers (Table 2). PCR amplications were standardized in 20 μl reaction mixture containing 40 ng genomic DNA, 1.5 units of Taq DNA polymerase, 0.2 mM of dNTP mix, 1.5 mM $MgCl_2$ and 0.2 μM of each primer. DNA amplifications were performed in T100™ Thermal cycler. The optimized PCR programme was as follows: an initial step of denaturation of 3 minutes at 95ºC, followed by 40 cycles of 1 minute at 94ºC, variable annealing temperature for 1 minute, extension of 1 minute at 72ºC, and a final extension step of 7 minutes at 72ºC. PCR products were separated on 6% polyacrylamide gel electrophoresis for better resolution. DNA ladder of 100 bp and 20 bp were used as molecular weight standard.

Data Analysis

Molecular Data Analysis

The data were scored for the presence (1) or absence (0) of bands and set in a binary matrix. Jaccard's coefficient (Jaccard, 1908) of similarity was measured and a dendrogram based on similarity coefficients was generated by using Unweighted Pair Group Method with Arithmetic Mean (UPGMA). Principal Component Analysis was done using the package NTSYS-PC (Rohlf, 1998).

Phenotypic Data Analysis

Phenotypic data were recorded on 20 different variables on five random plants for all the 96 germplasm lines grown in the field area. Observations were recorded on 14 quantitative traits i.e. number of lobes, days to flowering, number of primary branches/plant, number of secondary branches/plant, main shoot length (cm), plant height (cm), siliqua length (cm), siliqua number on main shoot, siliqua density, seeds/siliqua, days to maturity, oil content (%), seed yield/plant (g) and 1000seed weight (g) and six qualitative traits *i.e.* leaf angle, leaf hairiness, leaf color, leaf dentation, siliqua angle and seed color. Qualitative traits were recorded by giving scores in accordance with the standard DUS descriptor. The data were used for principal component and cluster analyses using SPSS 10.0. The dendrogram was cut to form the clusters based on the method suggested by Romesburg (1984).

RESULTS AND DISCUSSION

SSR marker based analysis

Table 1: List of 96 germplasm accessions of Indian mustard

Sr. No.	Accession No.	Sr. No.	Accession No.	Sr. No.	Accession No.	Sr. No.	Accession No.
1	RC2	25	RC34	49	RC91	73	RC152
2	RC5	26	RC35	50	RC93	74	RC154
3	RC6	27	RC36	51	RC95	75	RC161
4	RC7	28	RC37	52	RC96	76	RC162
5	RC8	29	RC38	53	RC99	77	RC163
6	RC12	30	RC46	54	RC102	78	RC164
7	RC13	31	RC47	55	RC104	79	RC165
8	RC14	32	RC48	56	RC105	80	RC166
9	RC15	33	RC49	57	RC106	81	RC171
10	RC18	34	RC50	58	RC107	82	RC174
11	RC20	35	RC51	59	RC108	83	RC175
12	RC21	36	RC52	60	RC110	84	RC185
13	RC22	37	RC53	61	RC111	85	RC195
14	RC23	38	RC54	62	RC112	86	RC260
15	RC24	39	RC57	63	RC114	87	RC261
16	RC25	40	RC61	64	RC116	88	RC263
17	RC26	41	RC74	65	RC118	89	RC264
18	RC27	42	RC77	66	RC127	90	RC265
19	RC28	43	RC78	67	RC129	91	RC268
20	RC29	44	RC81	68	RC134	92	RC270
21	RC30	45	RC85	69	RC135	93	RC273
22	RC31	46	RC86	70	RC142	94	RC275
23	RC32	47	RC87	71	RC148	95	RC280
24	RC33	48	RC89	72	RC150	96	RC283

Table 2: List of SSR primer sequences depicting annealing temperature and amplification profile

S. No.	Primer Name	Forward Primer	Reverse Primer	Ann. Temp.	Ampli.*
1	BG 1	GCTGGCTGCACAATAACAGA	GTACCACTGGAGGAGCTTCG	56.55	M
2	BG 2	GGCCTTTGGAGGTGACTGTA	CAGGGATATGCGGTCTTTCT	55.55	M
3	BG 12	TCGATTGAAACACTGAACATTGA	GCGTTTTCTGTTTTCCCAATAA	52.9	P
4	BG 33	CGTGTGTCCTCTCGTGTCTCA	TGCTCAGCAGTCAGCAATCA	57.8	P
5	BG 37	GCACCTAACCGAACCCCTTTAG	GAGAAGATCGTAGGGCACTGGA	57.9	M
6	BG 40	TCCATTGGCACAAATGAAGAAA	CTTTTTCCAAGCTTAAAAATATTGC	52.35	M
7	BG 45	TCCAAGAGACGAAACCACTTCC	GCTTGCTTATATCCTTCCTTGCC	56.45	P
8	BG 47	TGTCGACATGTGTTAAAATCTCTCAT	TTCTTTAGGATTTTAGAAAAGGGAAA	53.05	M
9	BG 50	CGCCAAATCAAATTAGGGTTTA	CCACGAATTTAACAAGAGACATCC	52.8	N
10	BG 54	ACAGGAGAAACGCAACACCA	GATGCAAACGCTAGCCCAAT	56.65	M
11	BG 55	TCCCAAGTGGTGCTCCTTT	GGTAATCCCTTTTCTGCAAGC	55.4	P
12	BG 56	ACTTGGGCGGTGAAACAGTAAA	GTTATGTGGTGGAGAGGCACAA	57.15	M
13	BG 58	TCCAAGTGAAGGATAATGCTCGT	TGTACAATGGGGATGTTGTGG	55.6	P
14	BG 68	AACCTTCATTTCATATACATACACAA	TTCAATCATTTTTATTGGTCATCA	51.05	M
15	BG 86	TCGCTAGGTTGATTATCCTAACACTCT	GCAAGCGTGTGAGTTTGGAC	57.0	P
16	BG 92	GACACGTGGCATTCATTAAACG	TCCTTCAGCCAAACCCAGAA	55.7	P
17	BG 103	TTTGTCCACCATTTCTTAAACATCTA	TCAATGAAATTGTTAAAATACAGCAA	52.15	P
18	BG 104	CATGCGGAAACCCGTTAAAA	AAAGCAACCACCCCATTCAA	54.65	M
19	BG 106	TTTTCCCTTTAAATTTCAATTGCTT	GGAAGTGAAGGTGAAGAGGAGTG	53.90	N
20	BG 107	TCCGAATCGAGACAGGAACA	AGGGCTTACGAAGCCAAACC	56.75	M
21	BG 108	TTTGGGCATCACGGATCTCT	CAAAAATAAGAAAGCGACAGCTGAA	55.45	M
22	BG 109	AAGCCGGTTCTGCAAGTGTT	CATGGCATCCTACGTGGACA	57.40	M
23	BG 111	ACCCGAAAAGAGAATATGGCCT	ACAGTGGCGTTAGGTGGGG	58.1	M
24	BG 118	TGTGCTTGCGTTTTAAAGGA	GCAAAACCCACAGGTCAGAT	54.35	M
25	BG 125	CGAACCGCGAACATAGTGTA	TAAGTGCCAGTCCATTGCAT	54.7	M
26	BG 129	CGGAGATAACCGGAATGGAA	GGATGCTCTGAGACACCCAAA	55.7	M
27	BG 137	TTGTTCGGTGTCCCCTTTGT	TGCTCCGGGAATGAACATCT	56.7	M
28	BG 138	GGGTCCAAAGCTAGAGGCTGT	GAAGTAATCCCGCCCCATAGA	57.65	M
29	BG 156	CATTGATGAGGCAAGACTTTGA	CACCAAAGCTTCTCAACTTTCTAA	53.4	P
30	BG 169	CTCCTTGCCGGAGATACAAA	CAGTTCGCTAGCAAGTCGTG	55.3	M
31	BG 176	CACGACGTGCCCAATGATTA	TGCTTGACTGTAAATCTGAGCAAA	55.25	M
32	BG 189	TGATCAGCCAATCTTGTTGCTT	CCACTCTGAAAATTTTGTGATCAAT	53.8	M
33	BG 195	TTCTGCACATGAGAGCACAAGA	TCGATAAAAGAAACTCAAATGACTGC	55.3	M
34	BG 204	CGGGACCATGCCTGTAAATA	AAAGGGAACATTAAGGTTTATAAAATG	52.65	M
35	BG 205	TCGGTTCCATAAAAATGTTGG	TTCAAATTTTGTGATCAATGGTT	50.5	M
36	BG 206	ACATATTGAATGATCTACCTCGACA	TGCATCTTACTTGAATATCTTACGAC	53.25	P
37	BG 213	CTGGCCGGTTCTCATGTAAT	TCCAAACCGGACTAGATCGT	55.1	P
38	BG 214	TGGACAAAATTAATTGCTACTCAC	GATATTTTTGGCCGGGTTC	51.7	P
39	BG 224	GGATGACGAGTGTCGAGGAT	CTCTTCATTCAACGCCTCAA	54.55	M
40	BG 225	GGGAGGAGAAAGAGGGAGAG	CTGTCTGCCACTCCACTTGA	56.6	M
41	BG 314	AGTTGTTGGCATTCTGGGTC	GATCAAACCTGAGGAAGCGA	55.15	M
42	BG 325	AAAAGGTTTCACTTGCACCG	GTCTGGAGGTGGTCGAGAAG	55.55	M
43	BG 333	CGTACAAAGTCCTCTTCCCC	GTTTGTTTTTGTCGGCGATT	53.7	M
44	BG 334	CCTTGCTGCATTCCTCTCTC	GCTCAGGGATGAGATGTTGG	55.5	M
45	SR 1	GTTTGGTTCAGAGGCAGAGG	CTATCGCTGCAGAAGAAGGG	63.85	N
46	SR 3	AAGAACGTCAAGATCCTCTGC	ACCACCACGGTAGTAGAGCG	63.7	M
47	SR 4	TCAAAGCCATAAAGCAGGTG	CATCTTCAACACGCATACCG	63.4	M
48	SR 5	GGTAAGCCAAAAACCCTTCC	GAAACCGGTAACAAAGTCGG	63.2	N
49	SR 6	TCCTCCTACTTTGATACTTGC	ACGTCAAATACTTCACTGCC	58.35	M
50	SR 7	TGCTGCTACAGACAGTGTTGG	AAAGGCTACACACTCATGAAACC	63.6	M
51	SR 8	GGAACCCAACAAGTGAGTCC	AGAGCTTGAGACACATAACACC	61.65	N
52	SR 9	ACTTCTTGCCCTCCTCACC	AAATACTCACTGCAATACCCAGG	63.1	M
53	SR 10	CCTCTTCAGTCGAGGTCTGG	AATTTGGAAACAGAGTCGCC	63.45	M
54	SR 11	GCAAAACAATAACTATTAGATTCGTCG	ACACACGCACATAAACACGC	64.1	M
55	SR 12	GCCAGAAACAGGAGAAATGG	GAAGCCGAAGAAAATAAGCG	63.05	M
56	SR 13	TCTATCAAATTTGTTTCTCAGGC	CATTGCGTTCTTCAATGGC	62.05	M
57	SR 14	CATTGGTTTAGTCATTTCGTCG	AATTCAAAAACTGCCGAACG	63.45	P
58	SR 15	GTTAAGTGTGGCGTTAGAGG	CCTTGGTACATGCCACTGAA	61.25	M
59	SR 16	TGGTGGCTTGAGATTAGTTC	ACTCGAAGCCTAATGAAAAG	58.5	M
60	SR 17	CATGGTTAAAACAATGGCCC	CAAGAAACACCATCATTTCTCA	62.6	M
61	SR 18	TCTTGAGCAAAGAAACTTGG	CAAACTGAGCCATACACAAAGG	61.45	M
62	SR 33	GGATGTTTTCACAGACCCTG	CTTTGCAGGTATGAACACGC	62.6	M
63	SR 37	TATGTACACATTCCTCATTTTC	CATTCGTCTCCACCTTCT	56.6	P
64	SR 42	CACCTATCAGATAAGAGTTTCCAT	TCATATGAAATGCTTTTCAAGCGAGTCTA	65.55	M
65	SR 46	AACCAAGGCACAGAAGCG	ACCCAAGTGATCGCCACC	65.15	P
66	SR 47	GGAGACTTTTTGAGATTGGCC	CCAAAACAATTAGCTCCCCA	63.65	M
67	SR 48	TCAGGAGGAACTAAAGTGAGGG	CACACTGAAGATGGTCTTGAGG	63.55	P
68	IM 4	CCCAAACGCTTTTGACACAT	GGCACAATCCACTCAGCTTT	55.0	N
69	IM 8	GCGATGTTTTTTCTTCAGTGTC	TTAATCCCTACCCACAATTTCC	52.62	N
70	BJ 2	GATGGTGATGGTGATAGGTC	GAAGAGAAGGAGTCAGAGATG	62.0	M
71	BJ 5	CTCGTCTTCTTCACCTACAAC	CTGACATCTTTCTCACCCAC	50.0	M
72	BJ 52	TATGAACAGCTACGTAGTCCCA	CCTGAGGATAACCAAATGAAAA	56.5	M
73	BJ 95	CGTAAGTTTCAATTGTCAACGG	TCGTACGAAACAATCAACGG	52.95	M
74	BJ 96	GATCTTCTCTCCAAAACTCTCT	AAAGTCCAAGCTAAATTACAAA	50.0	N

Table 2: Cont.....

S. No.	Primer Name	Forward Primer	Reverse Primer	Ann. Temp.	Ampli.*
75	BJ 97	AACTTTGCTTCCACTGATTTTT	TTGCTTAACGCTAAATCCATAT	50.8	M
76	BJ 103	TGCCTCCTCTCATTTTTTCTCT	TGACCGAGAGGTTCACAAGTAA	54.8	M
77	BJ 104	GCGATGTTTTTTCTTCAGTGTC	TTAATCCCTACCCACAATTTCC	52.65	N
78	BJ 106	ACGAATTGAATTGGACAGAG	CAGATGGGAGTCAAGTCAAC	51.65	N
79	BJ 107	ACCTCCTGCAGATTCGTGTC	GCTGACCTTTCTTACCGCTC	56.35	N
80	BJ 108	TGGTGGCTTGAGATTAGTTC	ACTCGAAGCCTAATGAAAAG	51.15	M
81	BJ 109	AGGACACCAGGCACCATATA	CATTGTTGTCTTGGGAGAGC	54.6	M
82	BJ 125	GCGTCGAGAGAGATCGAGAG	CTCACCGTCACTGCTTCATC	56.15	M
83	BJ 130	GGTAAGCCAAAAACCCTTCC	GAAACCGGTAACAAAGTCGG	54.0	N

*M-Monomorphic, P-Polymorphic, N-No amplification

Table 3: Polymorphic primers depicting band size, number of alleles and PIC values

Sr. No.	Primer	Amp. Range (bp)	Total no. of alleles	PIC
1	BG 12	280-300	3	0.659
2	BG 33	220-350	4	0.748
3	BG 45	140-350	4	0.116
4	BG 55	90-95	2	0.497
5	BG 58	100-200	2	0.489
6	BG 86	150-220	3	0.662
7	BG 92	240-260	3	0.554
8	BG 103	260-265	2	0.481
9	BG 156	170-200	4	0.623
10	BG 206	320-340	2	0.434
11	BG 213	290-320	3	0.619
12	BG 214	260-350	2	0.452
13	SR 14	200-350	3	0.403
14	SR 37	120-300	5	0.689
15	SR 46	240-250	2	0.496
16	SR 47	70-260	3	0.556

Table 4: Clustering pattern of 96 germplasm lines of Indian mustard on the basis of SSR marker analysis

Cluster No.	Germplasm lines	No. of lines
CI	RC2, RC5, RC6, RC7, RC8, RC12, RC14, RC21, RC22, RC23, RC25, RC27, RC28, RC29,RC30, RC32, RC51, RC52	18
CII	RC15, RC18, RC20, RC31, RC33, RC35, RC36, RC38, RC46, R47, RC49, RC53, RC85, RC86, RC87	15
CIII	RC13, RC34, RC48, RC50, RC74, RC78, RC95, RC99, RC102, RC263	10
CIV	RC24, RC54, RC57, RC61, RC89	5
CV	RC77, RC111, RC112, RC129	4
CVI	RC81, RC91, RC93, RC96, RC110, RC114, RC116, RC118, RC127, RC134, RC135,RC142,RC148,RC150, RC152,RC154, RC161,RC162, RC164, RC165, RC166, RC171, RC185, RC195, RC260, RC261, RC270, RC273, RC275, RC280, RC283	31
CVII	RC104, RC105, RC264, RC265, RC268	5
CVIII	RC163, RC174	2
CIX	RC106, RC107, RC108, RC175	4
CX	RC26, RC37	2

Molecular diversity analysis of 96 germplasm lines was carried out using SSR markers. Out of 83 primers, 12 primers gave no amplification, 55 primers were found to be monomorphic and 16 primers gave polymorphic bands. Sixteen primers amplified a total of 47 alleles which varied from 2 to 5 with an average of 2.9 alleles per primer. Parida *et al*. (2010) reported only 2 alleles per locus, where the fragment size varied from 100 to 2000 bp. The electrophoretic pattern with primer BG92 is depicted in Fig. 1. The overall size of PCR amplified products ranged from 70 bp (SR47) to 350 bp (BG33, BG214, BG45, SR14). The molecular size difference between the smallest and largest allele at a SSR locus varied from 5 bp (BG55) to 190 bp (BG45 and SR47). The highest PIC value was observed to be 0.748 with primer BG33 and lowest was 0.116 with primer BG45 (Table 3). The mean PIC value from all the polymorphic primers was 0.529. Gupta *et al*. (2014) reported low PIC value 0.281. Out of 16 polymorphic primers, 15 primers gave 100% polymorphism. Only one primer i.e. BG45 gave 75% polymorphism. The overall polymorphic percentage by 16 primers was 97.87%. Chandra *et al*. (2013) also reported a high level of polymorphism *i.e*. 97.56 %. Prajapat *et al*. (2014) reported nine primers out of 24 primers used to be 100% polymorphic in four *Brassica* species. The reason might be that the germplasm lines used in the present study are different in morphology and genomic constitution.

The SSR diversity data were used to determine genetic

Table 5: Clustering pattern of 96 germplasm lines of Indian mustard based on phenotypic variations

Cluster No.	Germplasm lines	No. of lines
CI	RC2, RC32, RC51, RC89	4
CII	RC5, RC24	2
CIII	RC6, RC8, RC13, RC20, RC22, RC23, RC31, RC33, RC34, RC91, RC104, RC107, RC108, RC114, RC127, RC150, RC162, RC165, RC175, RC185, RC195, RC260, RC264, RC265, RC270, RC273	26
CIV	RC7, RC27, RC37, RC46, RC48, RC57, RC61, RC93, RC99, RC102, RC105, RC110, RC111, RC112, RC116, RC129, RC142, RC283	18
CV	RC12, RC30, RC78, RC134, RC135, RC163, RC261, RC268, RC275	9
CVI	RC14, RC15, RC18, RC25, RC28, RC35, RC36, RC47, RC50, RC52, RC74, RC81, RC85, RC86, RC106, RC118, RC148, RC152, RC154, RC164, RC166, RC171, RC174, RC263, RC280	25
CVII	RC21, RC49, RC54, RC87	4
CVIII	RC26, RC53, RC161	3
CIX	RC29, RC38, RC77	3
CX	RC95, RC96	2

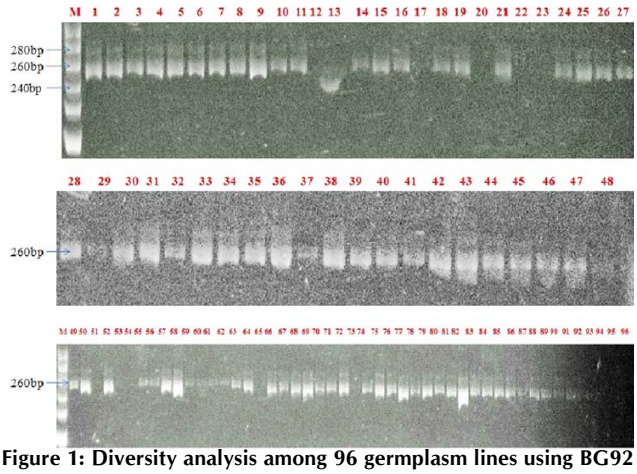

Figure 1: Diversity analysis among 96 germplasm lines using BG92 marker

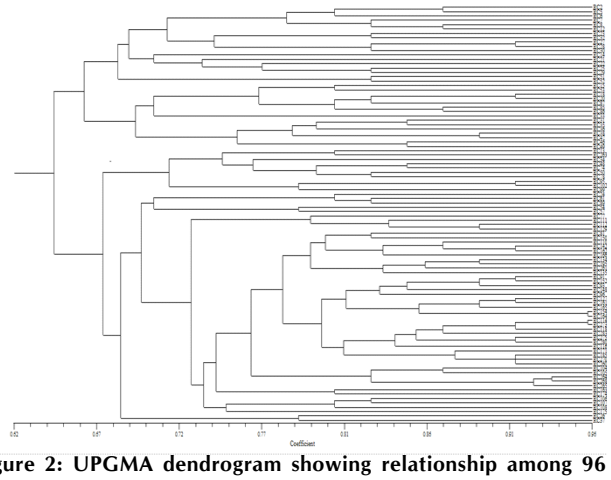

Figure 2: UPGMA dendrogram showing relationship among 96 germplasm lines of Indian mustard based on similarity matrix data using 16 SSR markers

relationship among the 96 germplasm lines using NTSYS- pc software version 2.02e. Dendrogram (Fig.2) was constructed based on the simple matching coefficient using SAHN sub-program which uses UPGMA algorithm to perform cluster analysis. The range of similarity coefficients was found between 0.38 to 0.96. Based on Jaccard's coefficient of similarity values, the maximum similarity (0.957) was observed between the accessions, RC273 and RC114; and RC164 and RC154. On the other hand minimum similarity value of 0.38 was found between RC106 and RC15. The accession RC15 was found to be the most diverse. It was having lesser similarity coefficient values with most of the accessions *viz*. 0.42 with RC174 and RC275; 0.44 with RC114 and RC268. The accession RC37 was having 0.51 similarity coefficient with RC54 and 0.55 with RC134. RC87 had 0.55 similarity coefficient with RC116 and 0.57 with RC270. Similarly, the accession RC26 had low similarity coefficient (0.55) with RC54. Vinu *et al*. (2013) also evaluated the genetic diversity among 44 Indian mustard (*Brassica juncea*) genotypes and found the range of Jaccard's similarity coefficients based on SSR data from 0.38 to 0.83. This wide range of similarity coefficients between different germplasm lines may be due to inclusion of large number of lines in the present study and use of SSR markers which are effective in identifying close pedigree relationships. Further, in dendrogram all the 96 germplasm lines were divided in to 2 main clusters at similarity coefficient of 0.65. At similarity coefficient of 0.67, cluster I is divided in to two sub clusters. Similarly, Pratap *et al*. (2015) obtained 2 main clusters in the UPGMA dendrogram while evaluating 20 Indian mustard genotypes for Alternaria blight tolerance using 25 SSR markers. Furthermore, simple matching matrix was subjected to Principal Component Analysis (PCA) for the three principal components. The grouping of all the 96 lines using PCA analysis in 2-D and 3-D scaling (Fig. 3) followed the same pattern as depicted in the dendrogram with minor differences. This could be attributed to the sensitivity of the PCA since this analysis is expected to be more informative about differentiation among major groups (Mwase *et al*., 2007). The accessions, RC26, RC37 and RC87 which had low similarity coefficients with other accessions; were placed on the periphery in 2-D and 3-D also. Results of the Principal Component Analysis showed that 74.70% of the total variation could be explained by three principal components based on first, second and third eigen vectors which accounted for 58.71, 11.80 and 4.19% variation, respectively. However, Soengas *et al*. (2008) reported that eight Principal components accounted for 94% of the total variation in a study of morphological and agronomic diversity of *Brassica napus*.

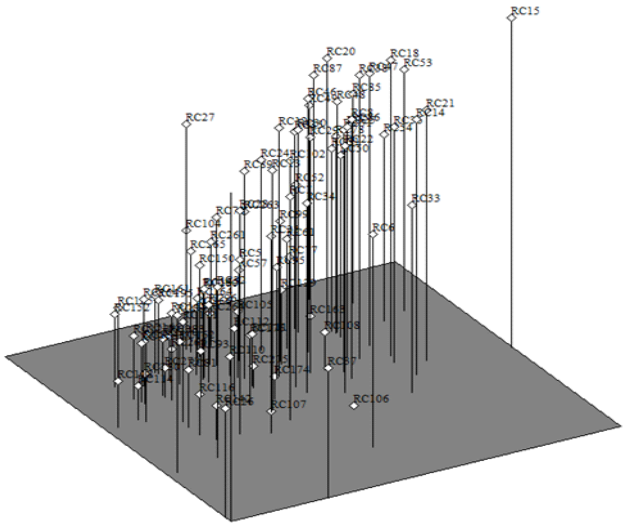

Figure 3: Three dimensional PCA (Principal Component Analysis) scaling of 96 germplasm lines of Indian mustard using similarty matrix data of 16 SSR markers

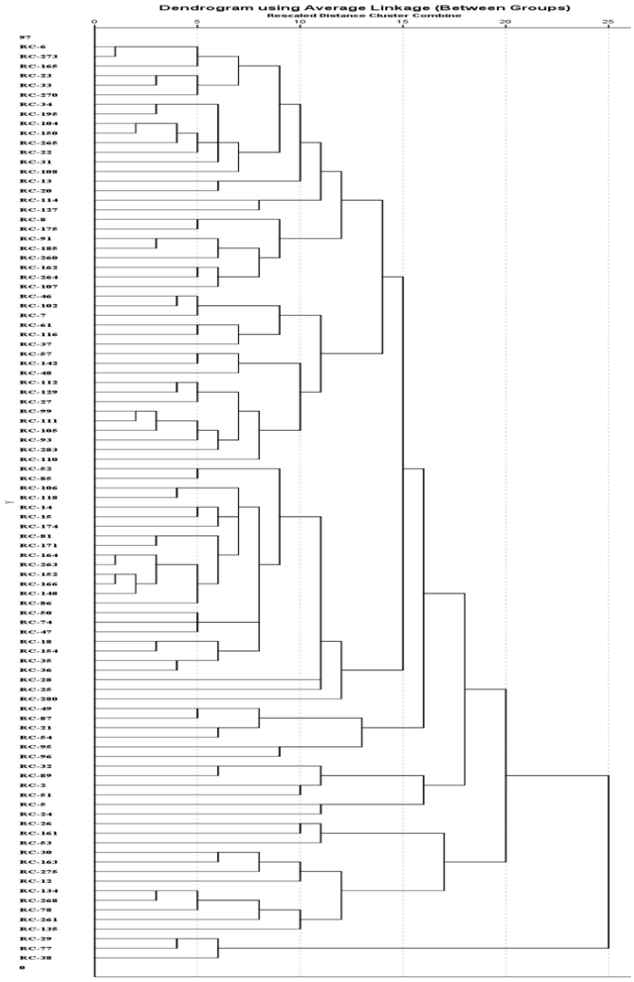

Figure 4: Dendrogram based on phenotypic variations among 96 germplasm lines of Indian mustard using average linkage between groups

They found that molecular and morphologic classifications were complementary and necessary to classify genetic relationships among cultivars.

Correlation between molecular and phenotypic analyses

To compare the phenotypic data with molecular data based on the dendrograms; all the germplasm accessions were divided into ten groups/clusters in both the cases. Total number of accessions in each cluster ranged from 2-31 lines (cluster VI to be the largest) in molecular data based analysis (Table 4) and 2-26 lines (cluster III to be the largest) in morphological data based analysis (Table 5 and Fig. 4). However, Kumar *et al*. (2013) observed eight clusters with a range of 1-18 genotypes in each cluster (cluster III being the largest) in the genetic diversity study on forty six Indian mustard genotypes. In the present study, none of the germplasm lines was found common in six clusters whereas, in rest of the four clusters (CI, CIII, CIV and CVI) some of lines were found to be common e.g. in cluster number six (CVI); lines RC81, RC118, RC148, RC152, RC154, RC164, RC166, RC171 and RC280 were found in common. The grouping of accessions based on SSR marker data seems to be more reliable than morphological (phenotypic) data as it is free from environmental fluctuations. Vieira *et al*. (2007) also suggested that a large portion of variation detected by molecular markers is non-adaptive and is, therefore, not subject to either natural or artificial selection as compared with phenotypic characters, which in addition to selection pressure are influenced by the environment. Principal component analysis (PCA) explained 74.70% of total variation based on the molecular data whereas 70.41% of total variability was explained by phenotypic data. Similar findings were recorded by Vinu *et al*. (2013) in *B. juncea* genotypes. In their study PCA revealed that the grouping of genotypes based on the SSR marker data was more convincing than the phenotypic data. Therefore, the genetic diversity analyzed in the present study based on SSR markers is more authentic than the phenotypic data and the germplasm lines RC15 and RC106 which were found to be the most diverse based on molecular data can be further utilized in hybridization programs.

REFERENCES

Alie, F. A., Singh, T., Tariq and Sharma, P. K. 2009. Genetic diversity analysis in Indian mustard *[Brassica juncea* (L.) Czern and Coss]. *Pro. Agri. Int. J.* **9:** 50-53.

Arunachalam, V., Prabhu, K. V. and Sujata, V. 1996. Efficiency of isozyme markers in genetic differentiation of *Brassica*. In 2nd International Rapeseed Congress. New Delhi, India.

Chandra, V., Usha Pant, Ram Bhajan and Singh, A. K. 2013. Studies on genetic diversity among Alternaria blight tolerant Indian mustard genotypes using SSR markers. *The Bioscan*. **8(4):** 1431-1435.

Diers, B. W. and Osborn, T. C. 1994. Genetic diversity of oilseed *Brassica napus* germplasm based on Restriction Fragment Length Polymorphisms. *Theor. Appl. Genet*. **88:** 662-668.

Falconer, D. and Mackay, T. F. 1997. Introduction to quantitative genetics. *J. Wiley & Sons,* New York, USA.

Gangapur, D. R., Prakash, B. G. and Hiremath, C. P. 2010. Genetic diversity analysis of Indian mustard (*Brassica juncea* L.). *Elect. J. Plant Breed.* **1:** 407-413.

Gupta, N., Zargar , S. M., Gupta , M. and Gupta, S. K. 2014. Assessment of Genetic Variation in Indian Mustard (*Brassica juncea* L.) Using PCR Based Markers. *Mol. Plant Breed.* **5(3):** 10-17.

Hopkins, C. J., Cogan, N.O.I., Hand, M., Jewell, E., Kaur, J., Xi, L., Lim, G. A. C., Ling, A. E., Love, C., Mountford, H., Todorovic, M., Vardy, M., Spangenberg, G. C., Edwards, D. and Batley, J. 2007. Sixteen new simple sequence repeat markers from *Brassica juncea* expressed sequences and their cross-species amplification, *Mol. Eco. Notes.* **7:** 697-700.

Jaccard, P. 1908. Nauvelles re churches sur la distribution florale. *Bull. Soc. Vaudoise Sci. Nat.* **44:** 223-270.

Khan, W. M., Munir, I., Farhatullah, Arif, M., Iqbal, A., Ali, I., Ahmad, D., Ahmad, M., Mian, A., Bakht, J., Inamullah. and Swati. Z. A. 2011. Assessment of genetic diversity of *Brassica juncea*germplasm using Randomly Amplified Polymorphic DNA (RAPD) markers. *Afr. J. Biotech.* **10:** 3654-3658.

Kumar, B., Pandey, A. and Singh, S. K. R. 2013. Genetic diversity for agro-morphological and oil quality traits in Indian mustard (*Brassica juncea* L. Czern & Coss). *The Bioscan.* **8(3):** 771-775.

Kumar, R. and Gupta, V. P. 1985. Isozyme studies in Indian mustard (*Brassica juncea* L.). *Theor. Appl. Genet.* **69:** 1-4.

Li, H., Younas, M., Wang, X., Li, X., Chen, L., Zhao, B., Chen, X., Xu, J., Hou, F., Hong, B., Liu, G., Zhao, H., Wu, X., Du, H., Wu, J. and Liu, K. 2013. Development of a core set of single-locus SSR markers for allotetraploid rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* **126:** 937-947

Mir, R. R., Kumar, J., Balyan, H. S. and Gupta, P. K. 2011. A study of genetic diversity among Indian bread wheat *(Triticum aestivum* L.) cultivars released during last 100 years. *Genet. Resor. Crop Evol.* **59:** 717-726.

Mwase, W. F., Erik-Lid, S., Bjornstad, A. Stedje, B., Kwapata, M. B. and Bokosi, J. M. 2007. Application of amplified fragment length polymorphism (AFLPs) for detection of sex-specific markers in dioecious *Upaca kirkiana* Muell. *Arg. Afr. J. Biotechnol.* **6:** 137-142.

Parida, S. K., Yadava, D. K. and Mohapatra, T. 2010. Microsatellites in *Brassica unigenes*: relative abundance, marker design, and use in comparative physical mapping and genome analysis. *Genome.* **53:** 55-67.

Prajapat, P., Sasidharan, N., Kumar, M. and Prajapati, V. 2014. Molecular characterization and genetic diversity analysis in four *Brassica* species using microsatellite markers. *The Bioscan.* **9(4):** 1521-1527.

Pratap, P., Thakur, A. K., Meena, P. D., Meena, H. S., Sharma, P., Singh, D. and Majumdar, R. 2015. Genetic diversity assessment in Indian mustard (*Brassica juncea* L.) for Alternaria blight tolerance using SSR markers**.** *J. Oilseed Brass.* **6(1):** 175-182.

Rahman, M. M., Rasaul, M. G., Hossain, M. A., Iftekharuddaula, K. M. and Hasegawa, H. 2012. Molecular characterization and genetic diversity analysis of rice (*Oryza sativa* L.) using SSR markers. *J. Crop Improv.* **26:** 244-257.

Ramchiary, N., Nguyen, V. D., Li, X., Hong, C. P., Dhandapani, V., Choi, S. R., Yu, G., Piao, Z. Y. and Lim, Y. P. 2011. Genic microsatellite markers in *Brassica rapa*: development, characterization, mapping, and their utility in other cultivated and wild *Brassica* relatives. *DNA Res.* **18:** 305-320.

Rohlf, F. J. 1998. NTSYS-pc Numerical taxonomy and multivariate analysis system, version 2.02. New York: Exeter Publication.

Romesburg, H. C. 1984. Cluster Analysis for Researchers. *Krieger Publishing Co., Malabar, Florida.*

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc. Natl. Acad. Sci.* **81:** 8014-8019.

Singh, D., Arya, R. K., Chandra, N., Niwas, R. and Salisbury, P. 2010. Genetic diversity studies in relation to seed yield and its component traits in Indian mustard (*Brassica juncea* L. Czern & Coss.). *J. Oilseeds Brass.***1:** 19-22.

Sivaranjani, R., Santha, I. M., Pandey, N., Vishwakarma, A. K., Nepolean, T. and Hossain, F. 2014. Microsatellite-based genetic diversity in selected exotic and indigenous maize (*Zea mays* L.) inbred lines differing in total kernel carotenoids. *Ind. J. Genet.* **74:** 34-41.

Soengas, P., Cartea, M. E., Velasco, P., Padilla, G. and Ordas, A. 2008. Morphologic and agronomic diversity of *Brassica napus* crops. *J. Amer. Soc. Horti. Sci.* **133(1):** 48-54.

Song, K. M., Osborn, T. C. and Williams, P. H. 1988. *Brassica* taxonomy based on nuclear Restriction Fragment Length Polymorphisms (RFLPs). *Theor. Appl. Genet.* **76:** 593-600.

Sun, D. L., Zhao, Q. C., Song, W. Q. and Chen, R. Y. 2001. Relationships analysis of Chinese cabbage species by AFLP. *Acta Horti. Sinica.* **28:** 331-335.

Suwabe, K., Iketani, H., Nunome, T., Kage, T. and Hirai, M. 2002. Isolation and characterization of microsatellites in *Brassica rapa* L. *Theor. Appl. Genet.* **104:** 1092-1098.

Vaishnava, A., Sachan, J. N. and Tewari, S. K. 2006. Genetic divergence for important quantitative traits in Indian mustard [*Brassica juncea* (L.) Czern and Coss.] *Agri. Sci. Digest.* **26:** 269-272.

Vieira, E., Carvalho, F., Bertan, I., Kopp, M., Zimmer, P., Benin, G., da Silva, J., Hartwig, I., Malone, G. and de Oliviera, A. 2007. Association between genetic distances in wheat (*Triticum aestivum* L.) as estimated by AFLP and morphological markers. *Genet. Mol. Biol.* **30:** 392-399.

Vinu, V., Singh, N., Vasudev, S., Yadava, D. K., Kumar, S., Naresh, S., Bhat, S. R. and Prabhu, K. V. 2013. Assessment of genetic diversity in *Brassica juncea* (Brassicaceae) genotypes using phenotypic differences and SSR markers. *Rev. Biol. Trop.* **61(4):** 1919-1934.

Wang, F., Wang, X.F., Chen, X., Xiao, Y., Li, H., Zhang, S., Xu, J., Fu, J., Huang, L., Liu, C., Wu, J. and Liu, K. 2012. Abundance, marker development and genetic mapping of microsatellites from *unigenes* in *Brassica napus*. *Mol. Breed.* **30:** 731-744.

Wang, J., Kaur, S., Cogan, N. O. I., Dobrowolski, M. P., Salisbury, P. A., Burton, W. A., Baillie, R., Hand, M., Hopkins, C., Forster, J. W., Smith, K. F. and Spangenberg, G. 2009. Assessment of genetic diversity in Australian canola (*Brassica napus* L.) cultivars using SSR markers. *Crop Pasture Sci.* **60:** 1193-1201.

Yildirim, E., Yildirim, N., Ercisli, S., Agar, G. and Karlidag, H. 2010. Genetic relationships among turnip (*Brassica rapa*var. *rapa*) genotypes. *Genet. Mol. Res.* **9:** 987-993.