Genome wide association studies for yield and its component traits under terminal heat stress in Indian mustard (Brassica juncea L.)

Article *in* Euphytica · October 2019

DOI: 10.1007/s10681-019-2489-z

CITATION
1

READS
320

4 authors, including:

Surinder K Sandhu
Punjab Agricultural University
71 PUBLICATIONS **485** CITATIONS

SEE PROFILE

Lalit Pal
Punjab Agricultural University
2 PUBLICATIONS **3** CITATIONS

SEE PROFILE

Dharminder Bhatia
Punjab Agricultural University
48 PUBLICATIONS **487** CITATIONS

SEE PROFILE

Euphytica (2019) 215:188
https://doi.org/10.1007/s10681-019-2489-z

Genome wide association studies for yield and its component traits under terminal heat stress in Indian mustard (*Brassica juncea* L.)

Surinder K. Sandhu · Lalit Pal · Jasneet Kaur · Dharminder Bhatia

Received: 20 December 2018 / Accepted: 3 September 2019 / Published online: 17 October 2019

Abstract Breeding for terminal heat stress (THS) in *Brassica juncea* L. Czern & Coss is recognized as an imperative objective for sustained productivity in contemporary climatic changes. A fixed diversity stock of 491 genotypes was documented for wide range of variations for seed yield under natural terminal heat stress. A set of top 20 genotypes comprising introgression lines from wild species *Erucastrum cardaminoides and B. tournefortii; derived B. juncea lines using B. carinata and B. napus*; land races;, commercial cultivars and breeding lines, having the lowest heat susceptibility index and the least yield reduction under heat stress, have been identified as potential heat tolerant donors. A panel of 96 genotypes was constituted from this stock on the basis of their differential response to heat susceptibility index and seed yield reduction under natural THS. The constituted panel was evaluated for validation under controlled conditions for ten seed yield-related traits. Moderate to low correlations between SY and its related traits were observed in NS and THS conditions. Double digest restriction site associated DNA sequencing of 71 genotypes identified 18,258 SNPs after filtration. Least square means of all the traits under NS and THS conditions and the best linear unbiased predictors along with identified SNPs were used for genome-wide association study. A total of 34 SNPs under NS, 24SNPs under THS and 30SNPs using BLUP values were found to be associated with all seed yield-related traits. Chromosome B05 harbored the maximum number of SNPs (nine) followed by chromosomes A07 and A09 (eight SNPs each). SNPs under NS conditions could not be associated with THS. This is the first report on the identification of 24 marker-traits associations detected for SY and its component traits under THS conditions. It may be possible to develop the molecular markers for significant SNPs after due validation. The constituted panel may also serve as a source of allelic diversity for genes controlling various economic traits. The derived introgression lines as potential heat tolerant donors indicated the possibility of using wild species to breed for abiotic stress tolerance in Indian mustard.

Keywords *Brassica juncea* · GWAS · BLUP · SNPs · Yield · Terminal heat stress

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10681-019-2489-z) contains supplementary material, which is available to authorized users.

S. K. Sandhu (✉) · L. Pal · J. Kaur · D. Bhatia
Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab 141 004, India
e-mail: surindersandhu@pau.edu

Introduction

Indian mustard (*Brassica juncea* L. Czern & Coss), an allotetraploid AABB; 2n = 36, is a major oilseed crop of South Asian countries. It has originated through several independent hybridization events between *B. rapa* (AA; 2n = 20) and *B. nigra* (BB; 2n = 16) in the areas encompassing Mediterranean, Irano-Turanian and Saharo-Sindian geographies (Nagaharu 1935; Kaur et al. 2014). It is an important winter season oilseed crop in India and contributes nearly 28.6% of edible oil supplies (GOI 2017). Being a C_3 plant, *B. juncea* is sensitive to climatic variables, affecting its productivity radically. It exhibits an efficient photosynthetic response at 15–20 °C and gets affected by both early (during germination and seedling stage) as well as by terminal (during flowering and seed ripening stage) heat stress. Early heat stress affects seed germination, increase seedling mortality and hence, results in poor crop stand (Azharudheen et al. 2013).

In India, delayed sowing of the mustard crop after harvesting of rice and cotton exposes the crop to high-temperature stress during the reproductive stage (Chauhan et al. 2009); referred to as terminal heat stress (THS). Due to delayed planting, THS occur during the post-anthesis/seed filling stage that not only affects the movement of photosynthates to the developing sinks and inhibits the synthetic processes, but also lead to an increase in disease incidence and insect-pest infestation, thereby causing lower seed weight and seed yield (Tuteja et al. 1996). Delayed planting also leads to shortening of the vegetative phase, advances flowering time and decreases seed development period (Srivastva and Balkrishna 2003 and Dhaliwal et al. 2007). The effect of heat stress on seed yield and its contributing traits viz. plant height, main shoot length, number of primary and secondary branches, number of pods on main shoot and seed size, has been well documented (Morrison and Stewart 2002; Alam et al 2014; Sharma 2014 and Sharma and Sardana 2016). Young et al. (2004) reported impaired seed filling on exposure of *B. juncea* crop to THS. Nutall et al. (1992) recorded a decline of 430 kg/ha in seed yield with a rise of 3 °C in maximum daily temperature (21–24 °C) during the flowering phase.

Terminal heat stress (THS) has now emerged as a major challenge in Indian mustard, because of climate change. Woods et al. (1991) and later Chauhan et al. (2009) favored the development of thermo-tolerance in *B. juncea* to expand its cultivation in relatively warm and dry regions of India as well as worldwide. Development of THS tolerant varieties would be an important approach in this direction. However, quantitative nature of seed yield and its component traits complicates direct phenotypic selection. In *B. juncea*, there are several reports on the identification of QTLs for seed yield and its component traits (Ramchiary et al. 2007; Yadava et al. 2012; Akhatar and Banga 2015 and Dhaka et al. 2017). But, there is no report on the identification of SNPs associated with SY and its component traits under heat stress conditions. It is important now to identify variation for these traits in germplasm and utilize them in breeding programmes.

With recent developments in next-generation sequencing (NGS) and genome complexity reduction methods, genotyping by sequencing (GBS) have emerged as powerful genotyping platforms. GBS can be used to discover thousands of markers across almost any genome of interest in number of individuals in a population in a few weeks (Davey et al. 2011). Utilizing the power of GBS, Genome wide association studies (GWAS) has emerged as a powerful approach for identifying QTLs/genes underlying complex traits (Zhu et al. 2008). The approach has generated wider interest in plant breeding community after its demonstration in maize (Thornsberry et al. 2001).

A fixed diverse germplasm stock (derived through continues selfing) of 491 Indian mustard genotypes comprising landraces; released Indian varieties/cultivars; accessions collected from different countries and developed introgression lines, is available with Punjab Agricultural University (PAU), Ludhiana. The germplasm collection was evaluated for seed yield and its component traits for two years at PAU- Regional Research Station, Abohar, Fazilka district of Punjab state, India under two planting regimes. This location is situated in that zone of Punjab which experiences higher temperature than rest of Punjab during onset of summer season. This region accompanied by delayed planting of germplasm was explored for natural screening of germplasm for terminal heat stress. From this stock, a panel of 96 genotypes, which represented differential response to seed yield based heat susceptibility index (HSI) and percent seed yield reduction under terminal heat stress was constituted and used for GWAS.

Materials and methods

Constitution of diversity panel for GWAS

A diversity panel of 96 genotypes, comprising introgression lines, land races, old cultivars, and fixed breeding material, was used in this study. To constitute this diverse panel, a germplasm collection of 491genotypes was evaluated for SY and its component traits at PAU Regional Research Station, Abohar, Fazilka (representing arid zone environment; located at 30.15°N, 74.19°E) for two years during 2015–16 and 2016–17 under two environments. Two environments represent two sowing regimes: set I sown on November 14 (represented normal sown conditions; abbreviated as NS); set II sown on December 2 (represented late sown conditions to expose the crop to terminal heat stress; abbreviated as THS). Each genotype was sown in a plot of two rows of two-meter length at a row to row distance of 30 cm and plant to plant spacing of 10 cm in two replications in an alpha lattice design. At maturity, the data for seed yield was recorded for ten random plants and averaged to seed yield per plant (g). Percent yield reduction (YD%) for each genotype, due to heat stress during the reproductive phase of crop, was computed using formula:

$$\mathrm{YD\%} = \frac{\mathrm{Y_{NS}} - \mathrm{Y_{THS}}}{\mathrm{Y_{NS}}} \times 100$$

$\mathbf{Y_{NS}}$: Seed yield under NS condition, $\mathbf{Y_{THS}}$: Seed yield under THS condition.

The heat susceptibility index (HSI) for seed yield per plant (g) for each genotype was calculated using formula given by Fischer and Maurer (1978).

$$\mathrm{HSI} = \frac{1 - Y_{THS}/Y_{NS}}{1 - \overline{Y_{THS}}/\overline{Y_{NS}}}$$

$\overline{Y_{NS}}$: Mean seed yield under NS; $\overline{Y_{THS}}$: Mean seed yield under THS.

Phenotyping and genotyping of the diversity panel

The diversity panel of 96 genotypes was planted at the research farm of Punjab Agricultural University, Ludhiana (30.90°N, 75.86°E) during the year 2017–18 under two different environments. In the environment I, the panel was sown under natural field conditions (NS) and in environment II, the same set of panel was sown under the same field conditions but covered with polytunnels during the reproductive phase i.e., the onset of flowering to seed ripening stage of the crop. Polytunnels were shaped by placing an iron rod frame in field and covering the crop by polythene sheets (3 mm thickness) from 10.00 am to 3.00 pm every day. In each environment, the panel set was sown in an alpha lattice design with two replications; with each genotype in 3-meter row length. Row to row distance of 30 cm and plant to plant distance of 10 cm was maintained. Weather data on temperature and relative humidity were recorded under normal and stress conditions (Table 1). All package and protection measures were taken to raise healthy crop. Data were recorded at physiological maturity on yield and yield components traits and were analyzed statistically. Based on five randomly taken plants of each genotype of each replication, plant height (PH; in centimeter) was determined on standing crop by measuring the height of plant from its base up to the end of main shoot. Main shoot length (MSL; in centimeter) was recorded by measuring the length of main raceme. Number of primary branches (NPB) recorded by counting the siliquae bearing branches emerging from main shoot of the plant, number of secondary branches (NSB) was recorded by counting siliquae bearing branches emerging from primary branches and number of siliquae on the main shoot was recorded by counting total siliquae attached to main shoot. Number of days to maturity (DM) was recorded as number of days from sowing to the date on which 75% of siliquae of a genotype, in plot, turned yellow. Seed yield (SY; in gm) was recorded on a plot basis after manual thrashing and drying of seeds of each genotype. 1000 seed weight (TSW; in gm) was measured by counting 1000 seeds from each genotype and weighted on an electronic balance. For number of seeds per siliquae (NSP), five siliquae of each were taken from main raceme of five randomly selected plants of each plot and after thrashing, seeds were counted and averaged.

Statistical analysis of phenotypic data

Phenotypic distributions of each trait were plotted as histograms with fitting normal curve. Analysis of variance and adjusted means for alpha lattice design were calculated by using *Proc GLM* procedure in SAS 9.3 (SAS Institute 2011) for each environment.

Reduction in the mean value of traits in THS as compared to NS environment was tested using Student's t-test statistics. Pearson's correlation coefficients were calculated to analyze the association of yield and yield contributing traits in two environments. BLUPs (Best Linear Unbiased Predictors) for yield and yield contributing traits were obtained by fitting a mixed linear model in R-package "lmer" (Robinson 1991; Douglas et al. 2015). These BLUP values were used in GWAS for QTL mapping. Mean broad sense heritability (h^2_{bs}) was calculated by variance components estimated in mixed linear model by using the formula:

$$h^2bs = \frac{\sigma^2 g}{\sigma^2 g + \frac{\sigma^2 gy}{k} + \frac{\sigma^2 e}{rk}}$$

where $\sigma^2 g$ is the genotypic variance, $\sigma^2 gy$ is the variance due to genotype and environment interaction, $\sigma^2 e$ is the error variance, k is the number of environments and r is the number of replications.

DNA extraction and GBS data

High-quality genomic DNA of 71 genotypes was isolated using CTAB (cetyl trimethyl ammonium bromide) method given by Doyle et al. (1990). Genotyping by sequencing (GBS) was outsourced from AgriGenome Labs Pvt Ltd, India using double digestion Restriction site Associated DNA sequencing (ddRAD-seq) platform. Quality clean GBS reads were aligned with the reference genome of *B. juncea* (www.brassicadb.org) using Bowtie2 (version 2–2.2.9) program with default parameters (Langmead and Salzberg 2012). Samtools (samtools version 0.1.18) was used for SNP calling (Li et al. 2009). SNPs, with missing data > 20% and MAF < 5%, were removed.

GWAS analysis of yield and its component traits

GWAS was done with FarmCPU (Fixed and Random Model Circulating Unification) (Liu et al. 2016) package in R statistical software (R Core Team 2018) to identify QTLs for eight test traits. FarmCPU is a multi-locus model that addresses the confounding problem of mixed linear models (MLM) by using population structure and multiple associated markers in fixed effect model. The multiple associated markers are estimated by incorporating kinship values in the random effect model (Liu et al. 2016). Population structure was estimated with principal component analysis (PCA) using GAPIT version 3.0 (Genomic Association and Prediction Integrated Tool) package (Lipka et al. 2012). The significantly associated SNPs based on their associated *p*-values were identified as putative genomic regions associated with yield and yield component traits in *B. juncea*.

Results

Field screening of germplasm stock for terminal heat tolerance

A fixed diversity stock of 491 genotypes was documented for wide range of variations for seed yield under natural terminal heat stress. This stock was

Table 1 Weekly maximum and minimum temperature during reproductive phase of *B. juncea* diversity panel in two environments

Week	Growth stage of crop	Normal sown conditions (NS)		Terminal heat stress conditions (THS)	
		Minimum temperature (°C)	Maximum temperature (°C)	Minimum temperature (°C)	Maximum temperature (°C)
II (8 to 14 Feb)	Siliquae initiation	6.4	21.6	17.9	30.6
III (15 to21 Feb)	Siliquae development	12.9	25.8	22.5	34.5
IV (22 to 28 Feb)	Siliquae development	8.4	24.5	20.9	36.4
I (1 to 7 March)	Seed development & filling	9.7	25.2	21.5	37.7
II (8 to 14 March)[a]	Seed filling	9.4	19.4	21.2	33.5
III (15 to 21 March)	Seed filling & onset of physiological maturity	11.6	26.2	22.3	38.8

[a]32.2 mm rainfall observed during this week

categorized into different groups based on HSI and YD% (Fig. S1a, S1b). The percent yield reduction (YD) under THS in comparison to optimum conditions ranged from − 26.7 to 21.4. The negative values of YD of some genotypes revealed yield augmentation under heat stress and thereby gave preliminary indication for their more adaptation to high temperature during maturity stage. This study also enabled to identify potential terminal heat stress donors for validation and utilization in commercial breeding. Out of 491, top 20 lines which expressed negative YD and low HSI (indicator of susceptibility of a genotype), were marked for their potency as terminal heat stress donor lines and hence, recommended for exploitation in commercial breeding after evaluation in larger plots (Table 2). Amongst these identified donors, MCP-12–211 and MCP-12–624 are introgression lines from *B. juncea x Erucastrum cardaminoides;* PTJ-3–100, PTJ-3–102 and PTJ-3–79 are introgessions from *B. tournefortii* x *B. juncea;* DAR-5 and MSC-5derived from *B.juncea* x *B. carinata* hybridizationand *DJ-149* is resynthesized *B. juncea* line. *Erucastrum cardaminoides,* an annual to biennial crucifer in the Micronesian region, is endemic to rocky places, fields, volcanicrocks and soil and hence, this species is expected to be a source of gene(s) for manybiotic and abiotic stresses (Warwick et al. 2000). *Brassica tournifortii* species, documented for drought resistance genes, has been identified to confer stress tolerance in cultivated Brassicas (Pratap and Gupta 2016). Cross-tolerance is a broadly described phenomenon in plants, whereby a response to one stress also helps to protect the plant from another coincident or subsequent environmental

Table 2 Heat Susceptibility Index (HSI) and percent yield reduction (YD%) of 20 (out of 491) top tolerant genotypes to terminal heat stress under field conditions

Genotype Name	Pedigree	HSI	% YD
RRN-598	Advanced breeding line	− 17.6	− 26.7
B-410	*B. juncea* germplasm line	− 17.4	− 26.5
RGN-253	Advanced breeding line	− 16.6	− 25.2
MCP-12–211	Derived *B. juncea* line from *B. juncea x Erucastrum cardaminoides*	− 16.1	− 24.5
PTJ-3–100	*B. juncea* introgression line from *B. tournefortii* x *B. juncea*	− 15.9	− 24.2
PTJ-3–79	-do-	− 15.9	− 24.2
PBR-378	Variety released for rainfed condition under AICRP	− 15.5	− 23.6
PTJ-3–102	*B. juncea* introgression line from *B. tournefortii* x *B. juncea*	− 15.3	− 23.3
DJ-149	Derived resynthesized *B. juncea*	− 15.2	− 23.1
Varuna	Widely adapted cultivar	− 14.8	− 22.6
MSC-5	Introgression *B. juncea* line from *B.juncea* x *B. carinata hybridization*	− 14.6	− 22.2
MCP-12–624	Derived *B. juncea* line from *B. juncea x Erucastrum cardaminoides*	− 14.5	− 22.0
RB-50	Variety released for rainfed condition under AICRP	− 14.2	− 21.5
PUSA BOLD DT	Determinate version of a commercial cultivar	− 13.8	− 21.0
Kranti	Widely adapted cultivar	− 13.4	− 20.3
PRG-2001–65	*B. juncea* germplasm line	− 13.0	− 19.7
RH-0644	Advanced breeding line	− 12.8	− 19.5
HUJM-02–01	*B. juncea* germplasm line	− 12.8	− 19.4
GMCN-1–2	-do-	− 12.4	− 18.8
DAR-5	Introgression *B. juncea* line from *B.juncea* x *B. carinata hybridization*	− 12.3	− 18.7

stress. The results by Rivero et al. (2014) indicate protection to salt stress following a mild heat stress treatment (35 °C).

Characterization of these distinctive genetic resources will enable their utilization in target oriented commercial breeding. It is pertinent to mention that two varieties viz. RB-50 and PBR 378 released for commercial cultivation under moisture stress (rainfed) conditions under All India Co-ordinated Research Program (AICRP), were also identified as heat stress donors. The widely distributed variability for HSI in stock of 491 lines (Fig. S1a) provided an opportunity to constitute a diverse panel for THS; displaying most susceptible (HSI = 14.1) to most tolerant (HSI = − 17.6) genotypes. This information was supported by YD% (though indirectly determined by HSI too).

Performance of diversity panel under normal and THS environments

The differential response of 96 genotypes to terminal heat stress, based on HSI and YD, was further validated by evaluating under controlled conditions. A significant increase in weekly temperature regimes during different reproductive stages of the crop under controlled conditions in comparison to NS conditions (control) authenticated the screening conditions of panel for THS using polytunnels (Table 1).

Significant variation was observed for seed yield and its component traits under both NS and THS environments (Table 3, Fig. 1). PH ranged from 105.71 cm to 184.63 cm with mean value 148.09 ± 1.96 under NS condition, while under THS condition, PH ranged from 92.04 cm to 170.05 cm with mean of 137.52 ± 1.57. Number of pods on main shoot ranged from 24.52 to 45.77 with mean value 35.34 ± 052 under NS condition, while it ranged from 23.52 to 43.76 with mean of 33.15 ± 0.52 under THS condition. TSW ranged from 2.71 g to 6.96 g with mean value4.27 ± 0.08

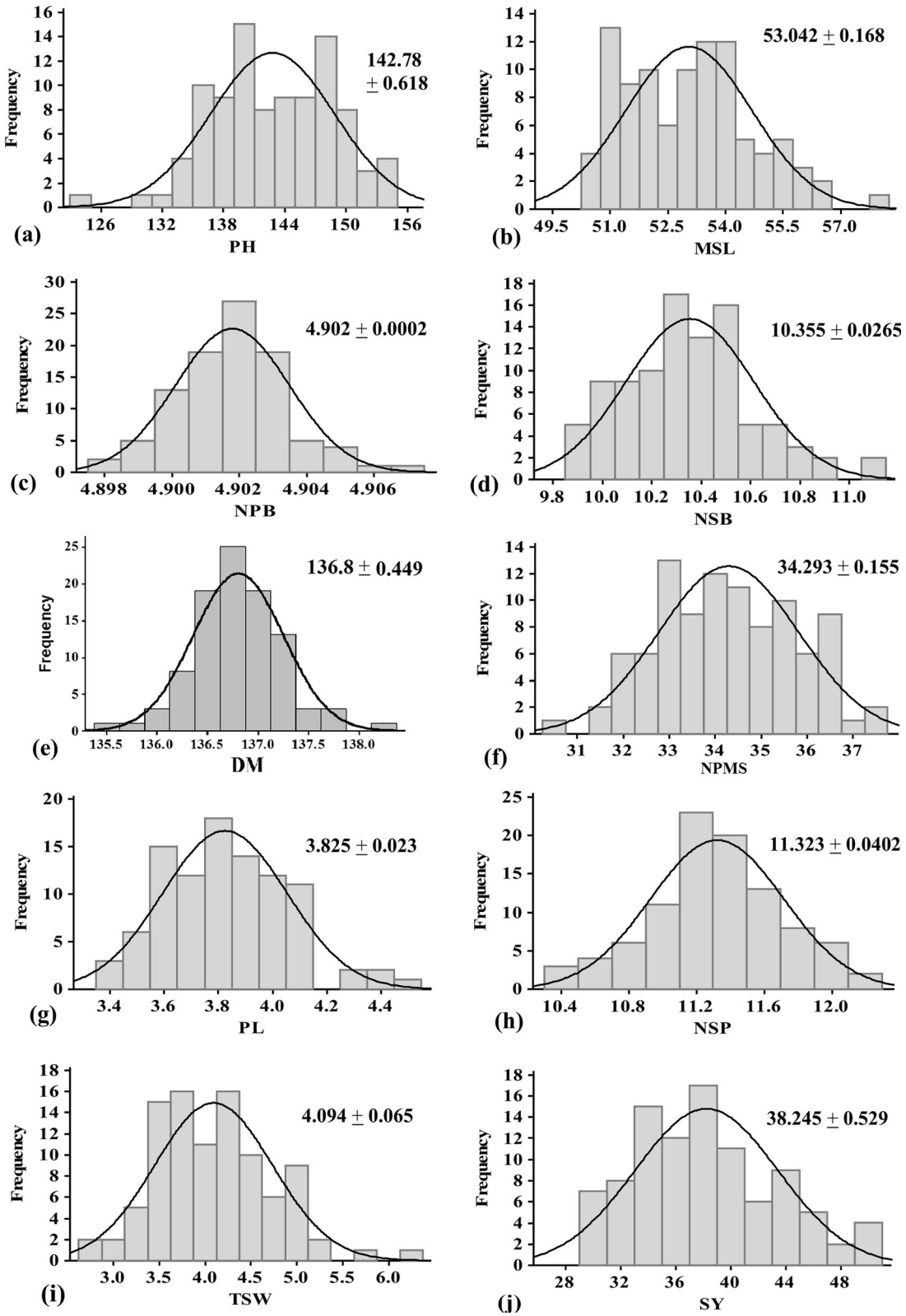
Fig. 1 Frequency distribution of phenotypic traits with normal distribution curve: **a** Plant height (PH) **b** Main shoot length (MSL) **c** Number of primary branches (NPB) **d** Number of secondary branches (NSB) **e** Days to Maturity (DM) **f** Number of pods on main shoot (NPMS) **g** Pod length (PL) **h** Number of seeds per pod (NSP) **i** Thousand seed weight (TSW) **j** Seed yield (SY). *Mean ± SE based on BLUPs for each trait are shown at the top-right

Table 3 Mean, standard error, range, mean square of genotypes and coefficient of variation (CV) of seed yield and its components in normal and stress environments

Traits	Normal sown conditions (NS)				Terminal heat stress conditions (THS)				h_{bs}^2
	Mean ± SE	Range	Mean square (genotype)	CV	Mean ± SE	Range	Mean square (genotype)	CV	
PH	148.09 ± 1.96	105.71–184.63	687.84***	3.89	137.52 ± 1.57	92.04–170.05	434.45***	3.44	0.44
MSL	53.75 ± 0.75	38.84–68.42	97.52***	7.14	52.12 ± 0.73	35.29–74.61	96.37***	10.88	0.30
NPB	4.86 ± 0.07	3.26–6.74	0.94***	9.47	4.94 ± 0.07	3.16–6.70	0.89**	14.06	0.004
NSB	10.38 ± 0.25	4.86–17.76	11.21***	14.08	10.36 ± 0.20	6.65–15.62	7.37***	13.75	0.16
DM	137.56 ± 0.81	134.61–140.55	2.93***	0.59	136.05 ± 0.99	132.01–140.13	5.18***	0.73	0.38
NPMS	35.34 ± 0.52	24.52–45.77	47.50***	6.73	33.15 ± 0.52	23.52–43.76	48.31***	7.65	0.38
NSP	11.50 ± 0.14	8.75–15.63	2.88***	7.98	11.15 ± 0.14	7.48–14.72	2.89***	7.598	0.38
PL	3.78 ± 0.04	3.12–4.71	0.26***	7.06	3.87 ± 0.04	3.02–4.72	0.23***	7.44	0.72
TSW	4.27 ± 0.08	2.71–6.96	1.15***	8.73	3.92 ± 0.07	2.25–6.27	0.96***	6.57	0.90
SY	40.60 ± 0.99	21.90–66.27	170.78***	8.04	35.72 ± 1.06	18.05–67.49	195.91***	9.37	0.63

PH plant height (cm), *MSL* main shoot length (cm), *NPB* number of primary branches, *NSB* number of secondary branches, *DM* days to maturity, *NPMS* number of pods on main shoot, *NSP* number of seeds/pod, *PL* pod length (cm), *TSW* thousand seed weight (g), *SY* seed yield (g/plant). Pod = Siliqua

**significance at $p < 0.01$, ***significance at $p < 0.001$


(a)
PH
142.78 ± 0.618
(b)
MSL
53.042 ± 0.168
(c)
NPB
4.902 ± 0.0002
(d)
NSB
10.355 ± 0.0265
(e)
DM
136.8 ± 0.449
(f)
NPMS
34.293 ± 0.155
(g)
PL
3.825 ± 0.023
(h)
NSP
11.323 ± 0.0402
(i)
TSW
4.094 ± 0.065
(j)
SY
38.245 ± 0.529
Frequency

under NS condition, while under THS, it ranged from 2.25 g to 6.27 g with mean value 3.92 ± 0.07. NPMS and TSW are major SY determining traits and hence any negative effect on these traits results decline in productivity (Sandhu et al. 2019). In NS condition, SY ranged from 21.90 g to 66.27 g with mean of 40.60 ± 0.99 whereas it ranged from 18.05 g to 67.49 g with mean value 35.72 ± 1.06 under THS condition (Table 3).The significant reduction in population means in PH, NPMS, TSW and SY were observed under THS as compared to NS condition (Table 4). Significant high to moderate correlations were obtained between NS and THS environment for seven traits viz. PH, DM, NPMS, PL, TSW, NSP, and SY with correlation coefficient (r) value of 0.28, 0.25, 0.25, 0.55, 0.81, 0.35 and 0.41, respectively (Table 5). BLUP based range of variation for test traits are given in supplementary Table S2.

Phenotypic data of all evaluated traits of 96 genotypes was used to draw a cluster dendrogram using Euclidean distance and hierarchical clustering method "ward.D2". Based on cluster dendrogram, 71 most diverse genotypes were selected for genotyping (Fig. 2, Table S1). Broad sense heritability (h^2_{bs}) estimates of test traits in this study ranged from 0.004 to 0.90 (Table 3). As h^2_{bs} of some traits (e.g. NPB = 0.004, NSB = 0.16) was low, a threshold of h^2_{bs-} = 0.30 was set for GWAS. Only traits having h^2_{bs} exceeding 0.30 (thereby excluding NPB and NSB) were used for GWAS analysis.

Table 4 Significant reduction in mean trait values from normal to heat stress environment

Trait	Reduction ($\overline{NS} - \overline{THS}$) [a]	95% CI	T-Value
PH	10.6	5.63–15.53	4.21***
NPMS	2.19	0.75–3.64	2.99**
TSW	0.35	0.13–0.56	3.13**
SY	4.88	2.01–7.74	3.36**

PH plant height (cm), *NPMS* number of pods on main shoot, *TSW* thousand seed weight (g), *SY* seed yield (g/plant)

[a]*NS* normal sown, *THS* terminal heat stress

**significance at $p < 0.01$, ***significance at $p < 0.001$

Genotyping by sequencing (GBS) data analysis

A total of 1.7 to 6.1 million raw reads was obtained through GBS of 71 genotypes. GC percent ranged from 39.8 to 44.5 with Q_{30} value ranging from 90.75 to 94.29. Overall, 72.23 to 82.42 percent of reads aligned to the reference genome for all the genotypes. A total of 30,673 SNPs were obtained with SNP calling using samtools version 0.1.18. Out of which, 18,258 SNPs having less than 20% missing data and MAF $\geq$ 5% were obtained after filtering. MAF in the filtered SNP data ranged from 0.05 to 0.5. These filtered SNPs were used for genome-wide association analysis incorporating phenotypic data. SNPs were uniformly distributed covering the whole genome of 18 chromosomes of *B. juncea* (Fig. 3). Chromosome-wise distribution of all SNPs is given in supplementary Table S3; with the highest SNPs obtained from B08 chromosome whereas the least numbers from A10 chromosome. Linkage disequilibrium (LD) decay in population was estimated to be $\sim$ 50 kb where r^2 value drops below 0.2 (Fig. S3).

GWAS for yield and its component traits

Preliminary GWAS using eight PC and three PC did not differ significantly for associated *p* values. As, hierarchical clustering based on phenotypic data of diversity panel suggested three clusters in the panel (Fig. 2), first three principal components were

Table 5 Pearson's correlation coefficient for seed yield and its components between NS and THS environments

Traits	PH	MSL	NPB	NSB	DM	NPMS	PL	TSW	NSP	SY
r^2	0.28**	0.15	0.04	0.10	0.25*	0.25*	0.55***	0.81***	0.35***	0.41***

MSL main shoot length (cm), *NPMS* number of pods on main shoot, *NPB* number of primary branches, *NSB* number of secondary branches, *NSP* number of seeds/pod, *PH* plant height (cm), *PL* pod length (cm), *TSW* thousand seed weight (g), *SY* seed yield (g/plant). Pod = Siliqua

*Significance at $p < 0.05$, **significance at $p < 0.01$, ***significance at $p < 0.001$

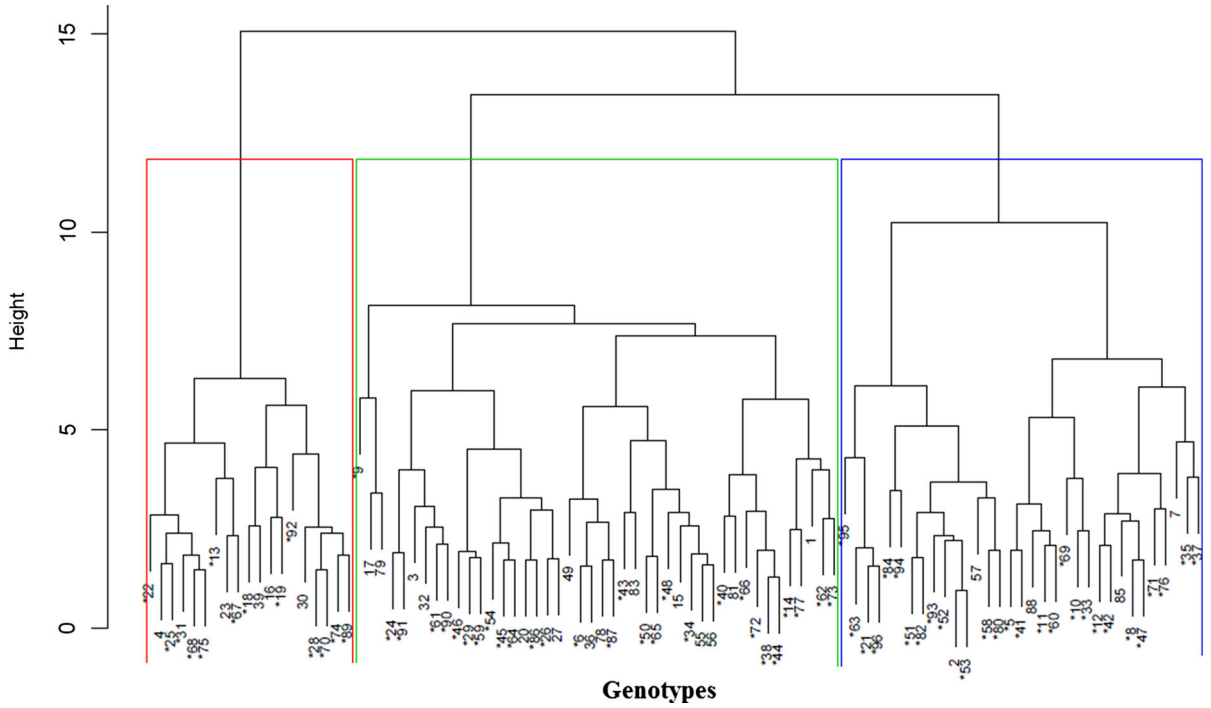


Fig. 2 Cluster dendrogram of 96 genotypes. Table S1 provides genotype names corresponding to the genotype id used in this figure. *Selected genotypes for GBS and GWAS

incorporated as a covariate in the GWAS model as fixed effect (Fig. S2). Kinship among the genotypes was incorporated as a random effect. Marker-trait associations (MTAs) for test traits are discussed as follows:

Plant height (PH)

A total of 15 significant SNPs were identified under THS, NS, and using BLUP values (Tables 6, 7, S4, S5). Under THS, significant associations were observed on five different chromosomes (Fig. 4a, S8a) whereas significant SNPs were located on four different chromosomes under NS (Fig. S4a, S6a). None of the SNP was found to be common in NS, THS and BLUP (Fig. S5a, S7a), thereby revealing association of specific genomic regions under THS (Table 7).

Main shoot length (MSL)

A total of 12 significant SNPs found to be associated with main shoot length (Table 7). Of these, three SNPs were identified under THS (Table 6, Fig. 4b, S8b), four SNPs under NS (Table S4, Fig. S4b, S6b) and five SNPs under BLUP values (Table S5, Fig. S5b, 7b).

Days to maturity (DM)

A total of nine SNPs associated with days to maturity were identified under THS, NS and BLUP (Tables 6, 7, S4, S5). Out of these, two SNPs were identified specifically under THS (Fig. 4c, S8c), while one SNP is identified both under THS and BLUP values (Table 6, S5; Fig. S5c, S7c). Under NS, five SNPs were identified on five different chromosomes (Fig. S4c, S6c).

Number of pods on main shoot (NPMS)

A total of seven SNPs were identified under THS, NS and BLUP values (Tables 6, 7, S4, S5). Of these, two different SNPs associated each under THS (Fig. 4d, S8d) and NS conditions (Fig. S4d, S6d). However, BLUPs identified three SNPs on different chromosomes (Fig. S5d, S7d).

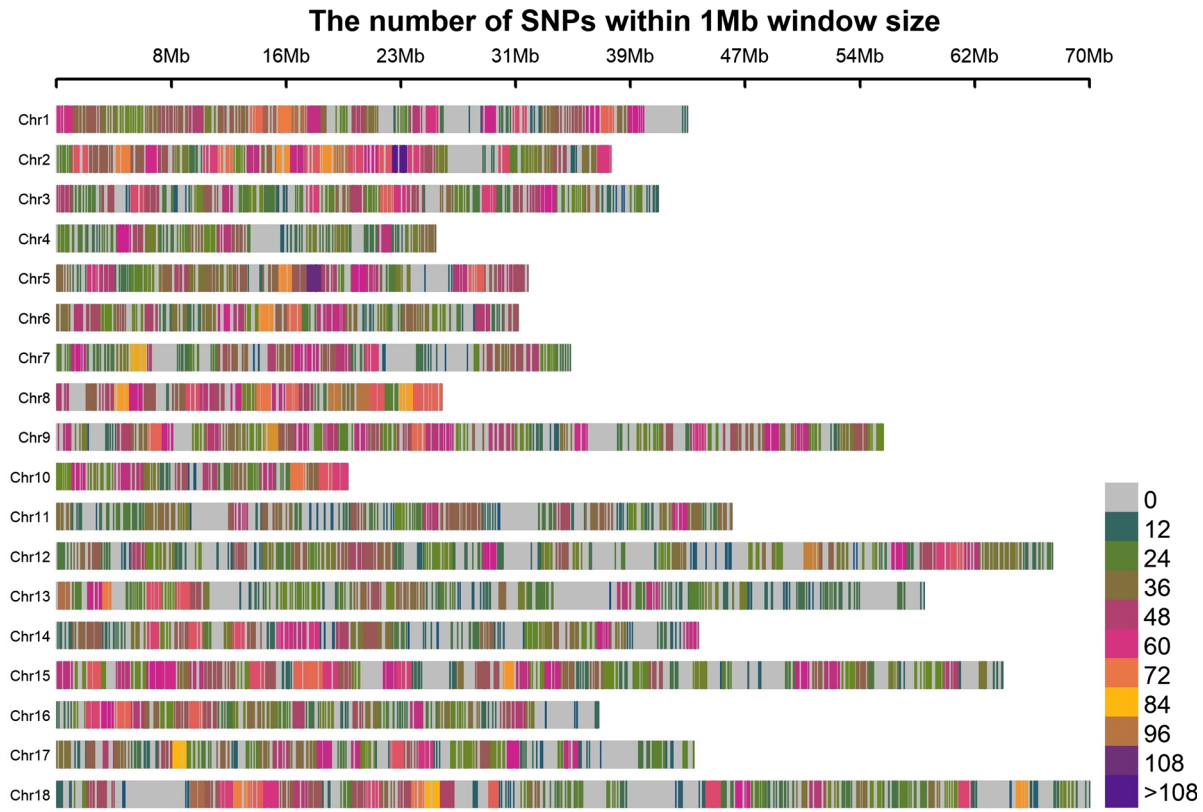


Fig. 3 SNP density plot chromosome wise representing number of SNPs within 1 Mb window size. *Chromosome 1 to 10 represents A genome mentioned as A01 to A10 and chromosome 11 to 18 represents B genome mentioned as B01 to B08 in manuscript

Pod length (PL)

For PL, out of total ten significant SNPs, two were identified under THS (Fig. 4e, S8e).Four SNPs were identified under NS (Fig. S4e, S6e) and BLUP each (Fig. S5e, S7e). All the conditions identified different associations (Tables 6, 7, S4, S5).

Number of seeds per pod (NSP)

A total of ten SNPs were found to be significantly associated under THS, NS and BLUP values (Tables 6, 7, S4 and S5). Under THS, three SNPs were identified (Fig. 4f, S8f) and five SNPs were identified under NS (Fig. S4f, S6f) whereas two SNPs were identified using BLUP values (Fig. S5f, S7f). One SNP (CB05.62648306) was identified under both NS and BLUP.

Thousand seed weight (TSW)

A total of nine significant SNPs mainly distributed on six chromosomes viz. A01, A09, B01, B02, B05 and B08 were found to be associated with TSW (Tables 6, 7, S4 and S5). Out of these, four SNPs were identified under THS (Table 6, Fig. 4g, S8g) and three under NS (Table S4, Fig. S4g, S6g). One SNP (CA09.51591158) located on chromosome A09 was observed under both NS and BLUP values (Table S4, S5, Fig. S4g, S5g).

Seed yield

For SY, a total of 15 significant SNPs were identified under THS, NS and BLUP values (Tables 6, 7, S4 and S5). Of these, one SNP was identified under THS on A09 (Fig. 4h, S8h), which was also identified in BLUP values (Table 6, S5, Fig. S5h, S7h).

Table 6 Significant SNP markers associated with seed yield and its components in *B. juncea* under THS conditions

Traits	SNPs associated	Chromosome	Position on genome (bp)	Alleles	MAF	Effect	$-\log_{10}(p)$ value	Previous reports
PH	CA07.21579885	A07	21,579,885	G/A	0.23	− 5.57	8.03	[1], [2]
	CA07.29396128	A07	29,396,128	G/A	0.45	− 4.59	6.42	[1], [2]
	CA08.29787	A08	29,787	C/T	0.20	5.77	9.10	[1], [2]
	CB05.30793004	B05	30,793,004	A/G	0.14	5.03	6.39	
	CB07.16582224	B07	16,582,224	G/A	0.20	− 6.21	8.57	[2]
	CB08.53250437	B08	53,250,437	A/T	0.07	7.41	6.95	
MSL	CA06.21401508	A06	21,401,508	C/A	0.34	− 2.02	7.65	
	CB01.25205837	B01	25,205,837	A/G	0.06	6.41	8.64	[1]
	CB02.18434222	B02	18,434,222	C/T	0.49	− 1.65	7.33	
DM	CA02.13632528[a]	A02	13,632,528	C/T	0.49	1.36	9.79	
	CB03.32442415	B03	32,442,415	T/G	0.14	− 0.83	7.26	
	CB07.23508366	B07	23,508,366	A/T	0.24	0.45	6.18	
NPMS	CA01.38595643	A01	38,595,643	T/C	0.21	− 2.63	6.02	
	CB05.32430484	B05	32,430,484	T/A	0.21	2.10	5.78	
NSP	CA04.19337826	A04	19,337,826	C/T	0.17	− 0.32	6.77	
	CA06.21230172	A06	21,230,172	T/G	0.11	0.48	6.01	[2], [3]
	CB07.8994243	B07	8,994,243	A/G	0.08	0.77	8.65	[2]
PL	CB07.34074606	B07	34,074,606	A/C	0.20	0.11	4.89	[2]
	CB08.9733991	B08	9,733,991	C/T	0.35	0.08	6.88	[1], [2]
TSW	CA01.14059125	A01	14,059,125	C/G	0.08	0.54	6.35	[2], [4]
	CA09.50886002	A09	50,886,002	C/A	0.34	0.23	5.98	[2], [4]
	CB05.8239837	B05	8,239,837	T/G	0.07	0.52	6.79	[4]
	CB08.24601593	B08	24,601,593	A/C	0.27	0.23	5.62	[1], [2]
SY	CA09.10419698[a]	A09	10,419,698	A/C	0.18	6.28	7.34	

PH plant height (cm), *MSL* main shoot length (cm), *DM* days to maturity, *NPMS* number of pods on main shoot, *NSP* number of seeds/pod, *PL* pod length (cm), *TSW* thousand seed weight (g), *SY* seed yield (g/plant). Pod = Siliqua

[*] [1] = Ramchiary et al (2007), [2] = Yadava et al (2012), [3] = Akhatar and Banga (2015), [4] = Dhaka et al (2017)

[a] SNPs consistently identified in BLUP values

Discussion

Due to low water requirement (80–240 mm), Indian mustard crops fit well in the rainfed cropping system and nearly 30.7% area (1.81 mha) is under rainfed farming in India. Under these condition, high temperatures at germination and terminal stages, drought, salinity and cold spells are major abiotic stresses, leading to low productivity. In addition, changing climatic conditions such as shortening of winter duration has further aggravated the problem of heat stress particularly during reproductive phase (terminal heat stress). The effect of THS on the performance of various yield contributing traits has been documented in Indian mustard (Baghel and Srivastava 2010; Sharma 2014). Delayed sowing is one of the major causes that expose the crop to THS. The effect of THS on maturity span in mustard has been reported by Nanda et al. (1996), Tripathi et al. (2007), Kumar et al. (2008) and Adak et al. (2011). THS led to forced maturity of the genotypes thereby reducing the time span for siliqua formation and seed filling.

In our study, significant reductions in population mean of PH, NPMS, TSW, and SY under THS condition verified the adverse effect of heat stress on the growth and development of *B.juncea.* Characterization of germplasm stock explored the genetic variation for economic traits under terminal heat

Table 7 Comparative chromosomal distribution of SNPs of eight yield associated traits in *B. juncea*

	Normal sown (NS)									Terminal heat stress (THS)								
Traits	PH	MSL	DM	NPMS	PL	NSP	TSW	SY	Total	PH	MSL	DM	NPMS	PL	NSP	TSW	SY	Total
A01		1	1		1			2	5				1			1		2
A02			1					1	2			1						1
A03						1			1									0
A04			1		1				2						1			1
A05	2							2	4									0
A06									0		1				1			2
A07		1		1	1	1			4	2								2
A08						2			2	1								1
A09	1						1		2							1	1	2
A10									0									0
B01							2		2		1							1
B02				1					1		1							1
B03	1								1			1						1
B04									0									0
B05		1	1			1		1	4	1			1			1		3
B06			1		1				2									0
B07		1							1	1		1		1	1			4
B08	1								1	1				1		1		3
Total	5	4	5	2	4	5	3	6	34	6	3	3	2	2	3	4	1	24

	BLUPs computed across environments (BLUPs)									
Traits	PH	MSL	DM	NPMS	PL	NSP	TSW	SY	Total	Grand total[$]
A01									0	**7**
A02			1	1					2	**5**
A03								2	2	**3**
A04									0	**3**
A05					1				1	**5**
A06		1			1			1	3	**5**
A07					1			1	2	**8**
A08									0	**3**
A09		1					2	1	4	**8**
A10	1	1		1					3	**3**
B01						1			1	**4**
B02				1			1	1	3	**5**
B03	1							1	2	**4**
B04								1	1	**1**
B05					1	1			2	**9**
B06									0	**2**
B07	1								1	**6**
B08	1	2							3	**7**
Total	4	5	1	3	4	2	3	8	30	**88**

PH plant height (cm), *MSL* main shoot length (cm), *DM* days to maturity, *NPMS* number of pods on main shoot, *NSP* number of seeds/pod, *PL* pod length (cm), *TSW* thousand seed weight (g), *SY* seed yield (g/plant). Pod = Siliqua

[*] *NS* normal sown, *THS* terminal heat stress, $BLUP_{env}$ BLUPs calculated across NS and THS environments

[$] Grand total represents total number of SNPs identified across environments

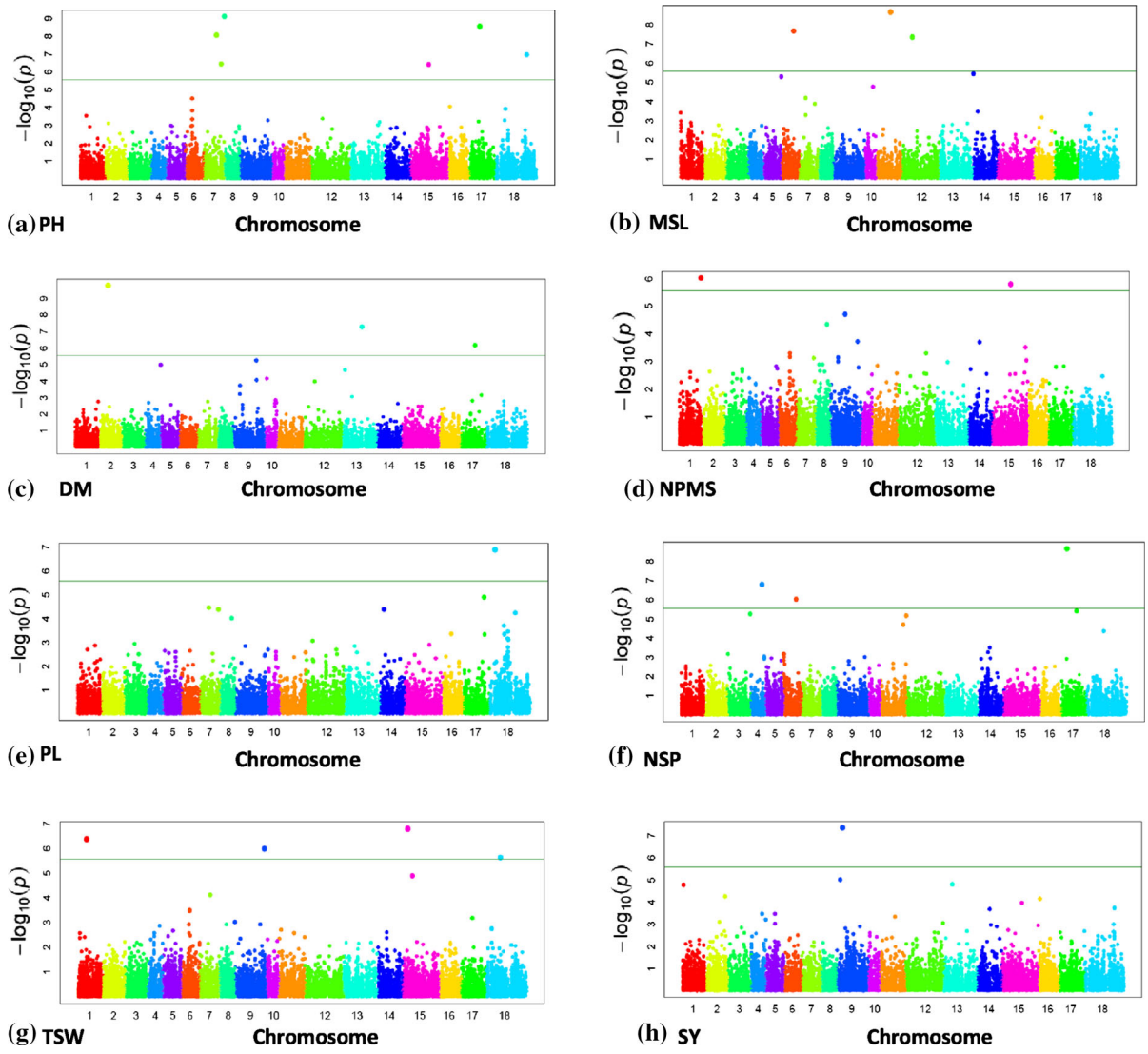


Fig. 4 Seed Yield component traits showing significant marker-trait associations (MTAs) for THS environment: **a** Plant height (PH) **b** Main shoot length (MSL) **c** Days to Maturity (DM) **d** Number of pods on main shoot (NPMS) **e** Pod length (PL) **f** Number of seeds per pod (NSP) **g** Thousand seed weight (TSW) **h** Seed yield (SY). *Chromosome 1 to 10 represents A genome mentioned as A01 to A10 and chromosome 11 to 18 represents B genome mentioned as B01 to B08 in manuscript

stress and identified potent heat tolerant donors. Yield augmentation under heat stress conditions in introgression lines developed from *Erucastrum cardaminoides* and *B. tournefortii* inferred the potential role of thesewild species to conferabiotic stress tolerancein *B. juncea*.

FarmCPU has been proposed as the best model to minimize false-positives and false-negatives in identifying significantly associated SNPs (Sanchez et al. 2018). FarmCPU is a multi-locus model that differs from the widely used mixed linear model (MLM) as it takes care of overwhelmed false positives and painful false negatives by dividing modified MLM into fixed effect and random effect model and using them iteratively (Liu et al. 2016). To minimize the effect of environmental variation, phenotypic BLUP values were used for association analysis in various studies (Henderson 1975; Kump et al. 2011; Hao et al. 2012). In this study, adjusted means for NS and THS separately and BLUP values over the environments

were used to identify QTLs (associated SNPs) for seed yield and its contributing traits under heat stress conditions. A comparison of QTLs in all the three conditions (NS, THS, and BLUP) helped to identify the QTLs that are associated under heat stress conditions.

Under THS, a total of 24SNPs were specifically detected that were not found associated under NS. This inferred the association of specific genomic regions/QTLs under THS and emphasized that SNPs detected under NS might not be associated in THS. Moderate to low correlation between SY and its components (except TSW) evaluated under NS and THS supported these findings. Further, phenotypic evaluation at stress specific location is mandatory to identify trait specific variables. One SNP (CA09.10419698 with SNP effect 6.28, $-\log_{10}(p)$ 7.34) was identified for SY under THS, which is also found to be co-localized with the SNP identified for SY using BLUP values. It may be possible to develop marker for effective selection for SY, a quantitative trait with high G x E interaction. Besides SY traits, three SNPs were identified for DM too. Out of three, one SNP located on chromosome A02 was also identified using BLUP values.

Several publications of earlier studies have identified SNPs associated with yield and yield component traits. As an example,QTLs for PH, PL, MSL, NPMS, TSW, SY have been reported by several workers (Ramchiary et al. 2007; Yadava et al. 2012; Akhatar and Banga 2015; Dhaka et al. 2017). Since all these studies were conducted in normal conditions utilizing variability of different traits in bi-parental or association mapping panels, genomic positions of these QTLs were different as compared to findings of our study. Similarly, earlier studies reported QTLs for multiple traits viz. PH, NSP, TSW (Ramchiary et al. 2007; Yadava et al. 2012; Dhaka et al. 2017) clustered on chromosome A07. We have identified SNPs for four traits viz., PH, MSL, TSW, and SY harbored on chromosome A09.

THS tolerance is an important trait to sustain the production of *B. juncea* in present state of climatic shocks and shortening of winters. The available variability for this trait can be commercially utilized. Development of heat stress tolerant varieties will facilitate to expand cultivation of *B. juncea*, which currently predominated in the Indian subcontinent, across arid and low-rainfall areas. Once validated, the putative SNPs detected in the present study can be used to select heat tolerant donor lines having favourable alleles for days to maturity, seed yield and its related traits. This will lay a strong foundation for marker assisted accelerated breeding programe for heat tolerance in Indian mustard.

Acknowledgements The Indian mustard germplasm, used in this study, was collected/developed and maintained by ICAR National Professor Dr. S. S. Banga. The authors duly acknowledge the receipt of germplasm and his inputs for the preparation of the manuscript.

References

Adak T, Narjary B, Chakravarty NVK (2011) Response of *Brassica* to micro environment modification under semi-arid agro ecosystem. Indian J Agric Sci 81(8):744–750

Akhatar J, Banga SS (2015) Genome-wide association mapping for grain yield components and root traits in *Brassica juncea* (L.) Czern&Coss. Mol Breed 35(1): 48.

Alam M, Begum F, Roy P (2014) Yield and yield attributes of rapeseed-mustard (*Brassica*) genotypes grown under late sown condition. BD J Agril Res 39:311–336

Azharudheen TPM, Yadava DK, Singh N, Vasudev S, Prabhu KV (2013) Screening Indian mustard [*Brassica juncea*(L.) Czern and Coss)] germplasm for seedling thermo-tolerance using a new screening protocol. Acad J 8:4755–4760

Baghel RS, Srivastava SBL (2010) Assessment of mustard genotypes for thermo-tolerance at seed development stage. Ind J Plant Physiol 15:36–43

Chauhan JS, Meena ML, Saini MK, Meena DR, Singh M, Meena SS, Singh KH (2009) Heat stress effects on morpho-physiological characters of Indian mustard (*Brassica juncea* L.). 16th Australian Research Assembly on Brassicas, Ballarat Victoria, 91–97.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genet 12(7):499

Dhaka N, Rout K, Yadava SK, Sodhi YS, Gupta V, Pental D, Pradhan AK (2017) Genetic dissection of seed weight by QTL analysis and detection of allelic variation in Indian and east European gene pool lines of *Brassica juncea*. Theor Appl Genet 130(2):293–307

DhaliwalLK Hundal SS, Chahal SK (2007) Agroclimatic indices of Indian mustard (*Brassica juncea*) under Punjab conditions. Indian J of Agri Sci 77:88–91

Douglas B, Martin M, Ben B, Steve W (2015) Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 67(1):1–48. https://doi.org/10.18637/jss.v067.i01

Doyle JJ, Doyle JL, Hortoriun LB (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fischer RA, Maurer R (1978) Drought resistance in spring wheat cultivars I grain yield responses in spring wheat. Aust J Agric Sci 29:892–912

GOI (Government of India). 2017. In: Third advance Estimates of Production of Food grains for 2016–17. Agricultural

Statistics Division, Department of Agriculture Cooperation & Farmers welfare, GOI, New Delhi. http:/agricoop.nic.in/statistics/advance-estimate-apy-english

Hao D, Chao M, Yin Z, Yu D (2012) Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. Euphytica 186(3):919–931

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 423–447.

Kaur H, Gupta S, Kumar N, Akhatar J, Banga SS (2014) Progression of molecular and phenotypic diversification in resynthesized *Brassica juncea* (L) gene pool with determinate inflorescence. Euphytica 199:325–338

Kumar R, Singh RP, Yeshpal (2008) Yield and quality of Brassica species as influenced by different dates of sowing and varieties. Pantnagar J Res 6(1):6–11

Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, Balint-Kurti PJ (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat Genet 43(2):163–169

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397–2399

Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS Genet 12(2):1005767

Morrison MJ, Stewart DW (2002) Heat stress during flowering in summer Brassica. Crop Sci 42:797–803

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

Nanda R, Bhargava SC, Mamta G (1996) Effect of flowering time on rate of dry matter accumulation, duration of filling period and total oil content in *Brassica* seeds. Indian J Plant Physiol 1(2):88–92

Nuttall WF, Moulin AP, Townley SLJ (1992) Yield responses of canola to nitrogen, phosphorous, precipitation and temperature. Agron J 84:765–768

Pratap A, Gupta SK (2016) Biology and ecology of wild crucifers. In Biology and breeding of crucifers (pp. 46–76). CRC Press

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramchiary N, Padmaja KL, Sharma S, Gupta V, Sodhi YS, Mukhopadhyay A, Arumugam N, Pental D, Pradhan AK (2007) Mapping of yield influencing QTL in *Brassica juncea*: implications for breeding of a major oilseed crop of dryland areas. Theor Appl Genet 115(6):807–817

Rivero RM, Mestre TC, Mittler RO, Rubio F, Garcia-Sanchez FR, Martinez V (2014) The combined effect of salinity and heat reveals a specific physiological, biochemical and molecular response in tomato plants. Plant, Cell Environ 37(5):1059–1073

Robinson GK (1991) That BLUP is a good thing: the estimation of random effects. Statistical Sci 6(1):15–32

Sanchez DL, Liu S, Ibrahim R, Blanco M, Lubberstedt T (2018) Genome-wide association studies of doubled haploid exotic introgression lines for root system architecture traits in maize (*Zea mays* L.). Plant Sci 268:30–38

Sandhu SK, Kang MS, Akash MW, Singh P (2019) Selection indices for improving selection efficiency in Indian mustard. Journal of Crop Improvement 33(1):25–41

SAS Institute Inc. (2011) Base SAS® 9.3 procedure guide. SAS Institute Inc, Cary

Sharma P (2014) Heat tolerance in Indian mustard at seedling and terminal stages. Ann Plant Soil Res 16:224–228

Sharma P, Sardana V (2016) Screening of Indian mustard (*Brassica juncea*) for thermo tolerance at seedling and terminal stages. J Oilseeds Brassica 1(1):61–67

Srivastava JP, Balkrishna (2003) Environmental parameters influencing phonological development of mustard in relation to yield. Indian J of Plant Physiol 8:349–353

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES IV (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28(3):286

Tripathi MK, Rao VUM, Singh D (2007) Effect of sowing time and in season growth manipulations on phenology of Indian mustard *(Brassica juncea)*. Indian J Agric Sci 77(4):253–257

Tuteja SS, Lakpale R, Tripathi RS (1996) Effect of date of sowing, nitrogen level and nitrogen splitting on mustard (*B. juncea*). Adv Plant Sci 9:167–168

Warwick SI, Francis A, La Fleche J (2000) Guide to wild germplasm of Brassica and allied crops (tribe Brassiceae. Brassicaceae), AAFC-ECORC Contribution no, p 40

Woods DL, Capcara JJ, Downey RK (1991) The potential of mustard (*Brassica juncea* (L.) Coss) as an edible oil crop on the Canadian Prairies. Can J of Plant Sci 71(1):195–198.

Yadava SK, Arumugam N, Mukhopadhyay A, Sodhi YS, Gupta V, Pental D, Pradhan AK (2012) QTL mapping of yield-associated traits in *Brassica juncea*: meta-analysis and epistatic interactions using two different crosses between east European and Indian gene pool lines. Theor Appl Genet 125(7):1553–1564

Young LW, Ron WW, Peta C (2004) High temperature stress of *Brassica napus* during flowering reduces micro and megagametophyte fertility, induces fruit abortion and disrupts seed production. J Exp Bot 55:485–495

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. The plant genome 1(1):5–20