Journal of Animal & Plant Sciences, 32(1): 2022, Page: 173-185
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2022.1.0413

ANALYSIS OF GENETIC STRUCTURE AND DIVERSITY IN INDIAN MUSTARD [*BRASSICA JUNCEA* (L.) CZERN & COSS.] ASSESSED BY SSR MARKERS

V. K. Singh, R. Avtar, Mahavir*, N. Kumari, Manjeet, R. Kumar and R. S. Khedwal

Oilseeds Section, Department of Genetics and Plant Breeding, CCS Haryana Agricultural University, Hisar, Haryana, 125 004, India
*Corresponding Author's Email: mahaveer.bishnoi9@gmail.com

ABSTRACT

In present endeavors, the genetic diversity and Bayesian structure of 88 genotypes of Indian mustard procured from four different centers *viz.,* ICAR-DRMR Bharatpur, ICAR-IARI New Delhi, CCS HAU Hisar and PAU Ludhiana were studied using 59 genomic SSR markers, and their genetic liaison was explored. A total of 209 repeatable alleles were detected by 59 SSR markers in a size range of 50-1000 bp with maximum (7) fragments intensified by BG6, BG32, BG46 and BG71 markers. The average value of PIC and the mean expected heterozygosity (*He*) value from all the polymorphic primers were 0.49 and 0.56, respectively, which is an indicative of the presence of ample amount of genetic diversity among Indian mustard genotypes. All the 88 genotypes were grouped into four distinct clusters based on Jaccard's dissimilarity coefficients and UNJ (Unweighted Neighbour Joining) methods, however, three subpopulations were predicted by bayesian structure analysis at delK = 3. The PCoA (Principal Coordinate Analysis) revealed 20.95% and 10.85% of variation, respectively, with 31.80% of cumulative variation. The present work indicates the presence of considerable genetic diversity among the Indian mustard genotypes, which could be used in future breeding programmes for developing mustard cultivars and germplasm management purposes.

Keywords: Indian mustard, genetic diversity, bayesian structure, SSRs.

Published first online June 14, 2021 Published final January 07, 2022.

INTRODUCTION

Global climate change and population explosion are the serious concerns for future food security. The world agriculture is predicted to be significantly affected by climate change, though the impact will vary by region and crop (Calicioglu *et al.,* 2019). Rapeseed-mustard (*Brassica*) is one of the most important oilseed crops all over the world. All these crops generally belong to four species which are diverse and originated from the family *Cruciferae*/*Brassicaceae* (Wani *et al.,* 2020). *Brassica juncea* (L.) Czern & Coss., usually known as 'Indian mustard' is an important member of the family *Cruciferae* and grown in major part of Indian subcontinent (Lakhanpal *et al.,* 2018). It takes over approximately 25% share among all the oilseed crops in India (Yadav *et al.,* 2019). Worldwide, it is the third largest oilseed crop after soybean and palm contributing 12.1 million tons of total consumable edible oil (Friedt and Snowdon, 2009). India is a major oilseed producer, consumer as well as importer globally but even then, about 57% of the total consumable vegetable oil is imported from other countries (Jat *et al.,* 2019). The major reasons behind this are the low productivity of Indian mustard at the farmer's fields. (Yadav *et al.,* 2019). The population explosion and changing dietary pattern led to the increased demand of edible oils (Tripathi *et al.,* 2019; Davis *et al.,* 2016); therefore, it is urgent need to break yield plateau so that enhanced oil demand can be copped with. Genetic diversity is a key factor in generating variability in segregating populations and this variability is the attribute of heterosis produced by hybrids which are generated from genetically diverse lines (Goulet *et al.*, 2017; Kaeppler, 2012). There are better chances to develop superior hybrids/cultivars through breeding programs if ample genetic variability is present in breeding material (Bhargava and Srivastava, 2019). Thus, breeders and geneticists can be assisted with better understanding of genetic diversity in *B. juncea* and this would be very helpful in accessing genetic variability present in germplasm and ultimately will lead the path to widen the genetic base of breeding material (Panjabi *et al.,* 2019). Diversity among living organisms is a consequence of variations in DNA sequences and of environmental effects (Carvalho *et al.,* 2019). Assessment of phenotype-based diversity is not so reliable due to the effect of environmental factors and G X E interactions (El-Soda *et al.,* 2014). Therefore, the DNA based diversity assessment is really vital and good enough for differentiating genotypes based on their genetic architecture. Molecular markers have proven to be invaluable tools for assessing the extent of genetic variation and its distribution within species. Microsatellites or Simple Sequence Repeats (SSR) are highly polymorphic, species-specific and co-dominant markers, considered as a most promising tool for

detecting genetic diversity in plants (Adhikari *et al.,* 2017). Keeping in view of the above fact, the present study was carried out for assessing genetic diversity among various genotypes of *B. juncea* in order to find out the most diverse genotypes which can further assist breeding program. These utmost diverse genotypes will be employed for crossing program which can ultimately lead to favorable allele combinations in segregating populations.

MATERIALS AND METHODS

A total of 88 diverse genotypes were procured from four different centers of India *viz.,* Directorate of Rapeseed-Mustard Research (DRMR), Bharatpur; Indian Agricultural Research Institute (IARI), New Delhi; CCS Haryana Agricultural University (CCS HAU), Hisar and Punjab Agricultural University (PAU), Ludhiana; including 14 genotypes of exotic origin (maintained at IARI). All the genotypes taken were grown in paired rows with standard package and practices at Research Area of Oilseeds Section, Department of Genetics and Plant Breeding, CCS HAU, Hisar during *rabi*, 2019-20. List of 88 genotypes used in the present study is given in Table 1.

Molecular marker evaluation: Fresh and young leaves were taken for genomic DNA isolation of 88 diverse genotypes and 2% CTAB (Cetyl Trimethyl Ammonium Bromide) was employed for the purpose (Saghai-Maroof *et al.,* 1984) with slight alterations. For removal of RNA contamination, isolated genomic DNA was subjected to RNase treatment. Quantity and integrity of DNA was assessed by 0.8% agarose gel while taking λ DNA (50ng/µl) as reference standard. DNA was finally diluted to 50ng/µL and further stored at -20°C for PCR amplification. Genomic DNA of all 88 genotypes was amplified using a set of 59 random primers (Table 2). The PCR amplification reaction for all SSR markers comprised a 20µL master mix containing 2.0µL of 10X buffer, 0.4µL of dNTP mix (10mM), 0.5µL of each forward and reverse primer (10µmol/L) (Integrated DNA Technologies), 0.5µL of *Taq* DNA polymerase (5U/µL) (New England Biolabs, Inc.), 2µL of template DNA (50ng/µL), and 13.1µL of ddH_2O employing a thermal cycler (Bio Rad T100, Thermal Cycler).

After an initial strand separation step of five minutes at 94°C, PCR was programmed for 35 cycles. Each cycle comprised of a denaturation step (one minute at 94°C), annealing step (variable annealing temperature for 30s), extension step (one minute at 72°C), followed by final extension (7 minutes at 72°C). The amplified products were resolved on 2% agarose gel. Band sizes of amplified products were analyzed by comparing with DNA ladder (50 bp) loaded in the gels for estimating the appropriate size of bands. The SSR bands obtained in gels were scored manually for absence and presence as zero and one, respectively in each genotype and binary data-set was created. The informativeness of the microsatellite markers used for differentiating Indian mustard genotypes was ascertained by expected heterozygosity (*He*) = $1-\sum_i^n p^2$; where *pi* is the frequency of the i^{th} allele and *i* is the total number of alleles at all loci as per Liu (1998). PIC (polymorphism information content) was calculated using the formula $PIC_j = 1-\sum_{i-1}^n p^2$; where *i* is i^{th} allele of the j^{th} marker, *n* is the number of the j^{th} marker's alleles, *pi* is allele frequency (Botstein *et al.,* 1980). Genetic dissimilarities/number of clusters and factorial analysis among the Indian mustard genotypes were analyzed with the help of DARwin 6.0 software (Perrier and Jacquemoud-Collet, 2006). Un-weighted Neighbor-joining (UNJ) tree was constructed by dissimilarity matrix generated by software. Evaluation of population structure and number of gene pool was performed with Bayesian model-based cluster analysis using the STRUCTURE version 2.3.4 software. (Pritchard *et al.,* 2000). For the analysis, number of presumed population (K) was set from 1 to 12; for each fixed K, 10 independent runs were assessed while taking that each run comprised of 50,000 burn-in period and 100,000 iterations. Structure Harvester v6.0 (Earl and vonHoldt, 2012) was employed for calculating optimum value of K by analyzing the delK statistics and L (K) (Evanno *et al.,* 2005).

Table 1: List of 88 genotypes of Indian mustard used in present study.

Sr. No.	Genotypes	Source Centre	Sr. No.	Genotypes	Source Centre	Sr. No.	Genotypes	Source Centre	Sr. No.	Genotypes	Source Centre
1	RH-222	HAU	23	M-156	PAU	45	M-10	PAU	67	DRMRB-17	DRMR
2	DRMRB-2	DRMR	24	M-12	PAU	46	DRMRB-18	DRMR	68	M-167	PAU
3	RH -0502	HAU	25	EC-28-18	IARI	47	EC-62-67-1	IARI	69	DRMRB-15	DRMR
4	DRMRB-13	DRMR	26	EC-27-9	IARI	48	DRMRB-14	DRMR	70	DRMRB-5	DRMR
5	DRMRB-16	DRMR	27	RH-1515	HAU	49	DRMRB-19	DRMR	71	DRMRB-20	DRMR
6	DRMRB-11	DRMR	28	M-7	PAU	50	RH-406	HAU	72	EC-62-46-1	IARI
7	NPJ-113	IARI	29	NPJ-124	IARI	51	EC-30-1	IARI	73	DRMRB-10	DRMR
8	DRMRB-3	DRMR	30	Heera	IARI	52	EC-61-67-1	IARI	74	EC-62-42-1	IARI
9	RH-923	HAU	31	DRMRB-4	DRMR	53	M-183	PAU	75	M-179	PAU
10	Pusa Mehak	IARI	32	DRMRB-8	DRMR	54	RH-1509	HAU	76	DRMRB-12	DRMR
11	DRMRB-9	DRMR	33	NPJ-156	IARI	55	EC-29-9-1	IARI	77	EC-28-1	IARI
12	RH -1490	HAU	34	EC-61-6-1	IARI	56	RH-832	HAU	78	M-21	PAU
13	EC-308575	IARI	35	M-160	PAU	57	MST-11-14-32	IARI	79	RH-0121	HAU
14	M-108	PAU	36	RB-50	HAU	58	RH-401 B	HAU	80	BioYSR	IARI
15	RH-119	HAU	37	RH-749	HAU	59	M-146	PAU	81	RH-8701	HAU
16	RH-1475	HAU	38	RH-1512	HAU	60	DRMRB-6	DRMR	82	M-178	PAU
17	NRCDR-02	IARI	39	RH-725	HAU	61	EC-29-5-4-2	IARI	83	M-82	PAU
18	M-40	PAU	40	Pusha Bahar	IARI	62	M-198	PAU	84	RH-555	HAU
19	M-27	PAU	41	EC-27-21	IARI	63	EC-61-5-2-1	IARI	85	M-4	PAU
20	RH-0305-1	HAU	42	M-187	PAU	64	RH-115	HAU	86	CS-52	HAU
21	DRMRB-1	DRMR	43	M-3	PAU	65	RH-630	HAU	87	Non waxy mutant	IARI
22	M-17	PAU	44	RH-501	HAU	66	DRMRB-7	DRMR	88	EJ-20	IARI

Table 2: List of 59 polymorphic SSR markers.

Sr No.#	Marker Name	Locus Name	Forward primer	Reverse primer	Ann. Temp. (°C)
1	BG1	cnu_m051a	GCTGGCTGCACAATAACAGA	GTACCACTGGAGGAGCTTCG	57
2	BG2	cnu_m054a	GGCCTTTGGAGGTGACTGTA	CAGGGATATGCGGTCTTTCT	56
3	BG4	cnu_m056a	CTGGTTTGGTTCGGTTTGAT	CCTGACAAATAGCAAGAAGTCG	54
4	BG5	cnu_m057a	TCACATGTGGGAAACATTCCT	TGTCATTTTTACTGCATTTTCCGTAT	54
5	BG6	cnu_m058a	TGAGGGTGAGGATGGTGATG	GCACAGTACACCGACGCCTA	58
6	BG7	cnu_m059a	GGGATATTGAAGACCCGCAAA	TCTCCCGGTGGCTTAAAGAA	56
7	BG8	cnu_m060a	TTGGATCAATCAAACTAAACCCTGA	CCAAAATGCCAACAAAAGCA	54
8	BG9	cnu_m061a	GGTGACCACCTCCGTCTTCTT	CTGTATGGAGCCCCAAGCTC	58
9	BG12	cnu_m066a	TTCGATTGAAACACTGAACATTGAA	GCGTTTTCTGTTTTCCCAATAA	53
10	BG15	cnu_m070a	CATAACCACACGGCCTCCTC	AAGTCATGCCCATTCGCCTA	57
11	BG16	cnu_m071a	CGAATCCGACGTGAATTTGA	ATTGAGAAGGTCCGCCATGA	55
12	BG17	cnu_m072a	TGTCATTGTTTCCGCCATTG	CTCCCTCCTCCGACAACAAC	56
13	BG19	cnu_m077a	CGTTGTGTGAAATCGCTCAAAT	TCCGAACTAGAAACCGAAAATATCC	55
14	BG20	cnu_m078a	TCACGTGGCAATATGCGAAC	CTCCGCCACTGGTTGAATCT	57

15	BG23	cnu_m081a	GAGGCAAAAGCGAAGGTGAA	AGCACCCAAACACTCCCAAA	57
16	BG28	cnu_m086a	TCACGCATGTCAGAGCCATT	AACCGCGCGTACGATACACT	58
17	BG30	cnu_m088a	GGATGTTCACGCCGTATGTG	CCATAAACTGCATTGTTTGAATTG	54
18	BG31	cnu_m089a	GCCAGTCGAAACAGATTAGCTAGG	CCACTTTGATTACCTTGCTTTTTCA	55
19	BG32	cnu_m091a	TCACCATGTGCGAGAGCCTA	CGGGCAGATCAAGAAACAGA	56
20	BG33	cnu_m092a	CGTGTGTCCTCTCGTGTCTCA	TGCTCAGCAGTCAGCAATCA	57
21	BG35	cnu_m094a	GAGAGAGAGAGAGAGAGAGAGAAAGA	CCGATACACAACCAGCCAAC	56
22	BG37	cnu_m096a	GCACCTAACCGAACCCCTTTAG	GAGAAGATCGTAGGGCACTGGA	57
23	BG38	cnu_m097a	AAATTCAGCGTTTTCGACCA	CTGAGGCGTGAGAGAAGAGAGA	55
24	BG39	cnu_m101a	TCAAACGCAAATTCAATAAGACAAA	ACTAGATTTCCACCCGCACAAC	54
25	BG41	cnu_m103a	TCCTCCGACAACAACAACTCAA	ATCTAACCCGTCTGCGAATCTG	56
26	BG42	cnu_m104a	CGTTTTTCCTTGGTTATTTGGA	TCGTTCAAATGTCGTATGGACAC	53
27	BG44	cnu_m107a	TGGACGTAACACCCATCTTGAA	AGCTGAGGAAGTGGCTGAGG	57
28	BG45	cnu_m108a	TCCAAGAGACGAAACCACTTCC	GCTTGCTTATATCCTTCCTTGCC	56
29	BG46	cnu_m109a	AGAGAGGGAAGGCGAAAGTGAT	GGTTAAATGAAACAGAGGGACCAA	56
30	BG48	cnu_m111a	CACGAAAGCTGTAGAGGCATGA	TCTTTTCCTGTCCATGAGATTCAA	56
31	BG49	cnu_m112a	CGGAAACGCACATCTCTCACT	TCGATCCATTAAGCCAAACTCA	55
32	BG50	cnu_m113a	CGCCAAATCAAATTAGGGTTTA	CCACGAATTTAACAAGAGACATCC	52
33	BG51	cnu_m115a	GGCGGTCCATCAAACTGGTA	CTGTCCCACAAGCAAAGATTCA	56
34	BG52	cnu_m117a	CAACAAAGGGTTTGAAAACTAAACTCA	GGCGCGGGTCTTAACCTAGT	57
35	BG53	cnu_m118a	TCCTTTTGCTTTCTCTCCATCC	AGACGCCGTCCAAGACAGAG	57
36	BG54	cnu_m121a	ACAGGAGAAACGCAACACCA	GATGCAAACGCTAGCCCAAT	56
37	BG56	cnu_m123a	ACTTGGGCGGTGAAACAGTAAA	GTTATGTGGTGGAGAGGCACAA	57
38	BG57	cnu_m124a	TGCTGGTTATGTTTGCTGATGG	TTCAATCCACGTTTTAGTGCCC	55
39	BG62	cnu_m130a	GGAGAGTACGCGGAGAGGAA	TTGCAAACGCACCACCAC	57
40	BG65	cnu_m134a	TCTCTTTGCCATCGTCGTTTC	CCCCTCAAACTGAGCAGTCAA	56
41	BG66	cnu_m135a	GAAAATACACCTCGCTTTTACACTCA	AAGAATTTAGGGTTCGAAAAGGAG	54
42	BG67	cnu_m136a	AAGCTTGCTTTCCCCGATTC	CCATATTAAGCTTCTATTTCTTTCACA	53
43	BG68	cnu_m137a	AACCTTCATTTCATATACATACACACA	TTCAATCATTTTTATTGGTCATCA	51
44	BG69	cnu_m138a	TTTTTCAAAATTGGTGGCTTAGG	CAATCGTACCTAACCGGTTCATAA	53
45	BG70	cnu_m140a	AGCTATAGCACATATTGAAACATATTG	AAGCGGGTACGTGTTGGAAG	54
46	BG71	cnu_m141a	TGGCAATGGTTTCAAGCTCA	ACTGCCTCGCAAGGAAAGAG	56
47	BG73	cnu_m144a	GCGTGCAGGGATTAGCTTGT	CCAACTCGCCCTTCTCTTCA	57
48	BG76	cnu_m151a	TGGACCACTTCCGTGGATCT	AGCATAATCGAAATGTCCCAAA	55
49	BG77	cnu_m152a	TCGAGAGAAGAAGATGGGATGA	CCGAACAAGTTGATAAAAAGTACAATG	54
50	BG79	cnu_m155a	CGTTTCCTCAGCCTCCTTCA	TGCCTACATCCACCGGAGTT	57
51	BG82	cnu_m159a	CGTATCCATGGCCTTGAATTTT	GGCGAGAACCTTGATGATCC	54
52	BG83	cnu_m160a	TGCATGCCATTGAAGCCTTA	TATGTCCGCATCAGCTCCAC	56
53	BG85	cnu_m163a	GGGGGAAGGTTCTTTGTTACAT	GCATTTGGGGATGGTGAGAG	55
54	BG88	cnu_m166a	CTCCTCCTCCAGCGTCTTCA	CGCGTTTGAAGGAGATTTGG	56
55	BG89	cnu_m170a	TGCCAACAAATCAAGGATGC	CCGAAGTTCACTTGTTATTCCAAC	54
56	BG94	cnu_m177a	CCTTCAAAAGAAAGGAGGGGAA	GAGAGAGAGAGAGGGCATAATAAAAGC	55
57	BG95	cnu_m178a	AGCTGCAAGAAAGCGCAAAA	ATTGCCGAACCTCACTTCCAT	56
58	BG99	cnu_m203a	CAGAGCGAGCTGCAAGACAG	CATTGCCGAACCTCACTTCC	57
59	BG105	cnu_m214a	TCGATCTTTTTGCGGTGGAT	TTGCAATGGGCATTACATCCT	55

RESULTS

SSR polymorphism: In present endeavors, all the 88 Indian mustard genotypes were inspected with 59 polymorphic SSR markers and merely steadfast and explicit fragments intensified by the primers were scored (Table 2). A total of 209 repeatable alleles were detected, which varied from 2 to 7 with an average of 3.54 alleles per locus. The overall fragment length of PCR amplified products ranged between 50 bp (BG5, BG19, BG42 and BG57) to 1000 bp (BG37). Among 59 polymorphic markers, 19 produced minimums of 2 alleles whereas, four primers namely, BG6, BG32, BG46 and BG71 produced a maximum of 7 alleles (Table 3). The percent polymorphism ranged between 33.33 to 100% with an average of 87.99%. Of 59 primers, 40 primers had highest percent polymorphism (100%). The PIC value indicates allelic disparity and frequency between the studied genotypes. In this study, the PIC value varied from 0.08 (BG20) to 0.84 (BG6 and BG32) with mean value of 0.49 for all the studied genotypes, demonstrating moderate discriminating capability of the SSR markers used for this study (Table 3). Twenty-eight SSR markers have reported PIC values of more than 0.50. Markers with high PIC value are considered as suitable markers for explaining the genetic diversity among genotypes. The average expected heterozygosity/gene diversity (*He*) was recorded to be 0.56 with maximum value for BG6 (0.86) and minimum values for BG1 and BG20 (0.09), respectively.

Genetic distance and cluster analysis: The dissimilarity coefficients of tested genotypes varied from 0.135 to 0.512 (Figure 1). Based on the genetic dissimilarity, MST-11-14-32 was found most dissimilar from DRMRB-15, DRMRB-20, M-108 and RH-401 B with dissimilarity coefficient of 0.512, 0.499, 0.498 and 0.492, respectively. On the other hand, a minimum dissimilarity value of 0.135 was found between genotypes, CS-52 and DRMRB-12. Dendrogram was constructed based on Jaccard's dissimilarity coefficient and UNJ method, which grouped all 88 genotypes of Indian mustard into four major clusters (Figure 2). Cluster-I, comprised of 49 genotypes which was further subdivided into two sub clusters *i.e., C-Ia* and *C-Ib*. Sub cluster *C-Ia* included 27 genotypes originating from all the four centers, without clear majority from any particular center whereas, sub cluster *C-Ib* comprised of 22 genotypes representing 7 each from Bharatpur and Ludhiana centers, 5 from Delhi center and 3 form Hisar center. Cluster-II includes 5 assorted genotypes from three centers, with no clear majority from any center. Cluster-III consisted of 32 genotypes, maintaining/originating from the entire four centers, but most of the genotypes were from Delhi (12) and Hisar (11) center. The smallest cluster-IV contained only 2 genotypes namely RH-0305-1 and NRCDR-02 from Hisar and Delhi center, respectively.

A PCoA had also been performed in order to better understand the distribution of variation and to categorize any genetic information associated with studied genotypes from the previous dendrogram (Figure 3). It clearly showed the presence of three major groups among the 88 genotypes of our study. The results of PCoA further supported results obtained by dendogram and STRUCTURE. The first two axes of PCoA demonstrated 20.95% and 10.85% of variation, respectively with a 31.80% of cumulative variation (Table 4).

Bayesian structure analysis: A model-based cluster analysis was carried out using STRUCTURE v2.3.3, to determine the genetic association among individual mustard genotypes. The magnitude of delK was obtained by comparing the logarithmic likelihood [L(K)] and estimates of the LnP(D) which ranged from 1 to 10 for the different delKs. The best value of delK was recorded at 3 (Figure 4). At this point, the grouping of genotypes largely supported the results of dendrogram and PCoA. The maximum likelihood and delK (delK = 3) values suggested that all the 88 genotypes of Indian mustard were predominantly allocated into three subpopulations (*SP*) *viz., SP1 (Red)*, *SP2 (Green)* and *SP3 (Blue)* (Figure 5). Of the 88 genotypes, approximately 23 genotypes had pure and 65 genotypes maintained their identity with admixture of alleles of other accessions.

Table 3: Genetic diversity indices of 59 SSR markers for 88 Indian mustard genotypes.

Sr. No.#	Marker	Band Size	Total no. of Alleles	Monomorphic	Polymorphic	% Polymorphism	PIC	*He*
1	BG1	125-135	2	0	2	100.0	0.09	0.09
2	BG2	120-220	3	0	3	100.0	0.43	0.50
3	BG4	140-850	4	1	3	75.0	**0.63**	0.69
4	BG5	50-800	3	1	2	66.7	0.43	0.48
5	BG6	190-800	7	0	7	100.0	**0.84**	0.86
6	BG7	210-320	3	0	3	100.0	**0.58**	0.66
7	BG8	150-250	3	0	3	100.0	0.22	0.24
8	BG9	260-360	3	0	3	100.0	0.36	0.43

9	BG12	90-305	3	1	2	66.7	0.42	0.47
10	BG15	250-275	2	0	2	100.0	0.38	0.50
11	BG16	60-700	3	1	2	66.7	**0.50**	0.58
12	BG17	152-550	3	0	3	100.0	**0.53**	0.61
13	BG19	50-420	4	1	3	75.0	**0.61**	0.67
14	BG20	250-275	2	0	2	100.0	0.08	0.09
15	BG23	120-170	2	0	2	100.0	0.10	0.11
16	BG28	175-600	2	0	2	100.0	0.37	0.49
17	BG30	200-225	2	0	2	100.0	0.37	0.50
18	BG31	60-200	4	1	3	75.0	**0.63**	0.69
19	BG32	200-600	7	0	7	100.0	**0.84**	0.85
20	BG33	220-900	6	2	4	66.7	**0.80**	0.82
21	BG35	350-900	4	0	4	100.0	**0.50**	0.75
22	BG37	310-1000	4	0	4	100.0	**0.69**	0.74
23	BG38	150-900	5	0	5	100.0	**0.73**	0.77
24	BG39	60-250	2	0	2	100.0	0.34	0.44
25	BG41	180-600	2	0	2	100.0	0.33	0.42
26	BG42	50-180	2	1	1	50.0	0.13	0.14
27	BG44	190-400	3	0	3	100.0	**0.59**	0.66
28	BG45	55-700	6	2	4	66.7	**0.69**	0.73
29	BG46	55-800	7	0	7	100.0	**0.75**	0.78
30	BG48	150-390	3	0	3	100.0	**0.51**	0.59
31	BG49	200-600	2	0	2	100.0	0.36	0.47
32	BG50	55-250	2	1	1	50.0	0.37	0.49
33	BG51	100-800	6	1	5	83.3	**0.74**	0.77
34	BG52	200-850	5	0	5	100.0	**0.74**	0.78
35	BG53	155-750	3	0	3	100.0	0.42	0.50
36	BG54	95-850	5	0	5	100.0	**0.74**	0.78
37	BG56	175-650	2	0	2	100.0	0.35	0.45
38	BG57	50-155	2	0	2	100.0	0.34	0.44
39	BG62	75-400	5	0	5	100.0	**0.73**	0.77
40	BG65	300-375	2	0	2	100.0	0.37	0.50
41	BG66	55-400	3	1	2	66.7	0.36	0.44
42	BG67	225-305	2	0	2	100.0	0.37	0.50
43	BG68	55-95	2	0	2	100.0	0.35	0.46
44	BG69	55-105	2	0	2	100.0	0.33	0.42
45	BG70	145-175	2	0	2	100.0	0.36	0.48
46	BG71	75-900	7	0	7	100.0	**0.73**	0.77
47	BG73	155-370	3	0	3	100.0	**0.54**	0.61
48	BG76	105-600	5	0	5	100.0	**0.70**	0.74
49	BG77	95-500	4	0	4	100.0	**0.58**	0.65
50	BG79	105-900	4	1	3	75.0	0.34	0.37
51	BG82	75-700	6	2	4	66.7	**0.64**	0.70
52	BG83	75-750	5	0	5	100.0	**0.72**	0.76
53	BG85	175-900	4	1	3	75.0	**0.70**	0.75
54	BG88	155-600	3	2	1	33.3	0.40	0.52
55	BG89	75-680	3	2	1	33.3	0.40	0.52
56	BG94	220-800	3	2	1	33.3	0.39	0.51
57	BG95	100-400	3	1	2	66.7	0.14	0.14
58	BG99	120-890	6	0	6	100.0	**0.81**	0.83
59	BG105	400-900	2	0	2	100.0	0.35	0.46
	Total		209	25	184	-	-	-
	Range		2 – 7	0 – 2	1 - 7	33.33 - 100	0.08 - 0.84	0.09 - 0.86
	Average		3.54	0.42	3.12	87.99	0.49	0.56

PIC: Polymorphic Information Content; He: Expected Heterozygosity

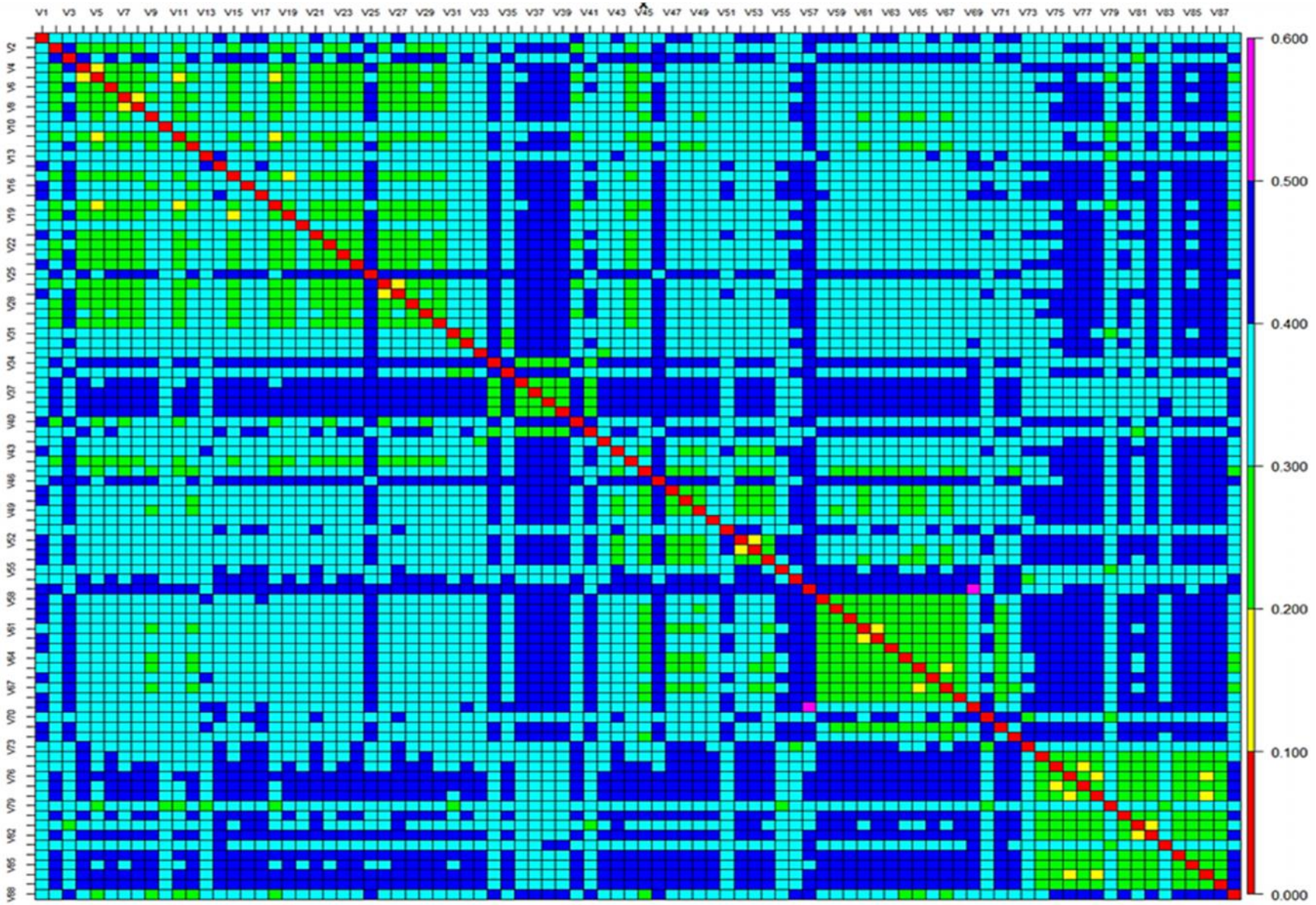

Figure 1: Graphical representation of Jaccard's dissimilarity coefficient matrix for 88 Indian mustard genotypes revealed by 59 polymorphic SSR markers. Legends *v1* to *v88* denotes the genotypes which are listed in Table 1.

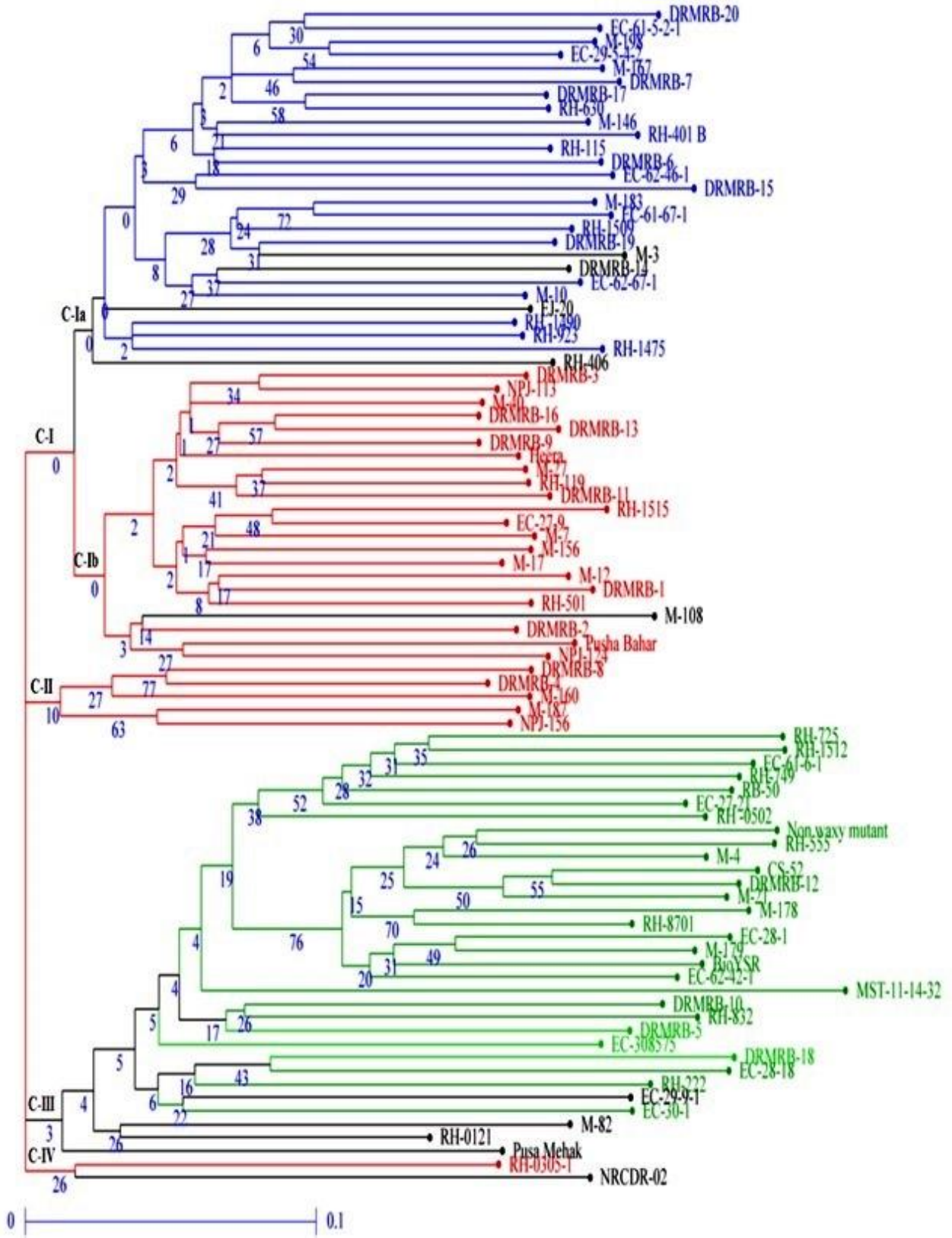

Figure 2: UNJ tree based on Jaccard's dissimilarity coefficient deliberate from the 59 polymorphic SSR markers across the 88 genotypes of Indian mustard. Colors reveal bayesian STRUCTURE subpopulation at K = 3: red for *SP1*, green for *SP2*, blue for *SP3*, and black for admixed.

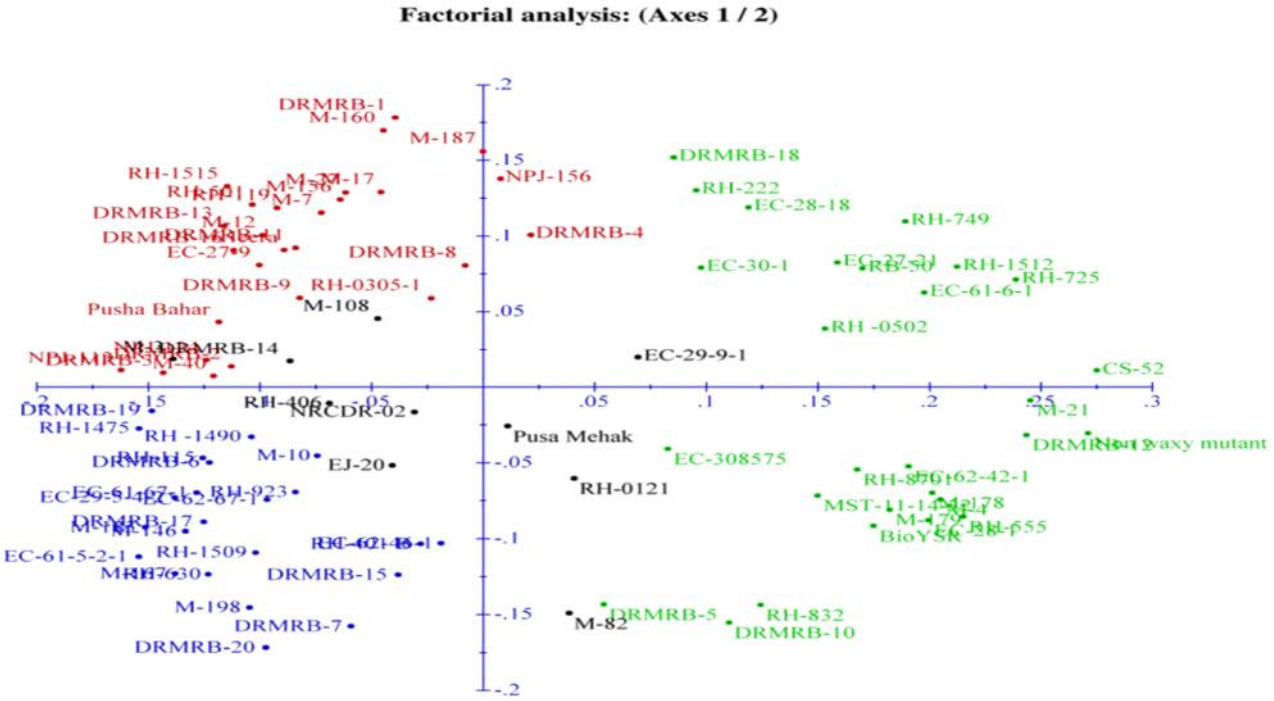

Figure 3: Scatter diagram of 88 genotypes of Indian mustard based on principal coordinates analysis superimposed with clustering. Colors reveal bayesian STRUCTURE subpopulation at K = 3: red for *SP1*, green for *SP2*, blue for *SP3*, and black for admixed.

Table 4: Percentage of variation explained by the first 3 axes in PCoA of 88 genotypes Indian mustard.

Per cent variation	1-axis	2-axis	3-axis
Variation (%)	20.95	10.85	4.53
Cumulative Variation (%)	20.95	31.80	36.33

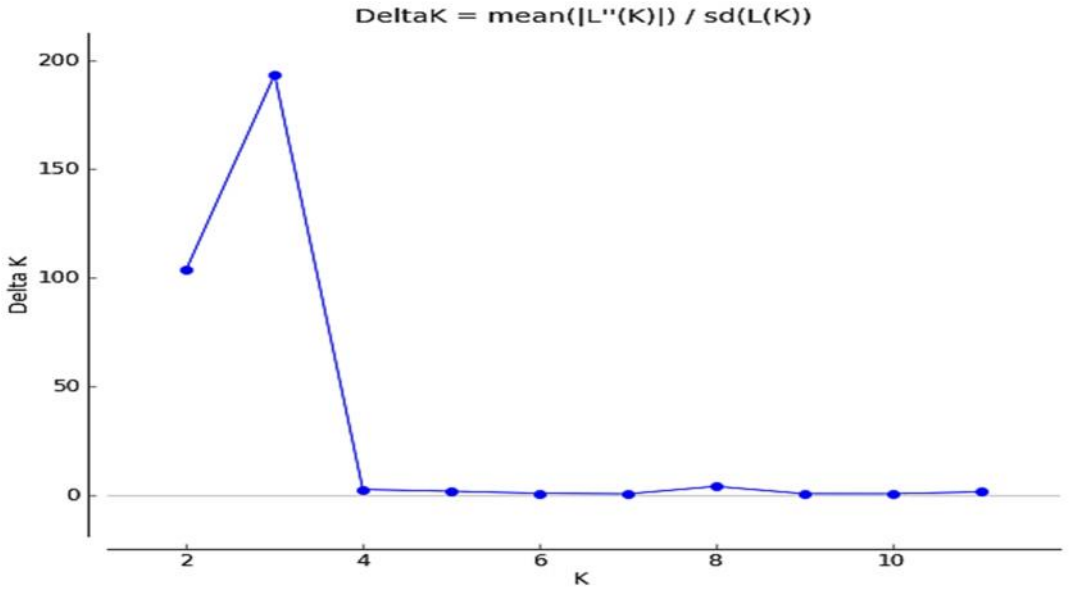

Figure 4: Estimation of populations in 88 genotypes of Indian mustard using LnP(D) derived delK for K from 1 to 12. The maximum value of delK was considered to be the value of K (subpopulations /groups). Results indicate the optimal partition is K=3.

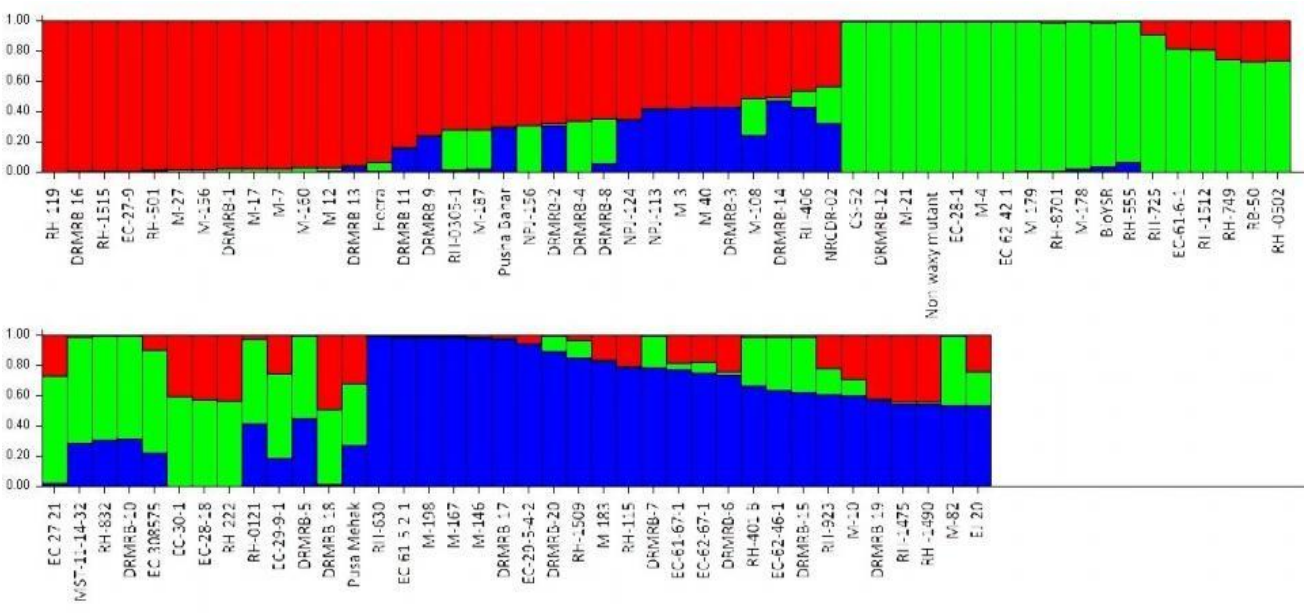

Figure 5: Bayesian STRUCTURE output for the genotypes of Indian mustard at K = 3; Red, Green and Blue colour represents subpopulation *SP1*, *SP2*, and *SP3*, respectively. X-axis denotes name of the genotypes and numbers on y-axis represents genetic proportion in different groups.

DISCUSSION

Assessment of the genetic diversity and population structure of a given crop species offers valuable information which is needed to broaden the narrow genetic base and selection of parental lines for initiation of crop improvement programme. Crossing of genotypes within the cluster will not lead to desirable segregates as compared to genotypes of different cluster sowing to high genetic similarity among the genotypes within a cluster (Kachare *et al.,* 2019). Due to co-dominant nature and high level of reproducibility, SSR markers have been frequently used in genetic diversity studies in Indian mustard (Vinu *et al.,* 2013; Panigrahi *et al.,* 2018). A total of 209 repeatable alleles were intensified by 59 polymorphic SSR primers in 88 genotypes with an average of 3.54 alleles per locus. The number of alleles detected in present study is significantly higher than previous reports (Sudan *et al.,* 2016; Avtar *et al.,* 2016). This difference arisen might be due to different sampling size, different genotypes being used in previous studies and different SSR primers taken for the study. In general, the PIC value point towards efficacy of markers in linkage analysis during inheritance study among parental lines and hybrids, whereas expected heterozygosity suggests average heterozygosity within the population (Osuman *et al.,* 2020). Mean PIC value of 0.49 with a range of 0.08 to 0.84 across 209 polymorphic loci showed nearly complete correspondence with the previous study (Avtar *et al.,* 2016; Singh *et al.,* 2017). It is reported that high diversity of locus is demonstrated in PIC value greater than 0.5 (Botstein *et al.,* 1980). In our study, 28 primers had a PIC value higher than 0.5, suggesting their efficacy in identification of genotypes along with extreme support in detecting the incidence of polymorphism at a particular SSR locus. Markers with high PIC values such as BG 6, BG 32, BG 33 and BG 99 (PIC >0.80) could be effectively used in genetic diversity studies of Indian mustard. The SSR polymorphism perceived in present study is in agreement with prior findings (Gupta *et al.,* 2014; Prajapat *et al.,* 2014). In our study, the average estimates of *He* (0.56) was higher than mean PIC value (0.49) which is closer to our expectations since *He* will always be greater than PIC value. In Indian mustard, limited reports are available on genetic diversity along with high polymorphism using SSR markers. Of 59 polymorphic primers, 40 primers gave 100% polymorphism. The average polymorphic percentage (87.99%) revealed the considerable polymorphism in the studied molecular markers though it was slightly lower than previous study (Avtar *et al.,* 2016 and Singh *et al.,* 2013) where 100 % polymorphism in Indian mustard genotypes has been reported. This difference might be arisen due to different genotypes and markers being used.

Genetic distance offers a measure of the degree of relatedness between genotypes of a species (Garcia *et al.,* 2004) which helps us in genetic improvement of a population (Liu *et al.,* 2019). A distance base tree was generated using UNJ method and Jaccard's dissimilarity coefficient which grouped all genotypes into four clusters. Each cluster contained genotypes from all four locations (except cluster IV). Cluster I was further subdivided into two subgroups C-Ia and C-Ib. Our

findings are in consistent with preceding studies where Thakur *et al.* (2020) also obtained five subgroups while evaluating 78 genotypes of *B. carinata* and Sun *et al.* (2018) where they also grouped 25 wild *B. juncea* populations into two major groups. Clustering of genotypes based on SSR markers revealed that genotypes procured/originated from a particular center did not group together in same cluster, however genotypes from same center strewn into different clusters. This suggested that grouping of genotypes was independent of geographic province. It could be due to the interchange of genotypes from one place to another. Prior reports of various researchers indicate that geographic diversity is not an inevitable criterion of its genetic diversity hence choice of elite lines based on their diversity estimates could be effective in mustard cultivars development programme (Singh *et al.,* 2013; Teklewold and Becker, 2006). Jaccard's dissimilarity coefficients ranged from 0.135 to 0.512 among 88 genotypes of Indian mustard. Based on dissimilarity coefficient, MST-11-14-32 was found most diverse from DRMRB-15, DRMRB-20, M-108 and RH-401 B genotypes. These diverse genotypes can be used effectively in the mustard breeding programme for selection of some desirable recombinants. Results obtained by molecular analyses revealed ample genetic diversity between studied genotypes of Indian mustard. Earlier reports further strengthened the present finding that the SSR markers can be used effectively to estimate genetic distances among Indian mustard genotypes (Abbas *et al.,* 2009). Similar results concerning efficiency of SSR markers regarding genetic diversity for yield and its component traits have been also reported by Sudan *et al.,* 2016; Vinu *et al.,* 2013; Tian *et al.,* 2017.

Next to delK (K = 3), bayesian structure analysis allocated the 88 genotypes of Indian mustard into subpopulations (*SP1, SP2 and SP3*) consisting of few genotypes of pure genetic make-up however others got admixtures of alleles from other subpopulations. This admixture may be the result of random mating involved in development of these breeding lines (Thakur *et al.,* 2020). This is in agreement with earlier finding of Thakur *et al.* (2020) who reported an admixture of alleles in addition to few pure forms in *B. carinata* A. Braun using 212 SSR markers. Likewise, Sun *et al.* (2018) also reported the two lineages among 25 wild *B. juncea* populations using 11 SSR markers with admixture of Northern and Southern lineages in few accessions. Of the few pure form genotypes, we found that CS-52, a salt tolerant variety and RH-119, a thermo-tolerant and high yielding variety recommended for cultivation in Zone-II were found to be 100% pure. The SSR markers clearly assigned genotypes into heterotic groups based on the source populations and individuals of similar genetic background. The outcomes of population structure analysis were further confirmed by the results of UNJ based clustering and PCoA. Parallel results were reported in *B. juncea* genotypes using SSR markers, where structure analysis and UNJ based dendrogram grouped all the 23 genotypes into three sub-populations and three clusters, respectively (Sudan *et al.,* 2016).

REFERENCES

Abbas, S.J., Farhatullah, K.B. Marwat, I. A. Khan and I. Munir (2009). Molecular analysis of genetic diversity in *Brassica* species. Pakistan J. Bot. 41(1):167-176.

Adhikari, S., S. Saha, A. Biswas, T.S. Rana, T.K. Bandyopadhyay and P. Ghosh (2017). Application of molecular markers in plant genome analysis: a review. The Nucleus. 60(3):283-297.

Avtar, R., B. Rani, M. Jattan, Manmohan, N. Kumari and A. Rani (2016). Genetic diversity analysis among elite gene pool of Indian mustard using SSR markers and phenotypic variations. The Bioscan. 11(4):3035-3044.

Bhargava, A. and S. Srivastava (2019). Plant Breeding. In: Participatory plant breeding: concept and applications. Springer, Singapore. pp. 29-68.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Calicioglu, O., A. Flammini, S. Bracco, L. Bellù and R. Sims (2019). The future challenges of food and agriculture: An integrated analysis of trends and solutions. Sustainability. 11(1):222.

Carvalho, Y.G., L.C. Vitorino, U.J. de Souza and L.A. Bessa (2019). Recent trends in research on the genetic diversity of plants: implications for conservation. Diversity. 11(4):62.

Davis, K.F., J.A. Gephart, K.A. Emery, A.M. Leach, J.N. Galloway and P. D'Odorico (2016). Meeting future food demand with current agricultural resources. Global Environ. Chang. 39:125-132.

Earl, D.A. and B.M. VonHoldt (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the evanno method. Conserv. Genet. Resour. 4:359-361.

El-Soda, M., M. Malosetti, B.J. Zwaan, M. Koornneef and M.G. Aarts (2014). Genotype × environment interaction QTL mapping in plants: lessons from Arabidopsis. Trends Plant Sci. 19(6):390-398.

Evanno, G., S. Regnaut and J. Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611-2620.

Friedt, W. and R. Snowdon (2009). Oilseed rape. In *Oil crops*. Springer, New York. Pp. 91-126.

Garcia, A.A.F., L.L. Benchimol, A.M.M. Barbosa, I.O. Geraldi, C.L. Souza and A.P. de Souza (2004). Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. Genet. Mol. Biol. 27(4):579-88.

Goulet, B.E., F. Roda and R. Hopkins (2017). Hybridization in plants: old ideas, new techniques. Plant Physiol. 173(1):65-78.

Gupta, N., S.M. Zargar, M. Gupta and S.K. Gupta (2014). Assessment of genetic variation in Indian mustard (*Brassica juncea* L.) using PCR based markers. Mol. Plant. Breed. 5(3):10-17.

Jat, R.S., V.V. Singh, P. Sharma and P.K. Rai (2019). Oilseed rassica in India: demand, supply, policy perspective and future potential. OCL. 26:8

Kachare, S., S. Tiwari, N. Tripathi and V.V. Thakur (2019). Assessment of genetic diversity of soybean (*Glycine max* L.) genotypes using qualitative traits and microsatellite markers. Agric. Res. 9:23-34.doi: 10.1007/s40003-019-00412-y.

Kaeppler, S. (2012). Heterosis: many genes, many mechanisms-end the search for an undiscovered unifying theory. Intl. Sch. Res. Notices. doi:10.5402/2012/682824.

Lakhanpal, N., D. Verma, R. Kaur and K. Singh (2018). Characterization of cold responsive uncoupling protein1 (UCP1) gene from *Brassica juncea* L. (Czern. and Coss.). J. Plant Biochem. Biot. 27(1):108-117.

Liu, B.H. (1998). Statistical genomics: Linkage, mapping and QTL analysis. CRC Press, Boca Raton, Florida, USA

Liu, F.M., Z. Hong, D.P. Xu, H.Y. Jia, N.N. Zhang, X.J. Liu, Z.J. Yang and M.Z. Lu (2019). Genetic diversity of endangered endemic *Dalbergia odorifera* revealed by SSR markers. Forests. 10:225.

Osuman, A.S., B. Badu-Apraku, B.E. Ifie, P. Tongoona, E. Obeng-Bio and A.L. Garcia-Oliveira (2020). Genetic diversity, population structure and inter-trait relationships of combined heat and drought tolerant early-maturing maize inbred lines from West and Central Africa. Agronomy. 10:1324 DOI: 10.3390/agronomy10091324.

Panigrahi, P., K.K. Panigrahi, and S. Bhattacharya (2018). SSR marker based DNA fingerprinting and diversity studies in mustard (*Brassica juncea*). Electron. J. Plant Breed. 9(1):25-37.

Panjabi, P., S.K. Yadava, N. Kumar, R. Bangkim and N. Ramchiary (2019). Breeding *Brassica juncea* and *B. rapa* for sustainable oilseed production in the changing climate: progress and prospects. In genomic designing of climate-smart oilseed crops. Springer, Cham. Pp. 275-369.

Perrier and Jacquemoud-Collet (2006). DARwin software [Internet]: Available from: http://http//darwin.cirad.fr.

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

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

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen and R.W. Allard (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dymnamics. Proc. Natl. Acad. Sci. 81: 8014-8018.

Singh, K.H., R. Shakya, A.K. Thakur, D.K. Chauhan and J.S. Chauhan (2013). Genetic diversity in Indian mustard [*Brassica juncea* (L.) Czern & Coss.] as revealed by agronomic traits and RAPD markers. Natl. Acad. Sci. Lett. 36(4)419-427.

Singh, S., V.V. Singh, S. Ambawat, M. Dubey and D. Singh (2017). Screening and estimation of allelic differentiation in Indian mustard using SSR markers for background selection. Int. J. Curr. Microbiol. Appl. Sci. 6(9):2506-2516.

Sudan, J., P. Khajuria, S.K. Gupta and R. Singh (2016). Analysis of molecular diversity in Indian and Exotic genotypes of *Brassica juncea* using SSR markers. Indian J. Genet. 76(3):361-364.

Sun, X.Q., Y.Q. Qu, M.M. Li, X.L. Song and Y.Y. Hang (2018). Genetic diversity, genetic structure and migration routes of wild *Brassica juncea* in China assessed by SSR markers. Genet. Resour. Crop Evol. doi: 10.1007/s10722-018-0628-3.

Teklewold, A. and H.C. Becker (2006). Geographic pattern of genetic diversity among 43 Ethiopian mustard (*Brassica carinata* A. Braun) accessions as revealed by RAPD analysis. Genet. Resour. Crop Evol. 53:1173-1185.

Thakur, A.K., K.H. Singh, N. Parmar, D. Sharma, D.C. Mishra, L. Singh, J. Nanjundan and S. Yadav (2020). Population structure and genetic diversity as revealed by SSR markers in Ethiopian mustard (*Brassica carinata* A. Braun): a potential edible and industrially important oilseed crop. Genet. Resour. Crop Evol. doi: 10.1007/s10722-020-00988-3.

Tian, H., J. Yan, S.A. Channa, R, Zhang, Y, Guo and S. Hu (2017). Analysis of the genome genetic diversity among *Brassica napus, B. rapa* and *B. juncea* accessions using specific simple

sequence repeat markers. Pakistan J. Bot. 49(1):125-132.

Tripathi, A.D., R. Mishra, K.K. Maurya, R.B. Singh and D.W. Wilson (2019). Estimates for world population and global food availability for global health. In the role of functional food security in global health. Academic Press. Pp. 3-24.

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

Wani, S.H., A.K. Thakur and Y.J. Khan (2020). Brassica improvement: molecular, genetics and genomic perspectives. Springer Nature.

Yadav, M.S., S. Godika, D.K. Yadava, N. Ahmad, N. Mehta, K. Bhatnagar, V.K. Agrawal, A. Kumar, L. Thomas and C. Chattopadhyay (2019). Prioritizing components of package of integrated pest management in Indian mustard (*Brassica juncea* L.) in India for better economic benefit. Crop Prot.120:21-29.