RESEARCH ARTICLE

Genetic diversity in leafy mustard (*Brassica juncea* var. *rugosa*) as revealed by agro-morphological traits and SSR markers

Deepika Sharma¹ · J. Nanjundan² · Lal Singh³ · S. P. Singh³ · Nehanjali Parmar³ ·
M. S. Sujith Kumar³ · K. H. Singh³ · A. K. Mishra⁴ · Ranbir Singh⁵ ·
Kumar Sambhav Verma¹ · Ajay Kumar Thakur³

Received: 18 May 2020 / Revised: 18 August 2020 / Accepted: 7 September 2020 / Published online: 21 September 2020
© Prof. H.S. Srivastava Foundation for Science and Society 2020

Abstract Leafy mustard (*B. juncea* var. *rugosa*) constitutes an important group of vegetable mustard crops in India and is mainly cultivated in home-backyard and hilly regions of Uttarakhand and some North-eastern states. In the present study, various agro-morphological traits, physiological and biochemical traits along with SSR markers were used for genetic diversity evaluation in a germplasm collection of leafy mustard. This study revealed a significant variation among 59 accessions of leafy mustard in both qualitative and quantitative agro-morphological traits indicating the accessions' promising potential for consumption purpose and for use in breeding programs. Maximum variability was recorded for leaf area elongation rate (CV = 53.12%), followed by total plant weight (TPW) (CV = 50.63%) and seed yield per plant (CV = 44.33%). In molecular analysis, 155 SSRs evaluated resulted in 482 alleles and the number of alleles varied from 1 to 8 with an average of 3.11 alleles per marker. A total of 122 (78.70%) SSRs resulted into polymorphic amplicons. PIC value varied from 0.32 to 0.77 with an average value of 0.44 per SSR locus. The unweighted neighbour-joining-based

dendrogram analysis divided all the 59 accessions into two major groups on the basis of both agro-morphological traits and SSR markers, whereas, three subpopulations/subgroups were predicted by population STRUCTURE analysis. AMOVA indicated the presence of more variability within population than among population. Overall, agro-morphologically better performing and genetically diverse genotypes have been identified which could be further used as donors for leafy mustard improvement programs.

Keywords *Brassica juncea* var. *rugosa* · Genetic diversity · Agromorphological traits · SSR markers

Introduction

Brassica juncea var. *rugosa*, also known as a leafy mustard/laipatta belongs to family *Brassicaceae* and is a tall, slow-growing mustard with lust green foliage varying from light green to deep purple in color (Pant et al. 2020). It is a native of central and Eastern Asia and is consumed as a green leafy vegetable around the world from China to South America. Asian countries including India, China and Japan are the top growers and suppliers of leafy mustard. In India, it is cultivated in small patches in home backyard, cultivated land and hilly regions of Uttarakhand and some of the Northeastern states including Arunachal Pradesh, Nagaland, Meghalaya, Mizoram and Manipur. Peppery, crispy leafy mustard has broad and soft leaves with high moisture content and thick tender stem, which can conveniently be used for saag preparation (Rauniyar and Bhat-tarai 2017). The green leaves can be eaten raw in salad form or cooked. They are enriched with a number of phytonutrients including vitamin A, B, C, E; also contains iron, calcium and protein in large quantities, and have

✉ Ajay Kumar Thakur
thakurak2010@gmail.com

¹ Amity Institute of Biotechnology, Amity University Rajasthan, Jaipur 302 006, India

² ICAR-Indian Agricultural Research Institute-Regional Station, Wellington, Tamilnadu 643 231, India

³ ICAR-Directorate of Rapeseed-Mustard Research, Bharatpur, Rajasthan 321 303, India

⁴ ICAR-National Bureau of Plant Genetic Resources Regional Station, Shillong, Meghalaya 793 103, India

⁵ ICAR-National Bureau of Plant Genetic Resources, New Delhi 110 012, India

health promotional and disease preventive properties. Regular consumption of leafy mustard as vegetable in diet helps to protect the consumers from iron deficiency, osteoporosis and various cardiovascular diseases. They play important roles in combating arthritis, asthma and nervous system disorders as well (Macready et al. 2014). Its leaves are used in a range of folk medicines as diuretics, stimulants and expectorants to cure various diseases. Leafy mustard is also consumed in the form of a fermented pickle product locally known as ‘Gundruk’, which is a very famous and nationally popular Nepali dish (Bhattarai et al. 2018).

Genetic diversity evaluation is an indispensable component of breeding programs for an effective and efficient management and utilization of plant genetic resources. At present, very scarce or little information is available about the genetic diversity present in leafy mustard as such studies have been ignored in this important vegetable crop, may be due to the lesser acreage or production. However, there is an urgent need to characterize the germplasm collections of leafy mustard for making their efficient utilization in breeding programs aimed at leafy mustard improvement. Evaluation of genetic diversity using both morphological traits and SSR markers is a very effective method for germplasm conservation and management (Ghaffari et al. 2014). Molecular markers are more stable, more in number, easily detectable in all the tissues and are free from the environmental influences (Yasin and Mayadevi 2010). The DNA-based markers have emerged as powerful tools for evaluating genetic diversity and determining genetic relationships in crop germplasm and cultivars. Molecular markers such as random amplified polymorphic DNA (RAPD) and inter simple sequence repeat (ISSR) suffer from the drawback of reproducibility, while other type of markers such as restriction fragment length polymorphism (RFLP) and amplified fragment length polymorphism (AFLP) are very much skill and cost demanding, and also time consuming. Among various type of molecular markers, SSR markers are the mostly preferred markers for evaluation of genetic diversity due to their co-dominant and multi-allelic inheritance, high reproducibility, genome coverage, abundance and easy scorability (Vieira et al. 2016).

The present research was aimed to evaluate genetic diversity and establish genetic inter-relationships between leafy mustard accessions in India using agro-morphological traits, physiological & biochemical traits, and SSR markers so that suitable donors can be identified for breeding programs for leafy mustard genetic improvement.

Materials and methods

Plant material

The plant material in this study comprised of 59 *B. juncea* var. *rugosa* accessions collected from Arunachal Pradesh, India during the year 2012–2013 (Table 1). All the 59 accessions of leafy mustard were grown at DRMR, Bharatpur field in augmented block design following the standard package of practices in 3 m row at a spacing of 45 cm from row-to-row and 10 cm from plant-to-plant for two consecutive crop seasons (Fig. 1). Young and tender leaves from five field grown plants for each genotype were collected, pooled and stored in -80°C deep freezer until further use.

Agro-morphological trait evaluation

All the 59 accessions of leafy mustard were evaluated for 17 agro-morphological traits including time of flowering initiation (IF, 50%), days to maturity (DM), plant height (PH, cm), fruiting zone length (FZL, cm), primary branches (PB), secondary branches (SB), main shoot length (MSL, cm), siliquae on main shoot (SMS), siliquae density on main shoot (SD), siliqua length (SL, cm), siliqua beak length (SBL, cm), seeds per siliqua (SS), 1000-seed weight (SW, g), total plant weight (TPW, g), seed yield per plant (SYP, g), harvest index (HI) and oil content (OC, %) following DUS guidelines (Singh et al. 2006). These germplasm accessions were evaluated for 6 physiological traits including leaf area elongation rate (LER, cm^2/day), leaf growth rate (LGR, $\text{cm}^2/\text{cm}^2/\text{day}$), absolute growth rate (AGR, $\text{g}/\text{day}/\text{plant}$), relative growth rate (RGR, $\text{g}/\text{g}/\text{day}$), net assimilation rate (NAR, $\text{mg}/\text{cm}^2/\text{day}$) and leaf area ratio (LAR, cm^2/g), and 2 biochemical traits as per the already available protocols (Hansen and Møller 1975 for total soluble sugar and Prieto et al. 1999 for total antioxidant capacity, respectively, Table 2).

Genomic-DNA extraction and purification

Genomic-DNA from the pooled leaf samples was extracted and purified as per the already standardized protocol in our laboratory (Thakur et al. 2013). DNA quantification was carried out on 0.8% agarose gel and its concentration was measured in a spectrophotometer. DNA was stored in TE buffer at 4°C till further use.

SSR analysis

A total of 155 Brassica species derived SSR markers as identified by Thakur et al. (2018) have been used in the

Table 1 Various germplasm accessions of *B. juncea* var. *rugosa* used in the present study

Code	Collector No.	IC No.	Vernacular Name	Nature of plant material (cultivar/wild)	Site of collection (District of Arunachal Pradesh)	Location (longitude/latitude/height)
1	RAN-01	IC597866	Ogiyi	Landrace	West Siang	N27° 57.889'/E 94° 43.021'/830 M
2	RAN-02	IC597867	Ogiyi	Landrace	West Siang	N27° 56.726'/E 94° 44.630'/747 M
3	RAN-03	IC597868	Ogiyi	Landrace	West Siang	N27° 57.867'/E 94° 43.043'/747 M
4	RAN-04	IC597869	Ogiyi	Landrace	West Siang	N27° 57.813'/E 94° 43.093'/817 M
5	RAN-05	IC597870	Ogiyi	Landrace	West Siang	N27° 57.813'/E 94° 43.093'/817 M
6	RAN-07	IC597872	Lai Patta	Landrace	West Siang	N28° 04.718'/E 94° 45.404'/417 M
7	RAN-08	IC597873	Ogiyi	Landrace	West Siang	N28° 04.891'/E 94° 46.437'/436 M
8	RAN-10	IC597875	Adi Ogiyi	Landrace	West Siang	N28° 10.562'/E 94° 48.117'/246 M
9	RAN-11	IC597876	Lai Patta	Landrace	West Siang	N28° 10.562'/E 94° 48.117'/246 M
10	RAN-13	IC597878	Lai Patta	Landrace	West Siang	N28° 10.562'/E 94° 48.117'/246 M
11	RAN-14	IC597879	Lai Patta	Landrace	West Siang	N28° 10.562'/E 94° 48.117'/246 M
12	RAN-15	IC597880	Adi Ogiyi	Landrace	West Siang	N28° 11.372'/E 94° 45.917'/273 M
13	RAN-16	IC597881	Tuk	Landrace	West Siang	N28° 17.816'/E 94° 40.816'/299 M
14	RAN-17	IC597882	Ogiyi	Landrace	West Siang	N28° 17.816'/E 94° 40.816'/299 M
15	RAN-18	IC597883	Adi Ogiyi	Landrace	West Siang	N28° 20.488'/E 94° 40.661'/332 M
16	RAN-19	IC597884	Puthu Ogiyi	Landrace	West Siang	N28° 20.488'/E 94° 40.661'/332 M
17	RAN-20	IC597885	Tuiling	Landrace	West Siang	N28° 25.243'/E 94° 40.661'/550 M
18	RAN-22	IC597887	Ogiyi Yit	Landrace	West Siang	N28° 01.213'/E 94° 40.833'/539 M
19	RAN-23	IC597888	Lai Patta	Landrace	West Siang	N28° 58.829'/E 94° 41.506'/539 M
20	RAN-24	IC597889	Lai Patta	Landrace	West Siang	N27° 55.880'/E 94° 42.359'/1004 M
21	RAN-27	IC597892	Lai Patta	Landrace	West Siang	N27° 49.125'/E 94° 43.157'/674 M
22	RAN-28	IC597893	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
23	RAN-29	IC597894	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
24	RAN-30	IC597895	Lai Patta	Landrace	East Siang	N27° 55.350'/E 95° 20.684'/140 M
25	RAN-38	IC597903	Lai Patta	Landrace	East Siang	N27° 53.039'/E 95° 18.920'/132 M
26	RAN-39	IC597904	Lai Patta	Landrace	East Siang	N27° 50.477'/E 95° 13.303'/137 M
27	RAN-40	IC597905	Lai Patta	Landrace	East Siang	N28° 04.543'/E 95° 19.948'/148 M

Table 1 continued

Code	Collector No.	IC No.	Vernacular Name	Nature of plant material (cultivar/wild)	Site of collection (District of Arunachal Pradesh)	Location (longitude/latitude/height)
28	RAN-42	IC597907	Lai Patta	Landrace	East Siang	N27° 51.020'/E 95° 21.860'/116 M
29	RAN-45	IC597910	Lai Patta	Landrace	East Siang	N27° 51.532'/E 95° 21.989'/122 M
30	RAN-49	IC597914	Lai Patta	Landrace	East Siang	N27° 54.810'/E 95° 20.636'/130 M
31	RAN-52	IC597917	Tusut	Landrace	East Siang	N28° 04.882'/E 95° 26.464'/154 M
32	RAN-53	IC597918	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
33	RAN-56	IC597921	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
34	RAN-57	IC597922	Lai Patta	Landrace	East Siang	N27° 50.477'/E 95° 13.303'/137 M
35	RAN-58	IC597932	Lai Patta	Landrace	East Siang	N27° 50.477'/E 95° 13.303'/137 M
36	RAN-59	IC597924	Lai Patta	Landrace	West Siang	N27° 41.402'/E 94° 49.351'/146 M
37	RAN-60	IC597925	Lai Patta	Landrace	West Siang	N27° 41.402'/E 94° 49.351'/146 M
38	RAN-62	IC597927	Lai Patta	Landrace	West Siang	N27° 41.402'/E 94° 49.351'/146 M
39	RAN-63	IC597928	Lai Patta	Landrace	West Siang	N27° 41.402'/E 94° 49.351'/146 M
40	RAN-64	IC597929	Lai Patta	Landrace	West Siang	N27° 41.402'/E 94° 49.351'/146 M
41	RAN-66	IC597931	Lai Patta	Landrace	L Subansiri	N27° 22.352'/E 93° 45.889'/627 M
42	RAN-69	IC597934	Lai Patta	Landrace	L Subansiri	N27° 32.361'/E 93° 48.911'/1511 M
43	RAN-71	IC597936	Lai Patta	Landrace	L Subansiri	N27° 30.923'/E 93° 50.364'/1511 M
44	RAN-73	IC597938	Lai Patta	Landrace	Papum Pare	N27° 04.595'/E 93° 33.732'/303 M
45	RAN-74	IC597939	Lai Patta	Landrace	Papum Pare	N27° 04.595'/E 93° 33.732'/303 M
46	RAN-78	IC597943	Lai Patta	Landrace	Papum Pare	N27° 06.423'/E 93° 41.674'/161 M
47	RAN-79	IC597944	Lai Patta	Landrace	Papum Pare	N27° 06.423'/E 93° 41.674'/161 M
48	RAN-80	IC597945	Lai Patta	Landrace	Papum Pare	N27° 06.423'/E 93° 41.674'/161 M
49	RAN-81	IC597946	Lai Patta	Landrace	Papum Pare	N27° 06.574'/E 93° 41.252'/165 M
50	RAN-82	IC597947	Lai Patta	Landrace	Papum Pare	N27° 07.346'/E 93° 42.127'/144 M
51	RAN-83	IC597948	Lai Patta	Landrace	Papum Pare	N27° 08.505'/E 93° 42.818'/149 M
52	RAN-21	IC597886	Lai Patta	Landrace	West Siang	N28° 25.243'/E 94° 40.661'/550 M
53	RAN-36	IC597901	Lai Patta	Landrace	East Siang	N27° 55.037'/E 95° 20.757'/137 M
54	RAN-46	IC597911	Lai Patta	Landrace	East Siang	N27° 55.455'/E 95° 22.192'/120 M

Table 1 continued

Code	Collector No.	IC No.	Vernacular Name	Nature of plant material (cultivar/wild)	Site of collection (District of Arunachal Pradesh)	Location (longitude/latitude/height)
55	RAN-54	IC597919	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
56	RAN-55	IC597920	Lai Patta	Landrace	East Siang	N28° 03.604'/E 95° 19.852'/150 M
57	RAN-77	IC597942	Lai Patta	Landrace	Papum Pare	N27° 06.423'/E 93° 41.674'/161 M
58	RAN-84	IC597949	Lai Patta	Landrace	Papum Pare	N27° 09.868'/E 93° 45.387'/144 M
59	RAN-85	IC597950	Lai Patta	Landrace	Papum Pare	N27° 09.868'/E 93° 45.387'/144 M

Fig. 1 Some representative accessions of leafy mustard; **a** IC597866, **b** IC597870, **c** IC597884, **d** IC597894, **e** IC597907, **f** IC597931

present study to investigate DNA polymorphisms. For PCR analysis, a reaction volume of 25 μ l was prepared which contained 50 ng genomic-DNA, 1XPCR buffer, 0.2 mM of each dNTP, 2.0 mM $MgCl_2$, 1.0 U *Taq* DNA polymerase (GCC Biotech, India) and 400 nM primers using Verity

96-w PCR machine (Invitrogen, USA). The first amplification cycle consisted of DNA denaturation at 94 °C temperature for 5 min, followed by a series of 45 cycles each of DNA denaturation at 94 °C for 30 s, primer annealing at 50–60 °C (depending upon the primer

Table 2 Range, mean and coefficient of variation for agro-morphological traits, physiological and biochemical traits among 59 germplasm accessions of leafy mustard

Sr. no.	Trait	Range		Mean	Standard deviation (SD)	Coefficient of variation (%)
		Minimum	Maximum			
A) Agro-morphological traits						
1	Time of flowering initiation (IF, 50%)	38	95	53.37	10.55	19.76
2	Days to maturity (DM)	136	162	142.3	5.16	3.62
3	Plant height (PH, cm)	144	227.5	181.2	19.78	10.92
4	Fruiting zone length (FZL, cm)	35	105	75.57	10.74	14.21
5	Primary branches (PB)	4	12.8	7.01	1.79	25.58
6	Secondary branches (SB)	5.6	28.2	13.48	4.46	33.08
7	Main shoot length (MSL, cm)	25	81	58.93	11.75	19.94
8	Siliquae on main shoot (SMS)	25.4	67.8	51.09	9.23	18.07
9	Siliquae density on main shoot (SD)	0.57	1.32	0.88	0.17	19.81
10	Siliqua length (SL, cm)	2.13	5.06	3.19	0.52	16.16
11	Siliqua beak length (SBL, cm)	0.27	0.64	0.45	0.07	15.63
12	Seeds per siliqua (SS)	7.1	18.4	12.88	2.01	15.62
13	1000-seed weight (SW, g)	0.99	4.53	1.65	0.45	27.6
14	Total plant weight (TPW, g)	17	200	50.03	25.33	50.63
15	Seed yield per plant (SYP, g)	1.52	20.02	8.62	3.82	44.33
16	Harvest index (HI)	5.2	26.35	17.59	5.19	29.56
17	Oil content (OC, %)	39.21	42.89	41.07	0.65	1.59
B) Physiological traits						
1	Leaf area elongation rate (LER, cm ² /day)	35.6	404.5	119.5	63.48	53.12
2	Leaf growth rate (LGR, cm ² /cm ² /day)	0.026	0.12	0.06	0.02	32.3
3	Absolute growth rate (AGR, g/day/plant)	0.12	1.09	0.48	0.21	43.1
4	Relative growth rate (RGR, g/g/day)	0.02	0.11	0.067	0.02	24.5
5	Net assimilation rate (NAR, mg/cm ² /day)	0.089	0.48	0.257	0.10	37.3
6	Leaf area ratio (LAR, cm ² /g)	166.3	476.2	278.7	73.2	26.3
C) Biochemical traits						
1	Total soluble sugar (TSS, mg/l)	0.001	0.225	0.04	0.05	1.14
2	Total antioxidant activity (TAA, mg/l)	0.004	0.088	0.02	0.02	0.75

combination) for 30 s, primer extension at 72 °C for 45 s and a final extension step at 72 °C or 7 min. PCR amplified products were run in a 3.5% Super Fine Resolution (SFR) agarose (Amresco, USA) at 100 V for their electrophoretic separation. A 50 bp DNA ladder was also run alongside as a reference mark for tracing the movement and size measurement. Further, the gels were visualized in a gel documentation unit (Syngene Gel Doc, Synoptic Ltd. UK).

Data analysis

Mean values of each of various agro-morphological traits, physiological and biochemical traits of 2 years observations were computed and further subjected to statistical analysis. Analysis of variance was calculated using the Fit Model of SAS JMP9.2. Dendrogram based on the

similarity coefficients on the basis of Euclidean distances was generated by Unweighted Neighbour-Joining method using DARwin 5 software (Perrier and Jacquemoud-Collet 2006).

The SSR marker data was scored in a binary format assigning '1' for the presence and '0' for the absence of a particular allele. The resulting binary matrix was subjected to Power Marker software v3.25 (Liu and Muse 2005) for computation of major allele frequency, PIC value and gene diversity for each individual SSR marker. UNJ-based dendrogram was constructed to demonstrate the genetic relationship among different accessions under investigation.

Population structure analysis

Population structure analysis was carried out to unravel the number of subgroups/subpopulations in the reference panel using STRUCTURE v. 2.3.4 software (Pritchard and Wen 2003) using admixture model with $K = 2$ –10. Three independent runs were carried out for each fixed K value with each run for 30,000 burn in period and 1,00,000 number of Markov Chain Monte Carlo (MCMC) repeats. The optimal value of K was determined by examining $\Delta L(K)$ statistic and $L(K)$ (Evanno et al. 2005) using Structure Harvester software (Earl and Von Holdt 2012). Analysis of molecular variance (AMOVA) was computed using GenALex6.5 software (Peakall and Smouse 2012).

Results and discussion

Genetic diversity based on agro-morphological traits

The descriptive statistics of each of the traits including mean, minimum, maximum, standard deviation and coefficient of variation (CV) obtained on the basis of averages of the data are summarized in Table 2. Analysis of variance demonstrated the presence of significant variations for various morpho-physio-biochemical traits except for days to maturity ($CV = 3.62\%$), oil content ($CV = 1.59$) and antioxidant activity ($CV = 0.75\%$) (Table 2). Maximum variability was recorded for leaf area elongation rate ($CV = 53.12\%$), followed by total plant weight (TPW) ($CV = 50.63\%$) and seed yield per plant (SYP) ($CV = 44.33\%$). Genotype 38 (IC597901) was found to induce flowering at the earliest i.e. in 38 days after sowing, whereas accession 1 (IC597866) started flowering at the last after 95 days of sowing. It is very interesting to know that all the genotypes were having almost similar oil content (%) in the range of 39.21 to 42.89%. Leafy mustard accessions thus possess almost the same quantity of oil as that of Indian mustard genotypes as reported by Singh et al. (2015). A high level of variation for 22 morpho-physio-biochemical traits existed in the 59 germplasm accessions which indicated the potential of these accessions for future breeding. Generally, most of the leafy mustard accessions bear small sized seeds, however, in this study, accession No. 52 (IC597886) was found to possess the bold seed size having the highest 1000-seed weight i.e. 4.53 g. Such genotypes may be employed as donors for improving seed size of leafy mustard accessions. Genotype 13 (IC597881) was found to have the highest leaf area elongation rate i.e. 404.5 cm^2/day . Such genotypes will highly suit the leafy mustard market as they have the ability to produce larger biomass per unit time and thus will fulfill the requirement of leafy vegetable.

Antioxidants play a major role in preventing various cancer related diseases by scavenging free radicals produced in the body as a result of various stresses. Pant et al. (2020) reported a very high antioxidative activity (20.09 mg/g) in one accession, EEC-5 of leafy mustard. Whereas, in our studies, genotype 37 (IC597925) exhibited the highest 0.088 mg/g AAE total antioxidant activity. UNJ-based dendrogram divided all the 59 accessions into two major groups, I and II (Fig. 2). Group I comprised of only two genotypes namely 13 (IC597881) and 15 (IC597883). Group II had a total of 57 accessions, which were further divided into two subgroups, IIA and IIB. Subgroup IIA was having only three accessions namely, 19 (IC597888), 12 (IC597880) and 21 (IC597892). The rest 54 leafy mustard accessions were grouped into subcluster IIB.

Genetic diversity based on SSR markers

The genetic diversity as observed by the morphological traits do not reflect the entire diversity inherent in the germplasm lines as the expression of many morphological traits is masked by different environmental factors. The genotypes having almost the same morphological features may show differences at DNA level. So, in the present study, SSR markers were employed as complementary tool for the analysis of genetic diversity in Indian mustard genotypes in addition to agro-morphological traits. Due to the lack of information on the application of SSR markers in leafy mustard, the SSR primer pairs used in this study were synthesized according to the reports of Thakur et al. (2018). They could identify a set of *Brassica species*-derived SSR markers in *B. carinata* using cross-transferability approach. Morphological markers and molecular markers such as RAPD and SSR have been used to evaluate genetic diversity in *B. juncea* (Singh et al. 2013; Vinu et al. 2013) and *B. rapa* (Thakur et al. 2017). However, to date there is no systematic record of genetic background, phylogenetic relationships and source of diversity for leafy mustard (*B. juncea* var. *rugosa*) accessions, which has restricted the optimal application of leafy mustard germplasm resources for crop improvement. In the present study, 155 SSR markers were evaluated for genetic diversity analysis in 59 accessions of leafy mustard which resulted into a total of 482 alleles and the number of alleles varied from 1–8 with an average of 3.11 alleles per marker. BrgMS234 marker amplified the highest number of alleles (8). A total of 122 (78.70%) SSRs resulted into polymorphic amplicons. Maximum allelic frequency (72%) was observed for Ra2-A01 and EJU4 marker showed minimum allele frequency (20%). PIC value, which indicates the informativeness of a marker, varied from 0.32 (cnu_m596a) to 0.77 (EJU4) with an average value of 0.44 per SSR locus (Table 3). A comparatively higher average

Fig. 2 UNJ-dendrogram showing genetic relationship among various leafy mustard accessions based on agro-morphological traits

PIC value (0.90) along with higher average number of alleles per marker (7.2) was reported by Yao et al. (2012) when they deciphered genetic diversity in 34 Chinese vegetable mustard landraces including root mustard, stem mustard, leaf mustard and seed stalk mustard using 69 SSR loci. In a similar study, Fang et al. (2013) carried out genetic diversity analysis in 133 tuber mustard (*B. juncea* var. *tumida*) cultivars using 81 SSR primer pairs and reported allele number ranging from 5 to 16 alleles with 10.0 average number of alleles per locus. In the present study, 48 SSR markers were reported to have PIC values higher than the average PIC value (0.44) in this reference set of germplasm, which infer that these 48 SSRs are highly

polymorphic in nature and can be used in trait mapping studies in leafy mustard. Gene diversity values ranged from 0.20 (Ni2F11) to 0.70 (BrGMS713) with average value of 0.53 per marker. The presence of low average PIC value in *B. juncea* var. *rugosa* accessions in the present investigation inferred the presence of low genetic diversity. The presence of narrow genetic diversity in leafy mustard may be attributed to the existence of same germplasm line with a different name or a dissimilar germplasm with the same name in the leafy mustard growing regions, particularly four districts of Arunachal Pradesh from where these germplasm collections were made. It has been found that most of the leafy mustard collections have been made from

Table 3 Various polymorphism parameters of SSR markers in reference set of *B. juncea* var. *rugosa*

S. no.	Marker ID	No. of alleles	Major allele frequency	PIC value	Gene diversity
1	SA0306	4	0.39	0.60	0.66
2	SJ3874I	4	0.38	0.62	0.68
3	SJ0338	2	0.50	0.37	0.50
4	SJ39119I	3	0.57	0.42	0.51
5	SJ3838	3	0.62	0.38	0.47
6	SB1752	4	0.36	0.62	0.67
7	SB3140	3	0.44	0.51	0.59
8	SJ6846	5	0.70	0.37	0.43
9	SJ3627R	3	0.64	0.35	0.45
10	SB2556	4	0.66	0.39	0.46
11	SJ1505	5	0.52	0.55	0.61
12	SB1937	2	0.50	0.37	0.50
13	SJ3412I	7	0.71	0.41	0.44
14	SB3751	2	0.50	0.37	0.50
15	SJ7079	2	0.50	0.37	0.50
16	SB0372	2	0.51	0.37	0.49
17	SJ1536	4	0.64	0.41	0.48
18	SB4817R	2	0.50	0.37	0.50
19	SB1935A	2	0.50	0.37	0.50
20	SJ7104	2	0.50	0.37	0.50
21	SJ13133	3	0.65	0.36	0.46
22	SB1728	2	0.52	0.36	0.48
23	SB5162	2	0.50	0.37	0.50
24	Ni2B03	2	0.51	0.37	0.49
25	Ni2B07	3	0.64	0.36	0.45
26	Ni4F09	3	0.45	0.51	0.58
27	Ni2D03	7	0.67	0.44	0.48
28	Ni2E05	6	0.68	0.42	0.47
29	Ni2F11	2	0.50	0.37	0.20
30	Ni3C08	7	0.69	0.43	0.46
31	Ni4A09	2	0.50	0.37	0.50
32	Ni4C02	5	0.55	0.54	0.59
33	Ni4C06	3	0.65	0.36	0.45
34	Ni4C09	4	0.47	0.57	0.64
35	Ni4C11	4	0.36	0.64	0.69
36	Ni4F11	6	0.64	0.47	0.52
37	Ni4G10	6	0.74	0.35	0.39
38	Ni2A12	2	0.51	0.37	0.49
39	nia_m066a	3	0.65	0.36	0.46
40	nia_m091a	2	0.50	0.37	0.50
41	BrgMS801	4	0.67	0.38	0.45
42	BrgMS1237	5	0.62	0.45	0.51
43	BrgMS782	2	0.50	0.37	0.50
44	BrgMS412	4	0.66	0.38	0.45
45	BrgMS388	3	0.57	0.41	0.49
46	BrgMS799	3	0.34	0.59	0.67
47	BrgMS590	3	0.59	0.40	0.49
48	BrgMS787	5	0.54	0.55	0.60

Table 3 continued

S. no.	Marker ID	No. of alleles	Major allele frequency	PIC value	Gene diversity
49	BrgMS778	3	0.59	0.40	0.49
50	BrgMS344	3	0.47	0.49	0.57
51	BrgMS4533	3	0.41	0.54	0.62
52	BrgMS780	2	0.50	0.37	0.50
53	BrgMS961	2	0.50	0.37	0.50
54	BrgMS234	8	0.59	0.53	0.57
55	BrgMS334	4	0.40	0.61	0.66
56	BrgMS4536	2	0.50	0.37	0.50
57	BrgMS732	3	0.38	0.55	0.63
58	BrgMS372	4	0.48	0.53	0.59
59	BrgMS430	4	0.65	0.39	0.47
60	BrgMS1238	3	0.34	0.58	0.65
61	BrgMS710	4	0.69	0.36	0.43
62	BrgMS429	2	0.51	0.37	0.49
63	BrgMS397	4	0.66	0.39	0.47
64	BrgMS751	2	0.50	0.37	0.50
65	BrgMS75	3	0.52	0.45	0.54
66	BrgMS421	2	0.50	0.37	0.50
67	BrgMS638	2	0.51	0.37	0.50
68	BrgMS68	2	0.67	0.34	0.44
69	BrgMS338	3	0.66	0.35	0.44
70	BrgMS4508	2	0.50	0.37	0.50
71	BrgMS794	4	0.63	0.42	0.48
72	BrgMS713	4	0.34	0.65	0.70
73	BrgMS804	2	0.50	0.37	0.50
74	BrgMS422	2	0.50	0.37	0.50
75	BrgMS4497	3	0.34	0.58	0.66
76	BrgMS426	2	0.50	0.37	0.50
77	BrgMS724	2	0.50	0.37	0.50
78	BrgMS135	2	0.50	0.37	0.50
79	BrgMS399	2	0.50	0.37	0.50
80	BrgMS802	2	0.50	0.37	0.50
81	BrgMS359	2	0.50	0.37	0.50
82	BRMS-003	3	0.47	0.49	0.57
83	BRMS-029	3	0.47	0.48	0.56
84	BRMS-005	6	0.71	0.39	0.44
85	BRMS-006	2	0.50	0.37	0.50
86	BRMS-011	6	0.67	0.44	0.48
87	BRMS-015	2	0.50	0.37	0.50
88	BRMS-017	4	0.66	0.38	0.45
89	BRMS-007	6	0.63	0.48	0.53
90	Ra1-F06	2	0.50	0.37	0.50
91	Ra2-A01	5	0.72	0.36	0.42
92	Ra2-A11	4	0.49	0.56	0.63
93	Ra2-D04	2	0.50	0.37	0.50
94	Ra2-E04	4	0.39	0.61	0.66
95	Ra2-E11	5	0.43	0.65	0.69
96	Ra2-E12	7	0.55	0.58	0.62

Table 3 continued

S. no.	Marker ID	No. of alleles	Major allele frequency	PIC value	Gene diversity
97	Ra2-F11	7	0.59	0.53	0.57
98	Ra2-G08	5	0.42	0.61	0.66
99	Ra3-H09	3	0.67	0.34	0.44
100	Ra3-H10	3	0.41	0.53	0.61
101	cnu_m584a	2	0.52	0.36	0.48
102	cnu_m587a	3	0.54	0.44	0.53
103	cnu_m594a	2	0.50	0.37	0.50
104	cnu_m596a	2	0.57	0.32	0.43
105	cnu_m597a	5	0.67	0.41	0.47
106	cnu_m600a	2	0.50	0.37	0.50
107	cnu_m604a	6	0.51	0.61	0.66
108	cnu_m626a	7	0.70	0.41	0.45
109	PW186	4	0.55	0.49	0.57
110	E120	3	0.34	0.59	0.66
111	E129	3	0.34	0.58	0.66
112	EJU1	5	0.46	0.62	0.66
113	EJU3	3	0.43	0.52	0.59
114	EJU4	5	0.20	0.77	0.80
115	EJU5	3	0.37	0.56	0.64
116	ENA15	3	0.67	0.34	0.44
117	ENA2	3	0.34	0.59	0.67
118	ENA20	3	0.53	0.44	0.53
119	ENA9	3	0.35	0.58	0.65
120	MB4	2	0.51	0.37	0.49
121	Ol10B11	6	0.69	0.42	0.46
122	sORA43	4	0.73	0.32	0.39
Mean			0.53	0.44	0.53

only one state. Genetic diversity was also evaluated among a panel of 15 Chinese vegetable mustard comprising of root mustard, stem mustard, leaf mustard and seed stalk mustard and one seed mustard using 14 pairs of AFLP markers (Qi et al. 2008). Wu et al. (2009) evaluated genetic diversity patterns in oil and vegetable mustard by SRAP markers and could conclude that remarkably great differences existed between both type of mustard.

UNJ-based dendrogram also grouped all the 59 leafy mustard accessions into two distinct groups (Fig. 3). Genotype IC597868 (4) did not show any resemblance with any other genotype at genomic level and formed a separate cluster i.e. cluster I. Cluster II was comprising of 58 genotypes which got further divided into two subclusters IIA and IIB. Subcluster IIA consisted of three accessions namely 53 (IC597901), 52 (IC597886) and 58 (IC597949), while rest of the 55 accessions got grouped into subcluster IIB. The clustering pattern of most of the genotypes inferred the presence of low genetic diversity in the reference set under investigation. They did not show any

grouping pattern based on their geographical location. It has been found that divergent genotypes have good breeding values and maximum variability for making genetic level selection may be created by hybridizing genetically divergent parents. Genotype IC597868 (4) was found to be the most genetically divergent accession and can thus be employed in breeding programmes aimed at increasing genetic diversity or leafy mustard improvement.

Population structure, AMOVA and principal coordinate analysis

The Evanno's method suggested three distinct groups ($\Delta K = 3$) i.e. three subpopulations in leafy mustard germplasm reference set (Fig. 4). These three subpopulations were designated as I, II and III and they had 13, 31 and 15 genotypes, respectively. These genotypes were further categorized as the pure and admixture type based on the membership fractions. Leafy mustard genotypes with probability score $\geq 80\%$ are considered as pure type, while

Fig. 3 UNJ-dendrogram showing genetic relationship among various leafy mustard accessions based on SSR markers data

Fig. 4 The population structure of 59 leafy mustard germplasm accessions estimated using 122 SSR loci

score $\leq 80\%$ are considered as admixture type. The subpopulation I has 4 pure and 9 admixture type genotypes. Similarly, subpopulation II has got 19 pure and 12 admixture type of genotypes, while subpopulation III had 8 pure and 7 admixture type of genotypes. The most plausible reason for getting admixture type of genotypes may be that these germplasm accessions had been collected from four adjoining districts of Arunachal Pradesh. As some of these lines might have been growing in close vicinity of each other, so there might have occurred allelic exchange among them leading to the development of admixture type of genotypes. The results of structure analysis were not in concordance with the UNJ clustering

patterns. However, it has been found that structure analysis gives more accurate estimate of genetic diversity than UNJ based dendrogram. A more clear picture of genetic diversity could be obtained by structure analysis in the present study. Similar results were reported in a coffee germplasm panel using ISSR markers, where UNJ tree grouped all the 159 tested accessions into two clusters, whereas STRUCTURE analysis predicted three clusters in the same germplasm panel (Yan et al. 2019). Similarly, different grouping pattern was observed for both UPGMA-based dendrogram and structure analysis in sponge gourd (*Luffa cylindrica*) accessions on the basis of ISSR and SCoT markers (Tyagi et al. 2020). Overall, the genetic diversity of leafy mustard

Fig. 5 Principal coordinate analysis showing clustering of various leafy mustard genotypes based on SSR data

Table 4 Summary of analysis of molecular variance (AMOVA) results

Source	Degree of freedom (df)	Sum of square (SS)	Mean sum of square (MSS)	Estimate of variance (Est)	%age of total variation
Among populations	2	451.92	225.96	3.60	7%
Among individuals	56	5380.19	96.07	48.03	93%
Total	58	5832.11		51.63	100%

is inferred narrow in the present study. PCoA analysis explained the clustering pattern of leafy mustard genotypes (Fig. 5). The analysis of molecular variance (AMOVA) based on three subpopulations explained 7% variation among populations, whereas, 93% of variation was observed within the individuals of leafy mustard (Table 4, Fig. 6).

Conclusion

In the present study, both agro-morphological traits and SSR markers used provided a comprehensive insight into the genetic diversity of leafy mustard accessions. Structure analysis provided a clearer estimate of genetic variability among leafy mustard accessions in comparison to UNJ-based dendrogram analysis. The study inferred the presence of narrow genetic variability in leafy mustard accessions. From this study, it is inferred that there is a need to

Fig. 6 AMOVA statistics for 59 leafy mustard accessions based on SSR data

broaden the germplasm base of leafy mustard by making more number of germplasm explorations from North-eastern states as well as from the high hills of Uttarakhand. Further, evaluation of genetic diversity inherent in those collections along with their nutritional profiling studies should be carried out, and efforts should be directed to create more variability and broadening the genetic base of leafy mustard by making hybridization/crosses between genetically diverse parents.

Acknowledgements The authors are highly thankful to the Director, ICAR-DRMR, Bharatpur for providing facilities to carry out this research work.

References

- Bhattarai BP, Shing MP, Shakya SM, Khatri-Chhetri GB, Khadka YG (2018) Effect of organic and conventional nutrient management on post-harvest status of broad leaf mustard (*Brassica juncea* var *rugosa*). *Int J Hort Agric* 3(2):1–4. <https://doi.org/10.15226/2572-3154/3/2/0012110.15226/2572-3154/3/2/00121>
- Earl dent A, Von Holdt Bridgett M (2012) STRUCTURE HAR-VESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons Genet Resour* 4:359–361
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. *Mol Ecol* 14:2611–2620
- Fang P, Chen FB, Yao QL, Yang KC, Zhou GF, Fan YH, Zhang ZR, Shen JJ (2013) Zhang H (2013) Analysis of genetic diversity in the tuber mustard (*Brassica juncea* var. *tumida* Tsen et Lee) in the Yangtze river basin of China. *Genet Resour Crop Evol* 60:129–143. <https://doi.org/10.1007/s10722-012-9821-y>
- Ghaffari P, Talebi R, Keshavarzi F (2014) Genetic diversity and geographical differentiation of Iranian landrace, cultivars, and exotic chickpea lines as revealed by morphological and microsatellite markers. *Physiol Mol Biol Plants* 20(2):225–233. <https://doi.org/10.1007/s12298-014-0223-9>

- Hansen J, Møller IB (1975) Percolation of starch and soluble carbohydrates from plant tissue for quantitative determination with anthrone. *Anal Biochem* 68(1):87–94
- Liu K, Muse M (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinfo* 21:2128–2129
- Macready AL, George TW, Chong MF (2014) Flavonoid-rich fruit and vegetables improve microvascular reactivity and inflammatory status in men at risk of cardiovascular disease—FLAVURS: a randomized controlled trial. *Am J Clin Nutr* 99: epub
- Pant U, Bhajan R, Singh A, Kulshrestha K, Singh AK, Punetha H (2020) Green leafy mustard: a healthy alternative electronic J. *Plant Breed* 11(1):267–270. <https://doi.org/10.37992/2020.1101.045>
- Peakall R, Smouse PE (2012) GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software. <http://darwin.cirad.fr/darwin>. Accessed 21 Apr 2020
- Prieto P, Pineda M, Aguilar M (1999) Spectrophotometric quantitation of antioxidant capacity through the formation of a Phosphomolybdenum complex: specific application to the determination of vitamin E. *Anal Biochem* 269(2):337–341
- Pritchard JK, Wen W (2003) Documentation for the structure software, version 2. Department of Human Genetics, University of Chicago, Chicago. <http://pritch.bsd.uchicago.edu/software>. Accessed 21 Apr 2020
- Qi XH, Yang JH, Zhang MF (2008) AFLP-based genetic diversity assessment among Chinese vegetable mustards (*Brassica juncea* (L.) Czern.). *Genet Resour Crop Evol* 55:705–711. <https://doi.org/10.1007/s10722-007-9278-6>
- Rauniyar K, Bhattarai BP (2017) Growth, yield and soil nutrient status of broad leaf mustard (*Brassica juncea* var. *rugosa*) under integrated nutrient management. *Nepalese J Agric Sci* 15:98–106
- Singh KH, Misra AK, Kumar A (2006) Draft guidelines to conduct the tests for distinctness, uniformity and stability in rapeseed and mustard (Brassica). National Research Centre on Rapeseed Mustard, Sewar
- Singh KH, Shakya R, Thakur AK, Chauhan DK, Chauhan JS (2013) Assessment of genetic diversity in Indian mustard (*Brassica juncea* L Czern & Coss) using morphological traits and RAPD markers. *Natl Acad Sci Lett* 36(4):419–427
- Singh KH, Shakya R, Singh KK, Thakur AK, Nanjundan J, Singh D (2015) Genetic enhancement of *Brassica carinata* through interspecific hybridization and population improvement. In: 14th international rapeseed congress, p 281
- Thakur AK, Singh BK, Verma V, Chauhan JS (2013) Direct organogenesis in *Brassica juncea* var. NRCDR-2 and analysis of genetic uniformity using RAPD markers. *Natl Acad Sci Lett* 36:403–409
- Thakur AK, Singh KH, Singh L, Nanjundan J, Khan YJ, Singh D (2017) Patterns of subspecies genetic diversity among oilseed *Brassica rapa* as revealed by agro-morphological traits and SSR markers. *J Plant Biochem Biotechnol* 26(3):282–292. <https://doi.org/10.1007/s13562-016-0390-6>
- Thakur AK, Singh KH, Lal S, Nanjundan J, Yasin JK, Singh D (2018) SSR marker variations in Brassica species provide insight into the origin and evolution of *Brassica* amphidiploids. *Hereditas* 155:6
- Tyagi R, Sharma V, Sureja AK, Munshi AD, Arya L, Saha D, Verma M (2020) Genetic diversity and population structure detection in sponge gourd (*Luffa cylindrica*) using ISSR, SCoT and morphological markers. *Physiol Mol Biol Plants* 26(1):119–131. <https://doi.org/10.1007/s12298-019-00723-y>
- Vieira ML, Santini L, Diniz AL, Munhoz Cde F (2016) Microsatellite markers: what they mean and why they are so useful. *Genet Mol Biol* 39:312–328. <https://doi.org/10.1590/1678-4685-GMB-2016-0027>
- Vinu V, Singh N, Vasudev S, Yadav DK, Kumar S, Naresh S, Bhat SR, Prabhu KV (2013) Assessment of genetic diversity in *Brassica juncea* (Brassicaceae) genotypes using phenotypic differences and SSR markers. *Rev Biol Trop* 61(4):1919–1934
- Wu XM, Chen BY, Lu GY, Wang HZ, Xu K, Gao GZ, Song YC (2009) Genetic diversity in oil and vegetable mustard (*Brassica juncea*) landraces revealed by SRAP markers. *Genet Resour Crop Evol* 56:1011–1022
- Yan L, Ogutu C, Huang L, Wang X, Zhou H, Lv Y, Long Y, Dong Y, Han Y (2019) Genetic diversity and population structure of coffee germplasm collections in China revealed by ISSR markers. *Plant Mol Biol Rep* 37:204–213. <https://doi.org/10.1007/s11105-019-01148-3>
- Yao QL, Chen FB, Fang P, Zhou GF, Fan YH, Zhang ZR (2012) Genetic diversity of Chinese vegetable mustard (*Brassica juncea* Coss) landraces based on SSR data. *Biochem Sys Ecol* 45:41–48
- Yasin JK, Mayadevi P (2010) Genetic diversity among commercial varieties of *Anthurium andreanum* Linden using RAPD markers. *Plant Genet Transgenics* 1(1):11–15

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.