



Unraveling the Relationship Between Seed Yield and Yield-Related Traits in a Diversity Panel of *Brassica juncea* Using Multi-Traits Mixed Model

Ranjit Saroj[†], S. L. Soumya[†], Satbeer Singh[‡], S. Mukesh Sankar, Rajat Chaudhary, Yashpal, Navinder Saini, Sujata Vasudev and Devendra K. Yadava^{*}

Division of Genetics, ICAR—Indian Agricultural Research Institute, New Delhi, India

OPEN ACCESS

Edited by:

Petr Smýkal,
Palacký University, Olomouc, Czechia

Reviewed by:

Yong Suk Chung,
Jeju National University, South Korea
Michael Benjamin Kantar,
University of Hawaii, United States

*Correspondence:

Devendra K. Yadava
dkygenet@gmail.com

[†]These authors have contributed
equally to this work and share first
authorship

‡Present address:

Satbeer Singh,
CSIR – Institute of Himalayan
Bioresource Technology, Palampur,
India

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 11 January 2021

Accepted: 16 March 2021

Published: 04 May 2021

Citation:

Saraj R, Soumya SL, Singh S,
Sankar MS, Chaudhary R, Yashpal,
Saini N, Vasudev S and Yadava DK
(2021) Unraveling the Relationship
Between Seed Yield and Yield-Related
Traits in a Diversity Panel of *Brassica
juncea* Using Multi-Traits Mixed
Model. *Front. Plant Sci.* 12:651936.
doi: 10.3389/fpls.2021.651936

The response to selection in any crop improvement program depends on the degree of variance and heritability. The objective of the current study was to explain variance and heritability components in Indian mustard *Brassica juncea* (L.) Czern & Coss to recognize promising genotypes for effective breeding. Two hundred and eighty-nine diverse accessions of Indian mustard belonging to four continents were analyzed for yield and yield-related traits (20 traits) over two seasons (2017–2018 and 2018–2019) using an alpha lattice design. The genetic variance was found to be significant ($P \leq 0.01$) for the individual and under pooled analysis for all of the evaluated traits, demonstrating the presence of significant genetic variability in the diversity panel, which bids greater opportunities for utilizing these traits in future breeding programs. High heritability combined with high genetic advance as percent of mean and genotypic coefficient of variation was observed for flowering traits, plant height traits, seed size, and seed yield/plant; hence, a better genetic gain is expected upon the selection of these traits over subsequent generations. Both correlation and stepwise regression analysis indicated that the main shoot length, biological yield, total seed yield, plant height up to the first primary branch, seed size, total siliqua count, days to flowering initiation, plant height at maturity, siliquae on the main shoot, main shoot length, and siliqua length were the most significant contributory traits for seed yield/plant. Also, promising genotypes were identified among the diversity panel, which can be utilized as a donor to improve Indian mustard further. These results indicated a greater scope for improving seed yield per plant directly through a selection of genotypes having the parsimonious combination of these nine traits.

Keywords: *Brassica juncea*, mixed model, stepwise regression, path analysis, heritability, germplasm

INTRODUCTION

Brassica juncea (L.) Czern & Coss (AABB) is the second most important edible oilseed crop in India after the soybean. *B. juncea* is a natural allotetraploid of two diploid species *viz.*, *Brassica rapa* (AA) and *Brassica nigra* (BB). Rapeseed is a key species from the *Brassica* genus and a high-value crop for oil and biofuel industries. In India, during 2018–2019, rapeseed–mustard was cultivated over an

area of 5.96 million hectares with production and productivity of 8.32 million tons and 1,397 Kg/ha, respectively (Directorate of Economics & Statistics, and Dac&Fw., 2019). Globally, India is the second-largest rapeseed–mustard-cultivating country after China and third in production next to China and Canada (Kumari et al., 2019). In addition to its use as edible oil, mustard oil has a spectrum of industrial utilities such as paint and printing ink additives, greases and lubricants, resins and polymers, plastics, cosmetics, and also in the pharmaceutical industries (Gupta, 2016).

For initiating any genetic improvement program, genetic variability is the prime criterion. Genetic parameters aid in recognizing the gene action and components of genetic variance identification and also facilitate the selection of an appropriate breeding technique. The genotypic and phenotypic variances generally influence the heritability and environmental factors (Bisne et al., 2009). Therefore, the information about heritability and predictability of genetic gains and the association between seed yield and yield-related traits in the base germplasm collection is vital for any genetic improvement program.

Yield is a complex trait and is greatly influenced by environmental factors. Hence, the selection of superior genotypes among a large set of genotypes based on their arithmetic mean performance may not be accurate (Piepho et al., 2008). In this context, the best linear unbiased prediction (BLUP) can provide a good predictive accuracy compared with other procedures for estimating the random effects due to genotypes in a mixed model. It gives a good fit for phenotypic effects to the nongenetic effects by the shrinkage effect toward the probable genetic values. Shrinkage effects by this model anticipate the regression to the mean observed in the selected genotypes, and the individuals having extremely high or low performance consequently got adjusted, thereby improving the accuracy of genotypic effects (Molenaar et al., 2018). Hu (2015) proved that BLUP was effective for calculating genetic parameters and predicting genotypic values and concluded that it could be applied in genetic improvement programs for rapeseed–mustard.

The yield of a crop is directly or indirectly influenced by various yield-contributing traits such as seed size, primary and secondary branches per plant, length of siliqua, seeds per siliqua, etc. Hence, plant breeders often focus on the selection of such traits in combination, each of which was assigned to have a certain level of economic weight based on their importance toward seed yield to form a selection index (Smith, 1936; Hazel, 1943). Multivariate analysis methods, such as genetic correlation analysis, stepwise multiple regression analyses, and path analysis, have been utilized in several crops, including mustard, to identify the causal traits having either direct or indirect effect on seed yield (Olivoto et al., 2016). The path coefficient analysis provides accurate information about the relationship of direct and indirect effects of variables by splitting the correlation coefficients. Therefore, the contribution of each character to yield could be assessed for selecting appropriate traits for indirect selection in any breeding technique (Rao et al., 2013). In contrast, studies using mixed models and sequential path analysis to identify the relationship of cause and effect considering genotypic values in *B. juncea* are still very scarce.

In this context, the current study was carried out with the objectives (i) to use restricted maximum likelihood/BLUP-based method to assess variance, genetic parameters, and genotypic performance of mustard genotypes in multiyear trials, (ii) to fit stepwise regression model for identifying highly significant traits to form a path diagram that explains the relationship of cause and effect among seed yield-related traits, and (iii) to group the germplasm lines sharing a common attribute based on Mahalanobis distance. This study unravels the nature of genetic variability in Indian mustard and would be helpful in the selection of superior genotypes for yield and related traits, which further augment the ongoing and future mustard breeding programs.

MATERIALS AND METHODS

Source of Germplasm and Experimental Location

Two hundred and eighty-nine diverse accessions of *B. juncea* germplasm obtained from Punjab Agricultural University, Ludhiana, under Indian Council of Agricultural Research (ICAR)—National Agricultural Science Fund-sponsored project, including varieties, cultivars, introgression lines, derived lines, and exotic and indigenous collections from the diverse origin such as from India, Australia, Europe, Germany, and Canada were evaluated for phenological and morphological traits under timely sown irrigated conditions. The details of the germplasm accessions used are presented in **Supplementary Table 1**. The experiments were conducted at ICAR—Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India (latitude—28.708°N, longitude—77.108°E, and altitude—219 m) during 2017–2018 and 2018–2019 winter seasons (October to March).

Experimental Layout and Observation Recording

Trials were laid out in a randomized alpha lattice design with two replications; each plot consisted of four rows of 2-m length in a plot size of 2.4 m². Five representative plants from each treatment were selected from the middle two rows for evaluation of agronomic performances. Recommended agronomic practices were followed in both seasons. Data of five characters *viz.* days to flowering initiation (DFI), days to 50% flowering (DFF), days to 100% flowering (DHF), days to flowering completion (DCF), and days to maturity (DMT) were recorded on a plot basis. Morphological data of 15 characters recorded on five plants each include plant height at flowering (PH_Fl), plant height up to the first primary branch (PH_FPb), plant height at maturity (PH_M), number of primary branches (PB), number of secondary branches (SB), main shoot length (MSL), siliquae on the main shoot (SMS), total siliquae count (TSC), siliqua length (SL), seeds per siliqua (SPS), seed size (SS), seed yield/plant (SY/Plant), total seed yield/plot (TSY/Plot), biological yield/plot (BY/Plot), and harvest index (HI).

Statistical Analysis of Phenotypic Data

Analysis of Variance

For each given trait, plot-level averages of both seasons were taken as the response variable in an iterative mixed linear model fitting procedure by the full model (Eq. 1) in lme4 R-package (Bates and Maechler, 2009). The best-fit model for each agronomic trait was attained by removing all random terms from the model that were not significant at $\alpha = 0.05$ in a likelihood ratio test (Littell et al., 2006). Three variance components (σ^2_g , σ^2_{gy} , and σ^2_e) for each of the 20 traits were calculated using the restricted maximum likelihood (Patterson and Thompson, 1971) estimation method. In the current study, the year was fitted as fixed effect, and genotypes, blocks, replications, and genotype relationship with year were fitted as random effects. The phenotypic results z_{ijkl} on accession m in replication k of block l and year i was displayed as:

$$z_{ijkl} = \mu + y_i + g_j + r_{ik} + b_{ikl} + (gy)_{ij} + \varepsilon_{ijkl} \quad (1)$$

where μ is the grand mean; y_i is the fixed effect of year i ; g_j is the random effect of genotype, j and is $\sim \text{NID}(0, \sigma^2_g)$; r_{ik} is the random effect of replication, k in year i and is $\sim \text{NID}(0, \sigma^2_r)$; b_{ikl} is the random effect of block l nested with replication k in year i and is $\sim \text{NID}(0, \sigma^2_b)$; $(gy)_{ij}$ is the random effect of the relations between genotype j and year i and is $\sim \text{NID}(0, \sigma^2_{gy})$; and ε_{ijkl} is random residual effect and $\sim \text{NID}(0, \sigma^2_\varepsilon)$. Diagnostic residual plots were used to check the normality and homogeneity of the response variable. If the residuals from the fitted model did not meet the assumptions, data were subjected to transformation. This final model was utilized to generate the BLUP for each genotype.

Estimation of Heritability and Genetic Parameters

The heritability parameter across the year was estimated by analysis of variance using the ratios of Hallauer and Miranda (1988). The genetic advance was estimated for traits using the formula given by Johnson et al. (1959). The genetic advance as percentage of mean was assessed, as defined by Souza et al. (2009). The phenotypic (PCV) and genotypic coefficients of variation (GCV) were calculated, according to Singh and Chaudhary (2004). Pearson's correlation coefficient (r) between BLUP values for each pair of genotypic traits was estimated using the "corr" package (Version 0.4.3; Ruiz et al., 2019) in R version 3.5.1 (R Core Team, 2018). To identify the most influential agronomic traits with respect to seed yield/plant as a dependent variable, a stepwise regression model was fitted. The independent variables with the highest share in explaining the variations of the dependent variable were recognized using PAST version 3.09 software. The significance level of a term in the regression model was 5%. R-programs—"agricolae" (de Mendiburu and de Mendiburu, 2019), "Hmisc" (Harrell and Harrell, 2019), and "diagram" (Soetaert, 2009) packages were utilized for path analysis. The divergence was estimated based on the predicted mean (BLUP_s) values of 20 characters pooled over the years, and the residual variance-covariance matrix generated using vcov function of lme4 package was subjected to grouping using the D^2 statistic according to Mahalanobis (1936) and extended by Rao

(1952). Mahalanobis's distance matrix thus obtained was further subjected to clustering by Ward2 hierarchical agglomerative clustering method (Murtagh and Legendre, 2014). R statistical software packages such as "biotools" (da Silva and da Silva, 2017), "dendextend" (Galili, 2015), "circlize" (Gu et al., 2014), "plotrix" (Lemon et al., 2015), "qgraph" (Epskamp et al., 2017), and "car" (Fox et al., 2007) were used for divergence studies based on 20 different agro-morphological traits.

RESULTS

Meteorological observations recorded during the season (October–March) are represented in **Figures 1A,B**. All weather parameters except rainfall were recorded as means over the crop growing period, October–March. Rainfall was recorded as cumulative rainfall received during the period. The average maximum and minimum temperatures were 25.6 and 23.8°C, and 8.9 and 9.2°C during the 2017–2018 and 2018–2019 crop seasons, respectively. During 2017–2018, rainfall and sunshine hours were 6.0 mm and 5.5 h, respectively, whereas, during the 2018–2019 crop season, the average rainfall and sunshine hours were 138.4 mm and 4.8 h, respectively.

Mean Performance and Variation for Phenotypic Traits

Extensive phenotypic variation was observed for the seed yield and related components under normal sown conditions during both seasons. A large range of variation was observed for most of the traits under study, with the coefficients of variation (CV) ranging from 1.80% for days to maturity to 30.0% for harvest index under normal sown conditions. High CV for some of the yield traits such as TSC, SY/P, TSY/Plot, BY/Plot, and HI was due to longer duration and photoperiod sensitivity of some Canadian accessions—CN-105305, CN-34005, and CN-34008. Most of the traits were approximately normally distributed apart from flowering traits and PH_FPB (**Figure 2A**). Also, the acceptable level for each trait was indicated by green color on histogram based on the ideotypic concept in mustard given by Bhargava et al. (1984); Thurling (1991), VijayaKumar et al. (1996), Yadav et al. (2017), and DUS guidelines given by Protection of Plant Variety and Farmers Right Authority of India. The boxplots obtained between seasons for each trait were compared using Wilcoxon statistic, and the corresponding level of significance was shown by p -values in figures. The analysis indicated a significant mean difference between seasons for each trait. The mean values of all traits except siliquae on the main shoot and harvest index were slightly higher in the season 2017–2018 than 2018–2019 (**Figure 2B**). The mean performance of SY/P during 2017–2018 was 20.38 ± 0.4 g, whereas, during 2018–2019, it was 15.38 ± 0.3 g. Based on the mean performance, IC-597867 yielded the highest seed yield/plant of 78.5 g per plant, and CN-34005 has no yield during 2017–2018. Similarly, IC-597867 remains the highest yielder of 59.5 g per plant, whereas CN-105364 yielded the lowest of 1 g/per plant. The best performing genotype across the year was IC-597867, whereas CN-105364 was the least performer.

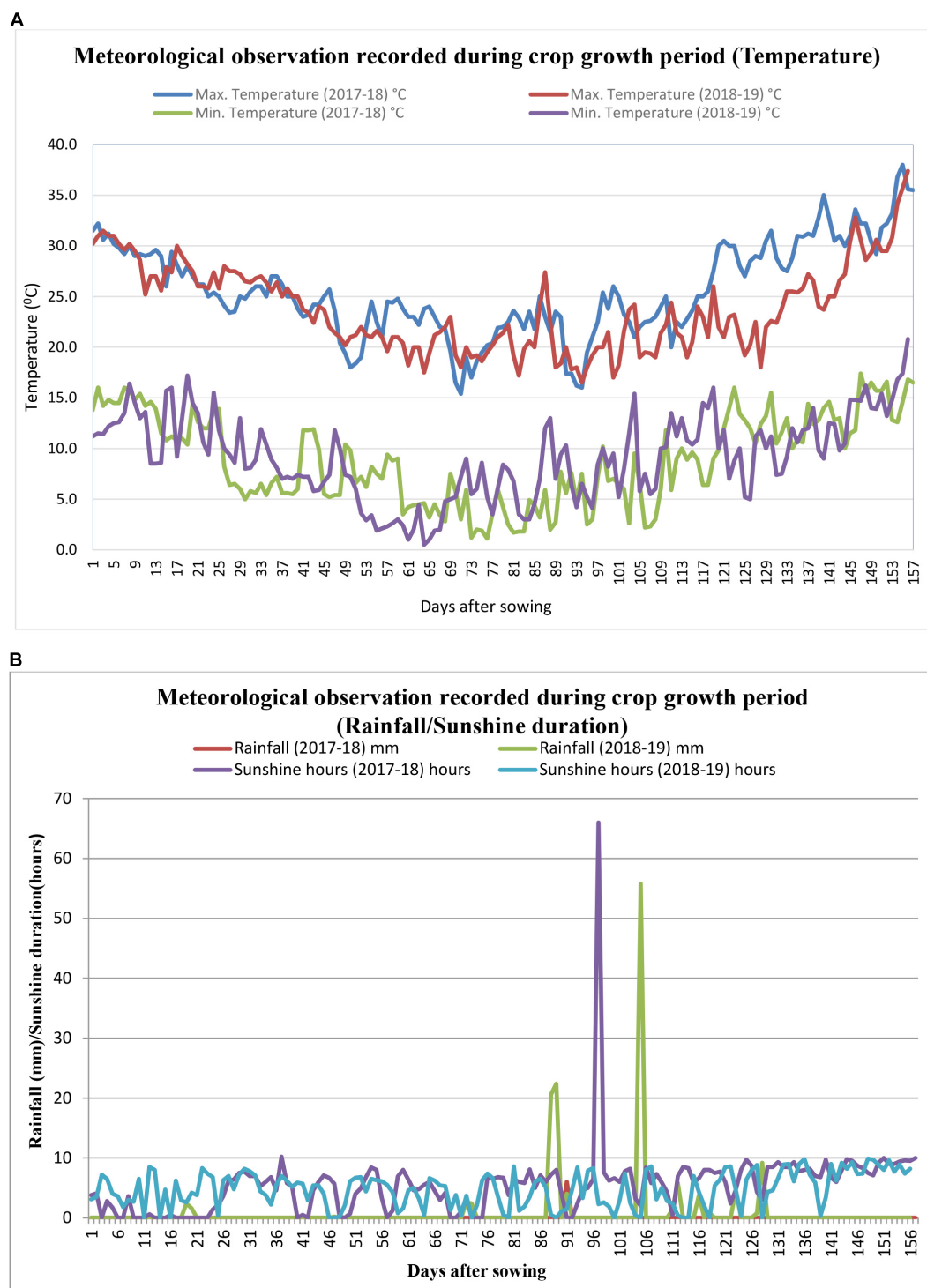


FIGURE 1 | (A) Daily maximum and minimum temperatures (°C) recorded during the crop grown period (2017–2018 and 2018–2019). **(B)** Rainfall (mm) and sunshine hours recorded during the crop grown period (2017–2018 and 2018–2019).

An overview of the agro-morphological traits recorded pooled over 2 years is shown in **Table 1**.

Significant effects of the sources of variation are presented in **Table 2**. The effect of genotypes was significant ($P \leq 0.01$)

for all the studied traits for both years, representing the presence of considerable genetic variability in the germplasm for all the traits. Significant $G \times E$ interactions were observed for all the traits during pooled analysis,

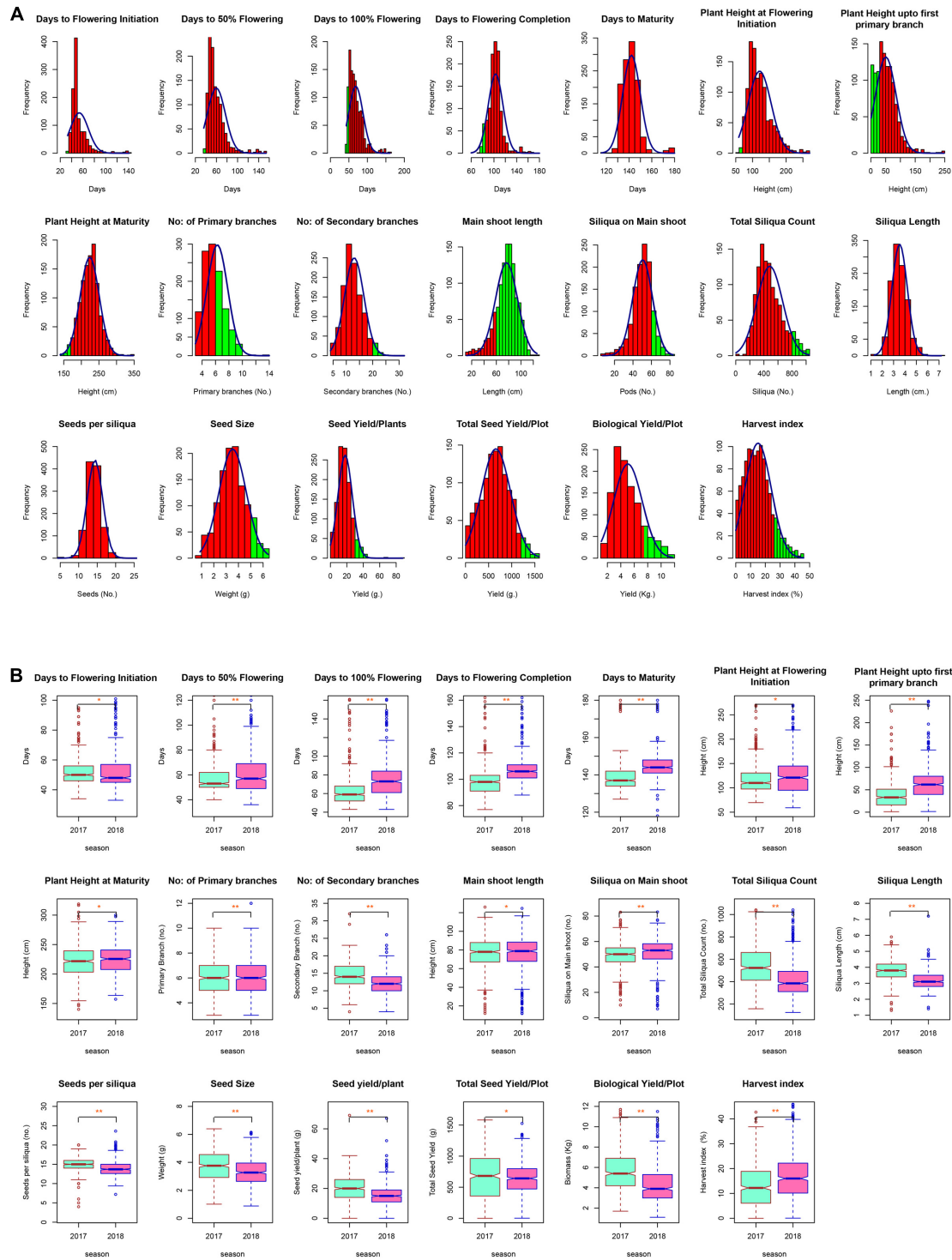


FIGURE 2 | (A) Histogram showing distribution of different agro-morphological traits across the diversity panel. Acceptable values for each trait were chosen based on various ideotype concepts given for oilseed *Brassica*, and acceptance region was indicated by green region on the histogram. **(B)** Notched box plots showing difference of 20 traits contributing to grain yield during 2017–2018 and 2018–2019 in New Delhi. Box edges represent upper and lower quartile, with median value shown as a bold line in the middle of the box. Whiskers represent 1.5 times the quartile of the data. Individuals falling outside the range of the whiskers shown as open dots. Boxplot obtained between seasons for each trait was compared using Wilcoxon statistic, and corresponding level of significance was shown by p -values in codes (** $p \leq 0.01$, * $p \leq 0.05$, NS, not significant).

TABLE 1 | An overview of the agro-morphological traits recorded pooled over 2 years.

Parameters	Mean	Range				SE(m)	CV	LSD	No. of genotypes above trial mean
		Min. value	Genotype	Max. value	Genotype				
Days to flowering initiation	53.5	33.5	DTM–25	139	CN–105364	0.47	5.23	10.67	94
Days to 50% flowering	59.71	38.5	DTM–25	153	CN–105364	0.51	4.85	11.17	107
Days to 100% flowering	68.69	44.0	DRMRIJ 17–46	160.5	CN–105364	0.57	5.41	12.47	120
Days to flowering completion	102.88	84.75	EJ–17	166	CN–105364	0.37	3.69	7.42	133
Days to maturity	142.12	126.5	IC–261627	179	CN–105364	0.23	1.80	4.63	121
Plant height at flowering	121.09	77.9	GR–325	268.9	CN–105379	1.01	9.45	24.99	131
Plant height up to the first primary branch	50.58	7.8	NRCHB–101	225.3	CN–105379	1.03	19.69	24.99	127
Plant height at maturity	223.74	165.56	IC–261627	320.3	CN–105379	0.80	5.25	20.4	151
Number of primary branches	6.27	4	PBR–378	11	AJ–2	0.05	14.53	1.07	138
Number of secondary branches	13	5	CN–105305	20.25	AJ–2	0.11	18.10	1.8	136
Main shoot length	76.53	14	CN–105257	107.06	RNN–505	0.53	11.26	12.61	151
Siliquae on the main shoot	50.46	12.85	CN–105257	74.3	IC–597885	0.31	13.19	7.68	158
Total siliquae count	483.56	173	CN–105305	904.5	AJ–11	5.39	24.12	116.7	135
Siliqua length	3.53	1.4	CN–105257	5.1	NRCDR–2	0.02	7.62	0.41	137
Seeds per siliqua	14.38	8.65	CN–105379	19.7	RE–35–4	0.06	11.19	1.25	152
Seed size	3.53	1.1	CN–113780	6.22	RH–406	0.03	6.52	0.76	143
Seed yield/plant	17.88	1.3	CN–34008	69.67	IC–597867	0.26	27.78	6.23	140
Total seed yield/plot	657.15	34.5	CN–34008	1275	RGN–73	9.36	23.13	220.94	148
Biological yield/plot	5.06	1.8	CN–113780	10.28	CS–54	0.06	25.95	1.28	148

SE(m) is the standard error of mean, CV is coefficient of variation, and LSD is least square difference.

suggesting different genotypes response to environmental conditions differentially.

Estimation of Heritability and Genetic Parameters

The phenotypic component of the variance was divided into genotypic variance (σ^2_g), $G \times E$ variance (σ^2_{gy}), and error variance (σ^2_e). Furthermore, genotypic and $G \times E$ variances were compared with total phenotypic variance to identify the magnitude of genotypic contribution for *Brassica* improvement. In the present study, most of the traits were highly heritable (>0.60) as per the scale of Robinson (1966) in combined environments except SB and HI, which showed low heritability (>0.30). Similarly, PB, TSC, SPS, TSY/Plot, and BY/Plot showed moderate heritability (Table 2). The estimates of broad-sense heritability in pooled data ranged from 0.22 (SB) to 0.86 (DCF). Traits such as flowering, plant height-related, siliquae on the main shoot, siliqua length, seed size, and seed yield/plant were found to be more heritable. For seed yield/plant, the heritability was high in first year (2017–2018) compared with second year (2018–2019), whereas combined analysis resulted in the lowest value for heritability, indicating the significant partitioning of $G \times E$ variance from genetic variance for these traits obtained in the individual environment.

The PCV and GCV, genetic advance, and GA as % mean were calculated along with heritability for all the traits (Table 2). The highest GCV and PCV were observed for plant height up to the first primary branch (46.43 and 64.38%, respectively), and the lowest GCV and PCV were recorded for days to maturity (4.09

and 5.01%, respectively). Results suggested a narrow difference between GCV and PCV for highly heritable phenological traits such as DFI, DFF, DHF, DCF, DMT, PH_FI, PH_M, MSL, etc., which can be improved directly by selecting genotypes having a higher trait value. However, the difference is more prominent for traits with low heritability, such as PB, SB, TS, BY, SY/P, and TSY/Plot, indicating the significant influence of environment and $G \times E$ interactions. In all these traits, PCV was considerably higher than GCV. The genetic advance as a percentage of the mean ranged from 8.6% in DMT to 94.0% in PH_FP. The results showed that selecting the top 5% of the genotypes could result in genetic improvement of 51.6% for DFI, 50.4% for DFF, and so on (Table 2). The current study found that among the parameters under study, high heritability ($\geq 60\%$), genetic advance as percentage mean ($>20\%$), PCV ($>20\%$), and GCV ($>20\%$) were observed in characters such as days to flowering initiation, days to 50% flowering, days to 100% flowering, plant height at flowering, plant height up to the first primary branch, seed size, and seed yield/plant. High PCV, GCV, and genetic advance values with low and moderate heritability were observed for total seed yield/plot, biological yield/plot, and harvest index.

Association Among Traits and Their Contribution Toward Seed Yield per Plant

The utility of independent traits in the selection can be expected by their significant association with seed yield (dependent trait). In the present study, genotypic correlations between 20 character pairs were studied in all possible combinations (Figure 3). The prime economic trait, seed yield per plant,

TABLE 2 | Genotypic variance (σ^2_g), genotypic–year interaction variance (σ^2_{gy}), standard error (SE), and heritability in broad-sense (H^2) estimated over each season and pooled data along with pooled mean values, genetic and phenotypic coefficients of variation (GCV and PCV), genetic advance (GA), and genetic advance as percentage of mean for the 20 traits studied.

Characters	2017–2018			2018–2019			Pooled analysis										
	σ^2_g	SE	H^2	σ^2_g	SE	H^2	σ^2_g	SE	σ^2_{gy}	SE	H^2	Mean	σ^2_p	GCV	PCV	GA	GA% mean
DFI	204.11**	8.49	0.98	292.47**	12.17	0.99	181.12**	5.33	66.07**	1.94	0.84	53.5	255.47	25.15	29.87	27.6	51.6
DFF	232.65**	9.68	0.98	338.66**	14.09	0.99	214.02**	6.29	71.29**	2.1	0.85	59.71	294.84	24.5	28.76	30.1	50.4
DHF	275.26**	11.45	0.98	384.12**	15.98	0.98	240.54**	7.07	89.13**	2.62	0.83	68.69	344.92	22.58	27.04	31.9	46.4
DCF	132.71**	5.52	0.93	113.30**	4.71	0.96	97.78**	2.88	25.62**	0.75	0.86	102.88	138.48	9.61	11.44	20.8	20.2
DMT	47.13**	1.96	0.94	40.01**	1.66	0.92	33.72**	0.99	9.77**	0.29	0.84	142.12	50.69	4.09	5.01	12.3	8.6
PH_FI	809.98**	33.69	0.92	1258.82**	52.36	0.95	680.44**	20.01	354.17**	10.42	0.76	121.09	1174.88	21.54	28.31	54	44.6
PH_FPB	740.72**	30.81	0.93	1170.72**	48.7	0.96	551.59**	16.22	404.01**	11.88	0.71	50.58	1060.49	46.43	64.38	47.5	94
PH_M	611.16**	25.42	0.87	559.88**	23.29	0.93	350.78**	10.32	236.86**	6.97	0.70	223.74	732.96	8.37	12.1	38.8	17.4
PB	0.99**	0.04	0.67	1.76**	0.07	0.84	0.56**	0.02	0.81**	0.02	0.48	6.27	2.31	11.94	24.23	1.5	23.8
SB	4.73**	0.2	0.55	7.08**	0.29	0.80	1.07**	0.03	4.82**	0.14	0.22	13	11.83	7.95	26.45	1.6	12
MSL	190.98**	7.94	0.78	314.70**	13.09	0.94	180.52**	5.31	68.76**	2.02	0.77	76.53	323.57	17.56	23.51	28.7	37.4
SMS	55.39**	2.3	0.68	78.67**	3.27	0.81	41.52**	1.22	25.48**	0.75	0.64	50.46	112.09	12.77	20.98	13.9	27.5
TSC	13826.75**	575.12	0.60	15146.49**	630.01	0.78	6653.75**	195.7	7807.28**	229.63	0.48	483.56	28794.07	16.87	35.09	166.6	34.5
SL	0.32**	0.01	0.89	0.23**	0.01	0.87	0.20**	0.01	0.07**	0	0.79	3.53	0.36	12.67	17.02	1	27.7
SPS	2.52**	0.1	0.81	0.71**	0.03	0.26	0.63**	0.02	0.99**	0.03	0.36	14.38	4.23	5.52	14.3	1.5	10.5
SS	1.19**	0.05	0.96	1.06**	0.04	1	0.78**	0.02	0.34**	0.01	0.81	3.53	1.18	25.03	30.8	1.8	51.4
SY/Plant	62.04**	2.58	0.81	29.99**	1.25	0.74	26.67**	0.78	19.64**	0.58	0.63	17.88	71.26	28.88	47.21	10.9	60.8
TSY/Plot	110225.70**	4584.78	0.88	42614.75**	1772.54	0.83	19844.99**	583.68	56780.60**	1670.02	0.37	657.15	101648.9	21.44	48.52	241.3	36.7
BY/Plot	2.16**	0.09	0.65	1.77**	0.07	0.76	0.70**	0.02	1.26**	0.04	0.40	5.06	3.93	16.54	39.21	1.6	32
HI	47.96**	1.99	0.81	45.42**	1.89	0.79	8.06**	0.24	38.96**	1.15	0.24	15.09	74.8	18.82	57.32	4.3	28.5

**is the level of significance at 1%.

DFI, days to flowering initiation; DFF, days to 50% flowering; DHF, days to 100% flowering; DCF, days to flowering completion; DMT, days to maturity; PH_FI, plant height at flowering (PH_FI), PH_FPB, plant height up to first primary branch; PH_M, plant height at maturity; PB, number of primary branches; SB, number of secondary branches; MSL, main shoot length; SMS, siliquae on the main shoot; TSC, total siliquae count; SL, siliqua length; SPS, seeds per siliqua; SS, seed size; SY/Plant, seed yield/plot; TSY/Plot, total seed yield/plot; BY/Plot, biological yield/plot; HI, harvest index.

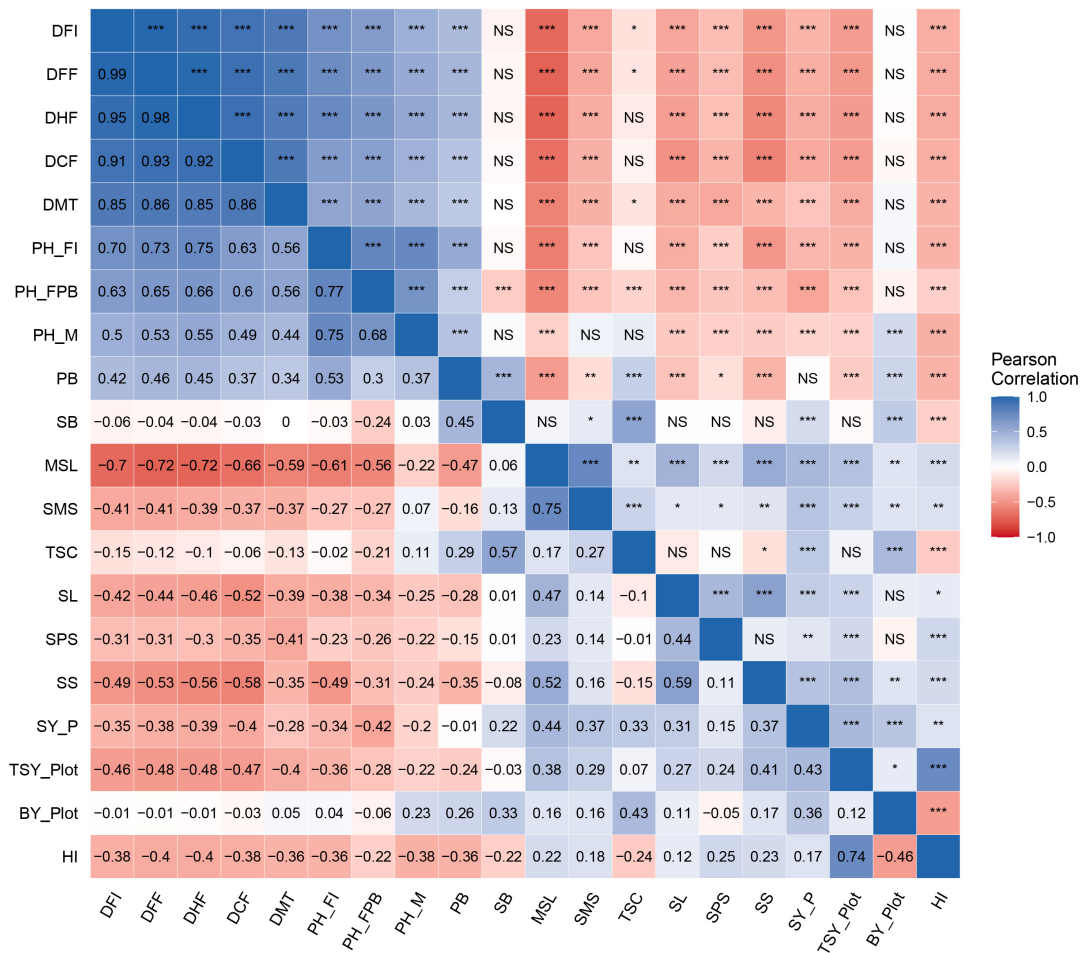


FIGURE 3 | Heat map of genotypic correlation matrix for yield and yield-contributing traits in *Brassica juncea* genotypes. Scale on the side of the figure indicates magnitude and direction of phenotypic correlations. Shades of blue from lighter to darker indicates strength of positive correlation between pairs of traits. Shades of red from lighter to darker indicates strength of negative correlation between pairs of traits. Lighter to white color indicates very weak or no correlation between pair of traits. Pearson's correlation coefficients (lower diagonal) and significance (upper diagonal) among different traits measured in this study for genotypic BLUP values. Significance codes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, NS, not significant.

was positively and significantly correlated ($p \leq 0.01$) with total seed yield per plot ($r = 0.43$), seed size ($r = 0.37$), main shoot length ($r = 0.44$), siliquae on main shoot ($r = 0.37$), siliqua length ($r = 0.31$), seeds per siliqua ($r = 0.15$), total siliquae count ($r = 0.33$), number of secondary branches ($r = 0.22$), biological yield per plot ($r = 0.36$), and harvest index ($r = 0.17$). Seed yield per plant exhibited a significant negative correlation with flowering characters such as days to 50% flowering ($r = -0.38$), days to 100% flowering ($r = -0.39$), days to flowering completion ($r = -0.40$), days to flowering initiation ($r = -0.35$), days to maturity ($r = -0.28$), and plant height characters such as plant height at flowering initiation ($r = -0.34$), plant height up to the first primary branch ($r = -0.42$), and plant height at maturity ($r = -0.20$), whereas number of primary branches showed a nonsignificant relationship.

A stepwise regression model with seed yield/plant as a dependent variable and other traits as the independent variable

identified BY/Plot, TSY/Plot, PH_FPB, SS, TSC, DFI, PH_M, SMS, and SL as highly significant traits. However, keeping seed yield per plant as a dependent variable, these traits with compounded effect explained approximately 46% of total variance with an R^2 value of 0.46 (Table 3).

The path analysis (Table 4) showed that the total seed yield/plot had the highest positive direct effect ($\rho_{X1} = 0.242$) followed by siliquae on the main shoot ($\rho_{X4} = 0.225$) and seed size ($\rho_{X5} = 0.215$). However, this trait had an important negative direct effect through plant height up to the first primary branch ($\rho_{X2} = -0.185$) and plant height at maturity ($\rho_{X9} = -0.150$). Siliqua length and biological yield/plot showed a positive indirect effect on yield ($\rho = 0.216$ and $\rho = 0.203$, respectively). However, days to flowering initiation ($\rho = -0.350$) and plant height up to the first primary branch ($\rho = -0.235$) contributed with negative indirect effects on seed yield. The present study also entrusted to identify promising genotypes for significantly associated traits identified by regression and

path analysis (Table 5). Some of the genotypes showed superior performance for more than one trait, viz. IC-597867 is found to be the potential donor for traits such as basal branching, total siliqua count, and seed yield per plant. Similarly, PBR-210 and RE-7-1 were found to be superior for total seed yield per plot and seed yield per plant. These trait donors can be further utilized for developing better genotypes through systematic hybridization.

Diversity Analysis

Cluster analysis based on phenotypic data revealed that genotypes were distributed into five different clusters (I to V) at Mahalanobis D^2 value of 60.0 (Figure 4A). D^2 distance reported in the present study ranges from 0 to 351.38, also indicated huge genetic diversity in the population. Cluster I had 281

genotypes followed by clusters II, III (three genotypes each), IV, and V (one genotype each). Again, cluster I was found to have around 26 subclusters (a to z) at Mahalanobis D^2 value of 20.0 (Figure 4 and Supplementary Table 2). Subcluster Ia contains three genotypes from the Australian gene pool; subcluster Ib contained two genotypes that were derivative of resynthesized *B. juncea*. Subcluster Ic consists of a breeding line from east European countries (2), and subcluster I (d-i) is dominated by high-yielding cultivars and advanced breeding lines of Indian origin. Most of the indigenous collection (IC) based on *B. juncea* var. *rugosa* and Canadian gene pool got together in subcluster I (j-z) (Figure 4B).

The relative contribution of each character indicated that seed size (13.14%), days to 50% flowering (10.85%), and siliqua length (7.88%) contributed maximum toward total divergence (Table 6). All the popular Indian cultivars got included in cluster I. Furthermore, it was explained that cluster I contained accessions producing the highest total seed yield per plot and seeds per siliqua, which are early and require a minimum for flowering traits. Genotypes belonging to cluster II were found to have a longer duration for flowering traits with higher harvest index and lowest values for several secondary branches, main shoot length, and total siliqua count. Cluster III is found to have genotypes having the higher plant height and biological yield per plot, whereas lower values for SL, SPS, and HI. Cluster IV contained a single genotype (IC-597867) having the highest seed yield per plant with higher MSL, SPS, and TSC with smaller seed size and shorter plant height up to the first primary branch. Similarly, cluster V contained a genotype (CN-34005) having maximum values for SS and SL and minimum values for PH_FL, PH_M, PB, SY/P, TSY/Plot, and BY/Plot.

Although genotypes in cluster I were from the different geographical locations, genotypes belonging to the same location or same breeding program that tend to fell together in the same subclusters such as DRMRIJ-17 series, i.e., 17-41,17-42, 17-43, and 17-52 derived from ICAR-Directorate of Rapeseed and Mustard Research, Bharatpur, fell on subcluster Id of cluster I. Similarly, the indigenous collections from Arunachal Pradesh,

TABLE 3 | Summary of stepwise regression analysis considering seed yield per plant as a dependent variable to determine the significance and relative contribution of other traits under normal sown condition.

Variable	Method	AIC	RSS	Sum Sq	R^2	Adj R^2
MSL	+	-59.96	231.62	56.38	0.20	0.19
BY/Plot	+	-89.87	207.41	24.21	0.28	0.27
TSY/Plot	+	-118.50	186.55	20.86	0.35	0.35
PH_FPB	+	-134.41	175.34	11.21	0.39	0.38
SS	+	-146.39	165.91	2.54	0.42	0.41
TSC	+	-151.25	162.02	3.89	0.44	0.42
DFI	+	-153.34	159.74	2.28	0.45	0.43
PH_M	+	-155.37	157.53	2.21	0.45	0.44
SMS	+	-156.83	155.66	1.88	0.46	0.44
MSL	-	-157.79	156.22	0.56	0.46	0.44
SL	+	-158.64	154.68	1.53	0.46	0.44

"+" means forward selection, and "-" means backward selection.

AIC, Akaike information criterion; RSS, residual sum of squares; R^2 , coefficient of determination; MSL, main shoot length; BY/Plot, biological yield/plot; TSY/Plot, total seed yield/plot; PH_FPB, plant height up to first primary branch, SS, seed size; TSC, total siliqua count; DFI, days to flowering initiation; PH_M, plant height at maturity; SMS, siliqua on the main shoot; SL, siliqua length.

TABLE 4 | A path coefficient showing total, direct, and indirect effect of various key traits identified on seed yield per plant in mustard.

Traits	Indirect effect to SY/Plant									Total effects		Total correlation to SY/Plant (r)
	TSY/Plot	PH_FPB	BY/Plot	SMS	SS	TSC	DFI	PH_M	SL	Direct	Indirect	
TSY/Plot (X_1)		0.05	0.02	0.07	0.09	0.01	-0.08	0.03	0.03	0.24	0.19	0.43
PH_FPB (X_2)	-0.07		-0.01	-0.06	-0.07	-0.04	0.10	-0.10	-0.03	-0.19	-0.24	-0.42
BY/Plot (X_3)	0.03	0.01		0.04	0.04	0.08	0.00	-0.03	0.01	0.16	0.20	0.36
SMS (X_4)	0.07	0.05	0.03		0.03	0.05	-0.07	-0.01	0.01	0.23	0.15	0.37
SS (X_5)	0.10	0.06	0.03	0.04		-0.03	-0.08	0.04	0.06	0.22	0.16	0.37
TSC (X_6)	0.02	0.04	0.07	0.06	-0.03		-0.03	-0.02	-0.01	0.19	0.14	0.33
DFI (X_7)	-0.11	-0.12	0.00	-0.09	-0.11	-0.03		-0.08	-0.04	0.16	-0.35	-0.19
PH_M (X_8)	-0.05	-0.13	0.04	0.02	-0.05	0.02	0.08		-0.02	-0.15	-0.05	-0.20
SL (X_9)	0.07	0.06	0.02	0.03	0.13	-0.02	-0.07	0.04		0.09	0.22	0.31

SY/P, seed yield/plant; TSY/Plot, total seed yield/plot; PH_FPB, plant height up to first primary branch; BY/Plot, biological yield/plot; SS, seed size; TSC, total siliqua count; DFI, days to flowering initiation; PH_M, plant height at maturity; SL, siliqua length.

such as IC-597870, IC-597871, and IC-597881 (subcluster Io) and IC-597904 and IC-597949 (subcluster In), and most of the Canadian gene pool from subclusters I (t-y) followed the same pattern as the DRMRIJ series. In contradictions to the observation discussed earlier, a few accessions collected from the same region, such as Canadian germplasm, did not fall in a single cluster or subcluster, indicating that geographical proximity does not always result in genetic similarity.

Based on the similarity matrix, the distance between clusters, inter-cluster *D*-values ranged from 8.30 to 15.68 (Figure 5). The higher inter-cluster distances than the intra-cluster distances designated wider genetic diversity of different groups among the genotypes. Clusters II and IV were strikingly diverse from the rest of the clusters (inter-cluster *D*-value = 15.68); therefore, intercrossing the genotypes from these two clusters may create wider variability and is estimated to throw high yielding transgressive segregants in the mustard breeding program. The minimum inter-cluster *D*-value (8.30) detected between clusters I and IV showed the higher genetic similarities between these clusters. Intra-cluster distance (*D*) revealed that cluster II showed maximum intra-cluster distance (6.29) followed by cluster III (4.68). Owing to solitary genotype, clusters IV and V did not show intra-cluster distance. The magnitudes of the intra-cluster distances were not always proportional to the number of genotypes in the cluster, as intra-cluster distance in cluster I was found to be moderate (4.12) regardless of maximum genotypes (281).

DISCUSSION

The exploitation of genetic diversity present in a species can lead to the improvement of traits of the economic importance of mandated crops with the intervention of plant breeders to benefit the farmers and consumers (Salgotra et al., 2015). Yield is one of the most important economic traits and is the product of multiplicative interactions of contributing characters (Kant and Gulati, 2001). To combat these complex interactions, we need to have a multipronged strategy by combining agronomical and breeding approaches. Hence, the major objective of the mustard improvement program is to develop varieties with high yield potential through the introgression of various yield component traits from the lines with high trait values. Hence, exploitation of germplasm lines to identify lines with higher trait values is of prime importance.

Brassica juncea is a major oilseed crop of *Rabi* (winter) season and is highly sensitive to weather parameters such as temperature and solar radiation, which affect growth, phenological events, and crop yield (Kumar, 2005). Intermittent rains during the flowering time of 2018–2019 caused substantial yield losses by physiological disorder and biotic stresses. Therefore, a large variation in yields from year to year can be attributed to the weather conditions. Mustard prefers moderate temperatures between 18 and 25°C with an optimum around 20°C and moderate rain of approximately 25–40 cm during the growth period (Bhatt et al., 2015). The sensitive periods for mustard crop growth signify emergence, flowering, siliqua formation, siliqua filling,

and physiological maturity. The analysis revealed that maximum and minimum temperatures had a positive effect on the yield during the sensitive period in both seasons, whereas total rainfall had a negative effect on the mustard yield during 2018–2019. The intermittent rainfall resulted in high RH (>92%) with T_{max} ranging from 18.7 to 24.4°C recorded from 4 to 10 standard meteorological weeks during the reproductive period of the plant (88–128 DOS) resulted in subsequent yield loss.

Mean Performance and Variation for Phenotypic Traits

Seed yield and related traits showed wide phenotypic variations during both seasons. The mean performance indicated the existence of enormous variability for the seed yield and related components, which offer greater opportunities for utilizing these traits in future breeding programs (Kumar et al., 2020). The genetic variation available for traits such as total seed yield per plot, total siliquae count, plant height at first branching, etc., can meet the breeding objective in evolving a high-yielding *B. juncea* cultivar. The greater variability observed in the present study could be due to the use of genotypes from diverse geographical origins. The lowest coefficient of variation for the number of days to maturity (1.80%) showed its best genetic potential and genetic influence, whereas the highest coefficient of variation for harvest index (30.0%) showed more influence of environmental fluctuations (Khan et al., 2008). Some of the yield traits such as total siliqua count, seed yield per plant, total seed yield per plot, biological yield, and harvest index showed a greater CV above 20%, which may be due to the longer duration and photoperiod sensitivity of some exotic lines, especially the Canadian lines. Characters with extensive genetic variability provide a better opportunity for selection instead of those with a narrow range of variability. Ali et al. (2003); Yadava et al. (2011), and Roy et al. (2018) had also found significant genetic variation as indicated by range for different seed yield-contributing characters in Indian mustard but comparatively lower than the present study.

The analysis of variance revealed highly significant differences for all the characters representing the presence of variability, which can be utilized through genetic improvements. Significant variance due to $G \times E$ interactions for all the 20 characters confirmed that the genotypes respond differently in diverse environments. Therefore, it is possible to exploit different environments by developing environmentally specific varieties from the diversity panel. The importance of $G \times E$ interactions had also been observed by Gunasekara et al. (2006); Kumar et al. (2012), Priyamedha and Haider (2017), and Kumari et al. (2019) in Indian mustard and canola (*B. napus*) for seed yield.

Significance of Heritability and Other Genetic Parameters for the Selection of Traits

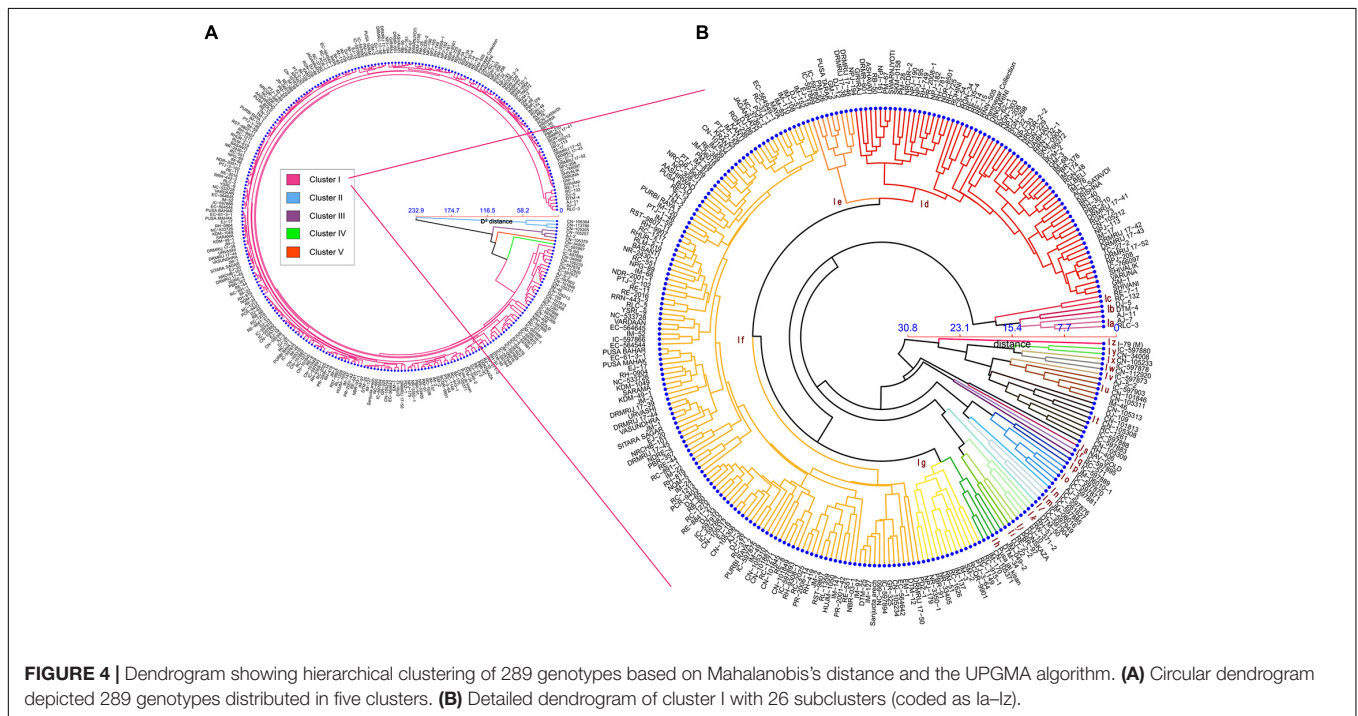
The observed variation in a population may be either due to genetic or environmental or both. Only those due to genetic components remain heritable. Heritability alone does not infer the estimate of the actual amount of genetic gain in the selection program, as it is also inclusive of

TABLE 5 | Promising genotypes and traits identified under study.

Favorable traits for breeding high yielding ideotype	Characters recorded	Acceptable limit*	Promising genotypes as donors with cluster information
Early flowering	DFI	<40 days	IC-766097 (I _d), DRMRIJ 17-46 (I _e), DTM-25 (I _f)
Basal branching	PH_FB	Productive branches from <30 cm from ground	EM-1 (I _f), NRCHB-101 (I _f), IM-46 (I _u), RNN-505 (I _d), RH-819 (I _f), EJ-22 (I _d), IC-597894 (I _f), NPJ-113 (I _d), EJ-20 (I _f), Pusa Mahak (I _f), IC-261627 (I _n), IC-597949 (I _m), NDR-8501 (I _d), Shivani (I _d), EC-61-3-1 (I _f), Sanjucta arech (I _f), RH-419 (I _f), DRMRIJ 17-38 (I _e), Pusa Mustard-28 (I _d), DRMRIJ 17-46 (I _e), DRMRIJ 17-45 (I _f), NRCD-2 (I _d), IC-597867, NQM-5 (I _f), IC-597873 (I _v), Sitara Sagar (I _f), CN-105234 (I _f), Pusa Mustard -26 (I _e), IC-766097 (I _e), GM-1 (I _e), IM-24 (I _f)
Medium plant stature	PH_M	151–175 cm	DRMRIJ 17-42 (I _d), IC-261627 (I _n)
No. of siliqua on main shoot	SMS	>60	IC-597885 (I _m), JM-06010-1 (I _f), RNN-505 (I _e), Purbi Raya-1 (I _f), NPJ-190 (I _d), RRN-443-2 (I _f), RLC-2 (I _f)
Bold seeds	SS	>6.0 g/1,000 seeds	RH-406 (I _t), RH-30 (I _d), RH-725 (I _d), RH-749 (I _d), Ashavathi (I _d), NAV-GOLD (I _s)
Total siliqua count	TSC	>800/plant	CN-101846 (I _u), IC-597867 (I _v), AJ-11 (I _a)
Siliquae length	SL	>5.25 cm	NRCD-2 (I _d)
Biological yield	BY/Plot	>30 t/ha	CS-54 (I _d), Dingahini (I _d), Pantnagar collection (I _d), RGN-73 (I _d), RNN-505 (I _d)
Seed yield in population/community	TSY/Plot	>35 Q/ha	RGN-73 (I _d), JM-3 (I _f), RH-725 (I _d), RCQR-9901 (I _g), RE-13 (I _f), Shivalik (I _d), PBR-210 (I _d), RE-7-1 (I _c), RNN-505 (I _d)
Seed yield per plant	SY/P	>30 g	IC-597867 (I _v), PBR-210 (I _d), RE-7-1 (I _c)

*Acceptable limit for each trait is fixed based on the various mustard ideotype concept given by Bhargava et al. (1984), VijayaKumar et al. (1996), DUS guidelines, Thurling (1991), and Yadav et al. (2017).

Parenthesis indicates the subcluster details.



nonadditive genetic factors (Shivanna, 2008). All the yield traits in the current study were highly heritable in individual seasons. Nevertheless, a partitioned genotype by environment interaction decreased the heritability across environments (pooled analysis). This type of reduced heritability

across the environment was also reported by Phuke et al. (2017) in Sorghum.

The study showed that high broad-sense heritability, genetic advance as percentage mean, and GCV were observed in flowering traits such as days to flowering initiation, days to 50%

TABLE 6 | Cluster means of 20 characters and its relative contribution evaluated for genetic divergence in 289 mustard genotypes in diversity panel.

Cluster	No. of genotypes	DFI	DF	DHF	DCF	DMT	PH_FI	PH_FP	PH_M	PB	SB	MSL	SMS	TSC	SL	SPS	SS	SYP	TSY/Plot	BY/Plot	HI
I	281	52.02	58.16	67.19	101.86	141.49	120.18	49.47	223.42	6.26	13.01	77.40	50.78	484.90	3.55	14.41	3.56	17.94	661.55	5.07	15.13
II	3	123.41	131.99	136.71	151.40	171.58	120.52	66.28	215.77	6.46	12.70	34.07	37.58	360.61	2.93	13.86	2.26	13.58	567.62	4.71	15.22
III	3	107.30	116.52	125.51	140.17	170.34	211.08	148.89	269.37	7.17	13.09	34.84	33.49	449.65	2.41	12.87	2.33	9.04	443.44	5.25	11.86
IV	1	64.65	71.62	82.68	109.97	143.32	129.35	28.14	218.77	6.93	13.08	79.91	56.16	601.30	3.32	14.35	1.97	49.77	550.72	4.22	14.75
V	1	88.77	95.60	99.46	125.23	143.87	99.28	44.95	206.41	6.12	12.93	79.02	44.38	459.56	4.07	13.89	4.63	8.24	437.83	4.22	11.96
Singh statistic		163.75	257.78	56.40	155.30	175.73	118.82	174.40	92.27	82.35	45.56	83.04	73.07	71.00	187.35	49.34	312.24	109.56	67.32	59.13	42.53
% Contribution		6.89	10.85	2.37	6.53	7.39	5.00	7.34	3.88	3.46	1.92	3.49	3.07	2.99	7.88	2.08	13.14	4.61	2.83	2.49	1.79
Rank		6	2	17	7	4	8	5	10	12	19	11	13	14	3	18	1	9	15	16	20

DFI, days to flowering initiation; DF, days to 50% flowering; DHF, days to 100% flowering; DCF, days to flowering completion; DMT, days to maturity; PH_FI, plant height at flowering (PH_FI); PH_FP, plant height up to first primary branch; PH_M, plant height at maturity; PB, number of primary branches; SB, number of secondary branches; MSL, main shoot length; SMS, silique on the main shoot; TSC, total silique count; SL, silique length; SPS, seeds per silique; SS, seed size; SYP/Plant, seed yield/plot; TSY/Plot, total seed yield/plot; BY/Plot, biological yield/plot; HI, harvest index.

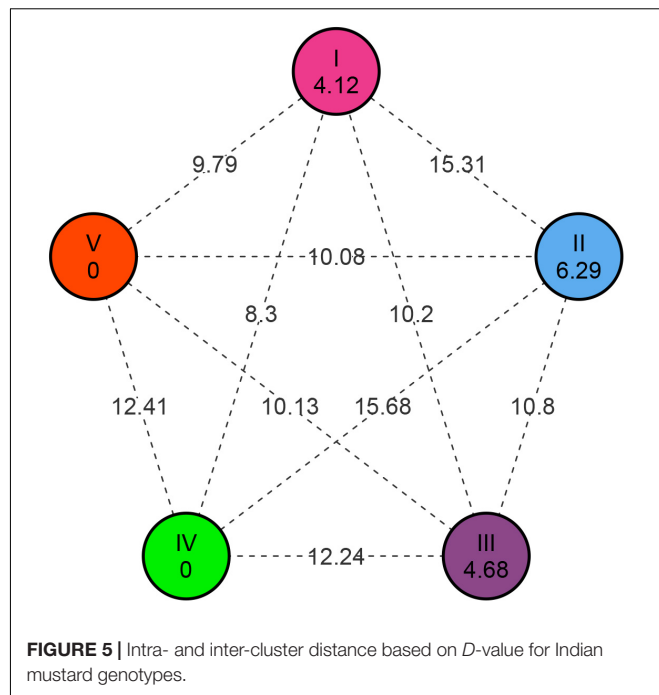


FIGURE 5 | Intra- and inter-cluster distance based on *D*-value for Indian mustard genotypes.

flowering, days to 100% flowering, plant height at flowering initiation, plant height up to the first primary branch, seed size, and seed yield/plant. The very high heritability of seed yield per plant (63%) in pooled analysis with a high GA of 60.8% indicated that the results would be repeatable and rewarding over generations of selection cycles, which is a boost for the breeding program. This concurred with previous studies (Kumar and Misra, 2007; Yadava et al., 2011; Tripathi et al., 2013; Meena et al., 2017; Kumar et al., 2019). These results indicate a greater scope for selection to improve seed yield *per se* in the *Brassica* breeding program (Tiwari et al., 2017).

Estimates of heritability for yield component traits varied from low (22%) to high (85%). There is a need to identify the traits that should be targeted for improving the seed yield in mustard. Flowering traits, plant height traits, and seed size showed high heritability and high GA, as few of them are governed by a few major quantitative trait loci reported earlier by Kaur and Banga (2015) and Akhatar et al. (2021). The high value of heritability and moderate genetic advance for plant height at maturity indicated that improvement in this trait could be made through the selection to some extent. High genetic advance and moderate heritability were shown in the number of primary branches, total silique count, total seed yield/plot, and biological yield/plot in which both additive and nonadditive gene actions may be expressed. A parameter having low heritability coupled with high genetic advance revealed the additive gene effects of traits (Belete et al., 2011). The low heritability is due to high environmental effects, and selection *per se* may be ineffective for such traits as harvest index in the present study. None of the traits exhibited low heritability with low genetic advance.

Higher PCV values than their corresponding GCV for most of the traits showed the considerable role of environment in the

expression of these traits; therefore, the variation in the genotypes is due to both genotype and the environment (Kumar et al., 2015). The wide difference between PCV and GCV was detected for plant height up to the first primary branch, the number of secondary branches, total siliquae count, seed yield/plant, total seed yield/plot, biological yield/plot, and harvest index, which indicated the high contribution of environmental variance to the phenotypic variance.

Association Among Traits

Correlation analysis indicated that seed yield per plant was significantly correlated with biological yield/plot, total seed yield/plot, seed size, plant height up to the first branch, total siliqua count, days to flowering initiation, plant height at maturity, siliqua length, and siliqua on the main shoot, which implies that selection in improving these traits would lead to improved seed yield (Rauf and Rahim, 2018). Genotypic correlations involving flowering characters such as days to flowering initiation, days to 50% flowering, days to 100% flowering, days to flowering completion, days to maturity, and plant height characters such as plant height at flowering initiation, plant height up to the first primary branch, and plant height at maturity with seed yield for plants in studied genotypes were negative, indicating selection for these traits would decrease seed yield (Joshi et al., 2009; Yadava et al., 2011). Reduction in flowering days prevents plant exposure toward high temperature that builds up during the late growth periods and consequent reduction in yield due to sterility and shriveling of seeds (Azharudheen et al., 2013). Also, reduced plant height makes plant architecture more compact to utilize the source toward increment in yield. These attributes can serve as marker characters for seed yield improvement in mustard. According to Kardam and Singh (2005), characters such as height up to the first branch, seed yield/plant, number of primary branches, number of siliquae per plant, and seed size have been reported as main yield contributing traits.

For instance, the number of secondary branches with very low heritability was significantly positively correlated to high heritability traits *viz.* total siliquae count ($r = 0.57$) and biological yield/plot ($r = 0.33$). Therefore, the selection of genotypes with higher siliqua count and biological yield/plot would indirectly improve the number of secondary branches per plant in successive generations. This is in accordance with the findings of Prasad et al. (2001), Swarnkar et al. (2002), and Singh et al. (2002).

Regression is a method for automatic selection in a stepwise manner based on partial correlations of a dependent variable with the independent variables near to optimal in the sense of maximizing the squared multiple correlations coefficient (R^2) of the dependent variable (Card et al., 1988). Based on regression studies, biological yield/plot, total seed yield/plot, plant height up to the first primary branch, seed size, total siliquae count, number of days to flowering initiation, plant height at maturity, siliquae on the main shoot, and siliqua length were the most contributing traits for seed yield per plant. However, these independent traits with individual R^2 of less than 20% had only a negligible direct contribution to the seed yield per plant, although they had a significant association with the dependent variable. This indicates that those traits that had less direct influence cannot be ignored

because their cumulative contribution to seed yield/plant could be highly influential (Maphumulo et al., 2015). These identified traits in combination could be used as effective indicators in Indian mustard for the calculation of yield performance; hence, a selection index based on identified influential traits could realize higher genetic advances than selecting seed yield alone (Hussain et al., 2004; Sandhu et al., 2019).

Contribution Toward Seed Yield per Plant

Specific direct and indirect effects of traits and relative importance of each trait in determining the key goal, i.e., seed yield, was accompanied through path coefficient analysis (Albayrak and Tongel, 2006). Path analysis that showed total seed yield/plot, siliquae on the main shoot, and seed size had a highly positive correlation and moderate direct effect on seed yield per plot, which suggested that selection for these traits would be quite effective for improving seed yield in mustard. Traits such as plant height up to the first primary branch and plant height at maturity had a negative moderate direct effect. Similar conclusions were reported by Kardam and Singh (2005) and Sandhu et al. (2019). Indirect effects were ranked similar to those of Lenka and Mishra (1973), as follows: 0.00–0.09 = negligible, 0.10–0.19 = low, 0.20–0.29 = moderate, and >0.30 = high path coefficients. Plant height at maturity toward dwarf plant type exhibited negligible indirect effects on seed yield, indicating that they had little contribution to seed yield. The rest of the traits were low to moderate, showing that indirect selection for these traits would improve the yield of the mustard. Hence, for improving the seed yield per plant in mustard, one should focus on “selecting for” traits such as total seed yield per plot (under crop community), more siliquae on the main shoot, bold seed size, and highest total siliquae count per plant and “selecting against” plant height up to the first primary branches (i.e., selection for basal branching genotypes). Traits such as days to flowering initiation, plant height up to the primary branches, siliqua length, and biological yield have significant indirect effects *via* component traits toward seed yield per plant. The traits as mentioned earlier with high direct effects inferred from path analysis such as total seed yield/plot, seed size, siliqua on the main shoot, plant height up to the first primary branch, and plant height at maturity had moderate to high heritability ($\geq 30\%$) coupled with high GA% of the mean ($>20\%$). These traits also showed a highly significant ($p \leq 0.001$) correlation with seed yield/plant. **Table 5** summarizes promising genotypes identified based on the cumulative performance for yield and associated traits that can be a guide in bringing high-yielding ideotypes in mustard for the entire mustard breeding community.

Diversity Analysis

The advantage of genetic diversity analysis based on Mahalanobis D^2 distance over the Euclidian distance is that it can take account of the correlation between a highly correlated variable and can scale the contribution of individual variables to the distance value according to the variability of each variable (Ghorbani, 2019). The Mahalanobis D^2 distance among genotypes in the diversity panel ranged from 0.0 to 351.38, which was huge and higher than previous reports by Bind et al. (2015); Gupta et al. (2015), and Chandra et al. (2018). The huge extent of

genetic diversity in the present study was due to the involvement of genetic material from four continents across the globe. Furthermore, the number of genotypes studied was higher compared with the previous reports. Most of the genotypes (281/289) got included in a single cluster and the rest of them in four different clusters. The cluster forming point was having a very high D^2 -value = 232. It suggests that the eight genotypes included in clusters II to V were more diverse than cluster I. Also, a detailed analysis of cluster I also suggested that there were 26 subclusters with a cluster-forming point at a D^2 value of 20. Intra-cluster distance of cluster I ($D^2 = 17$) was much higher than earlier reports related to D^2 clustering studies by Yadava et al. (2011) and Kumari and Kumari (2018), indicating a wider genetic base of materials within cluster I in the present study. The subcluster I (d-i) included most of the high-yielding cultivars and breeding lines of Indian origin. The rest of the subclusters consisted of exotic and indigenous gene pools received in various bilateral collaborative projects utilized in mustard improvement by various researchers (Chauhan et al., 2011). Subcluster I (j-z) consisted of lines having more height and longer duration with various oil quality traits, which remain unexploited due to lack of synchrony in flowering time. Still, there was no obvious clustering pattern related to geographic proximity and use types among mustard. Grouping of certain improved varieties and cultivars along with Canadian and Australian genetic stocks and indigenous collections from Arunachal Pradesh (India) indicated that the geographical distribution need not necessarily be the indicator of genetic divergence as reported by Verma and Sachan (2000) and Jeena and Sheikh (2003). The possible reason could be the common ancestry of these genotypes, which permitted the free exchange of germplasm among the breeders of different locations and/or the unidirectional selection experienced by breeders in tailoring the promising cultivars for different locations (Yadava et al., 2011; Mukesh Sankar et al., 2014).

Broad variability in the current material holds great promise to use these genotypes from different clusters (such as II and IV) for future breeding programs. The highest-yielding genotype (IC-597867) identified in the present study was present in cluster IV, whereas genotypes with a high harvest index were contained in cluster II. So, the improved cultivars from cluster I can be utilized to exploit the genotypes present in clusters II and IV for further yield increments and genetic diversification through hybridization.

CONCLUSION

The study assessed a diversity panel representing 289 genotypes across four continents for the existence of genetic variability for seed yield and yield-related traits over two seasons. The results revealed an enormous genetic variability for the traits under study, which can be exploited to acquire further breeding gains. The use of BLUP values for genotypes provided higher selection accuracy by reducing residual error, which permitted the identification of potential genotypes for the *Brassica* improvement program. Flowering and plant height-related traits

were found to be more heritable, although these were negatively correlated with yield. High PCV, GCV, and genetic advance values with low to moderate heritability were observed for total seed yield, which indicated that the yield performance still needs to be improved to produce superior varieties. Moreover, the current study could aid breeders in enhancing the seed yield by considering the traits that have a good correlation with seed yield. Cluster analysis revealed that genotypes under study were more diverse, which could be utilized for future hybridization programs, and it can release transgressive segregants for economic trait improvement. The greater variability among the studied genotypes and the association between seed yield and secondary traits in the current study suggested the exploitation of superior genotypes in the near future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

DY conceptualized, supervised the research, and contributed to reviewing the manuscript. RS and SLS performed the field trials and prepared the original draft manuscript. SS performed first-year field trial. SM conducted the data analysis, curated the data, and contributed in reviewing the manuscript. RC helped in field trials and data entry. YP and NS coordinated the study and revised the manuscript. SV administered the project and revised the manuscript. All authors read the manuscript and agreed with its content.

FUNDING

This study was funded by ICAR—National Agricultural Science Fund, New Delhi, with project number NASF/GTR-6012/2016-17/330.

ACKNOWLEDGMENTS

The authors pay their sincere gratitude to the Director, ICAR-IARI, New Delhi, for his consistent support and for providing all the experimental field facilities for the research work. RS is thankfully acknowledging ICAR-IARI New Delhi for providing IARI Merit Fellowship. Thanks to the Department of Agricultural Physics, ICAR-IARI, New Delhi, for providing the meteorological data of the experimental sites.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.651936/full#supplementary-material>

REFERENCES

- Akhtari, J., Goyal, A., Kaur, N., Atri, C., Mittal, M., Singh, M. P., et al. (2021). Genome wide association analyses to understand genetic basis of flowering and plant height under three levels of nitrogen application in *Brassica juncea* (L.) Czern & Coss. *Sci. Rep.* 11, 1–14. doi: 10.1038/s41598-021-83689-w
- Albayrak, S., and Tongel, O. (2006). Path analyses of yield and yield-related traits of common vetch (*Vicia sativa* L.) under different rainfall conditions. *Anadolu Tarım Bilimleri Dergisi*. 21, 27–32.
- Ali, N., Javidfar, F., Elmira, J. Y., and Mizra, M. Y. (2003). Relationship among yield components and selection criteria for yield improvement in winter rapeseed (*Brassica napus* L.). *Pakistan J. Bot.* 35, 167–174.
- Azharudheen, T. M., Yadava, D. K., Singh, N., Vasudev, S., and Prabhu, K. V. (2013). Screening Indian mustard [*Brassica juncea* (L.) Czern and Coss)] germplasm for seedling thermo-tolerance using a new screening protocol. *Afr. J. Agr. Res.* 8, 4755–4760. doi: 10.5897/AJAR2013.7681
- Bates, D., and Maechler, M. (2009). *lme4: Linear Mixed-Effects Models Using Stochastic Gradient Descent*. R Package Version 0.999375-32.
- Belete, Y. S., Kebede, S. A., and Gemelal, A. W. (2011). Heritability and genetic advance in Ethiopian mustard (*Brassica carinata* A. Brun). *Int. J. Pl. Breed.* 6, 42–46. doi: 10.3923/ijar.2011.494.503
- Bhargava, S. C., Tomar, D. P. S., and Sinha, S. K. (1984). *Physiological Basis of Productivity in Brassica ecotypes. In: Research and Development Strategies for Oil Seed production in India*. New Delhi: ICAR Publication, 103–110.
- Bhatt, K., Gill, K. K., and Sandhu, S. S. (2015). Comparison of different regression models to predict mustard yield in central Punjab. *Vayu Mandal*. 41, 28–38.
- Bind, D., Singh, S. K., and Dwivedi, V. K. (2015). Assessment of genetic diversity and other genetic parameters in Indian mustard [*Brassica Juncea* (L.) Czern & Coss.]. *Indian J. Agr. Res.* 49, 554–557. doi: 10.18805/ijare.v49i6.6685
- Bisne, R., Sarwgi, A. K., and Verulka, S. B. (2009). Study of heritability, genetic advance and variability for yield contributing characters in rice. *Bangladesh J. Agr. Res.* 34, 175–179. doi: 10.3329/bjar.v34i2.5788
- Card, D. H., Peterson, D. L., and Matson, P. A. (1988). Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sens. Environ.* 26, 123–147. doi: 10.1016/0034-4257(88)90092-2
- Chandra, K., Pandey, A., and Mishra, S. B. (2018). Genetic diversity analysis among Indian mustard [*Brassica juncea* (L.) Czern and Coss] genotype under rainfed condition. *Int. J. Microbial. App* 7, 256–268. doi: 10.20546/ijcmas.2018.703.030
- Chauhan, J. S., Singh, K. H., Singh, V. V., and Kumar, S. (2011). Hundred years of rapeseed–mustard breeding in India: accomplishments and future strategies. *Indian J. Agr. Sci.* 81, 1093–1109.
- da Silva, A. R., and da Silva, M. A. R. (2017). *Package 'biotools'*. Available online at: <https://CRAN.R-project.org/package=biotools> (accessed December 27, 2020).
- de Mendiburu, F., and de Mendiburu, M. F. (2019). *Package 'Agricolae'*. R Package, Version, 1-2. Available online at: <https://CRAN.R-project.org/package=agricolae> (accessed December 27, 2020).
- Directorate of Economics & Statistics, and Dac&Fw. (2019). Available online at: https://eands.dacnet.nic.in/Advance_Estimate/3rd_Adv_Estimates2019-20_Eng.pdf (accessed June 19, 2020).
- Epskamp, S., Rhemtulla, M., and Borsboom, D. (2017). Generalized network psychometrics: combining network and latent variable models. *Psychometrika* 82, 904–927. doi: 10.1007/s11336-017-9557-x
- Fox, J., Friendly, G. G., Graves, S., Heiberger, R., Monette, G., Nilsson, H., et al. (2007). *The car package*. Vienna: R Foundation for Statistical Computing.
- Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* 31, 3718–3720.
- Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Univ. Ser. Math. Inform.* 34, 583–595. doi: 10.1093/bioinformatics/btv428
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812. doi: 10.1093/bioinformatics/btu393
- Gunasekara, C. P., Martin, L. D., Siddique, K. H. M., and Walton, G. H. (2006). Genotype x environment interaction of Indian mustard (*B. juncea* L.) and canola (*B. napus* L.) in Mediterranean type environments, crop growth and seed yield. *Eur. J. Agron.* 25, 1–12. doi: 10.1016/j.eja.2005.08.002
- Gupta, R. K., Dwivedi, V. K., and Meena, B. (2015). Genetic diversity analysis for seed yield parameters and oil content in Indian mustard. *Indian Res. J. Genet. Biotech.* 7, 84–86.
- Gupta, S. K. (2016). *Biology and Breeding of Crucifers*. Boca Raton FL: CRC Press.
- Hallauer, A. R., and Miranda, J. B. (1988). *Quantitative Genetics in Maize Breeding*, 2 Edn. IA, Ames: Iowa State University Press, 468.
- Harrell, F. E. Jr., and Harrell, M. F. E. Jr. (2019). Package 'rms'. *CRAN2018* 2019, 235–236.
- Hazel, L. N. (1943). Genetic basis for constructing selection indices. *Genetics* 28, 476–490.
- Hu, X. (2015). A comprehensive comparison between ANOVA and BLUP to evaluate location-specific genotype effects for rape cultivar trials with random locations. *Field Crops Res.* 179, 144–149. doi: 10.1016/j.fcr.2015.04.023
- Hussain, S., Hazarika, G. N., and Barua, P. K. (2004). Selection Indices in Indian Rapeseed (*Brassica campestris*) and Mustard (*B. Juncea*). *J. Oilseeds Res.* 21, 143–145.
- Jeena, A. S., and Sheikh, F. A. (2003). Genetic divergence analysis in gobhi sarson (*Brassica napus* L.) J. *Oilseed Res.* 20, 210–212.
- Johnson, H. W., Rodinson, H. F., and Cronstrook, R. E. (1959). Estimation of genetic and environmental variability in soybeans. *Agron. J.* 47, 314–318. doi: 10.2134/agronj1955.00021962004700070009x
- Joshi, V., Pathak, H. C., Patel, J. B., and Haibatpure, S. (2009). Genetic variability, correlation and path analysis over environments in mustard. *GAU Res. J.* 34, 14–19.
- Kant, L., and Gulati, S. C. (2001). Genetic analysis for yield and its components and oil content in Indian mustard [*B. juncea* (L.) Czern & Coss.]. *Indian J. Genet.* 61, 37–40.
- Kardam, D. K., and Singh, V. V. (2005). Correlation and path analysis in Indian mustard [*Brassica juncea* (L.) Czern & Coss.] grown under rainfed condition. *J. Spices Aromatic Crops.* 14, 56–60.
- Kaur, H., and Banga, S. S. (2015). Discovery and mapping of *Brassica juncea* Sdt 1 gene associated with determinate plant growth habit. *Theor. Appl. Genet.* 128, 235–245. doi: 10.1038/s41598-021-83689-w
- Khan, S., Farhatullah, I., Khalil, H., Munir, I., Khan, M. Y., and Ali, N. (2008). Genetic variability for morphological traits in F3:4 Brassica populations. *Sarhad J. Agric.* 24, 217–222.
- Kumar, A. (2005). “Rapeseed–mustard in India: current status and future prospects,” in *Winter School on Advances in Rapeseed Mustard Research Technology for Sustainable Production of Oilseeds, National Centre on Rapeseed–Mustard, Sewar, Bharatpur, December 15 to January 04, 2005* (Karnal: Agricultural Research Communication Centre), 278–288.
- Kumar, A., Pandey, A., Cochen, C., and Pattanayak, A. (2015). Evaluation of genetic diversity and interrelationships of agro morphological characters in soybean (*Glycine max*) genotypes. *Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci.* 85, 397–405. doi: 10.1007/s40011-014-0356-1
- Kumar, A., Sah, R. P., Haider, Z. A., Ghosh, J., Kumar, R., Mahato, C. S., et al. (2012). “G x E interaction for seed yield and its components in Indian mustard,” in *Proceedings of 1st National Brassica Conference on Production Barriers and Technological Options in Oilseed Brassica Held at CCS Haryana Agricultural University (Hisar)*, 56.
- Kumar, M. S., Mawlong, I., and Rani, R. (2020). “Biofortification of brassicas for quality improvement,” in *Brassica Improvement*, eds S. H. Wani, A. K. Thakur, and Y. J. Khan (Cham: Springer), 127–145. doi: 10.1007/978-3-030-34694-2_7
- Kumar, S., and Misra, M. N. (2007). Study on genetic variability, heritability and genetic advance in populations in Indian mustard [*Brassica juncea* (L.) Czern & Coss.]. *Int. J. Plant Sci.* 2, 188–190.
- Kumar, S., Ram, S., Chakraborty, M., Ahmad, E., Verma, N., Lal, H. C., et al. (2019). Role of genetic variability for seed yield and its attributes in linseed (*Linum usitatissimum* L.) improvement. *J. Pharmacogn. Phytochem.* 8, 266–268.
- Kumari, A., and Kumari, V. (2018). Studies on genetic diversity in Indian mustard (*Brassica juncea* Czern & Coss) for morphological characters under changed climate in midhills of Himalayas. *Pharma Innov. J.* 7, 290–296. doi: 10.20546/ijcmas.2018.707.393
- Kumari, V., Jambhulka, S., Chaudhary, H. K., Sharma, B. K., Sood, P., Guleria, S. K., et al. (2019). Phenotypic stability for seed yield and related traits in Trombay mustard genotypes under North western Himalayas. *J. Oilseed Brassica* 10, 33–37.

- Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlingson, B., Wickham, H., et al. (2015). *Package 'plotrix'*. Vienna: R Development Core Team.
- Lenka, D., and Mishra, B. (1973). Path coefficient analysis of yield in rice varieties. *Indian J. Agric. Sci.* 43, 376–387.
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Oliver, S. (2006). *SAS for mixed models*. Cary, NC: SAS publishing.
- Mahalanobis, P. C. (1936). On generalized distance in statistics. *Proc. Natl. Inst. Sci.* 2, 49–55.
- Maphumulo, S. G., Derera, J., Qwabe, F., Fato, P., Gasura, E., and Mafongoya, P. (2015). Heritability and genetic gain for grain yield and path coefficient analysis of some agronomic traits in early-maturing maize hybrids. *Euphytica* 206, 225–244. doi: 10.1007/s10681-015-1505-1
- Meena, H. S., Kumar, A., Singh, V. V., Meena, P. D., Ram, B., and Kulshrestha, S. (2017). Genetic variability and interrelation of seed yield with contributing traits in Indian mustard (*B. juncea*). *J. Oilseed Brassica*. 8, 131–137.
- Molenaar, H., Boehm, R., and Piepho, H. P. (2018). Phenotypic Selection in Ornamental Breeding: It's better to have the BLUPs than to have the BLUEs. *Front. Plant Sci.* 9:1511. doi: 10.3389/fpls.2018.01511
- Mukesh Sankar, S., Satyavathi, C. T., Singh, S. P., Singh, M. P., Bharadwaj, C., and Barthakur, S. (2014). Genetic diversity analysis for high temperature stress tolerance in pearl millet [*Pennisetum glaucum* (L.) R. Br. *Indian J. Plant Physiol.* 19, 324–329. doi: 10.1007/s40502-014-0099-2
- Murtagh, F., and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J. Classif.* 31, 274–295. doi: 10.1007/s00357-014-9161-z
- Olivoto, T., Nardino, M., Carvalho, I. R. C., Follmann, D. N., Szareski, V. J., Ferrari, M., et al. (2016). Pearson correlation coefficient and accuracy of path analysis used in maize breeding: a critical review. *Int. J. Curr. Res.* 8, 37787–37795.
- Patterson, H. D., and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554. doi: 10.1093/biomet/58.3.545
- Phuke, R. M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., et al. (2017). Genetic Variability, Genotype × Environment Interaction, Correlation, and GGE Biplot Analysis for Grain Iron and Zinc Concentration and Other Agronomic Traits in RIL Population of Sorghum (*Sorghum bicolor* L. Moench). *Front. Plant Sci.* 8:712. doi: 10.3389/fpls.2017.00712
- Piepho, H. P., Mohring, J., Melchinger, A. E., and Buchse, E. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*. 161, 209–228. doi: 10.1007/s10681-007-9449-8
- Prasad, L., Singh, M., and Dixit, R. K. (2001). Analysis of heritability and genetic advance in Indian Mustard [*Brassica juncea* (L.) Czern and Coss]. *Adv. Plant Sci.* 14, 577–581.
- Priyamedha, K. A., and Haider, Z. A. (2017). Stability for seed yield and components traits in Indian mustard (*B. juncea* L.) under Jharkhand conditions. *J. Oilseed Brassica*. 8, 37–42.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org>
- Rao, C. R. (1952). *Advanced Statistical Method in Biometrical research*, 1 Edn. New York NY: Wiley and sons Inc.
- Rao, V. T., Bharathi, D., Mohan, Y. C., Venkanna, V., and Bhadr, D. (2013). Genetic variability and association analysis in sesame (*Sesamum indicum* L.). *Crop Res.* 46, 122–125.
- Rauf, M. A., and Rahim, M. A. (2018). Genetic variability studies among yield and its contributing traits in mustard (*Brassica napus* L.). *Adv. Zool. Bot.* 6, 101–108. doi: 10.13189/azb.2018.060402
- Roy, S. K., Hijam, L., Chakraborty, M., Vishal, N., Das, A., Kundu, A., et al. (2018). Cause and effect relationship in yield and its attributing traits in early segregating generations of mustard crosses under terai agro-climatic zone of West Bengal. *India. J. Homepage* 7:2018. doi: 10.20546/ijcmas.2018.703.024
- Ruiz, E., Jackson, S., and Cimentada, J. (2019). *corr: Correlations in R*. Available online at: <https://CRAN.R-project.org/package=corr> (accessed December 12, 2020).
- Salgotra, R. K., Gupta, B. B., Bhat, J. A., and Sharma, S. (2015). Genetic diversity and population structure of basmati rice (*Oryza sativa* L.) germplasm collected from northwestern Himalayas using trait linked SSR markers. *PLoS One* 10:e0131858. doi: 10.1371/journal.pone.0131858
- Sandhu, S. K., Kang, M. S., Akash, M. W., and Singh, P. (2019). Selection indices for improving selection efficiency in Indian mustard. *J. Crop. Improv.* 33, 25–41. doi: 10.1080/15427528.2018.1539689
- Shivanna, S. (2008). *Genetic diversity, combining ability and stability analysis of selected castor lines*. Ph. D thesis. Bangalore: University of Agricultural Sciences, 28–45.
- Singh, M., Swarnkar, G. B., Prasad, L., and Rai, G. (2002). Genetic variability, heritability and genetic advance for quality traits in Indian mustard (*B. juncea*). *J. Pl. Archives*. 2, 27–31.
- Singh, R. K., and Chaudhary, B. D. (2004). *Biometrical Method in Quantitative Genetics Analysis*. Ludhiana: Kalyani Publishers.
- Smith, H. F. (1936). A discriminant function for plant selection. *Ann. Eugen.* 7, 240–250. doi: 10.1111/j.1469-1809.1936.tb02143.x
- Soetaert, K. (2009). *diagram: Functions for Visualising Simple Graphs (networks)*, *Plotting flow Diagrams*. R package version 1.4.
- Souza, A. R. R., Miranda, G. V., Pereira, M. G., and Souza, L. V. D. (2009). Predicting the genetic gain in the Brazilian white maize landrace. *Ciênc Rural*. 39, 19–24. doi: 10.1590/S0103-84782009000100004
- Swarnkar, G. B., Prasad, S. M., and Lallu. (2002). Analysis of heritability and genetic advance in relation to yields and its contributing traits in Indian mustard (*Brassica juncea*). *Plant Arch.* 2, 305–308.
- Thurling, N. (1991). Application of the ideotype concept in breeding for higher yield in the oilseed brassicas. *Field Crops Res.* 26, 201–219. doi: 10.1016/0378-4290(91)90036-U
- Tiwari, A. K., Singh, S. K., Tomar, A., and Singh, M. (2017). Heritability, genetic advance and correlation coefficient analysis in Indian mustard (*Brassica juncea* (L.) Czern & Coss). *J. Pharmacogn. Phytochem.* 6, 356–359.
- Tripathi, N., Kumar, K., and Verma, O. P. (2013). Genetic variability, heritability and genetic advance in Indian mustard (*Brassica juncea* L. Czern and Coss.) for seed yield and its contributing attributes under normal and saline/alkaline condition. *Int. J. Sci. Res.* 4, 983–985.
- Verma, S. K., and Sachan, J. N. (2000). Genetic divergence in Indian mustard (*Brassica juncea* (L.) Czern & Coss.). *Crop Res.* 19, 271–276.
- VijayaKumar, C. H. M., Arunachalam, V., Chakraborty, S. K., and Kesava Rao, P. S. (1996). Ideotype and relationship between morpho-physiological characters and yield in Indian mustard (*Brassica juncea*). *Indian J. Agr. Sci.* 66, 633–637.
- Yadav, R., Singh, R., Kumar, S., Prasad, T. V., Bharadwaj, R., Kaur, V., et al. (2017). Genetic diversity among indigenous germplasm of *Brassica juncea* (L.) Czern and Coss, using agro-morphological and phenological traits. *Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci.* 87, 1125–1131. . <https://doi.org/10.1007/s40011-015-0689-4>,
- Yadava, D. K., Giri, S. C., Vignesh, M., Vasudev, S., Yadav, A. K., Dass, B., et al. (2011). Genetic variability and trait association studies in Indian mustard (*Brassica juncea*). *Indian J. Agric. Sci.* 81, 712–716.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Saroj, Soumya, Singh, Sankar, Chaudhary, Yashpal, Saini, Vasudev and Yadava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.