# Development of a set of SSR markers for characterization of Indian mustard germplasm and varieties

**Article** · November 2021

DOI: 10.1007/s13562-021-00737-2(

CITATIONS
2

READS
251

**8 authors**, including:

Joghee Nanjundan
Indian Agricultural Research Institute
**50** PUBLICATIONS **341** CITATIONS
SEE PROFILE

Kunwar HARENDRA Singh
Indian Council of Agricultural Research
**128** PUBLICATIONS **695** CITATIONS
SEE PROFILE

Deepika Sharma
Tata Institute of Fundamental Research
**10** PUBLICATIONS **172** CITATIONS
SEE PROFILE

Nehanjali Parmar
ICAR-Directorate of Rapeseed-Mustard Research, Bharatpur, Rajasthan
**24** PUBLICATIONS **242** CITATIONS
SEE PROFILE

Journal of Plant Biochemistry and Biotechnology
https://doi.org/10.1007/s13562-021-00737-2

ORIGINAL ARTICLE

# Development of a set of SSR markers for characterization of Indian mustard germplasm and varieties

Lal Singh[1] · J. Nanjundan[2] · K. H. Singh[3] · Deepika Sharma[3] · Nehanjali Parmar[3] · Anshul Watts[4] · Rohit Jain[1] · Ajay Kumar Thakur[3]

Received: 10 February 2021 / Accepted: 14 October 2021


## Abstract

Simple sequence repeat (SSR) markers can serve as an important tool for characterization and genetic diversity evaluation in Indian mustard germplasm and varieties. For DNA fingerprinting of various Indian mustard varieties, it is necessary that a common set of SSRs be used by various laboratories so that the developed genetic profiles of various genotypes can be compared with one another. In this study, we evaluated the polymorphic potential of 350 SSR markers to derive a set of SSR markers for characterization of Indian mustard germplasm and varieties. Out of a total of 350 SSR markers evaluated, 310 (88.57%) SSRs produced polymorphic amplicons, while remaining 40 (11.43%) SSRs resulted into monomorphic products. The allele number varied from 2 to 7 with 3.22 average number of alleles per locus. Polymorphism information content (PIC) value ranged from 0.24 (Ol09A01) to 0.75 (nia-m141a) with an average PIC value of 0.40 per locus. A total of 95 (31%) SSR markers evaluated were having PIC values more than the average PIC value, which constitute the representative set of SSR markers. Unweighted pair group method with arithmetic averages (UPGMA)-dendrogram grouped all the 46 genotypes into two main clusters, while STRUCTURE analysis formed three subpopulations having admixture of alleles. This SSR marker set will facilitate systematic characterization and classification of various Indian mustard germplasm accessions and varieties.

**Keywords** *Brassica juncea* · SSR markers · Genetic diversity · Molecular characterization

## Abbreviations

| | |
|---|---|
| SSR | Simple sequence repeat |
| ISSR | Inter simple sequence repeat |
| RAPD | Random amplified polymorphic DNA |
| RFLP | Restriction fragment length polymorphism |
| AFLP | Amplified fragment length polymorphism |
| PIC | Polymorphic information content |
| UPGMA | Unweighted pair group method with arithmetic averages |

## Introduction

Indian mustard (*Brassica juncea*. L. Czern & Coss., 2n = 38, AABB) is an economically important oilseed crop of family *Brassicaceae*, which is being cultivated mainly in the Indian subcontinent, Canada, Russia, China and Australia. It originated thousands of years ago from natural cross-hybridization between *B. nigra* (2n = 16, BB) and *B. rapa* (2n = 20, AA) with subsequent chromosome duplication (U 1935). In India, *B. juncea* has emerged as the dominant crop of Rapeseed-Mustard (oilseed Brassica) group and is presently being cultivated over 6.23 m ha and constitutes > 85% of total RM acreage (Govt. of India 2018). However, due to the increasing population pressure, it would become a herculean task to meet the demand of edible oil in India by 2050 (Singh et al. 2018). This necessitated the development of high yielding varieties of Indian mustard. The success of any breeding program relies

✉ Ajay Kumar Thakur
thakurak2010@gmail.com

[1] Manipal University Jaipur, Jaipur, Rajasthan 303 007, 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 Institute for Plant Biotechnology, New Delhi 110 012, India

Published online: 09 November 2021

**Table 1** Reference panel of *Brassica juncea* germplasm used for developing a set of SSR markers

| Code No | Name of accession | National identity | Seed coat color | Nature/Type of accessions | Locality/institute & Province of origin | Country |
|---|---|---|---|---|---|---|
| 1 | B 15 | IC 609,530 | Brown | Others | Unknown | India |
| 2 | B/K/S 72 | IC 310,771 | Brown | Indigenous collection | Morena, Madhya Pradesh | India |
| 3 | Bio Q-108 | IC 609,329 | Brown | Advanced breeding line | IARI, New Delhi | India |
| 4 | BT 15 | IC 609,452 | Brown | Others | Unknown | India |
| 5 | EC 447,055 | EC 447,055 | Yellow | Exotic | – | Germany |
| 6 | EC 482,983 | EC 482,983 | Brown | Exotic | – | Canada |
| 7 | EC 511,589 | EC 511,589 | Yellow | Exotic | – | Canada |
| 8 | EC 511,690 | EC 511,690 | Brown | Exotic | – | Canada |
| 9 | EC 511,718 | EC 511,718 | Brown | Exotic | – | Canada |
| 10 | EC 557,025 | EC 557,025 | Brown | Exotic | – | China |
| 11 | Glossy CCH | IC 609,583 | Brown | Others | Unknown | India |
| 12 | HP 11 | IC 347,667 | Brown | Indigenous Collection | Palampur, Himachal Pradesh | India |
| 13 | SKCV 09–6 | IC 571,629 | Brown | Indigenous Collection | Bidar, Karnataka | India |
| 14 | SKCV 09–23 | IC 571,646 | Brown | Indigenous Collection | Dharwad, Karnataka | India |
| 15 | JAIPUR 12 | IC 511,592 | Brown | Others | Unknown | India |
| 16 | JCR 914 | IC 427,136 | Brown | Indigenous Collection | Bilaspur, Himachal Pradesh | India |
| 17 | NDRS 2017–1 | IC 609,623 | Brown | Advanced Breeding Line | NDUAT, Faizabad, Uttar Pradesh | India |
| 18 | AKMSS 68 | IC 522,378 | Brown | Indigenous Collection | Dimapur, Nagaland | India |
| 19 | P 14 | IC 346,700 | Brown | Indigenous Collection | Sangroor, Punjab | India |
| 20 | P 58 | IC 346,744 | Brown | Indigenous Collection | Faridkot, Punjab | India |
| 21 | PBR 2004–06 | IC 511,487 | Brown | Advanced Breeding Line | PAU, Ludhiana, Punjab | India |
| 22 | PHR 2 | IC 609,873 | Brown | Others | Unknown | India |
| 23 | RGN 157 | IC 511,523 | Brown | Advanced Breeding Line | ARS, SKRAU, Sriganganagar, Rajasthan | India |
| 24 | RH 0114 | IC 609,440 | Brown | Advanced Breeding Line | CCSHAU, Hisar, Haryana, | India |
| 25 | RH 704 | IC 609,631 | Brown | Advanced Breeding Line | CCSHAU, Hisar, Haryana, | India |
| 26 | SN 55 | IC 426,343 | Brown | Indigenous Collection | Karimnagar, Andhra Pradesh | India |
| 27 | SN 56 | IC 426,344 | Brown | Indigenous Collection | Karimnagar, Andhra Pradesh | India |
| 28 | TM 1 | IC 296,829 | Brown | Registered genetic stock | BARC, Mumbai, Maharashtra | India |
| 29 | UP 1–70 | IC 346,178 | Brown | Indigenous Collection | GB Nagar, Uttar Pradesh | India |
| 30 | UP 1–87 | IC 346,195 | Brown | Indigenous Collection | Faridabad, Haryana | India |
| 31 | UP 11–119 | IC 345,973 | Brown | Indigenous Collection | Ghaziabad, Uttar Pradesh | India |
| 32 | UP 11–123 | IC 345,977 | Brown | Indigenous Collection | Meerut, Uttar Pradesh | India |
| 33 | UP 11–22 | IC 345,876 | Brown | Indigenous Collection | Muzafarnagar, Uttar Pradesh | India |
| 34 | UP 11–28 | IC 345,882 | Brown | Indigenous Collection | Saharanpur, Uttar Pradesh | India |
| 35 | B/K/S 127 | IC 310,799 | Brown | Indigenous Collection | Gawalior, Madhya Pradesh | India |
| 36 | B/K/S 67 | IC 310,767 | Brown | Indigenous Collection | Ferozepur, Haryana | India |
| 37 | B/K/S 19 | IC 310,730 | Brown | Indigenous Collection | Bharatpur, Rajasthan | India |
| 38 | RH 30 | IC 73,217 | Brown | Cultivar | CCSHAU, Hisar, Haryana, | India |
| 39 | NRCDR 02 | IC 511,526 | Brown | Cultivar | DRMR, Bharatpur, Rajasthan | India |
| 40 | DRMRIJ 31 | - | Brown | Cultivar | DRMR, Bharatpur, Rajasthan | India |
| 41 | Rohini | IC 113,100 | Brown | Cultivar | CSAUA & T, Kanpur, Uttar Pradesh | India |
| 42 | NRCHB 101 | IC 559,588 | Brown | Cultivar | DRMR, Bharatpur, Rajasthan | India |
| 43 | Kranti | IC 494,384 | Brown | Cultivar | GBPUA&T, Pantnagar, Uttarakhand | India |
| 44 | Pusa Jai Kisan | IC 305,064 | Brown | Cultivar | IARI, New Delhi | India |

**Table 1** (continued)

| Code No | Name of accession | National identity | Seed coat color | Nature/Type of accessions | Locality/institute & Province of origin | Country |
|---|---|---|---|---|---|---|
| 45 | RH 0749 | IC 594,689 | Brown | Cultivar | CCSHAU, Hisar, Haryana, | India |
| 46 | Pusa Bold | IC 73,216 | Brown | Cultivar | IARI, New Delhi | India |

upon the extent of genetic variability inherent in the donors. However, genetic diversity within Indian mustard germplasm collections and varieties has been very much limited. At present, around 2100 and 9720 germplasm accessions of Indian mustard along with > 110 cultivated varieties are being conserved in gene banks at ICAR-DRMR, Bharatpur and ICAR-National Bureau of Plant Genetic Resources (NBPGR), New Delhi, respectively (Nanjundan et al. 2014). Molecular markers assist in the discovery of a number of novel alleles which otherwise remain hidden in genome background (Tanksley and McCouch 1997). The development of a portable set of molecular markers is urgently needed for molecular characterization of Indian mustard germplasm accessions and varieties, and in mustard improvement programmes to lend precision and speed.

Among various types of molecular markers, SSR markers are the most preferred markers for molecular characterization and genetic diversity evaluation due to their co-dominant and multi-allelic inheritance, reproducibility and abundance (Viera et al. 2016). Over the last few years, using cross-transferability approach, various Brassica species-derived SSRs have been evaluated in *B. juncea* background to develop markers for this species (Yadava et al. 2009; Thakur et al. 2015, 2018; Singh et al. 2016). Scattered efforts had been made for genetic diversity evaluation in Indian mustard (Vinu et al. 2013; Nanjundan et al. 2015; Pratap et al. 2015; Singh et al. 2018). All of these studies utilized a very limited number of SSR markers in a limited panel of different Indian mustard genotypes. So far, no comprehensive study has been conducted using a large and a common set of SSR markers to characterize Indian mustard accessions and varieties. So, the results of those labs always remained uncomparable.

The present study was undertaken with the objective to develop a portable set of SSR markers in a reference set of Indian mustard for characterization of germplasm and varieties. This SSR marker set can further be utilized for various purposes like gene/QTL mapping, linkage map development and marker assisted breeding in Indian mustard improvement programs.

## Materials and methods

### Plant material

The plant material in this study comprised of 46 Indian mustard genotypes including 37 local and and exotic germplasm accessions, representing mini-core set of germplasm accessions (Nanjundan et al. 2015), and 9 popular varieties developed under the aegis of All India Coordinated Crop Improvement Program (Table 1).

### Isolation and purification of genomic-DNA

Young leaves from five plants per genotype were collected, pooled and used for genomic-DNA isolation as per already available protocol (Thakur et al. 2013). DNA was quantified on 0.8% agarose gel and further stored in TE buffer at 4 °C.

### Genotyping

SSR primers used in the present study were developed by different Brassica research groups. *B. rapa* (AA genome) specific primer sequences were obtained from Xu et al. (2010), *B. nigra* (BB genome) specific primers were procured from PAU, Ludhiana, and other primer sequences (*B. oleracea*, CC genome) were procured from Yadava et al. (2009).

For genotyping, a 25 µl reaction volume containing 50 ng DNA, 1XPCR buffer, 2.0 mM $Mgcl_2$, 0.2 mM of each dNTP, 1.0 U *Taq DNA polymerase* (GCC Biotech, India) and 400 nM primers was prepared. The first amplification cycle involved DNA denaturation at 94 °C temperature for 5 min, followed by a series of 45 cycles each of DNA denaturation for 30 s at 94 °C, primer annealing at a temperature of 50–60 °C (depending upon the primer combination) for 30 s, primer extension at 72 °C for 45 s and a final extension step for 7 min at 72 °C. PCR amplicons were run in a 3.5% Super Fine Resolution (SFR) agarose (Amresco, USA) gel at 100 V along with a 50 bp DNA ladder alongside as a reference point for tracing

**Table 2** List of selected set of SSR markers along with their allelic parameters

| S. No | Marker ID | Origin | Number of alleles | Major allele frequency | PIC value | Gene diversity |
|---|---|---|---|---|---|---|
| 1 | Ni2A08 | *B. nigra* | 3.00 | 0.49 | 0.47 | 0.56 |
| 2 | Ni2A09 | *B. nigra* | 5.00 | 0.60 | 0.49 | 0.55 |
| 3 | Ni4G06 | *B. nigra* | 4.00 | 0.60 | 0.45 | 0.52 |
| 4 | Ni2C12 | *B. nigra* | 6.00 | 0.67 | 0.45 | 0.50 |
| 5 | Ni2D07 | *B. nigra* | 4.00 | 0.35 | 0.64 | 0.70 |
| 6 | Ni2E05 | *B. nigra* | 4.00 | 0.58 | 0.46 | 0.53 |
| 7 | Ni2E12 | *B. nigra* | 3.00 | 0.34 | 0.59 | 0.66 |
| 8 | Ni2H03 | *B. nigra* | 6.00 | 0.65 | 0.46 | 0.51 |
| 9 | Ni3F02 | *B. nigra* | 3.00 | 0.36 | 0.57 | 0.65 |
| 10 | Ni3H02 | *B. nigra* | 3.00 | 0.54 | 0.44 | 0.53 |
| 11 | Ni4A03 | *B. nigra* | 6.00 | 0.58 | 0.54 | 0.59 |
| 12 | Ni4C02 | *B. nigra* | 5.00 | 0.60 | 0.50 | 0.56 |
| 13 | Ni4C09 | *B. nigra* | 4.00 | 0.46 | 0.57 | 0.64 |
| 14 | Ni4D08 | *B. nigra* | 3.00 | 0.55 | 0.43 | 0.52 |
| 15 | Ni4D09 | *B. nigra* | 3.00 | 0.37 | 0.56 | 0.64 |
| 16 | Ni4D10 | *B. nigra* | 4.00 | 0.41 | 0.60 | 0.65 |
| 17 | Ni4E01 | *B. nigra* | 4.00 | 0.54 | 0.51 | 0.58 |
| 18 | Ni4H06 | *B. nigra* | 5.00 | 0.57 | 0.52 | 0.58 |
| 19 | SJ4933 | *B. nigra* | 4.00 | 0.69 | 0.44 | 0.48 |
| 20 | SJ3302RI | *B. nigra* | 4.00 | 0.61 | 0.44 | 0.51 |
| 21 | SĂ | *B. nigra* | 3.00 | 0.38 | 0.56 | 0.63 |
| 22 | SB0372 | *B. nigra* | 3.00 | 0.58 | 0.41 | 0.50 |
| 23 | SB1935A | *B. nigra* | 4.00 | 0.68 | 0.43 | 0.45 |
| 24 | SJ8033 | *B. nigra* | 3.00 | 0.58 | 0.41 | 0.50 |
| 25 | SB3140 | *B. nigra* | 3.00 | 0.43 | 0.51 | 0.59 |
| 26 | SJ3874I | *B. nigra* | 6.00 | 0.67 | 0.45 | 0.50 |
| 27 | SJ6842 | *B. nigra* | 4.00 | 0.60 | 0.45 | 0.51 |
| 28 | SJ1505 | *B. nigra* | 4.00 | 0.63 | 0.42 | 0.48 |
| 29 | SJ13133 | *B. nigra* | 3.00 | 0.46 | 0.50 | 0.58 |
| 30 | nia_m042a | *B. nigra* | 6.00 | 0.63 | 0.49 | 0.54 |
| 31 | nia_m053a | *B. nigra* | 4.00 | 0.50 | 0.55 | 0.63 |
| 32 | nia_m140a | *B. nigra* | 4.00 | 0.50 | 0.55 | 0.62 |
| 33 | nia_m141a | *B. nigra* | 5.00 | 0.22 | 0.75 | 0.78 |
| 34 | Ol10B01 | *B. oleracea* | 4.00 | 0.42 | 0.60 | 0.66 |
| 35 | BRMS-003 | *B. rapa* | 3.00 | 0.36 | 0.57 | 0.64 |
| 36 | BRMS-006 | *B. rapa* | 3.00 | 0.33 | 0.59 | 0.67 |
| 37 | BRMS-007 | *B. rapa* | 5.00 | 0.65 | 0.44 | 0.50 |
| 38 | BRMS-014 | *B. rapa* | 3.00 | 0.54 | 0.44 | 0.53 |
| 39 | BRMS-018 | *B. rapa* | 4.00 | 0.45 | 0.58 | 0.65 |
| 40 | BRMS-027 | *B. rapa* | 3.00 | 0.46 | 0.50 | 0.58 |
| 41 | BRMS-008 | *B. rapa* | 4.00 | 0.50 | 0.55 | 0.62 |
| 42 | BrgMS397 | *B. rapa* | 3.00 | 0.54 | 0.43 | 0.52 |
| 43 | BrgMS388 | *B. rapa* | 3.00 | 0.55 | 0.43 | 0.52 |
| 44 | BrgMS2766 | *B. rapa* | 4.00 | 0.74 | 0.42 | 0.39 |
| 45 | BrgMS638 | *B. rapa* | 4.00 | 0.50 | 0.55 | 0.63 |
| 46 | BrgMS787 | *B. rapa* | 4.00 | 0.51 | 0.55 | 0.62 |
| 47 | BrgMS732 | *B. rapa* | 6.00 | 0.69 | 0.41 | 0.46 |
| 48 | BrgMS713 | *B. rapa* | 4.00 | 0.49 | 0.56 | 0.63 |

**Table 2** (continued)

| S. No | Marker ID | Origin | Number of alleles | Major allele frequency | PIC value | Gene diversity |
|---|---|---|---|---|---|---|
| 49 | BrgMS1474 | *B. rapa* | 5.00 | 0.61 | 0.48 | 0.54 |
| 50 | BrgMS383 | *B. rapa* | 4.00 | 0.26 | 0.70 | 0.75 |
| 51 | BrgMS465 | *B. rapa* | 4.00 | 0.53 | 0.50 | 0.57 |
| 52 | BrgMS4539 | *B. rapa* | 3.00 | 0.34 | 0.59 | 0.66 |
| 53 | BrgMS10 | *B. rapa* | 4.00 | 0.62 | 0.44 | 0.51 |
| 54 | BrgMS653 | *B. rapa* | 4.00 | 0.65 | 0.41 | 0.48 |
| 55 | BrgMS643 | *B. rapa* | 4.00 | 0.63 | 0.42 | 0.49 |
| 56 | BrgMS216 | *B. rapa* | 3.00 | 0.35 | 0.58 | 0.66 |
| 57 | BrgMS4513 | *B. rapa* | 3.00 | 0.45 | 0.51 | 0.59 |
| 58 | BrgMS166 | *B. rapa* | 4.00 | 0.58 | 0.48 | 0.55 |
| 59 | BrgMS139 | *B. rapa* | 3.00 | 0.53 | 0.45 | 0.54 |
| 60 | BrgMS190 | *B. rapa* | 4.00 | 0.25 | 0.70 | 0.75 |
| 61 | BrgMS66 | *B. rapa* | 4.00 | 0.45 | 0.58 | 0.65 |
| 62 | BrgMS361 | *B. rapa* | 3.00 | 0.36 | 0.58 | 0.65 |
| 63 | BrgMS233 | *B. rapa* | 4.00 | 0.57 | 0.48 | 0.55 |
| 64 | BrgMS841 | *B. rapa* | 4.00 | 0.54 | 0.51 | 0.58 |
| 65 | BrgMS36 | *B. rapa* | 4.00 | 0.50 | 0.55 | 0.63 |
| 66 | BrgMS502 | *B. rapa* | 3.00 | 0.51 | 0.46 | 0.55 |
| 67 | BrgMS639 | *B. rapa* | 4.00 | 0.53 | 0.53 | 0.60 |
| 68 | BrgMS3322 | *B. rapa* | 3.00 | 0.54 | 0.44 | 0.53 |
| 69 | Ra1-F06 | *B. rapa* | 3.00 | 0.38 | 0.56 | 0.63 |
| 70 | Ra2-A01 | *B. rapa* | 4.00 | 0.53 | 0.52 | 0.59 |
| 71 | Ra2-A11 | *B. rapa* | 3.00 | 0.57 | 0.42 | 0.51 |
| 72 | Ra2-C09 | *B. rapa* | 4.00 | 0.53 | 0.52 | 0.59 |
| 73 | Ra2-D04 | *B. rapa* | 3.00 | 0.33 | 0.59 | 0.67 |
| 74 | Ra2-E04 | *B. rapa* | 4.00 | 0.53 | 0.51 | 0.57 |
| 75 | Ra2-E11 | *B. rapa* | 5.00 | 0.32 | 0.70 | 0.74 |
| 76 | Ra2-E12 | *B. rapa* | 4.00 | 0.42 | 0.60 | 0.67 |
| 77 | Ra2-F09 | *B. rapa* | 3.00 | 0.56 | 0.43 | 0.52 |
| 78 | Ra2-G04 | *B. rapa* | 3.00 | 0.33 | 0.59 | 0.67 |
| 79 | Ra2-G10 | *B. rapa* | 3.00 | 0.56 | 0.43 | 0.52 |
| 80 | Ra2-H04 | *B. rapa* | 5.00 | 0.60 | 0.49 | 0.55 |
| 81 | cnu_m583a | *B. rapa* | 5.00 | 0.60 | 0.50 | 0.56 |
| 82 | cnu_m584a | *B. rapa* | 5.00 | 0.59 | 0.51 | 0.57 |
| 83 | cnu_m585a | *B. rapa* | 4.00 | 0.64 | 0.41 | 0.48 |
| 84 | cnu_m619a | *B. rapa* | 4.00 | 0.65 | 0.41 | 0.47 |
| 85 | cnu_m623a | *B. rapa* | 3.00 | 0.33 | 0.59 | 0.67 |
| 86 | cnu_m626a | *B. rapa* | 4.00 | 0.58 | 0.46 | 0.53 |
| 87 | EJU1 | *B. rapa* | 3.00 | 0.38 | 0.55 | 0.62 |
| 88 | ENA10 | *B. rapa* | 4.00 | 0.46 | 0.56 | 0.63 |
| 89 | ENA17 | *B. rapa* | 3.00 | 0.48 | 0.49 | 0.57 |
| 90 | ENA23 | *B. rapa* | 4.00 | 0.27 | 0.69 | 0.74 |
| 91 | ENA26 | *B. rapa* | 4.00 | 0.50 | 0.54 | 0.61 |
| 92 | ENA28 | *B. rapa* | 4.00 | 0.54 | 0.52 | 0.59 |
| 93 | ENA3 | *B. rapa* | 3.00 | 0.46 | 0.50 | 0.58 |
| 94 | ENA4 | *B. rapa* | 3.00 | 0.34 | 0.59 | 0.66 |
| 95 | E039 | *B. rapa* | 5.00 | 0.64 | 0.45 | 0.51 |

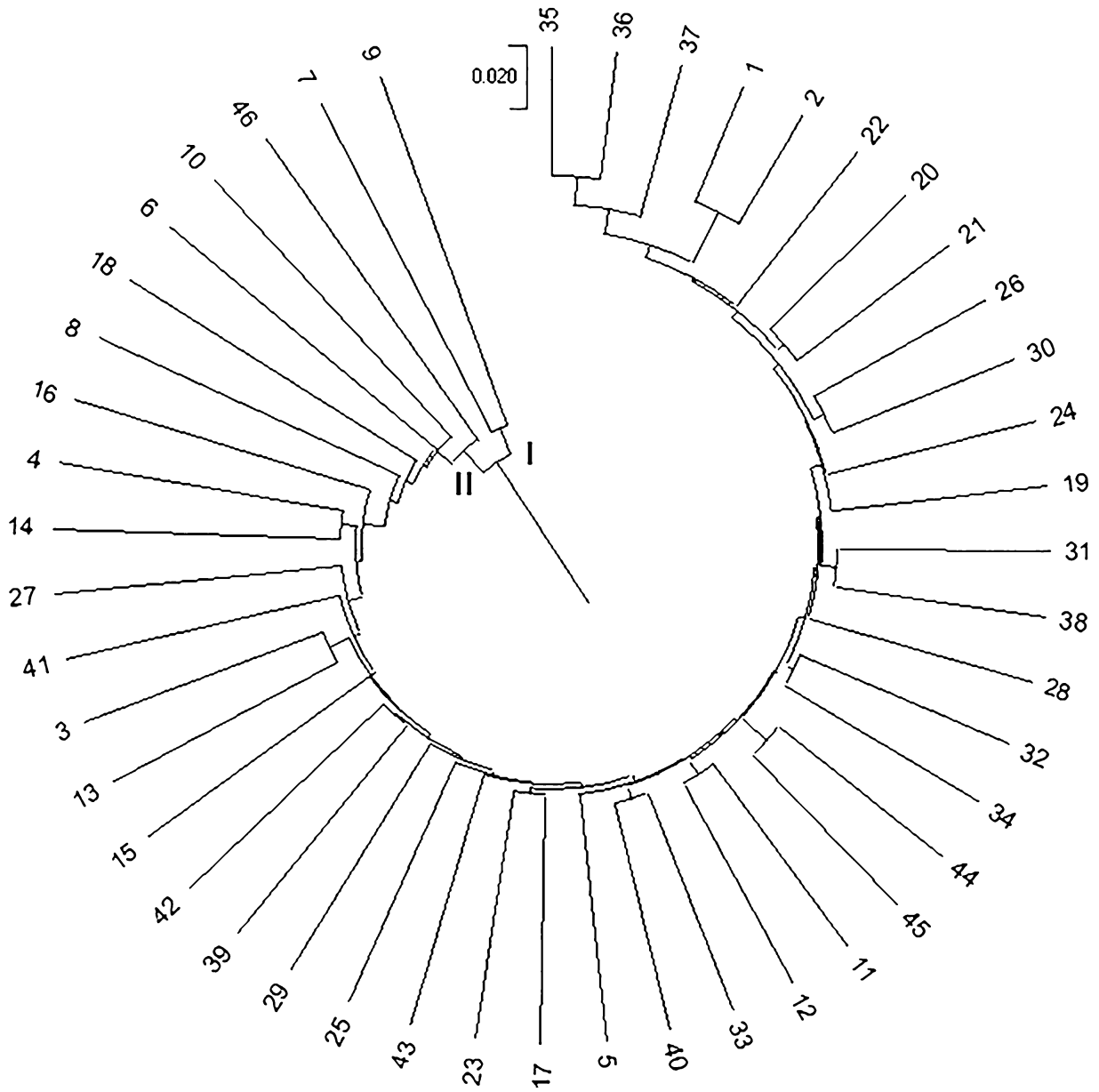


**Fig. 1** UPGMA dendrogram depicting the genetic inter-relationship among different accessions of Indian mustard based on the allelic data of 310 SSR markers

movement and size measurement. Further, the gels were photographed in a gel doc unit (Syngene Gel Doc, UK).

## Data analysis

Only single locus SSR markers were considered for data analysis. SSR markers were scored for amplified allele size and data was analyzed using Power Marker software (Liu & Muse 2005) for computation of major allele frequency (MAF), PIC value and gene diversity values for individual SSR. SSR markers having PIC values above the average PIC value were selected for developing a reference set of markers for characterization of Indian mustard germplasm and varieties. BLASTN analysis was done using forward and reverse primers in the BRAD database (http://brassicadb.cn/#/) against the Braju tum V 1.5 genome (Yang et al. 2016) to identify the location of all the primers in the *B. juncea* genome. UPGMA-dendrogram was developed to demonstrate the genetic relationship among different accessions under investigation using MEGA version 5.03 software (Tamanna and Khan 2005).

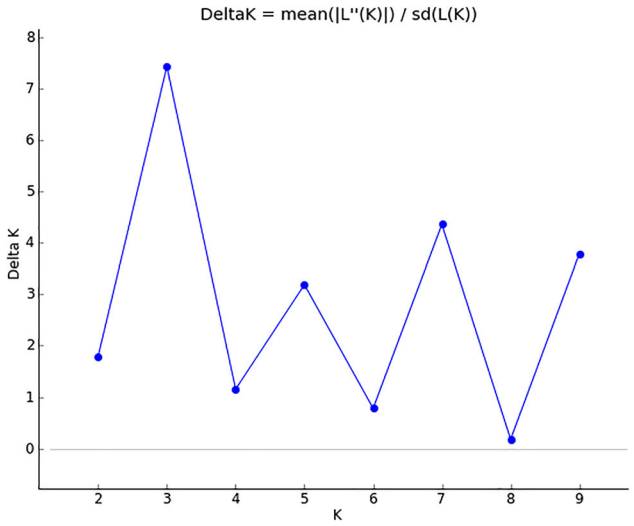


**Fig. 2** Population genetic structure and relationships among 46 Indian mustard accessions based on the allelic data of 310 SSR markers. Values of delK with its modal value detecting a true K of three groups (K = 3)

## 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 delK statistic and L(K) (Evanno et al. 2005) using Structure Harvester software (Earl and VonHoldt 2012). Analysis of molecular variance (AMOVA) and principal component analysis were carried out using GenALex6.5 (Peakall and Smouse 2012).

# Results and discussion

## SSR marker analysis

A total of 350 SSR markers derived from various Brassica species (*B. nigra*, *B. oleracea* and *B. rapa*) were evaluated for their polymorphic potential in a reference set of Indian mustard germplasm and varieties in the present study. Three hundred and ten (88.57%) SSRs produced polymorphic products, whereas, the remaining 40 (11.43%) SSR markers amplified single band. The polymorphic SSRs amplified a total of 998 alleles and allele number varied from 2 to 7 with an average number of 3.22 alleles per marker. *B. juncea* is actually an amphidiploid crop having AABB genome derived from *B. rapa* and *B. nigra*. Evidence for genome triplication has also been reported for this species (Cheng et al. 2014). This might have resulted in amplification of more number of alleles (> 5 alleles). This is to mention that the multiple amplicons observed in this study were indeed alleles coming from the same locus and we have considered only single locus SSR markers for data analysis. Thakur et al. (2015) reported a comparatively lesser average number of alleles (2.17) per SSR locus when they characterized twelve popular Indian mustard varieties using SSR markers. The most plausible reason may be that we have included genetically diverse germplasm lines (both exotic and indigenous) in our study, which might have resulted in higher number of alleles. The allelic data of these 310 polymorphic SSR markers have been used for computation of various diversity parameters including major allele frequency (MAF), gene diversity and polymorphism information content (PIC) value, which have been represented in Table S1. The major allele frequency ranged from 0.22 (nia-m141a) to 0.84 (Ol09A01) with an average value of 0.57 per locus. Gene diversity values were in the range of 0.27 (Ol09A01) to 0.78 (nia-m141a) with an average value of 0.49 per SSR marker.

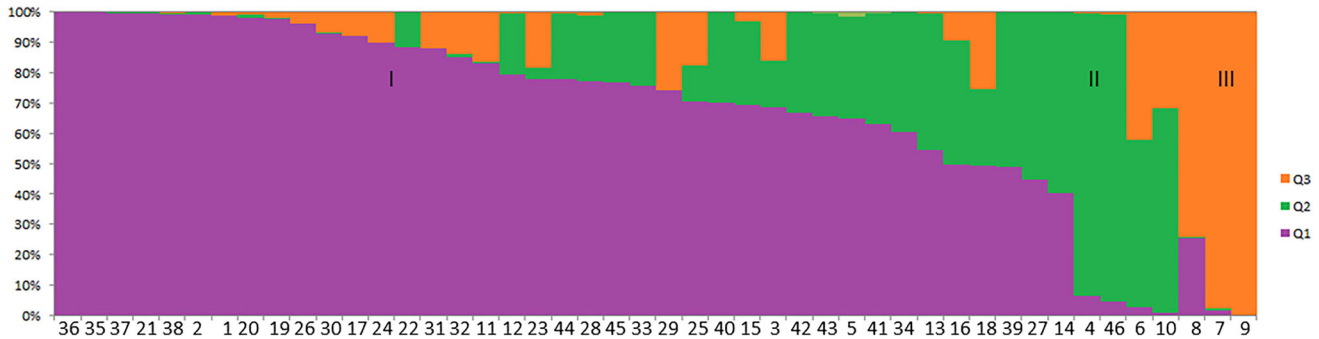


**Fig.3** The three subgroups inferred from the STRUCTURE analysis using the allelic data of 310 microsatellite loci

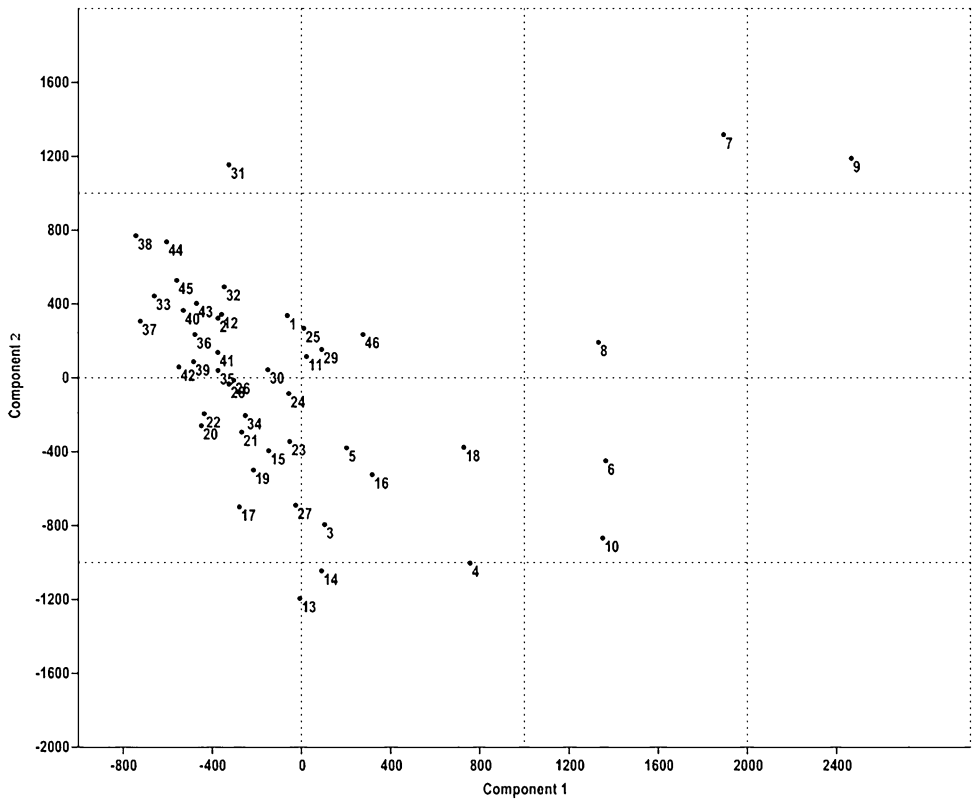


**Fig. 4** Principal co-ordinate analysis (PCoA) of 46 Indian mustard accessions base on 310 microsatellite loci

**Table 3** 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. var.) | %age of total variation (%) | Fst (Genetic differentiation among populations | *P* value |
|---|---|---|---|---|---|---|---|
| Among populations | 2 | 489.65 | 244.82 | 3.90 | 4 | 0.042 | 0.01 |
| Among individuals | 43 | 7747.17 | 180.16 | 90.08 | 96 | 1.00 | 0.01 |
| Total | 45 | 8236.82 | | 93.98 | 100 | | |

PIC value indicates the informativeness of a molecular marker and is an index of its polymorphic potential (Botstein et al. 1980). In the present study, PIC values varied from 0.24 (Ol09A01) to 0.75 (nia-m141a) with an average value of 0.40 per SSR locus. PIC value is influenced by various factors including the nature of germplasm, number of SSR markers used and the number of accessions genotyped. It has been found that majority of the SSR markers (66%) resulted in PIC values in the range of 0.3–0.4; whereas, 47 (15%) SSRs were having PIC values in the range of 0.4–0.5 and 48 (16%) SSR markers resulted in PIC values above 0.5. A total of 95 (31%) SSRs resulted into PIC values higher than the average PIC value of 0.4 in this study (Table 2). A comparatively lower average PIC value (0.30) was reported by Singh et al. (2016) when they evaluated SSR marker polymorphism in six genotypes of *B. juncea*. All the higher PIC values obtained in the present study suggest that these SSRs have deep resolution power and high discriminatory ability. Such type of highly polymorphic SSRs might have resulted from multiple copies of PCR amplicons and will be helpful in identification of unique DNA profiles and potentially novel genes with in Indian mustard germplasm (Thakur et al. 2018). These 95 SSR loci can be designated as highly polymorphic markers as inferred from their PIC values (Table 2) and they constitute the representative set of SSR markers, which can be

deployed for molecular characterization of Indian mustard germplasm accessions and varieties. The allele size of the amplicons of this selected set of SSR markers has been provided in Table S2, so that the researchers working in different labs can use this marker set and compare their results. In silico analysis to find out the probable location of the selected set of SSR markers using available *B. juncea var. tumida* genome sequence (Yang et al. 2016) has been carried out. BLASTN analysis was carried out using forward and reverse primers in the BRAD database (http://brassicadb.cn/#/) against the Braju tum V 1.5 genome (Yang et al. 2016) to identify the location of all the primers in the *B. juncea* genome. The SSR primer sequence information and their location in *B. juncea* genome has been provided in Supplementary Table S3. However, it has been observed that only 66 SSR marker sequences showed alignment to *B. juncea var. tumida* genome sequence, and 29 SSR marker sequences could not be assigned to any chromosome. These 66 SSR sequences were distributed on all the eighteen chromosomes of *B. juncea*. Dettori et al. (2015) successfully developed a universal set of 26 polymorphic SSRs distributed in the eight chromosomes of peach, almond, apricot, Japanese plum and sweet cherry for genetic analysis in these five Prunus species. Nguyen et al. (2018) also evaluated 521 SSR markers to develop a set of SSRs comprising of 43 SSRs for cultivar identification and seed purity testing in oriental melon. Li et al. (2019) developed a set of KASP (Kompetitive Allele Specific PCR) markers for evaluation of genetic diversity in *B. rapa* ssp. *Chinensis* Makino. Yang et al. (2019) also developed a set of 16 perfect SSRs for genetic diversity evaluation and identification of 382 cucumber varieties.

## Genetic diversity analysis

Evaluating genetic distances, determining the relatedness among genotypes, assists in selection of diverse genotypes for their utilization in deciding crossing programmes for pureline and hybrid breeding. The UPGMA-based dendrogram based on allelic data of 310 SSR markers grouped all Indian mustard accessions into two main clusters, which can be designated as cluster I and II. Cluster I comprised of only two exotic germplasm accessions from Canada viz. EC511589 (7) and EC511711 (9); whereas cluster II consisted of the rest of 44 genotypes. In cluster II, Indian mustard variety Pusa Bold (46) did not show affinity to any other sub-cluster and formed a separate sub-cluster IIA, while all the other 43 accessions were grouped in subcluster IIB (Fig. 1). Overall, the genotypes did not display any geographical patterns of diversity as two exotic accessions from Canada grouped in cluster I, while accessions belonging to Germany, Canada and China (EC 447,055, EC 482,983, EC 511,690 and EC 557,025) were grouped in cluster IIB.

However, the results of analysis using allelic data of 95 selected SSR markers were at variance with the results obtained using 310 markers. When the allelic data of the selected set of 95 SSR markers was used for UPGMA-based dendrogram construction, all the 46 genotypes grouped into two major clusters i.e. cluster I and II (Fig. S1). However, cluster I comprised of only one indigenous Indian mustard genotype i.e. Pusa Bold variety (46); while cluster II was further divided into two subclusters viz. IIa and IIb. Subcluster IIa comprised of only two genotypes—BT 15 (4), an indigenous germplasm line and EC 557,025 (10), which is an exotic germplasm line from China. Further, rest of the five exotic accessions showed clustering with Indian genotypes in subcluster IIb. It has been observed that the exotic germplasm lines could not be differentiated well from the indigenous genotypes using even this selected set of 95 SSRs. It may be inferred from the variation obtained in grouping patterns that there is a need to expand and refine the search for more number of suitable markers to develop a comprehensive set of markers for use by the wider Brassica community.

## Population structure analysis, PCoA and AMOVA

Indian mustard genotypes comprising of both Indian and exotic germplasm accessions and popular varieties of India were evaluated for the estimation of population structure using a panel of 310 SSR loci spread across all the 18 chromosomes of Indian mustard. The highest value of delK was at K = 3, inferring the presence of three subgroups or subpopulations into which these 46 accessions were distributed (Fig. 2). These genotypes were further categorized as the pure and admixture type based on the membership fractions. Indian mustard genotypes with probability score $\geq$ 80% are considered as pure type, while with score $\leq$ 80% are considered as admixture type. A high extent of admixture among the subpopulations was found (Fig. 3). It occurs often when the genotypes of one subpopulation share the alleles with other subpopulation (Tyagi et al. 2020). Population structure analysis using the allelic data of selected 95 SSRs also divided all the 46 genotypes into three subgroups with the highest value of delK at K = 3 (Fig. S2). However, there was observed variation in the grouping pattern of genotypes as that obtained using 310 SSR marker data (Fig. S3).

However, it has been found that structure analysis gives more accurate estimate of genetic diversity than UPGMA-based dendrogram. A more clear picture of genetic diversity could be obtained by structure analysis in the present study. In another study carried out by Yan et al. (2019), similar trend of result was reported where they

characterized a coffee germplasm set with ISSR markers. In their study, all the coffee accessions were divided into two major clusters based on UNJ tree, while three subpopulations were obtained using STRUCTURE analysis. Similarly, different grouping pattern was observed for both UPGMA-based dendrogram and structure analysis in sponge guard (*Luffa cylindrica*) accessions on the basis of ISSR and SCoT markers (Tyagi et al. 2020). The principal component (PCoA) analysis contributed to the interpretations of results for genetic diversity and population structure. Figure 4 shows the accessions distribution in a two-dimensional space on the basis of genetic distances obtained using allelic data of 310 SSR markers. On the contrary, PCoA analysis based on the allelic data of 95 selected SSRs exhibited some variations in the distribution of genotypes (Fig. S4). The results of AMOVA showed that 4% of the variation was among the three subpopulations as observed by the analysis of the STRUCTURE program. It was further observed that there was 96% variation among individuals within the subpopulations (Table 3).

## Conclusion

In the present study, we developed a set of 95 polymorphic SSR markers, which can be utilized for molecular characterization of Indian mustard germplasm accessions and varieties. The researchers associated with this crop across the labs can take up the set of SSR markers developed in the present study readily. This marker set will enable the comprehensive analysis of genetic diversity inherent in national and international Indian mustard gene banks and porting of the results of various labs will provide an opportunity to integrate molecular characterization of Indian mustard germplasm collections around the world into a single cohesive international effort to evaluate Indian mustard variability. Further, it has been observed that the selected SSR markers are distributed well across all the eighteen linkage groups of Indian mustard; but still the research efforts should be continued to enrich this marker set by testing more number of markers for various genomic interventions in Indian mustard crop improvement. After the whole genome sequencing of *B. juncea* var. *tumida* and var. Varuna, a big array of SNPs would be available for genomic studies. This set of SSR markers may be augmented with SNPs in future for various genetic mapping, association mapping, DNA fingerprinting and comparative genomics studies.

**Supplementary Information**The online version contains supplementary material available at https://doi.org/10.1007/s13562-021-00737-2.

**Acknowledgements** Authors are thankful to the Science & Engineering Research Board (SERB), New Delhi for providing funding for this work.

**Funding** Funds for this research work was provided under the 'Start-up Research Grant No. SB/YS/LS-86/2014' received by AKT from Science & Engineering Research Board (SERB), New Delhi.

**Availability of data and material** All the data and material is available with the authors.

## Declarations

**Conflict of interest** The authors do not have any conflict of interest.

## References

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

Cheng F, Wu J, Wang X (2014) Genome triplication drove the diversification of Brassica plants. Hort Res 1:14024. https://doi.org/10.1038/hortres.2014.2

Dettori MT, Micali S, Giovinazzi J, Scalabrin S, Verde I, Cipriani G (2015) Mining microsatellites in the peach genome: development of new long-core SSR markers for genetic analyses in five *Prunus species*. Springerplus 4:337. https://doi.org/10.1186/s40064-015-1098-0

Earldent A, Von Holdt Bridgett M (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons Genet Resour 4(359–361):224

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

Govt. of India (2018) Agricultural statistics at a glance. Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, Govt. of India, New Delhi. http://cands.daenet.nic.in/PDF

Li P, Su T, Wang H, Zhao X, Wang W, Yu Y, Zhang D, Wen C, Yu S, Zhang F (2019) Development of a core set of KASP markers for assaying genetic diversity in *Brassica rapa* subsp. *chinensis* Makino. Plant Breed 1:16. https://doi.org/10.1111/pbr.12686

Liu K, Muse M (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinfo 21:2128–2129

Nagaharu U (1935) Genomic analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn J Bot 7:389–452

Nanjundan J, Singh K, Singh KH, Singh D (2014) Catalogue on rapeseed-mustard germplasm. Directorate of Rapeseed-Mustard Research, Bharatpur, Rajasthan, pp 180

Nanjundan J, Singh KH, Thakur AK, Meena KN, Singh D (2015) Development of core collection and trait-specific reference sets in Indian mustard [*Brassica juncea* (L.) Czern & Coss.] germplasm. Abstract 64, 14th International Rapeseed Congress, held during July 5–9, 2015, at Saskatoon, Saskatchewan, Canada

Nanjundan J, Thakur AK, Singh KH, Mishra DC, Singh K, Verma V (2015) Assessment of genetic diversity in Indian mustard [Brassica juncea (L.) Czern & Coss.] for high temperature tolerance using SSR markers. Vegetos 28(4):122–130

Nguyen NN, Kwon YS, Park JR, Sc S (2018) Development of a core set of SSR markers for cultivar identification and seed purity T tests in Oriental Melon (*Cucumis melo* L. var makuwa). Hort Sci

Technol 37(1):119–129. https://doi.org/10.12972/kjhst.20190011

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539

Pratap P, Thakur AK, Meena PD, Meena HS, Sharma P, Singh D, 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

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

Singh BK, Choudhary SB, Yadav S, Malhotra EV, Rani R, Ambawat S, Priyamedha PA, Kumar R, Kumar S, Sharma SK, Singh DK, Rai PK (2018) Genetic structure identification and assessment of interrelationships between Brassica and allied genera using newly developed genic-SSRs of Indian Mustard (*Brassica juncea* L.). Ind Crops Prod 113:111–120

Singh BK, Mishra DC, Yadav S, Ambawat S, Vaidya E, Tribhuvan KU, Kumar A, Kumar S, Kumar S, Chaturvedi KK, Rani R, Yadav P, Rai A, Rai PK, Singh VV, Singh D (2016) Identification, characterization, validation and cross-species amplification of genic-SSRs in Indian mustard (*Brassica juncea*). J Plant Biochem Biotechnol 25:410. https://doi.org/10.1007/s13562-016-0353-y

Tamanna A, Khan AU (2005) Mapping and analysis of simple sequence repeats in the *Arabidopsis thaliana* genome. Bioinformation 1:64–68

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Sci 277(5329):1063–1066

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, 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

Thakur AK, Singh KH, Singh L, Nanjundan J, Rana MK, Singh D (2015) Transferability of SSR markers of Brassica species to some popular varieties of *Brassica juncea*. Proc Natl Acad Sci India Sect B Biol Sci 85(4):1001–1010

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

Xu J, Qian X, Wang X, Li R, Cheng X, Yang Y, Fu J, Zhang S, King GJ, Wu J, Liu K (2010) Construction of an integrated genetic linkage map for the A genome of Brassica napus using SSR markers derived from sequenced BACs in *B. rapa*. BMC Genomics 11:594

Yadava DK, Parida SK, Dwivedi SK, Varshney A, Ghazi IA, Sujata V, Mohapatra T (2009) Cross- transferability and polymorphic potential of genomic STMS markers of Brassica species. J Plant Biochem Biotechnol 18(1):29–36

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

Yang et al (2016) The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. Nat Genet 48:1225. https://doi.org/10.1038/ng.3657

Yang J, Zhang J, Han R, Zhang F, Mao A, Luo J, Dong B, Liu H, Tang H, Zhang J, Wen C (2019) Target SSR-Seq: A novel SSR genotyping technology associated with perfect SSRs in genetic analysis of cucumber varieties. Front Plant Sci 10:531. https://doi.org/10.3389/fpls.2019.00531