Theor Appl Genet (2015) 128:221–234
DOI 10.1007/s00122-014-2423-7

ORIGINAL PAPER

Population structure and breeding value of a new type of *Brassica juncea* created by combining A and B genomes from related allotetraploids

Mehak Gupta · Shilpa Gupta · Hitesh Kumar · Nitin Kumar · S. S. Banga

Received: 23 January 2014 / Accepted: 2 November 2014 / Published online: 16 November 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract

Key message **Derived amphiploidy helped to resynthesize agronomically superior *B. juncea* germplasm which showed high heterosis in crosses with natural *B. juncea*. This new procedure facilitates a seamless flow of variation across *Brassica* digenomics.**

Abstract *Brassica* digenomics, artificially resynthesized by hybridizing extant genome donor diploids, show poor breeding value due to the linkage drag associated with diploid donors. We recently developed a method that involves resynthesis through hybridization between related allotetraploids. Derived *B. juncea* was created by combining A and B genomes extant in *B. napus* and *B. carinata*, respectively. Large genomic and agronomic modifications resulted. Population structure analysis based on the DNA polymorphism generated using 108 locus-specific SSR primers helped to identify three pools of allelic diversity. Thirteen progenies with determinate plant growth habit were discovered, and these aligned closely with B genome of the donor species like *B. nigra* and *B. carinata.* The indeterminate group showed greater genetic affinity with extant *B. juncea*. Derived genotypes possessed high agronomic potential. Importantly, high heterosis was observed in crosses between derived and natural *B. juncea*. Some derived *juncea* progenies figured in heterotic combinations during both the years of F_1 hybrid evaluation. In essence, the hybrids between derived *B. juncea* and natural *B. juncea* can be considered as interspecific hybrids between *B. juncea* and *B. napus* for A genome and between *B. juncea* and *B. carinata* for B genome. This possibly explains their high heterosis-inducing potential. Integrating genetic diversity with the inherent breeding value allowed more efficient prediction of heterosis. Besides generation of new novel variability of huge economic importance and operational simplicity, the method of derived amphiploidy allows a seamless flow of heritable variation across *Brassica* digenomics.

Communicated by Heiko C. Becker.

M. Gupta · S. Gupta · H. Kumar · N. Kumar · S. S. Banga (✉)
Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana 141001, India
e-mail: nppbg@pau.edu

Introduction

Indian mustard (*Brassica juncea*, AABB; $2n = 36$) is a natural amphiploid between *B. rapa* (AA; $2n = 20$) and *B. nigra* (BB; $2n = 16$). It arose several times through independent hybridization events in the sympatric areas of the diploid progenitor species (Prakash et al. 2009; Kaur et al.2014). Its genetic base, restricted by the dual bottlenecks of polyploidy and domestication, stands further eroded by intensive plant breeding activities (Chauhan et al. 2011; Banuelos et al. 2013). Narrow genetic base coupled with a lack of architectural variations is now undermining the ability of plant breeders to improve productivity in this premier oilseed crop of Indian sub-continent. *B. juncea* also possesses adaptive potential for drier ecologies of Australia and Canada (Wijesundera et al. 2008). Like other alloploid crops, Indian mustard expresses low heterosis (Bansal et al. 2012). As allelic diversity is considered central to the expression of heterosis (Falconer and MacKay 1996), intensive efforts have been made to enlarge its genetic base by exploiting germplasm variation (Banga and Labana 1984; Jain et al. 1994) or by artificial resynthesis (Bansal et al. 2009, 2012). Experiences have shown that the inclusion of unadapted and widely diverged germplasm can

broaden the genetic base, but may not influence the breeding outcomes in terms of improved selection efficiency and heterosis. In the related crop of *B. napus*, concept of sub-genome was introduced to extend the limits of *B. napus* gene pool by incorporating allelic diversity in all *Brassica* species harboring A and C genome(s) (Zou et al. 2010). According to this, A genome of *B. rapa* and C genome of *B. carinata* we considered diverse from A and C genome(s), respectively of *B. napus* (Song et al. 1995). Partial introgression of sub-genomic components from A genome of *B. rapa* and C genome of *B. carinata* into *B. napus* helped to increase heterosis (Zou et al. 2010). Cytogenetic complexities, phenotypic instabilities and linkage drag, have not allowed the wider adaptation of this method.

We have developed a new procedure, termed derived amphiploidy (Banga and Kaur 2009; Gupta et al.2014). It envisages an alternate pathway to polyploidy by sourcing desired diploid genomes from related allotetraploids, instead of the extant diploid progenitors. We hypothesized that obtaining genetic diversity of high breeding value from elite strains of the non-parental allotetraploids will improve the outcomes in terms of the higher selection efficiency. We also expected to overcome limitations associated with traditional methods of resynthesis from diploid progenitors or by exploiting sub-genomic variation. To create *B. juncea* ($A^nA^nB^cB^c$); *B. napus* ($A^nA^nC^nC^n$) cv. Surpass 400 was hybridized as female with *B. carinata* ($B^cB^cC^cC^c$) cv. PC5. Both of these are OP varieties and were under commercial cultivation in Australia and India, respectively in the recent past. Chromosome doubling was induced in the F_1 ($A^nB^cC^nC^c$) to create octoploid plants ($A^nA^nB^cB^cC^cC^nC^n$) or octoploid sectors on F_1 plants. Such octoploid plants or octoploid sectors showed an aberrant meiosis and impaired seed fertility, but some selfed or OP seed could be obtained. The chromosomes in the tetrasomic dose ($C^cC^cC^nC^n$) formed many multivalents and as a consequence had poor transmission frequency. This led to the preferential elimination of C genome chromosomes in the selfed progenies. Occurrence of quadrivalents with mean frequencies ranging between 1.5 and 4.0 was also reported in the previous studies with *B. oleracea* autotetraploids (Prakash and Tsunoda 1983). In contrast, the chromosomes in disomic dose (A^nA^n/B^cB^c) formed bivalents which were retained. *B. juncea* ($A^nA^nB^cB^c$) segregates could be scored easily due to their phenotype in A2/A3 generations. These were further confirmed by meiotic studies and only plants with euploid ($2n = 36$) chromosome number were retained and maintained by selfing. The new method is conceptually very different from the concept of intersub-genome as this method aims at recombining two entire genomes from different but related allotetraploids to recreate a non-parental allotetraploid. In the case of intersub-genome, aim is only the partial introgression of genetic components from different species into one or two genomes of an existing species e.g., *B. napus* (Zou et al.2010; Fu et al.2012).

In this communication, we present results from substantive polyploid events to reflect on phenotypic and population structural changes in the derived *B. juncea* amphiploids after 5–6 generations of selfing using single-seed descent method. These results demonstrate that besides mobilizing superior genetic variation available in related allotetraploids, the new method also capitalizes on enormous de novo variation that arise following resynthesis. The agronomic superiority of newly evolved genotypes and their application to improve predictability and expression of heterosis are also demonstrated. This idea can be extended to other polyploid crops.

Materials and methods

Phenotypically normal derived *B. juncea* plants with high pollen grain and seed fertility, identified in A2/A3 generations following resynthesis, were advanced through single-seed descent method (Fig. 1). Only 62 A_5/A_6-derived *B. juncea* allopolyploid progenies with euploid chromosome number were evaluated to infer genetic diversity and morphological variation. Experiments were conducted at the Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana during winters of 2010–2011 and 2011–2012.

Population structure and genetic diversity

Molecular investigations involved extant diploid genome donors [*B. rapa* (AA), *B. nigra* (BB), *B. oleracea* (CC)]; amphiploid species *B. napus* (AACC) and *B. carinata* (BBCC); and resynthesized derived *B. juncea* ($A^nA^nB^cB^c$). Natural *B. juncea* was used as a control. For genotyping, 60 A genome (Kim et al. 2009) and 48 B genome (Isobel Parkin pers. comm.) chromosome-specific polymorphic SSR markers were selected. These amplified only A or B alleles, respectively as checked by including *B. rapa* (AA) and *B. nigra* (BB) along with natural *B. juncea*, derived *B. juncea* bulk, *B. napus* and *B. carinata* for initial polymorphism assays. DNA was isolated from young leaves of test genotypes using the standard procedure (Doyle and Doyle 1990). PCRs were carried out on an AB Gene Amp PCR system 9,700. The reaction volume was 20 µl containing 50 ng genomic DNA, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 3 mM MgCl2, 400 µM dinucleotides, 1 µM each primer, and 1U of *Taq* DNA polymerase (Roche Applied Science). The temperature cycles were programmed as 95 °C for 5 min followed by 35 cycles at 94 °C for 30 s., 55 °C for 45 s, 72 °C for 50 s, and finally 7 min at 72 °C for a final extension. PCRs were analyzed by a

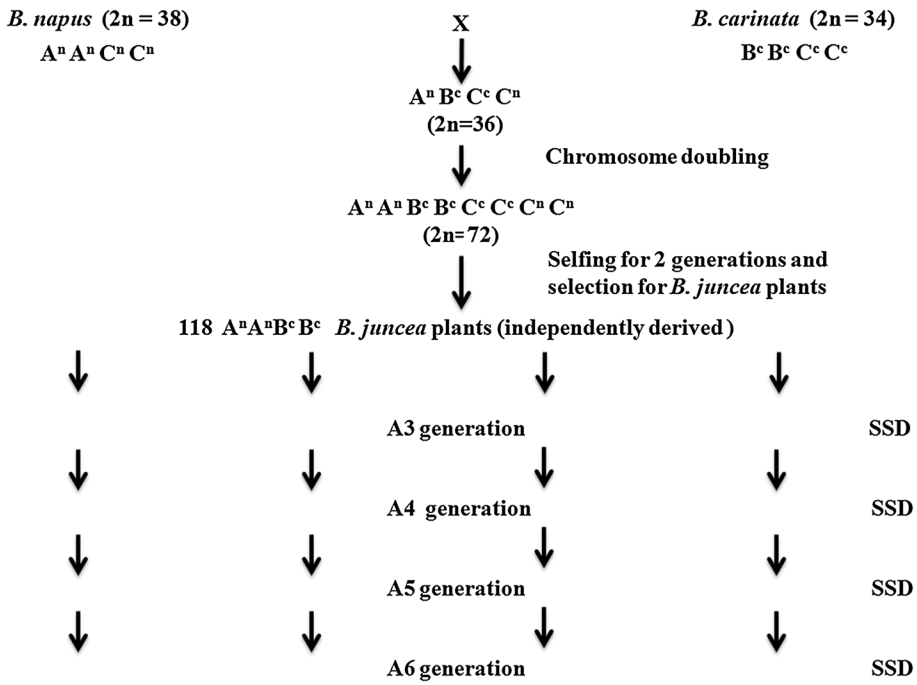

Fig. 1 Procedure for the development of derived *Brassica juncea*. *B. juncea* ($A^nA^nB^cB^c$) was developed by hybridizing *B. napus* ($A^nA^nC^nC^n$) as a female with *B. carinata* ($B^cB^{c-}C^cC^c$), followed by chromosome doubling, selfing and selection for *B. juncea* type plants

high-throughput capillary electrophoresis system (Caliper LabChip GX version 3.0.618.0). The PCR products were automatically sized relative to the internal standard. Each band position was considered a single allele.

The program STRUCTURE (Pritchard et al. 2000) was used to identify *K* discrete subpopulations based on models characterized by non-correlated gene frequencies in the absence of admixture. We used two Bayesian Markov Chain Monte Carlo programme STRUCTURE, to infer lineages that show clusters of similar genotypes (Pritchard et al. 2000). Although STRUCTURE assumes Hardy–Weinberg equilibrium and linkage equilibrium within populations, it can be also applied to partially inbred genotypes by randomly choosing one allele. The program allows differentiation of populations based on their unique allele frequency profiles. The population membership was adjusted until the point of the highest goodness of fit as measured by Pr (*X/K*). A complete scoring matrix was used to examine population structure using both admixture and no admixture commands. Values of $K = 1$ through $K = 10$ were tested for each data set by identifying a minimum of three numerical solutions for each combination. Each solution was optimized using 1,000,000 iterations (including 500,000 'burn in' iterations). A value of *K* was selected as the minimum *K* at which, in most numerical situations, Pr (*X/K*) no longer increased with appreciating values of *K*. A representative solution was selected, and the membership of each variety in one of the *K* subpopulations was designated by a numeric index, as the 'Structure Coefficient'. Results of structure analysis were confirmed by BAPS software (Corander et al. 2008). The Bayesian method available in the BAPS software was also used to estimate the number of clusters. Clustering of individual accessions was carried out using the model for non-linked markers and 100 replicate runs of the algorithm with the upper-bound values (*K*) for the number of clusters ranging between 2 and 10. The ideal genetic structure obtained from the BAPS program was also analyzed with respect to inherent genetic diversity and relation between the genotypes across the clusters. PCA of 62 derived *B. juncea* genotypes was also carried out using software DARwin (Perrier and Jacquemoud-Collet 2006) based on the number of SSR alleles.

Phenotyping derived *B. juncea* lines for morphological and agronomic traits

Derived *B. juncea* lines were evaluated against two prevalent OP cultivars, PBR210 and RLC1 during 2010–2011 as per alpha-lattice field design with two replications. The experiment was repeated during 2011–2012 with a bigger set of 58 (including 42 genotypes from 2010 to 11). Each derived *B. juncea* genotype was sown in a plot size of 4.5 m^2 along with commercial genotypes like PBR210, RLC1, PBR91, PBR357 and hybrid CORAL. A Chinese *B. juncea* genotype, AMH 100, was also used as an additional check. The data for morphological characteristics were recorded from five random plants/genotypes/replications. Investigated traits include days to 50 % flowering (FLW50), days to 100 % flowering (FLW100), days to maturity (DM), plant height (cm) (PH), main shoot length (cm) (MSL), number of primary branches per plant (PB), number of secondary branches per plant (SB), number of

pods on the main shoot (PMS), pod length (cm) (PL) and yield (g) (SY). All the plants in the test plots were bulked harvested for yield determination. Oil and protein content in seeds was recorded on FOSS NIRS system model 6,500 using the standard protocols.

Estimating heterosis in a set of hybrids involving derived and natural *B. juncea*

Forty derived *B. juncea* genotypes were crossed as lines with four natural *B. juncea* lines (RLC 1, PBR 210, CBJ 002 and AMH) as testers to generate 160 F_1 hybrids. RLC1 and PBR 210 were the best commercial OP genotypes of the region while CBJ 002 and AMH were the newly developed high-yielding fertility restorers for *ogura* CMS in *B. juncea*. The test genotypes along with the commercial checks were sown in an alpha-lattice design with two replications during 2010–2011. Each genotype was grown on a plot size of 2.7 m^2/replication. The experiment was repeated, with a bigger plot size of four rows for each genotype per replication, during 2011–2012 for a larger number of F_1 hybrids to reconfirm the results as well as to evaluate new hybrids. Morphological data were collected as described earlier. Standard heterosis was estimated as an increase or decrease of hybrid over a commercial variety (testers) expressed as its percent. In most of the F_1 hybrids, testers were also the better parent(s).

$$\text{Commercial Heterosis} = \frac{\overline{F_1} - \overline{\text{variety}}}{\overline{\text{variety}}} \times 100$$

where variety = mean of tester/standard check genotype.

The standard error of the difference, wherever applicable, was calculated as follows:

$$\text{SE} = \sqrt{\frac{2\text{ems}}{r}}$$

Critical difference was computed by multiplying S.E. with respective '*t*' value at error d.f. at 5 % level of significance.

Statistical analysis

Analysis of variance was conducted as per the alpha-lattice or randomized block design using SAS software with PROC GLM as a nested model. Regression analysis was carried out using Minitab. Statistical significance was considered at the $p < 0.05$ level. *R* package was used to draw frequency curves.

Results

Meiotic analysis was first carried out for all the SSD-derived A_5 plants to reconfirm their chromosome number. Majority of the genotypes showed typical 18II configuration (Fig. 2a) with 18-18 anaphase distribution (Fig. 2b). All such plants were also crossed with natural *B. juncea*. Eleven combinations (derived × natural *B. juncea*) showed 17II + 2I (Fig. 2c) meiotic configuration at metaphase and aberrant anaphase distribution. These were not included for present studies as each one of such lines was expected to carry a substituted chromosome. Only 62 A_5/A_6-derived *B. juncea* allopolyploid progenies ($A^nA^nB^cB^c$) with euploid chromosome number were then assayed for genetic diversity and morphological variation.

Population structure and genetic diversity

Analysis of DNA polymorphism data using STRUCTURE software allowed allocation of individuals into clusters by estimating the probable membership coefficients (*Q*) for individuals in each cluster. In this analysis, a model was assumed in which there are *K* populations (where *K* may be unknown), each of which reflects the allele frequencies at each locus. Owing to the nature of origin of the studied germplasm, we preferred not to assume that clusters of individuals represent natural populations. For all the analyses conducted, KSD and DK resulted in a different number of clusters being most likely, but $K = 3$ presented a high likelihood value in all the surveys. Membership coefficients (*Q*) for assigning individuals into clusters were almost similar for each *K* value, resulting in analogous clustering and assignment patterns for a given *K*. Clustering patterns were verified, and these appeared same for a particular *K* value. At $K = 3$, three allelic pools were formed. These are distinguishable in Fig. 3 by three different colors: indeterminate derived and natural *B. juncea* (yellow), determinate derived *B. juncea*, *B. nigra*, *B. carinata* and *B. oleracea* (blue) and admixture of red and blue for *B. napus* and *B. rapa*.

The function 'Changes of log likelihood' in the BAPS graph menu provided a model-based investigation of the 'genetic shapes' of the perceived populations (clusters). For this, a 'source' cluster was selected and compared with a set of 'target' clusters. This helped in discerning changes of the log-marginal likelihood of the mixture clustering model when an individual is re-allocated from the source cluster to a target cluster (Fig. 4a, b, c). Negative changes in the log-marginal likelihood close to zero indicate that the mixture model identifies both assignments (source cluster and target cluster) to be statistically reasonable for an individual. In contrast, values further away from zero showed the relative reduction in the genetic affinity of the clusters. A bimodal curve in case of cluster two (green) revealed that the estimated population comprised two parts with some member's showing distinct genetic affinity toward cluster three. The values of the log-marginal likelihood changes were concentrated over a rather short interval as reflected by higher peak size. In contrast, affinities of the members of cluster 3 (blue curve) are more evenly

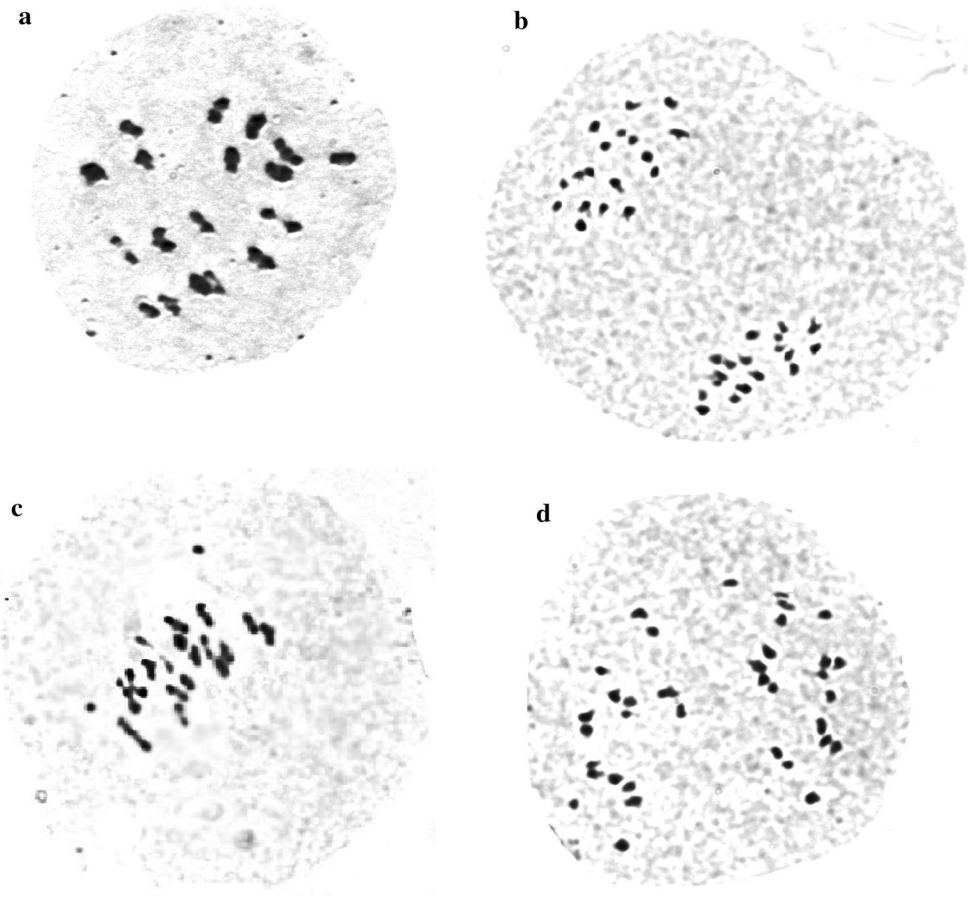

Fig. 2 Meiotic configuration of derived *B. juncea* genotypes. **a** 18II during meiotic metaphase-1, **b** perfect 18I–18I distribution at anaphase-1, **c** 17II+2I at metaphase-1, **d** 19–17 distribution at anaphase-1

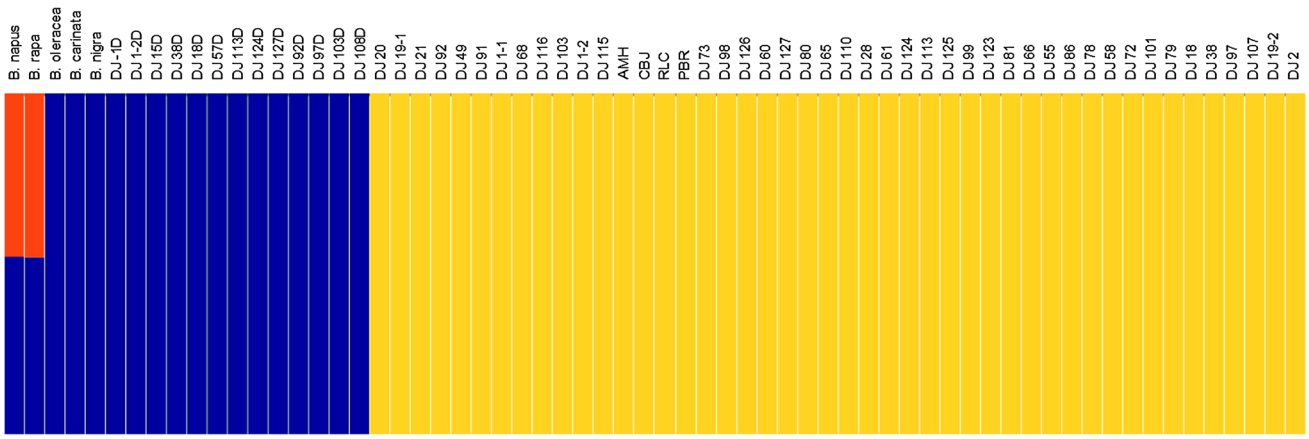

Fig. 3 Population structure in derived *Brassica juncea* genotypes showing three allelic pools namely, indeterminate derived and natural *B. juncea* (*yellow*), determinate derived *B. juncea*, *B. nigra*, *B.carinata* and *B.oleracea* (*blue*) and admixture of *red* and *blue* for *B. napus* and *B. rapa*

distributed. Members of cluster 1 seemed to have nearly similar level of affinity with those of clusters 2 and 3. Diversity tree generated by BAPS suggested closer affinity between cluster two (derived determinates) and cluster three (derived indeterminates and natural *B. juncea*). Cluster one (*B. napus, B. rapa* and *B. oleracea*) was distinct (Fig. 4d). Genotypes included in three BAPS defined cluster are presented in Table 1.

Principal Component analysis (PCA) was also used as an alternative way of visualizing the genotype data. It was calculated from SSR data and was used in the factorial analysis performed by the DARwin software for generating the PCoA

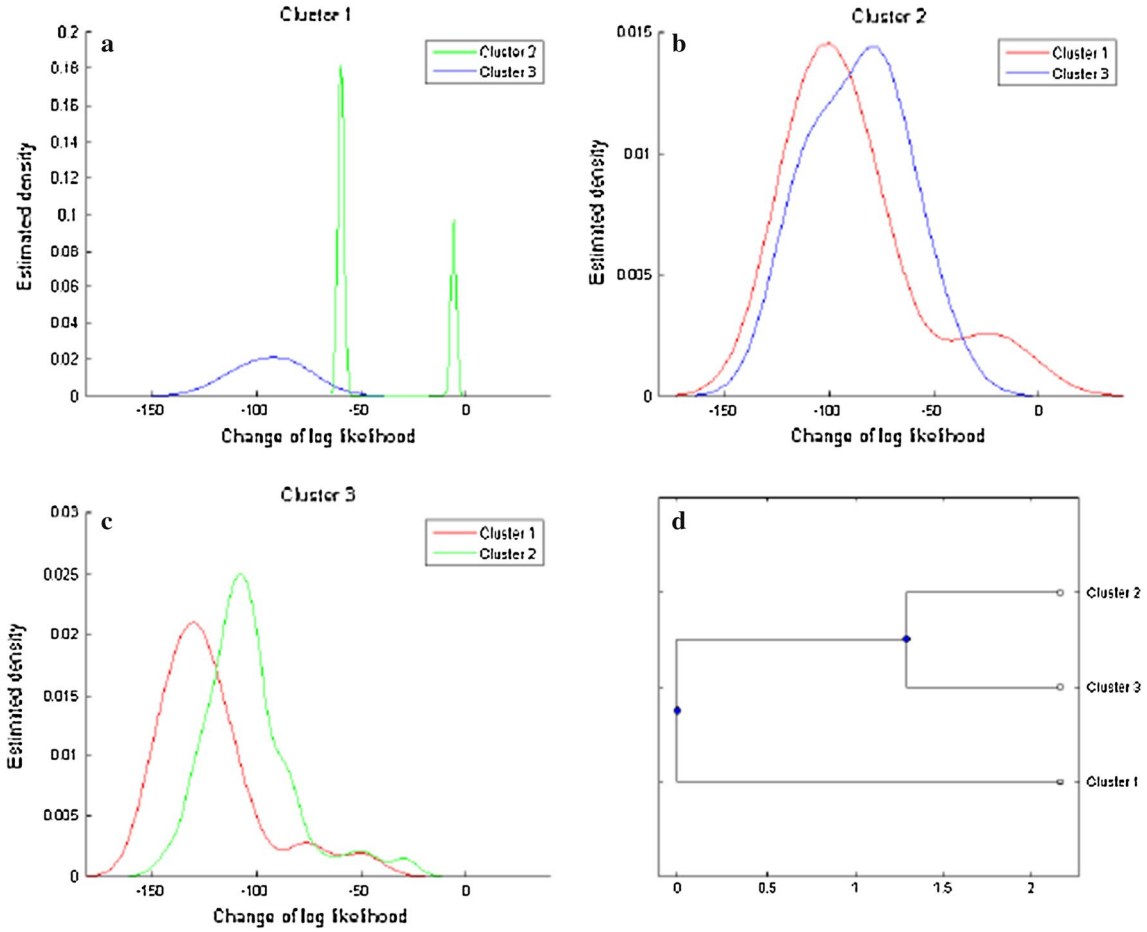

Fig. 4 Population clustering and changes of log likelihoods in BAPS in terms of 'genetic shapes' of the perceived populations (*clusters*). Bimodal curve in case of cluster two (*green*) indicates distinct genetic affinity of its members toward cluster three. Diversity tree showing closer affinity between cluster two (derived determinates) and cluster three (derived indeterminates and natural *B. juncea*) with cluster one (*B. napus, B. rapa* and *B. oleracea*) being distinct (d) (color figure online)

Table 1 Clustering of genotypes based on BAPS

Cluster	Genotype
1	*B. napus*, *B. rapa*, *B. oleracea*
2	*B. carinata*, *B. nigra*, DJ-1D, DJ1-2D, DJ15D, DJ18D, DJ38D, DJ57D, DJ92D, DJ97D, DJ103D, DJ108D, DJ113D, DJ124D, DJ127D
3	AMH, RLC, CBJ, PBR, DJ92, DJ20, DJ49, DJ115, DJ91, DJ68, DJ1-1, DJ98, DJ116, DJ19-1, DJ1-2, DJ103, DJ18, DJ21, DJ38, DJ97, DJ107, DJ73, DJ19-2, DJ2, DJ58, DJ101, DJ79, DJ72, DJ126, DJ55, DJ28, DJ86, DJ78, DJ60, DJ61, DJ65, DJ124, DJ125, DJ110, DJ113, DJ123, DJ99, DJ81, DJ66, DJ127, DJ80

plots. The resynthesized determinate *B. juncea* progenies and B genome carrying species (*B. nigra* and *B. carinata*) formed a compact cluster (Fig. 5) in the same quadrant of the plot. In contrast resynthesized indeterminate accessions were dispersed and present in two quadrants. These were associated with natural accessions of *B. juncea*; *B. rapa, B. oleracea* and *B. napus* were very distinct. Contours of two potential lineages were evident from factorial analysis.

Phenotyping derived *B. juncea* progenies

Forty-two A_5 progenies were evaluated during 2010–2011. The same genotypes were repeated for testing along with 16 new progenies during 2011–2012, taking total A_6 progenies to 58. Alpha-lattice design was used to take advantage of controlling more than one sources of variation. Analysis of variance for different morphological data in A_5 and A_6 progenies revealed significant treatment differences for all the test characteristics (Table 2). Significant blocking effects were observed for days to 50 % flowering (FLW50), 100 % flowering (FLW100), plant height (PH) in 2011–12 and FLW100 during 2010–11. The pooled analysis of variance over 2 years was conducted for a common set of 42 progenies (Table 2). Significant year and year × treatment interactions occurred for all the traits. Replications × year

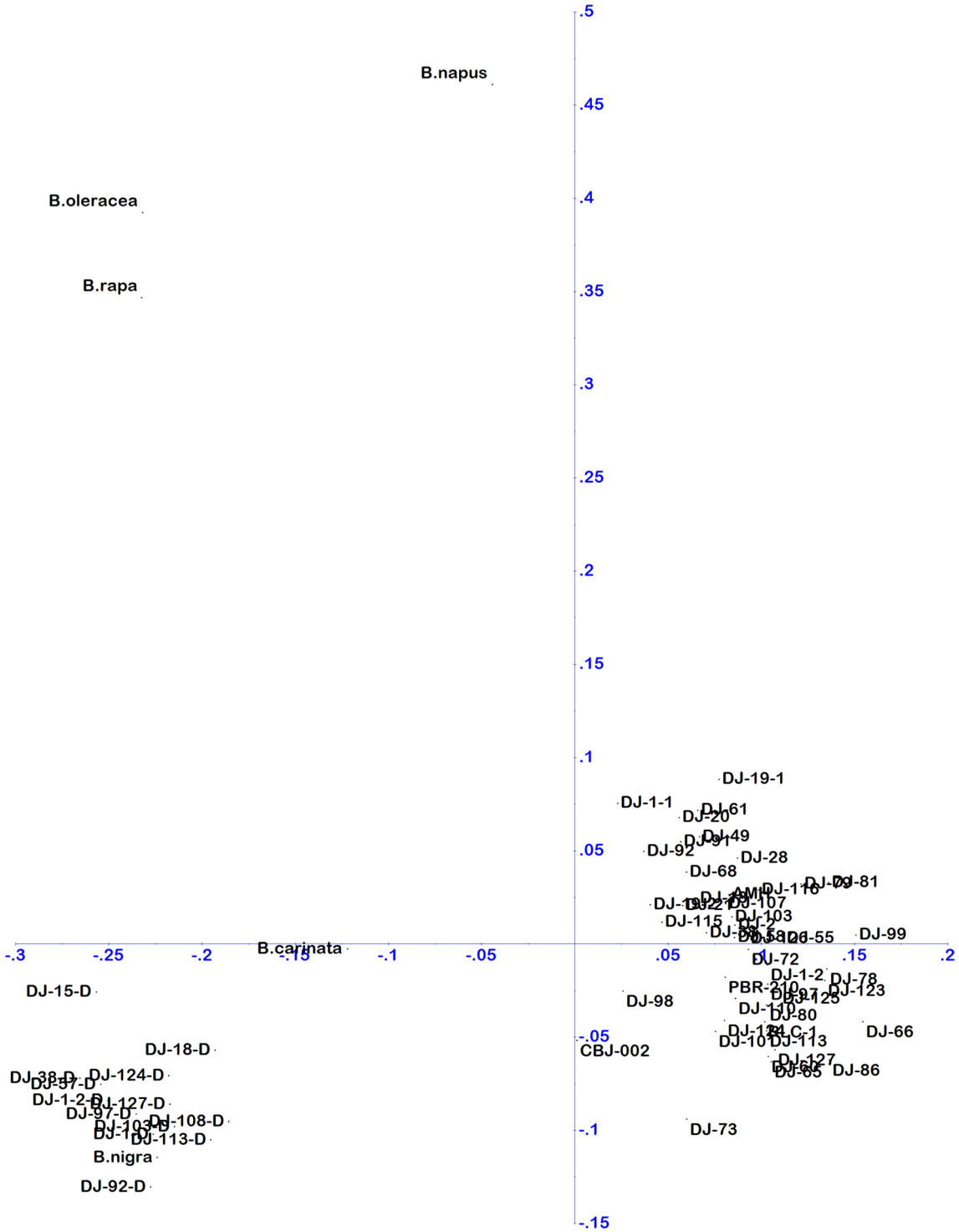

Fig. 5 Principal component analysis of 62 derived and natural *B. juncea* genotypes distinct clustering of determinate derived *B. juncea* (alphabet D after the accession number) along with *B. nigra* as compared to separate clustering of indeterminate derived and natural *B. juncea*

interactions were significant for PH, SB and SY. Distribution of variation is depicted in density plots (Fig. 6). These show wide variations, but the curve appeared to normalize during A_6. There was relative fixation of variation following one generation of selfing. The variation spread seemed nearly bimodal during A_6 for FLW, MSL, PB, SB and SY. Distribution was skewed toward higher trait values for PH and oil content (OIL). Pods on the main shoot

Table 2 Combined analysis of variance for key morphological traits in derived *B. juncea* genotypes

Source	DF	Flw50	Flw100	DM	PH	PB	SB	MSL	PMS	PL	SV	SY	GLS	OIL	PT
Year	1	645.28[b]	1656.82[b]	1718.75[b]	13100.34[b]	347.77[b]	2603.37[b]	135.84[a]	6.67	5.66[b]	1.72	56947.71[b]	5448.36[b]	61.38[b]	51.52[b]
REP(year)	2	0.05	0.02	0.11	214.22[a]	1.04	39.06[b]	44.78	9.34	0.05	2.85	9733.89[b]	14.29	1.17	0.30
TRT	43	195.52[b]	82.77[b]	70.36[b]	366.33[b]	3.02[b]	28.14[b]	158.76[b]	133.42[b]	0.23[b]	7.92[b]	51075.16[b]	184.35[b]	3.72[b]	2.51[b]
Year*Trt	43	126.49[b]	60.63[b]	57.20[b]	298.86[b]	4.68[b]	24.24[b]	149.00[b]	108.02[b]	0.19[b]	9.95[b]	36064.73[b]	159.29[b]	6.40[b]	1.84[b]
Error	74	0.10	0.16	0.16	67.58	0.80	6.44	26.30	14.38	0.05	1.15	1101.94	20.20	0.83	0.82
Total	163														

[a] significance at 5 %, [b] significance at 1 %

(PMS) and SPAD values behaved similarly in both the generations. Coefficient of variation (CV) declined during A_6, except for PL and SPAD values. DJ 2 was the earliest to flower (62 days) among the derived progenies. Flowering started 54 days after sowing in commercial hybrid, Coral, whereas, for DJ 13, it started as late as 92 days. Distribution was skewed toward late flowering. For days to maturity, however, frequency distribution was more aligned with natural mustard cultivars. Derived progenies were taller than commercial mustard. Average main shoot length (MSL) in derived *B. juncea* progenies (67.26 ± 6.93) was higher than the commercial mustard. Derived progenies had more branch number and longer main shoots than natural *B. juncea*. Seed yield per plot ranged from 85.47 to 561.3 g with a mean of 276.46 ± 17.32 in A_5. During A_6, it varied from 195.5 to 766 g, averaging 498.07 ± 15.16. Twelve progenies outperformed commercial standard, PBR210 (598.48 g). Three progenies, DJ114 (766 g), DJ1-2 (705.50 g) and DJ127 (700 g), were superior in yield than even the Hybrid, Coral (699.50 g). Fifteen A_6 generation progenies had higher oil content than PBR210 (40.97), whereas nine genotypes showed oil content higher than hybrid Coral.

Heterosis for yield and component traits

Hybrids were developed by hand crossing 40 derived *B. juncea* genotypes with four diverse natural *B. juncea* testers namely PBR210, RLC1, CBJ002 and AMH2. Analysis of variance for different characteristics, including yield, indicated highly significant variation in the test hybrids (Table 3). During the first year of evaluation, 59 hybrids involving PBR210 as a male parent were evaluated. Frequency histogram showing performance of various hybrids is presented in Fig. 7. The performance over the years was consistent. Heterosis over best commercial check, PBR 210, ranged from −76.5 to 103 % with an average of −8.66 %. The corresponding ranges for RLC1, CBJ002 and AMH2 were −68.4 to 82.7, −11.09 (60 hybrids); −83.6 to 79.6, −3.96 (70 hybrids) and −78.3 to 59.0, −14.53 (64 hybrids). During 2011–12, the ranges and means for hybrids were PBR210 (−42.8 to 102.7; 16.87 for 35 hybrids); RLC1 (−52.7 to 104.8; 14.97 for 34 hybrids); CBJ002 (−55.9 to 117; 23.73 for 35 hybrids) and AMH2 (−51.8 to 69.6; 8.23 for 34 hybrids). Productive hybrids in general were tall, had more branches and took longer to mature. Pods on the main shoot were major and consistent contributors to yield heterosis in hybrids. There was no heterosis for early maturity or dwarf stature. With some exceptions, heterosis was low for pod length and oil content. Some hybrids such as DJ125 × PBR210 and DJ15 × PBR210 expressed heterosis for oil and protein content, respectively. Derived *B. juncea* progenies namely

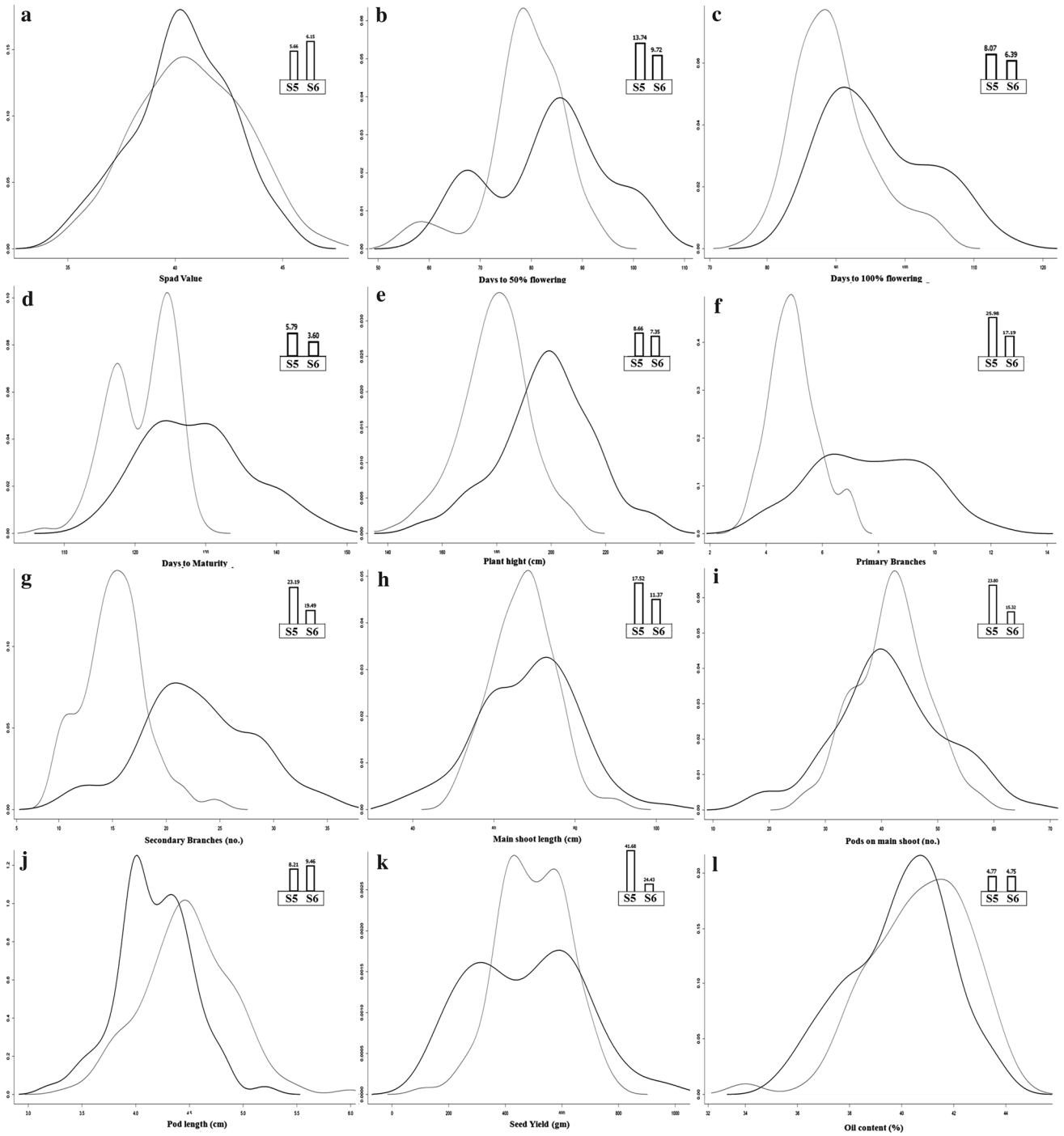

Fig. 6 Density plots showing variation among the genotypes for A5 and A6 generations for various traits. **a** SPAD value, **b** 50 % flowering, **c** 100 % flowering, **d** days to maturity, **e** plant height, **f** number of primary branches, **g** number of secondary branches, **h** main shoot length, **i** pods on the main shoot, **j** pod length, **k** seed yield, **l** oil content

DJ 1, DJ 15, DJ 18, DJ 20 and DJ 21 produced heterotic hybrids with all the four natural *B. juncea* testers during both the years of testing. Although the estimates of heterosis significantly regressed toward the genetic distance of the parents, it was not enough for confident prediction of heterosis through linear regression (Fig. 8). R^2 values were non-significant. The correlation between average genetic diversity and mean heterosis was moderate but significant (Fig. 9). Same was true for proportions of hybrids showing positive heterosis over all the testers.

Discussion

Oilseed brassicas, especially Indian mustard, are highly vulnerable and non-resilient, possibly due to their narrow genetic base. Their resynthesis is considered a long-term option to augment germplasm by mobilizing novel alleles from extant progenitors. Despite lower yields, resynthesized amphiploids formed rich reservoirs of genetic diversity (Seyis et al. 2003; Srivastava et al. 2004). However, resynthesized amphiploids are generally unproductive and show poor breeding value. This is possibly because two (*B. nigra* and *B. oleracea*) of the three monogenomics used for resynthesis of *Brassica* digenomics were not domesticated or selected as an oilseed crop; *B. rapa* being the only diploid with history of domestication and selection as an oilseed. To overcome these lacunae, we resynthesized *B. juncea* through hybridizing elite genotypes from related digenomic species, *B. napus* and *B. carinata*. Total allelic patterns, established in the test progenies, revealed greater within population and relatively lower between population's variations, indicating that family characters were still evolving. Polymorphism in the derived progenies was higher than in the extant diploid donors or natural *B. juncea*. Since all the progenies arose from a common gene pool, this variation may accrue from allopolyploidization or/and non-homologous crossing over during initial development processes. Enhanced genetic variation as a consequence of deletions (Ma and Gustafson 2006), gene conversions (Kovarik et al. 2005), transposon activation (Madlung et al. 2005), chromosomal rearrangements (Lim et al. 2006), intergenomic exchanges (Prakash 1973) or increased meiotic recombinations (Pecinka et al. 2012) and DNA methylation (Lukens et al. 2006) have been reported in nascent polyploids.

Among different groups, determinate progenies had lower variation (34–41 %) than in the indeterminate genotypes. Apparently, determinacy arose only in a small subset of the population. Determinate group as a whole appeared to array closely with B genome carrying species, *B. nigra* and *B. carinata.* A relatively higher divergence between indeterminate and determinate groups was probably due to the divergent trajectory of change in determinate vs. indeterminate group. Sibling line-specific chromosome number variations and rapidly evolving phenotype divergence have been demonstrated in a population from synthesized *Arabidopsis* allohexaploids (Matsushita et al. 2012). Derived *B. juncea* carried cytoplasm from *B. napus*, which is most likely to be distinct from extant A and C genome donor species (Allender and King 2010). While both the genomes in *B. napus* (A and C) share a common lineage, *B. carinata* genomes carry genomes (B and C) from a very diverged lineage (B and C). The role of cytoplasm in

Table 3 Combined analysis of variance for key morphological traits in the hybrids involving derived *B. juncea* genotypes

Source	DF	F50	F100	DM	PH	PB	SB	MSL	PMS	PL	SY	OIL	PT	GLS
Year	1	209.19[b]	162.31[b]	2814.56[b]	30525.27[b]	196.77[b]	9477.02[b]	1922.35[b]	1439.72[b]	0.01	187237.13[b]	36.77[b]	251.96[b]	20461.94[b]
REP (year)	2	0.02	0.66	0.19	1042.59[b]	0.62	5.52	57.73	45.3	0.05	4158.28[b]	0.02	0.4	0.62
TRT	65	125.02[b]	72.46[b]	69.13[b]	618.79[b]	3.53[b]	80.65[b]	208.57[b]	145.97[b]	0.52[a]	39973.03[b]	10.21[b]	3.97[b]	210.76[b]
Year*TRT	65	113.99[b]	57.44[b]	57.04[b]	590.1[b]	4.84[b]	97.68[b]	169.37[b]	95.64[b]	0.46[a]	55230.96[b]	11.67[b]	4.76[b]	307.02[b]
BLK	5	0.14	0.49	0.36	128.23	1.04	5.11	11.22	16.33	0.02	374.12	0.44	0.21	23.3
Error	125	0.33	0.26	0.33	57.53	0.62	6.9	26.52	18.34	0.04	848.18	0.5	0.64	20.2
Total	263													

[a] significance at 5 %, [b] significance at 1 %

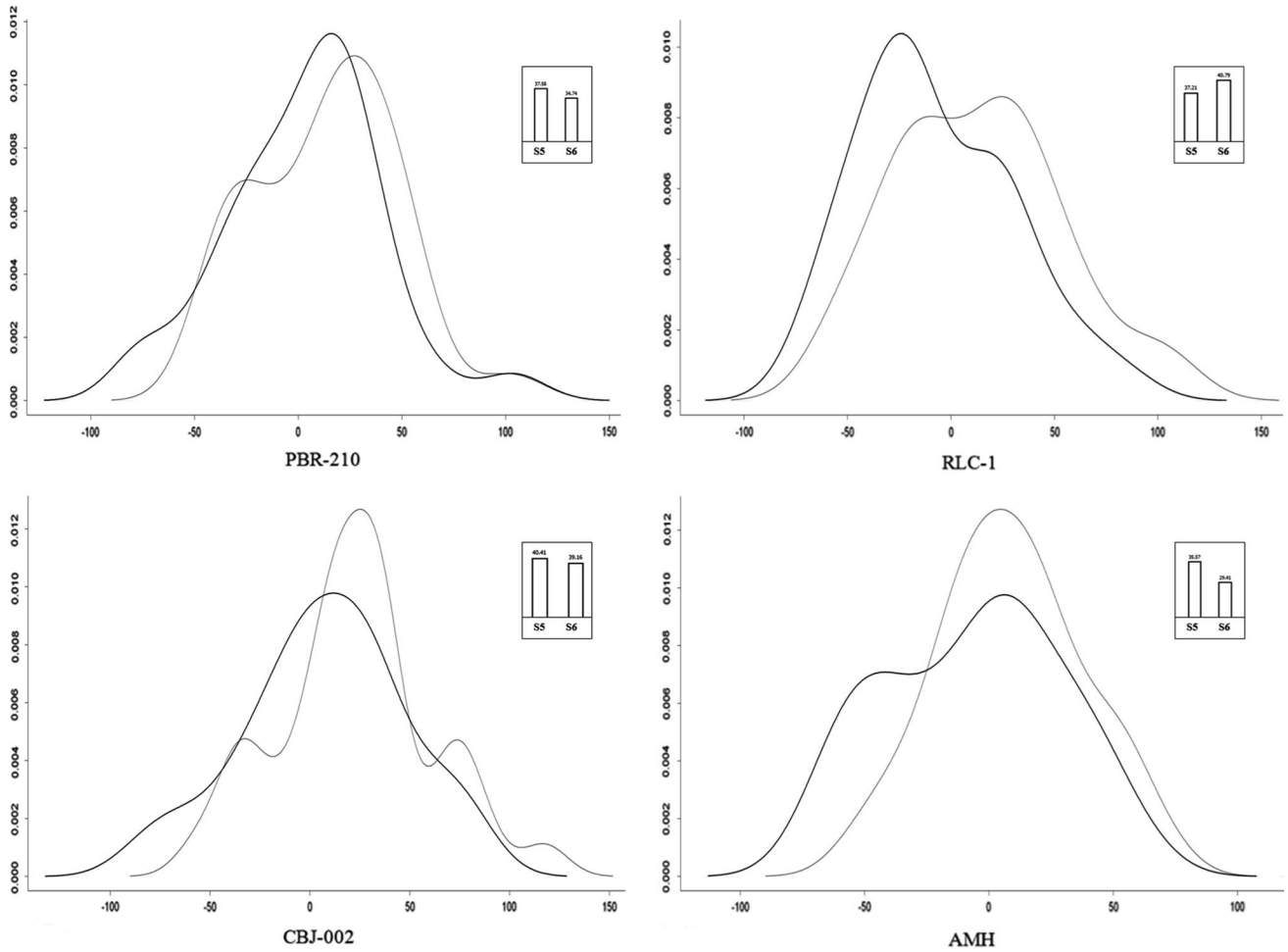

Fig. 7 Frequency histogram showing patterns of heterosis over commercial genotype PBR210

molding responses of nuclear genomes is well known (Cui et al. 2012).

Derived *B. juncea* progenies revealed very high phenotypic variation. Bimodal distribution of variation for many traits was indicative of increasing phenotypical divergence. As postulated, variation in derived genotypes was closer to the agronomic requirements of natural *B. juncea*. The derived *B. juncea* genotypes showed late onset but rapid completion of the flowering. In many species, immediate effect of allotetraploidization on phenotypes has been demonstrated in both natural and resynthesized allopolyploids, arguably due to increased heterozygosity obtained through intergenomic crossing over. These included *B. carinata* (Prakash et al. 1984), *B. juncea* (Olsson 1960a, b; Prakash et al. 1984; Bansal et al. 2009, 2012) and *B. napus* (Olsson 1960a, b; Seyis et al. 2003). However, resynthesized genotypes almost always showed poor breeding value. In comparison, as many as 12 progenies of derived *B. juncea* outperformed standard check and three were higher yielding than even a hybrid, Coral. As pods on the main shoot are

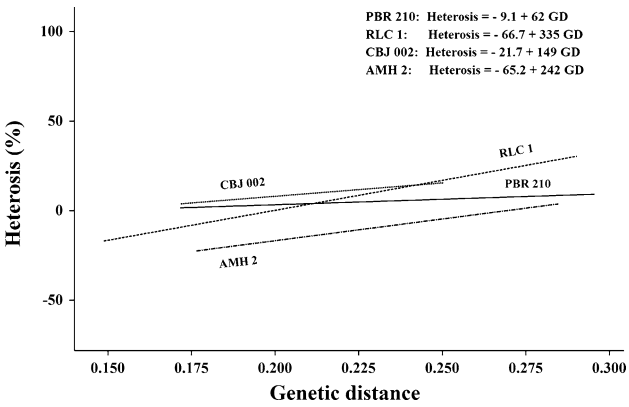

Fig. 8 Correlation between genetic distance and heterosis in cross-combinations involving natural and derived *B. juncea* genotypes

a major contributor to seed yield, genotypes having more than 50 pods on the main shoot may be of interest to commercial mustard breeders. Oil content in natural *B. juncea*

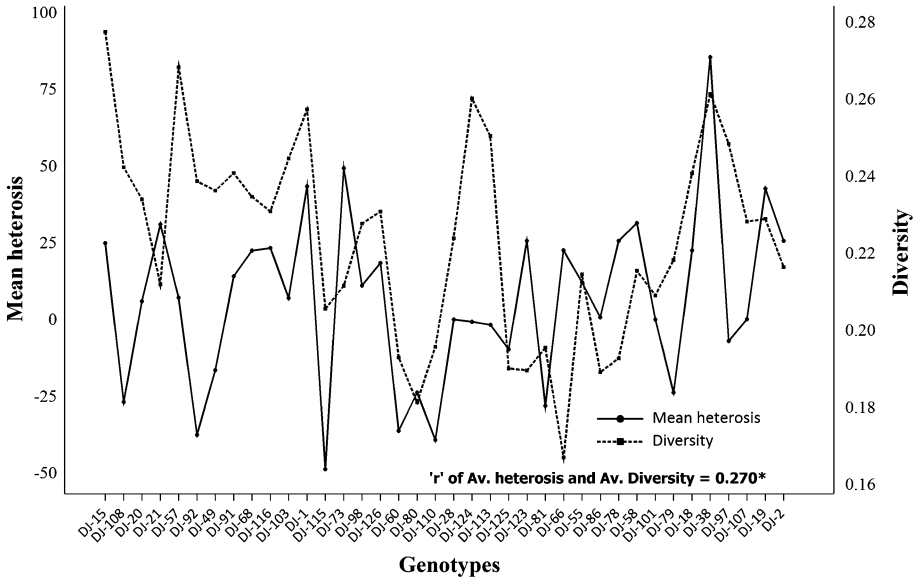

Fig. 9 Association between diversity (av. genetic distance) and mean heterosis in cross-combinations involving natural and derived *B. juncea* genotypes

hovers around 39–40 %. Fifteen derived progenies had oil content higher than observed for check cultivar, PBR 210 (40.97 %), maximum being 43.67 %. *B. napus* cv. Surpass, the A genome donor to derived *B. juncea* also possessed high seed oil content. This agronomic valuation supports our argument that derived amphiploidy allows a seamless flow of breeding gains across *Brassica* digenomics.

A high level of heterosis is considered critical for the commercial success of hybrids. Current mustard hybrids generally show low heterosis. We used four natural *B. juncea* genotypes as testers to study heterosis-inducing potential of derived progenies. For the tester PBR210, average heterosis over all test hybrids was 3.4 and 13.7 % during two test years, respectively. Sixty percent of test hybrids showed positive heterosis during 2010–2011. The corresponding figure for 2011–2012 was 70 %. Almost similar trend was indicated for RLC1, but values were lower than those obtained for PBR210. Among Chinese *B. juncea* genotypes, CBJ002 was the best tester. Almost 60 % and a high of 80 % hybrids involving CBJ002 showed positive heterosis during 2010–2011 and 2011–2012, respectively. Derived *B. juncea* genotypes namely, DJ 1, DJ 15, DJ 18, DJ 20 and DJ 21, produced heterotic hybrids with all the four tester parents during both the years. It may be useful to investigate them in a large number of combinations to confirm their heterosis-inducing potential. There may be some heterosis-inducing genetic factors, as well. Productive hybrids in general were tall, had more branches and took longer to mature. Positive heterosis for dwarf stature or early flowering/maturity, pod length and oil content was rare. Number of pods on the main shoot was a consistent component of heterosis. Superior agronomic potential of derived progenies and higher heterosis in crosses between derived and natural *B. juncea* genotypes confirmed our hypothesis that breeding value of used germplasm is as important as genetic diversity. Same was true for proportions of hybrids showing positive heterosis over all the testers. Although the genetic distance was linearly related to yield heterosis in the present studies, the relationship was not strong enough to optimally predict the heterosis solely on the basis of genetic distance between the parents.

Studies conducted in different crops have revealed that interspecific hybrids are, usually, more heterotic than intraspecific hybrids if reproductive barriers are not a limitation to forming compatible crosses between the species or genera. Notable example is the success of *Gossypium hirsutum* × *G.barbadense* hybrids in cotton (Basbag and Gencer 2007). The means by which intersubspecific and wide-hybridization heterosis can be exploited has been a key area of interest for enhancing heterosis. The concept of sub-genome aimed at introgressing variation from related species in resynthesized *B. napus* constitutes only a partial exploitation of interspecific heterosis (Qian et al. 2005). In contrast, derived amphiploidy overcomes the problem of hybrid sterility associated with interspecific heterozygosity (Moehring 2011) while maintaining the genetic distance between genomes. In essence, the hybrids between derived *B. juncea* and natural *B. juncea* were interspecific hybrids between *B. juncea* and *B. napus* for A genome and between *B. juncea* and *B. carinata* for B genome. Derived amphiploidy envisages the creation of seamless diversity conduits for the transfer of breeding gains across all *Brassica* amphiploids. Moreover, 'synthesis' of new polyploid plants by derived amphiploid route led to immediate and striking reactions in the form of new variation of high breeding value.

Author contributions SSB developed basic genetic resources and supervised the research. SSB and MG

designed the experiments. MG, SG and HK performed the experiments. NK and MG analyzed the data. SSB and MG wrote the manuscript.

Acknowledgments Thanks are due to Dr. Isobel Parkin, Plant Gene Resources of Canada, Agriculture and Agri-Food Canada for providing access to chromosome-specific SSR's for *Brassica* B genome. Helpful suggestions from anonymous reviewers are also gratefully acknowledged. This work was supported by Indian Council of Agricultural Research under ICAR National Professor Project "Broadening the genetic base of Indian mustard (*Brassica juncea*) through alien introgressions and germplasm enhancement" awarded to S. S. B.

Conflict of interest The authors declare that they have no conflict of interest.

References

Allender CJ, King GJ (2010) Origins of the amphiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers. BMC Plant Biol 10:54

Banga SS, Kaur N (2009) An alternate procedure for resynthesis of Brassica juncea. Proceedings of 16th Australian research assembly on Brassicas. Ballarat Vic 106:1–4

Banga SS, Labana KS (1984) Heterosis in Indian mustard (*Brassica juncea* (L.) Coss.). Z Pflanzenzuecht 92:61–70

Bansal P, Kaur P, Banga SK, Banga SS (2009) Augmenting genetic diversity in *Brassica juncea* through its resynthesis using purposely selected diploid progenitors. Int J Plant Breed 3:41–45

Bansal P, Banga S, Banga SS (2012) Heterosis as investigated in terms of polyploidy and genetic diversity using designed *Brassica juncea* amphiploid and its progenitor diploid species. PLoS One 7:e29607

Banuelos GS, Dhillon K, Banga S (2013) Biofuel crops: production, physiology and genetics. Oilseed Brassicas. In: BP Singh (ed). Wallinford: CABI International, pp 339–368

Basbag S, Gencer O (2007) Investigation of some yield and fibre quality characteristics of interspecific hybrid (*Gossypium hirsutum* L. × *G. barbadense* L.) cotton varieties. Hereditas 144:33–42

Chauhan JS, Singh KH, Singh VV, Kumar S (2011) Hundred years of rapeseed–mustard breeding in India: accomplishments and future strategies. Indian J Agric Sci 81:1093–1109

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinform 9:539

Cui C, Ge XH, Gautam M, Kang LZ, Li Y (2012) Cytoplasmic and genomic effects on meiotic pairing in *Brassica* hybrids and allotetraploids from pair crosses of three cultivated diploids. Genetics 191:725–738

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

Falconer DS, MacKay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman, London

Fu D, Qian W, Zou J, Meng J (2012) Genetic dissection of intersub genomic heterosis in *Brassica napus* carrying genomic components of *B. rapa*. Euphytica 184:151–164

Gupta M, Atri C, Banga SS (2014) Cytogenetic stability and genome size variations in newly developed derived *Brassica juncea* allopolyploid lines. J Oilseed Brassica 5:118–127

Jain A, Bhatia S, Banga SS, Prakash S, Lakshmikumaran M (1994) Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. Theor Appl Genet 88:116–122

Kaur P, Banga S,Kumar N,Gupta S,Akhtar J,Banga S (2014) Polyphyletic origin of *Brassica juncea* with *B. rapa* and *B. nigra* (Brassicaceae) participating as cytoplasm donor parents in independent hybridization events. Am J Bot 2014 0: ajb.1400232v1-0

Kim H, Choi SR, Bae J, Hong CP, Lee SY et al (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. BMC Genomics 10:432

Kovarik A, Pires JC, Leitch AR, Lim KY, Sherwood AM et al (2005) Rapid concerted evolution of nuclear ribosomal DNA in two Tragopogon allopolyploids of recent and recurrent origin. Genetics 169:931–944

Lim YP, Plaha P, Choi SR, Uhm T, Hong CP, Bang JW, Hur YK (2006) Toward unraveling the structure of *Brassica rapa* genome. Physiol Plant 126:585–591

Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L, Osborn T (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. Plant Physiol 140:336–348

Ma XF, Gustafson J (2006) Timing and rate of genome variation in Triticale following allopolyploidization. Genome 49:950–958

Madlung A, Tyagi AP, Watson B, Jiang HM, Kagochi T, Doerge RW, Martienssen R, Comai L (2005) Genomic changes in synthetic *Arabidopsis* polyploids. Plant J 41:221–230

Matsushita SC, Tyagi AP, Pires JC, Thornton GM, Madlung A (2012) Allopolyploidization lays the foundation for evolution of distinct populations: evidence from the analysis of synthetic Arabidopsis allohexaploids. Genetics 191:535–547

Moehring AJ (2011) Heterozygosity and its unexpected correlations with hybrid sterility. Evolution (N Y) 65:2621–2630

Olsson G (1960a) Species crosses within the genus *Brassica*. I. Artificial *Brassica juncea* Coss. Hereditas 4:171–222

Olsson G (1960b) Species crosses within the genus *Brassica*. II. Artificial *Brassica napus* L. Hereditas 46:351–386

Pecinka A, Fang W, Rehmsmeier M, Levy AA, Mittelsten Scheid O (2012) Polyploidization increases meiotic recombination frequency in Arabidopsis. BMC Biol 10:33

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. http://darwin.cirad.fr/

Prakash S (1973) Non-homologous meiotic pairing in the A and B genomes of Brassica: its breeding significance in the production of variable amphidiploids. Gene Res Camb 2:133–137

Prakash S, Tsunoda S (1983) Cytogenetics of *Brassica*. In: Swaminathan MS, Gupta PK, Sinha U (eds) Cytogenetics of crop plants. Macmillan India, New Delhi, pp 481–513

Prakash S, Gupta S, Raut RN, Kalra A (1984) Synthetic *Brassica carinata*. Crucif Newslett 9:36

Prakash S, Bhat SR, Quiros CF, Kirti PB, Chopra VL (2009) *Brassica* and its close allies: cytogenetics and evolution. Plant Breed Rev 31:21–187

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Qian W, Chen X, Fu D, Zou J, Meng J (2005) Intersubgenomic heterosis in seed yield potential observed in a new type of *Brassica napus* introgressed with partial *Brassica rapa* genome. Theor Appl Genet 110:1187–1194

Seyis F, Snowdon RJ, Luhs W, Friedt W (2003) Molecular characterization of novel resynthesized rapeseed (*Brassica napus*) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars. Plant Breed 12:473–478

Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc Natl Acad Sci USA 92:7719–7723

Srivastava A, Mukhopadhyay A, Arumugam M, Gupta V, Verma JK, Pental D, Pradhan AK (2004) Resynthesis of *Brassica juncea* through interspecific crosses between *B. rapa* and *B. nigra*. Plant Breed 123:204–206

Wijesundera C, Ceccato C, Fagan P, Shen Z, Burton W, Salisbury P (2008) Canola quality Indian Mustard Oil (*Brassica juncea*) is more stable to oxidation than conventional canola oil (*Brassica napus*). J Am Oil Chem Soc 85:693–699

Zou J, Zhu J, Huang S, Tian E, Xiao Y, Fu D, Tu J, Fu T, Meng J (2010) Broadening the avenue of intersubgenomic heterosis in oilseed *Brassica*. Theor Appl Genet 120:283–290