CrossMark

ORIGINAL ARTICLE

Genetic dissection of seed weight by QTL analysis and detection of allelic variation in Indian and east European gene pool lines of *Brassica juncea*

Namrata Dhaka[1] · Kadambini Rout[2] · Satish K. Yadava[2] · Yaspal Singh Sodhi[2] ·
Vibha Gupta[2] · Deepak Pental[2] · Akshay K. Pradhan[1,2]

Abstract

Key message **Seed weight QTL identified in different populations were synthesized into consensus QTL which were shown to harbor candidate genes by *in silico* mapping. Allelic variation inferred would be useful in breeding *B. juncea* lines with high seed weight.**

Abstract Seed weight is an important yield influencing trait in oilseed Brassicas and is a multigenic trait. Among the oilseed Brassicas, *Brassica juncea* harbors the maximum phenotypic variation wherein thousand seed weight varies from around 2.0 g to more than 7.0 g. In this study, we have undertaken quantitative trait locus/quantitative trait loci (QTL) analysis of seed weight in *B. juncea* using four bi-parental doubled-haploid populations. These four populations were derived from six lines (three Indian and three east European lines) with parental phenotypic values for thousand seed weight ranging from 2.0 to 7.6 g in different environments. Multi-environment QTL analysis of the four populations identified a total of 65 QTL ranging from 10 to 25 in each population. Meta-analysis of these component QTL of the four populations identified six 'consensus'

QTL (*C-QTL*) in A3, A7, A10 and B3 by merging 33 of the 65 component *Tsw* QTL from different bi-parental populations. Allelic diversity analysis of these six *C-QTL* showed that Indian lines, Pusajaikisan and Varuna, hold the most positive allele in all the six *C-QTL*. *In silico* mapping of candidate genes with the consensus QTL localized 11 genes known to influence seed weight in *Arabidopsis thaliana* and also showed conserved crucifer blocks harboring seed weight QTL between the A subgenomes of *B. juncea* and *B. rapa*. These findings pave the way for a better understanding of the genetics of seed weight in the oilseed crop *B. juncea* and reveal the scope available for improvement of seed weight through marker-assisted breeding.

Introduction

Seed weight is one of the most important traits in crops in which seeds are the direct product to be used, like cereals or oilseeds. An understanding of genetics of seed weight is of particular importance regarding to not only the economic and nutritional aspects, but also physiological, biochemical and molecular aspects. Since the beginning of plant domestication, plant breeders have been selecting varieties with bold seeds and increased seed weight. Genetic improvement of Indian mustard (*Brassica juncea*) for high seed weight has led to the development of bold-seeded varieties using the natural variation in the crop. It has generated a preference among the Indian farmers for the bold-seeded varieties.

Brassica juncea, an allotetraploid (AABB) with diploid genomes of *B. rapa* (AA) and *B. nigra* (BB), is an important oilseed crop of the Indian subcontinent, having two distinct gene pools—Indian and east European (Pradhan et al. 1993; Srivastava et al. 2001). The major breeding objectives in the *B. juncea* target yield enhancement,

✉ Akshay K. Pradhan
pradhancgmcp@gmail.com

1 Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India

2 Centre for Genetic Manipulation of Crop Plants, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India

oil quality and resistance against pathogens and tolerance for abiotic stresses (Pradhan and Pental 2011). It has been shown that F_1 hybrids between the lines of these two gene pools are more productive than those involving genotypes within a gene pool (Pradhan et al. 1993). In *B. juncea*, the first commercial F_1 hybrid DMH-1 released in India has been developed from a cross between an Indian gene pool line Pusa bold and an east European gene pool line EH-2 that shows about 30 % higher yield than the best check varieties (Sodhi et al. 2006). However, the farmers' preference for the hybrid DMH-1 is relatively low because the seed weight of the hybrid is less than 4.0 g per thousand seeds due to the smaller seed size of the pollinator parent, EH-2 (thousand seed weight ~2.5 g).

The current challenge in the improvement of seed size, however, lies in the identification of genomic regions quantitative trait locus/quantitative trait loci (QTL) directly influencing seed size and yields. QTL mapping studies identifying seed weight QTL have been carried out in various plant species like *Arabidopsis* (Alonso-Blanco et al. 1999; Herridge et al. 2011), tomato (Doganlar et al. 2000), chickpea (Hossain et al. 2010), soybean (Sun et al. 2012) and various *Brassica* species (Ramchiary et al. 2007; Fan et al. 2010; Yang et al. 2012; Yadava et al. 2012; Li et al. 2014; Qi et al. 2014). Seed weight, a highly heritable trait, has been reported to be governed by a relatively large number of QTL with small effects and a few QTL with large effects controlling the trait.

Seed weight in dicots is determined by regulating the growth of the embryo, endosperm, integuments and cell number/cell size or both in the female gametophyte (Sundaresan 2005; Alonso-Blanco et al. 1999). Different studies on factors influencing seed size in *Arabidopsis* have implicated several important genes. *HAIKU (IKU)*, a transcription factor, was the first to be shown to affect seed size by influencing endosperm development (Garcia et al. 2005). *IKU1*, *IKU2* and *MINI3* and *SHB1* act in the same developmental pathway affecting endosperm size (Luo et al. 2005; Zhou et al. 2009). Genes involved in hormone signaling such as the Brassinosteroid hormone play a central role by regulating endosperm and embryo development (acting upstream to *SHB1*, *IKU1, MINI3* and *IKU2*) and growth of integuments [acting upstream of *AP2, MNT/ARF2* and cytochrome P450 *KLUH (KLU)/CYP78A5*] (Jiang et al. 2013; Jiang and Lin 2013). Cytokinins and ethylene-mediated signaling also play a major role in seed size regulation (Li et al. 2013; Bennett et al. 2014). Flavonoids communicate signals from seed coat to endosperm, thus affecting its proliferation (Doughty et al. 2014). The ubiquitin pathway controls the seed size by targeting proteins for degradation (Li and Li 2015). A polycomb (PcG) group of genes and other epigenetic regulators also affect seed size (Li and Li 2015).

In this article, we have made concerted attempts on genetic analysis of seed weight in the oilseed crop *B. juncea*. By encapsulating maximum genetic variability from four different bi-parental mapping populations and QTL analyses, we have defined major QTL for seed weight. Using meta-analysis, consensus QTL for seed weight have been identified highlighting the underlying allelic variations. The candidate genes for seed weight from various *Brassica* species and the model crucifer *Arabidopsis* overlapping the consensus QTL have also been identified and their relevance in breeding programs for increasing seed weight in *B. juncea* is described.

Materials and methods

Plant materials

Four bi-parental microspore-derived doubled-haploid (DH) mapping populations covering maximum variability for seed weight in *B. juncea* were used in the present study. These four populations have been developed from six parents including three Indian (Pusajaikisan, Varuna and TM-4) and three east European lines (Heera, Donskaja-IV and EH-2). Two of these populations, namely, VH (derived from F_1 between Varuna and Heera) and TD (derived from F_1 between TM-4 and Donskaja-IV) used in the study have been earlier reported by Ramchiary et al. (2007) and Yadava et al. (2012), respectively. Two new DH populations, DE (derived from F_1 of Donskaja-IV and EH-2) and EPJ (derived from F_1 of EH-2 and Pusajaikisan), were developed in the present study. Detailed phenotypic contrasts for seed weight among the parental lines of these four populations are shown in Table 1.

The EPJ and DE map used for the QTL analysis of seed weight consisted of 182 and 184 DH lines, respectively. The EPJ map was constructed with a total of 860 markers consisting of 241 intron polymorphisms (IP), 462 genic SSRs and 157 BAC-derived SSRs with a total genetic length of 2073.6 cM (Dhaka 2015). The DE linkage map with 653 markers has 344 IPs, 278 SSRs (both genic and BAC-derived), 26 SNPs and 5 Indels that cover a total genetic length of 1598.5 cM (Rout 2014). Details of these two linkage maps will be reported elsewhere.

Field experimental design and trait measurements

The EPJ population along with the parents, EH-2 and Pusajaikisan, was planted in three environments of northwest India in Delhi, Bharatpur and Alwar during the winter growing season. The field trials of the DE population along with the parents Donskaja-IV and EH-2 was conducted at

Table 1 Mean thousand seed weight (*Tsw*) of the parental and doubled-haploid lines of the four *B. juncea* mapping populations

Mapping population	Number of lines	Environment	*Tsw* value of parents (g)				Mean *Tsw* of mapping population (g)	Range of *Tsw* in mapping population (g)	References
			Female parent (gene pool[a])		Male parent (gene pool)				
DE	184	Delhi year 1	Donskaja-IV (E)	2.8	EH-2 (E)	2.4	2.4	1.7–3.3	Present study
		Delhi year 2		2.8		2.5	2.8	1.6–3.9	
		Delhi year 3		2.6		2.3	2.6	1.6–3.7	
EPJ	182	Delhi	EH-2 (E)	2.8	Pusajaikisan (I)	7.6	4.3	2.5–8.6	Present study
		Alwar		2.4		5.9	3.6	2.2–6.2	
		Bharatpur		2.4		6.2	3.5	1.9–7.2	
TD	100	Delhi	TM-4 (I)	2.7	Donskaja-IV (E)	2.3	2.3	1.3–3.2	Yadava et al. (2012)
		Bharatpur		2.0		2.2	1.7	0.7–2.5	
		Leh		3.1		3	3.0	2.0–4.1	
VH	123	Delhi	Varuna (I)	6.3	Heera (E)	2.3	3.3	1.9–6.4	Ramchiary et al. (2007)
		Gwalior		5.3		2.3	3.1	2.2–5.6	
		Leh		4.5		2.5	3.2	1.9–4.6	

[a] *Brassica juncea* gene pool: Indian (I)/east European (E)

Delhi for 3 years. The experiments were conducted in a randomized block design with three replications. Each line was planted in a single row with a row-to-row distance of 40 cm and a plant-to-plant distance of 10 cm. Five representative plants from the middle of the row were used for phenotyping. Seed weight trait was phenotyped as thousand seed weight (*Tsw*) measured in grams. The mean weight of thousand seeds from the mixed seeds of 15 plants, with 5 plants from each of the three replications was used as the trait value for statistical and QTL analyses.

QTL analysis for seed weight

Marker genotyping data of the four DH populations were scored as 'a' (female parent allele) and 'b' (male parent allele) with an exception to EPJ population where male parent allele (Pusajaikisan) was scored as 'a' and female parent allele as 'b' (EH-2). All the statistical analyses were performed using SPAR 2.0 (2005). Composite interval mapping (CIM) using Windows QTL Cartographer 2.5 (Wang et al. 2006) was used for analyzing QTL for seed weight (*Tsw*). A window size of 10 cM and a walk speed of 1 cM with five control markers were used. For declaring the presence of a QTL, genome-wide threshold values ($P = 0.05$) were estimated from 1000 permutations of trait data across all genetic intervals (Churchill and Doerge 1994; Doerge and Churchill 1996). QTL detected with more than 10 % of the phenotypic variance (R^2) were considered as major QTL.

For identification of consensus QTL by meta-analysis, QTL maps of the four bi-parental populations were used. In addition to QTL mapping of seed weight trait of EPJ and DE populations in the present study, a QTL re-analysis of

the trait was also undertaken for VH and TD populations, using the multi-site trait data of VH (Ramchiary et al. 2007) and TD (Yadava et al. 2012) (Table 1). The VH and the TD genetic maps used for the purpose consisted of 2205 and 970 markers, respectively. Prior to meta-analysis, a consensus map was developed using the genetic maps of TD, EPJ and DE projected on to the VH map using the 'Map projection' command of BioMercator 2.1 (Arcade et al. 2004; Goffinet and Gerber 2000). *Arabidopsis* genomic blocks (Schranz et al. 2006) were assigned to the consensus map based on the corresponding IDs of the intron polymorphisms and the genic-SSR markers based on synteny with *B. rapa* available in the Brassica database.

The *Tsw* QTL identified were then integrated into consensus QTL using the meta-analysis method followed by Yadava et al. (2012). QTL obtained from different populations which merged into a single QTL upon meta-analysis were regarded as 'consensus' QTL (*C-QTL*), while those which did not merge by meta-analysis were designated as 'population-specific' QTL. The population-specific QTL were further classified into either 'stable' QTL (*S-QTL*) when at least two QTL from different environments merged into a single QTL or 'environment' QTL (*E-QTL*) which did not merge. The allelic effects at these QTL merging into a consensus QTL were estimated in each of the component bi-parental populations by calculating the phenotypic difference between the two parental genotypes at the marker closest to the consensus QTL following the procedure of Chardon et al. (2014). Similar exercise of allelic effects estimation was also undertaken for 'population-specific' QTL using the respective bi-parental QTL maps.

The epistasis and QTL × environment interactions influencing seed weight (*Tsw*) were identified by QTL Network

2.0 (Yang et al. 2007). It simultaneously investigates the additive QTL, QTL × environment, QTL × QTL (epistasis) as well as epistasis × environment interactions. The analyses were based on a mixed linear model approach with 1 cM walk speed and a testing window of 10 cM. To detect epistatic effects among loci with or without individual additive effects, 2D genome scans were carried out. 1000 permutations were performed for generating thresholds for the presence of QTL.

Results

Phenotypic variation in seed weight

The phenotypic variation for the trait among the six parents of the four bi-parental populations ranged from 2.0 to 7.6 g in different environments (Table 1). The maximum parental contrast was recorded for EPJ population ranging from 2.4 g in EH-2 to 7.6 g in Pusajaikisan. The parents of the VH population also showed high contrast in seed weight ranging from 2.3 g in Heera to 6.3 g in Varuna. The parents of the other two bi-parental populations (DE and TD) were not contrasting for seed weight trait. The frequency distributions of the trait in all the four populations over different environments showed continuous variation indicating the quantitative mode of inheritance for *Tsw* (Fig. 1). Transgressive segregants were observed in all of the populations with *Tsw* beyond the parental values, wherein DE and TD populations showed a higher frequency of transgressive segregants as compared to the EPJ and VH populations (Fig. 1). The trait showed high broad-sense heritability (ranging from 73.0 to 94.0 %) in all the four populations (Online Resource 1).

QTL mapping of seed weight (*Tsw*) in four bi-parental populations

QTL analysis of the trait in the EPJ population showing very high parental contrast for the phenotype identified a total of 25 QTL with LOD scores greater than 2.5 from the three environments (Table 2). These QTL were detected over 12 linkage groups, wherein 12 QTL mapped to A and 13 to the B subgenome. Pusajaikisan contributed the positive alleles for all the 25 QTL. The additive effects of these QTL ranged from 0.1 to 0.4 g and the phenotypic variance explained (R^2) by each QTL ranged from 3.3 to 16.5 %. The analysis identified four major QTL with R^2 greater than 10 % located on LGs A3, B3 and B8.

The DE population derived from the parents Donskaja-IV and EH-2 lacking much variability in the phenotype identified a total of ten QTL from 3 years of trials. The QTL were distributed over six linkage groups, three each in

the A and B subgenomes (Table 2). Unlike the EPJ population, where all the positive alleles were contributed by the high parent (Pusajaikisan), the DE population identified positive alleles from both the parents. While Donskaja-IV contributed positive alleles for seven QTL, EH-2 contributed positive alleles for three QTL. Except two major QTL in LG A3, all others were detected as minor QTL.

The pattern of QTL re-analysis of the TD population developed from TM-4 and Donskaja-IV with no phenotypic contrast for seed weight was also similar to that of the DE population with both the parents contributing the positive alleles. Of the 14 *Tsw* QTL detected in the population, 8 had trait enhancing alleles from TM-4 and 6 had trait enhancing alleles from Donskaja-IV. Of these, five QTL in the LGs A7, A8, A9 and B7 were detected as major QTL with R^2 greater than 10 % (Table 2).

The QTL re-analysis of the VH population identified a total of 16 QTL for seed weight. In all the 16 detected QTL, the parent Varuna with high trait phenotype contributed positive alleles to all the QTL showing a similar genetic architecture as that of the EPJ population. The QTL were distributed over eight LGs with five major QTL in LGs A3, A10 and B3 (Table 2).

The QTL analysis in the four bi-parental populations identified a total of 65 *Tsw* QTL, of which 39 and 26 QTL were detected in the A and B subgenomes, respectively. Of these 65 QTL, 25 % of the QTL (=16) were detected as major QTL explaining more than 10 % of the phenotypic variance.

Identification of consensus *Tsw* QTL

The 65 QTL for seed weight identified in the four mapping populations phenotyped in different environments were projected onto the consensus genetic map of *B. juncea* for the identification of 'consensus' QTL (*C-QTL*). A meta-analysis was performed with the hypothesis that the *Tsw* QTL identified in different environments and located in a given region of a chromosome could be different estimations of the same QTL. Meta-analysis of 65 QTL resulted in six consensus *Tsw* (*Tsw-C*) QTL distributed on LGs A3, A7, A10 and B3 (Fig. 2; Table 3). These six *Tsw-C* QTL were formed by integrating 33 component *Tsw* QTL, while the remaining 32 *Tsw* QTL remained as population-specific QTL. The number of component QTL that merged into each of these six *Tsw-C* QTL ranged from three to nine. Of the six *Tsw-C* QTL, LGs A3 and B3 included two *C-QTL* each and LGs A7 and A10 harbored one *C-QTL* each. One *C-QTL* each in LGs A3 (*Tsw-C-A3-1*) and B3 (*Tsw-C-B3-1*) were formed by merging the maximum number of major component QTL, while the second *C-QTL* in these LGs (*Tsw-C-A3-2* and *Tsw-C-B3-2*) were formed by merging minor component QTL. The remaining two *C-QTL* in
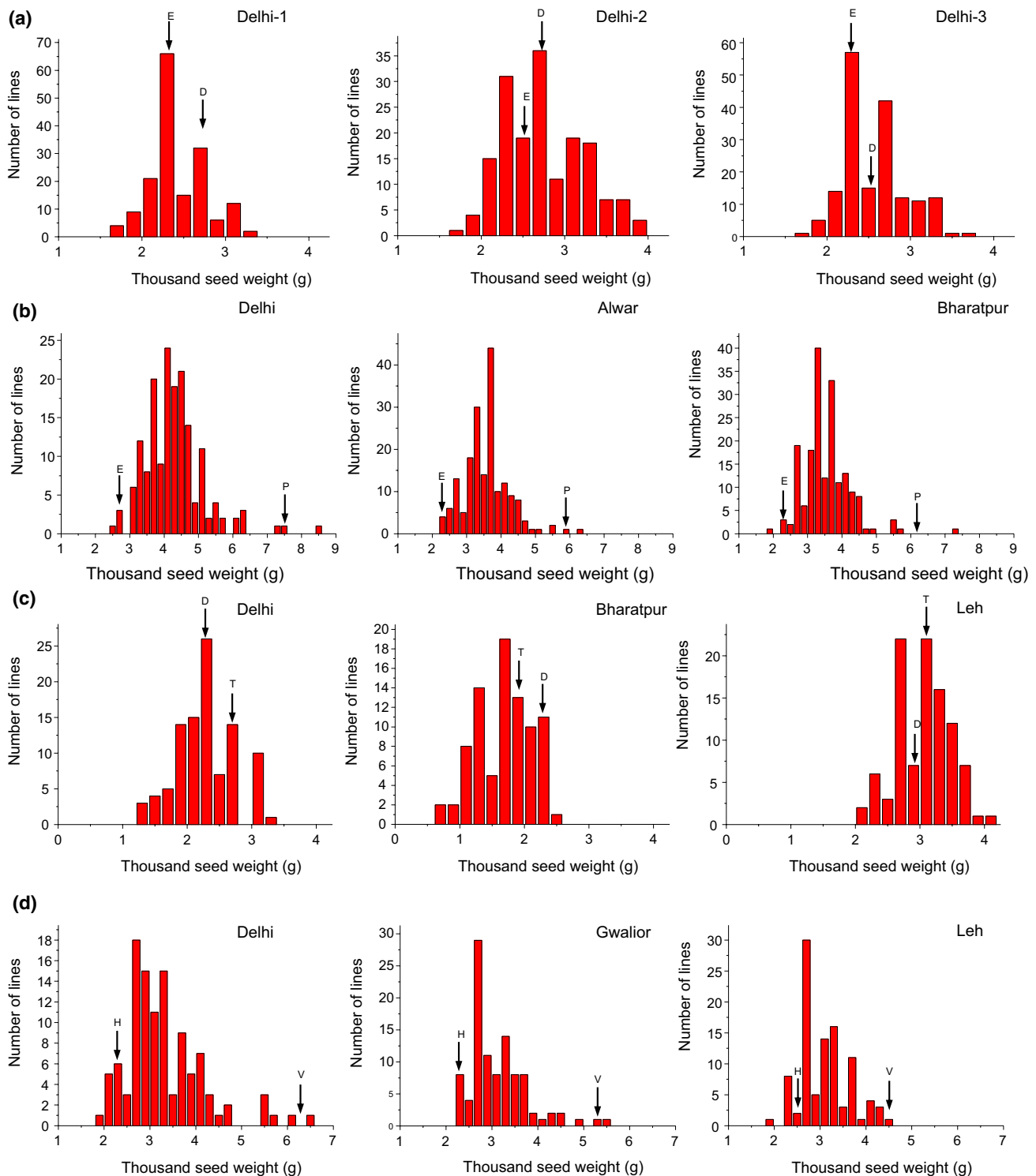
Fig. 1 Frequency distributions of thousand seed weight in four mapping populations of *Brassica juncea*. **a** DE, **b** EPJ, **c** TD and **d** VH in multiple environments

LGs A7 and A10 were formed by merging the maximum number of minor component QTL.

Of the 32 *Tsw* QTL from the four bi-parental populations that did not merge to a consensus QTL, 15 QTL merged into population-specific stable QTL (*S-QTL*) as they were detected in more than one environment. The remaining 17 QTL which did not merge were detected as environment QTL (*E-QTL*) (Online Resource 2). The maximum number

Table 2 *Tsw* QTL data of DE, EPJ, TD and VH populations used for meta-QTL analysis

QTL	LG	Location	LOD score	Additive effect	R^2 (%)	Interval (cM)
EPJ population						
EPJ-Tsw-A1-1	A1	Delhi	6.2	0.3	9.7	62.9–64.6
EPJ-Tsw-A1-2	A1	Bharatpur	2.7	0.1	3.3	67.6–68.6
EPJ-Tsw-A3-1	A3	Delhi	10	0.4	16.5	28.8–41.5
EPJ-Tsw-A3-2	A3	Alwar	8.7	0.2	14.6	46.2–49.2
EPJ-Tsw-A3-3	A3	Alwar	2.8	0.2	3.8	63.4–64.7
EPJ-Tsw-A3-4	A3	Bharatpur	6.7	0.2	9.2	56.0–58.3
EPJ-Tsw-A6-1	A6	Delhi	5.5	0.3	8.5	51.9–67.5
EPJ-Tsw-A7-1	A7	Delhi	3.1	0.2	4.2	31.6–34.4
EPJ-Tsw-A7-2	A7	Delhi	3	0.2	4.2	61.4–70.4
EPJ-Tsw-A7-3	A7	Alwar	2.5	0.1	3.7	51.0–57.4
EPJ-Tsw-A7-4	A7	Bharatpur	5.5	0.2	7.6	65.5–82.5
EPJ-Tsw-A10-1	A10	Delhi	3.5	0.2	5.7	67.7–78.3
EPJ-Tsw-B1-1	B1	Alwar	2.6	0.1	3.6	50.7–52.0
EPJ-Tsw-B2-1	B2	Delhi	2.9	0.2	5.6	66.7–71.3
EPJ-Tsw-B3-1	B3	Delhi	4.3	0.2	6.5	70.4–75.7
EPJ-Tsw-B3-2	B3	Alwar	4.8	0.2	8.4	62.2–77.4
EPJ-Tsw-B3-3	B3	Bharatpur	10.1	0.3	14.5	37.8–39.3
EPJ-Tsw-B4-1	B4	Alwar	2.5	0.2	5.4	57.1–60.5
EPJ-Tsw-B4-2	B4	Alwar	2.8	0.1	4.6	72.1–83.6
EPJ-Tsw-B4-3	B4	Bharatpur	6.7	0.2	9.2	44.2–46.5
EPJ-Tsw-B5-1	B5	Delhi	3.2	0.2	5.6	44.3–61.7
EPJ-Tsw-B7-1	B7	Alwar	5.7	0.2	9.2	81.5–87.7
EPJ-Tsw-B7-2	B7	Bharatpur	3.4	0.1	4.2	101.8–108.3
EPJ-Tsw-B8-1	B8	Alwar	3.8	0.2	6.5	110–113.6
EPJ-Tsw-B8-2	B8	Bharatpur	8.4	0.2	12.6	110–124.3
DE population						
DE-Tsw-A3-1	A3	Delhi3	3.2	0.1	6.4	72.6–74.5
DE-Tsw-A3-2	A3	Delhi1	8.9	0.2	18.6	37.7–50.1
DE-Tsw-A3-3	A3	Delhi2	7.3	0.2	13.1	30.8–42.4
DE-Tsw-A7-1	A7	Delhi1	5.5	0.1	9.5	18.4–21.4
DE-Tsw-A7-2	A7	Delhi2	3.4	0.2	5.7	21.4–23.4
DE-Tsw-A10-1	A10	Delhi3	2.9	−0.1	5.3	31.3–32.7
DE-Tsw-A10-2	A10	Delhi1	4.2	−0.1	9.6	0–17.9
DE-Tsw-B3-1	B3	Delhi2	3.2	−0.1	5.4	15.3–33
DE-Tsw-B4-1	B4	Delhi2	3	0.1	5.1	9.7–19.7
DE-Tsw-B7-1	B7	Delhi3	4	0.1	7.4	85.6–93.3
TD population						
TD-Tsw-A1-1	A1	Delhi	3.8	−0.1	8.6	72.3–75.1
TD-Tsw-A1-2	A1	Leh	2.6	−0.1	6.6	59.4–68.5
TD-Tsw-A2-1	A2	Bharatpur	2.5	−0.1	7.7	0–6
TD-Tsw-A7-1	A7	Leh	4.3	0.1	11.8	39.2–44.9
TD-Tsw-A8-1	A8	Delhi	3.8	0.1	8.3	35.7–61.4
TD-Tsw-A8-2	A8	Leh	5.1	0.2	14.7	61.4–77.4
TD-Tsw-A9-1	A9	Delhi	4.5	−0.2	10.7	46.4–49.6
TD-Tsw-A9-2	A9	Bharatpur	4.1	−0.2	13.1	51.9–53.1
TD-Tsw-A9-3	A9	Leh	2.8	−0.1	7.2	51.7–52.5
TD-Tsw-A10-1	A10	Delhi	3.3	0.1	7.5	9.1–14.7
TD-Tsw-A10-2	A10	Bharatpur	2.9	0.1	8.7	4.9–14.7

Table 2 continued

QTL	LG	Location	LOD score	Additive effect	R^2 (%)	Interval (cM)
TD-Tsw-B1-1	B1	Leh	2.7	0.1	8.4	62.5–87.2
TD-Tsw-B7-1	B7	Delhi	2.8	0.1	7.7	11.3–30.8
TD-Tsw-B7-2	B7	Bharatpur	4.1	0.2	13.3	22.4–41.6
VH population						
VH-Tsw-A3-1	A3	Gwalior	2.8	0.2	6.1	17.6–23.3
VH-Tsw-A3-2	A3	Delhi	6.8	0.3	12.6	22.2–38.9
VH-Tsw-A4-1	A4	Gwalior	3.1	0.2	7	44.2–48.3
VH-Tsw-A4-2	A4	Leh	3.1	0.2	8.1	42.0–48.5
VH-Tsw-A7-1	A7	Delhi	2.5	0.2	3.8	31.5–46.5
VH-Tsw-A7-2	A7	Gwalior	3.3	0.2	6.3	34.3–52.2
VH-Tsw-A10-1	A10	Leh	3.8	0.3	9.8	1.4–10.9
VH-Tsw-A10-2	A10	Gwalior	4.7	0.2	10.7	26.6–36.1
VH-Tsw-A10-3	A10	Delhi	7.4	0.3	13.9	25.4–46.2
VH-Tsw-B1-1	B1	Gwalior	2.5	0.2	4.9	62.1–86.9
VH-Tsw-B3-1	B3	Leh	4.7	0.3	12.5	33.2–48.5
VH-Tsw-B3-2	B3	Delhi	6.8	0.3	12.5	45.1–66.8
VH-Tsw-B3-3	B3	Gwalior	2.5	0.2	5.6	70.8–77.6
VH-Tsw-B3-4	B3	Delhi	3.9	0.2	6.8	99.2–118.1
VH-Tsw-B4-1	B4	Delhi	5	0.3	8.9	1.5–26.4
VH-Tsw-B6-1	B6	Gwalior	3	0.2	6.7	0.0–9.6

of three *S-QTL* were detected in the TD population wherein one *S-QTL* was formed by merging the component QTL from three environments (*Tsw-S-A9-1-TD*), while the other two *S-QTL* were formed by merging the component QTL from two environments. In the other three populations, thec EPJ population identified two *S-QTL* (*Tsw-S-A1-1-EPJ* and *Tsw-S-B8-1-EPJ*) and the VH population identified one *S-QTL* (*Tsw-S-A4-1-VH*). All the population-specific QTL in the DE population were detected as *E-QTL*.

Tsw-C-A3-1 in LG A3 was formed by merging six component QTL and all included positive alleles from both Indian and east European gene pool lines in the 'J' Brassica genomic block. This *C-QTL* that merged the QTL from three populations (DE, EPJ and VH) showed the highest allelic effects from Pusajaikisan, showing an increase in seed weight by 0.6 g followed by Donskaja-IV by 0.3 g and Varuna by 0.2 g in the respective populations (Table 4). In the second *C-QTL*, *Tsw-C-A3-2* in F block, formed by merging minor QTL from two populations (DE and EPJ), revealed the maximum allelic effects in the EPJ population, wherein the Pusajaikisan allele increased the trait phenotype by 0.4 g as compared to Donskaja-IV allele in the DE population showing an increase by 0.2 g (Table 4).

The *Tsw-C-A7-1* that merged QTL from all the four populations and located at the junction of E and N blocks showed contribution of positive alleles from both Indian (Pusajaikisan, TM-4 and Varuna) and east European (Donskaja-IV) lines. Among the Indian parents, Varuna increased the phenotype by 0.5 g followed by Pusajaikisan

(0.3 g) and TM-4 (0.2 g) in the respective populations. In the DE population, the parent Donskaja-IV increased the phenotype by 0.2 g (Table 4).

Tsw-C-A10-1 also merged QTL from all the four populations. The contribution of trait enhancing alleles to this *C-QTL* was observed to be from both female (TD and VH populations) and male parents (EPJ and DE populations). The allelic effect of Varuna allele was found to be maximum, as it increased the phenotype by 0.6 g while the other three parents, Pusajaikisan, EH-2 and TM-4, increased the trait phenotype by 0.2 g each in the respective mapping populations (Table 4).

The two *C-QTL* in LG B3, *Tsw-C-B3-1* and *Tsw-C-B3-2* were formed by merging QTL from three and two populations, respectively. In *Tsw-C-B3-1*, Varuna allele increased the phenotype by 0.6 g followed by Pusajaikisan (0.5 g) and EH-2 (0.1 g). In *Tsw-C-B3-2*, both Varuna and Pusajaikisan increased the phenotype by 0.4 g each in the respective populations (Table 4).

On the basis of the diversity in allelic effects of the six parents utilized in the development of the four bi-parental populations, a comparative allelic status has been deciphered for these parental lines. For the purpose, it has been hypothesized that Heera/EH-2 is the common parent, as EH-2 is the EMS-induced early mutant of Heera (Rout et al. 2015) and would therefore possess the same alleles since both these lines have the same phenotype for the trait (Table 1). The exercise revealed that Indian parents—Pusajaikisan or Varuna—possess the most positive allele in all
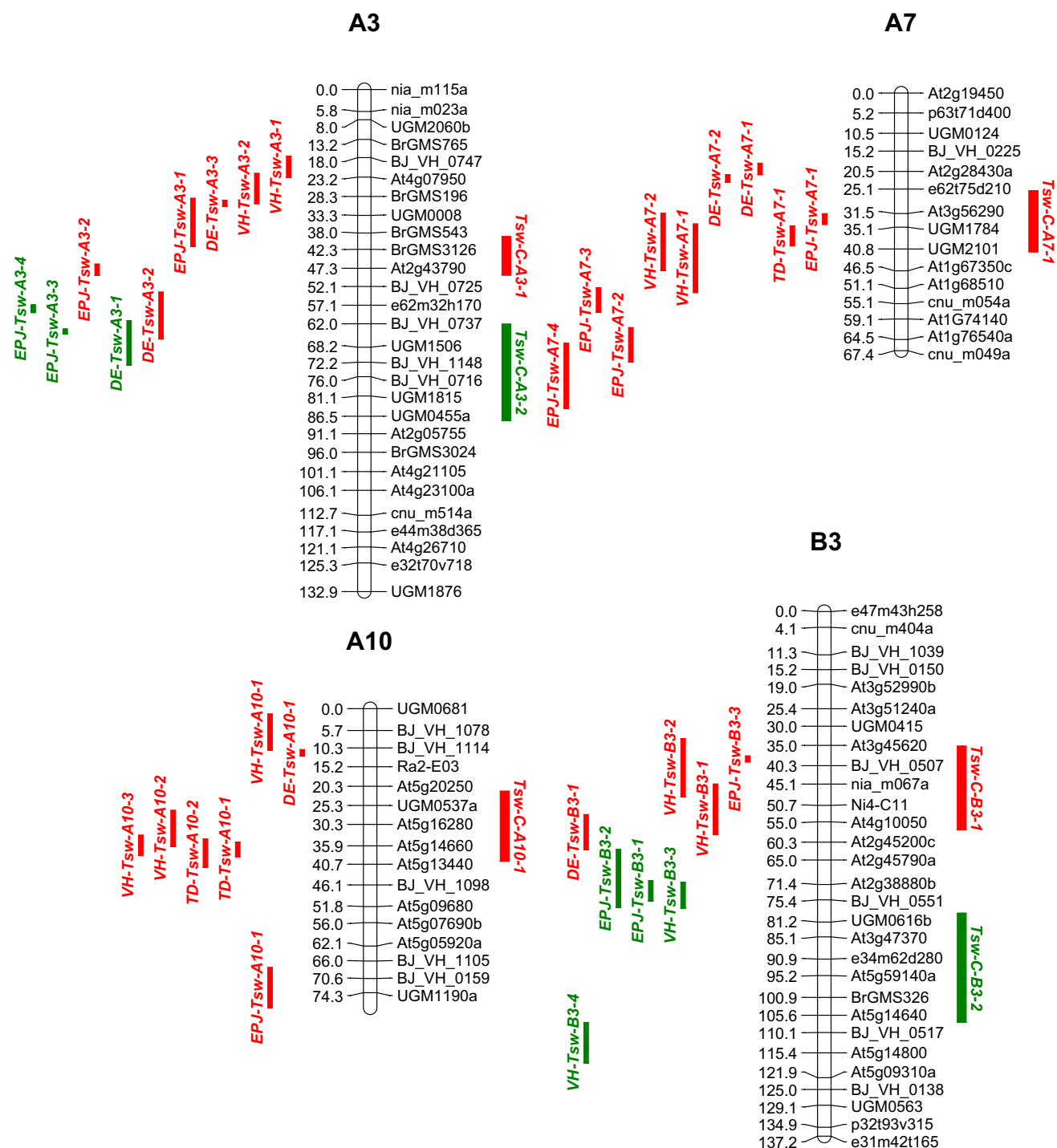
A3

cM	Marker
0.0	nia_m115a
5.8	nia_m023a
8.0	UGM2060b
13.2	BrGMS765
18.0	BJ_VH_0747
23.2	At4g07950
28.3	BrGMS196
33.3	UGM0008
38.0	BrGMS543
42.3	BrGMS3126
47.3	At2g43790
52.1	BJ_VH_0725
57.1	e62m32h170
62.0	BJ_VH_0737
68.2	UGM1506
72.2	BJ_VH_1148
76.0	BJ_VH_0716
81.1	UGM1815
86.5	UGM0455a
91.1	At2g05755
96.0	BrGMS3024
101.1	At4g21105
106.1	At4g23100a
112.7	cnu_m514a
117.1	e44m38d365
121.1	At4g26710
125.3	e32t70v718
132.9	UGM1876

A3 left labels: *EPJ-Tsw-A3-4*, *EPJ-Tsw-A3-3*, *EPJ-Tsw-A3-2*, *DE-Tsw-A3-1*, *DE-Tsw-A3-2*, *EPJ-Tsw-A3-1*, *DE-Tsw-A3-3*, *VH-Tsw-A3-2*, *VH-Tsw-A3-1*
A3 right labels: *Tsw-C-A3-1*, *Tsw-C-A3-2*, *EPJ-Tsw-A7-4*, *EPJ-Tsw-A7-3*, *EPJ-Tsw-A7-2*

A7

cM	Marker
0.0	At2g19450
5.2	p63t71d400
10.5	UGM0124
15.2	BJ_VH_0225
20.5	At2g28430a
25.1	e62t75d210
31.5	At3g56290
35.1	UGM1784
40.8	UGM2101
46.5	At1g67350c
51.1	At1g68510
55.1	cnu_m054a
59.1	At1G74140
64.5	At1g76540a
67.4	cnu_m049a

A7 left labels: *DE-Tsw-A7-2*, *DE-Tsw-A7-1*, *VH-Tsw-A7-2*, *VH-Tsw-A7-1*, *TD-Tsw-A7-1*, *EPJ-Tsw-A7-1*
A7 right label: *Tsw-C-A7-1*

A10

cM	Marker
0.0	UGM0681
5.7	BJ_VH_1078
10.3	BJ_VH_1114
15.2	Ra2-E03
20.3	At5g20250
25.3	UGM0537a
30.3	At5g16280
35.9	At5g14660
40.7	At5g13440
46.1	BJ_VH_1098
51.8	At5g09680
56.0	At5g07690b
62.1	At5g05920a
66.0	BJ_VH_1105
70.6	BJ_VH_0159
74.3	UGM1190a

A10 left labels: *VH-Tsw-A10-3*, *VH-Tsw-A10-2*, *TD-Tsw-A10-2*, *TD-Tsw-A10-1*, *VH-Tsw-A10-1*, *DE-Tsw-A10-1*, *EPJ-Tsw-A10-1*
A10 right label: *Tsw-C-A10-1*

B3

cM	Marker
0.0	e47m43h258
4.1	cnu_m404a
11.3	BJ_VH_1039
15.2	BJ_VH_0150
19.0	At3g52990b
25.4	At3g51240a
30.0	UGM0415
35.0	At3g45620
40.3	BJ_VH_0507
45.1	nia_m067a
50.7	Ni4-C11
55.0	At4g10050
60.3	At2g45200c
65.0	At2g45790a
71.4	At2g38880b
75.4	BJ_VH_0551
81.2	UGM0616b
85.1	At3g47370
90.9	e34m62d280
95.2	At5g59140a
100.9	BrGMS326
105.6	At5g14640
110.1	BJ_VH_0517
115.4	At5g14800
121.9	At5g09310a
125.0	BJ_VH_0138
129.1	UGM0563
134.9	p32t93v315
137.2	e31m42t165

B3 left labels: *DE-Tsw-B3-2*, *EPJ-Tsw-B3-1*, *VH-Tsw-B3-3*, *DE-Tsw-B3-1*, *VH-Tsw-B3-2*, *VH-Tsw-B3-1*, *EPJ-Tsw-B3-3*, *VH-Tsw-B3-4*
B3 right labels: *Tsw-C-B3-1*, *Tsw-C-B3-2*

Fig. 2 Positions of six consensus *Tsw* (*C-Tsw*) QTL distributed on LGs A3, A7, A10 and B3 in *Brassica juncea* consensus map. Marker names are on the *right* of the linkage group bar and the positions in centiMorgans (cM) are on the *left*. The component QTL identified in EPJ, DE, TD and VH populations integrating into the consensus QTL are shown on the *left side* of the linkage group bar

the six *C-QTL* among the *B. juncea* lines studied (Table 4). In three *C-QTL* (*Tsw-C-A3-1*, *Tsw-C-A3-2* and *Tsw-C-B3-2*), Puasjaikisan had the most positive alleles and in the other three *C-QTL* (*Tsw-C-A7-1*, *Tsw-C-A10-1* and *Tsw-C-B3-1*) Varuna had the most positive alleles.

Epistatic and QTL × environment interactions

Genome-wide analysis for epistatic (additive × additive) and QTL × environment interactions for *Tsw* trait in the four bi-parental populations revealed no significant

Table 3 Characteristics of consensus *Tsw* (*C-Tsw*) QTL identified by meta-analysis through merging four component QTL maps (VH, TD, DE and EPJ)

LG	Component QTL	Donor parent	R^2 (%)	Consensus QTL	Peak position (interval)[a] (cM)	Crucifer genomic block
A3	*DE-Tsw-A3-2*	Donskaja-IV	18.6	*Tsw-C-A3-1*	44 (38.9–49.1)	J
	DE-Tsw-A3-3	Donskaja-IV	13.1			
	EPJ-Tsw-A3-1	Pusajaikisan	16.5			
	VH–Tsw-A3-1	Varuna	6.1			
	VH-Tsw-A3-2	Varuna	12.6			
	EPJ-Tsw-A3-2	Pusajaikisan	14.6			
	DE-Tsw-A3-1	Donskaja-IV	6.4	*Tsw-C-A3-2*	74.4 (61.2–87.6)	F
	EPJ-Tsw-A3-3	Pusajaikisan	3.8			
	EPJ-Tsw-A3-4	Pusajaikisan	9.2			
A7	*DE-Tsw-A7-1*	Donskaja-IV	9.5	*Tsw-C-A7-1*	33.5 (25.6–41.5)	E, N
	DE-Tsw-A7-2	Donskaja-IV	5.7			
	EPJ-Tsw-A7-1	Pusajaikisan	4.2			
	EPJ-Tsw-A7-2	Pusajaikisan	4.2			
	EPJ-Tsw-A7-3	Pusajaikisan	3.7			
	EPJ-Tsw-A7-4	Pusajaikisan	7.6			
	TD-Tsw-A7-1	TM-4	11.8			
	VH-Tsw-A7-1	Varuna	6.3			
	VH-Tsw-A7-2	Varuna	3.8			
A10	*DE-Tsw-A10-1*	EH-2	5.3	*Tsw-C-A10-1*	30.1 (21–39.3)	R
	EPJ-Tsw-A10-1	Pusajaikisan	5.7			
	TD-Tsw-A10-1	TM-4	7.5			
	TD-Tsw-A10-2	TM-4	8.7			
	VH-Tsw-A10-1	Varuna	9.8			
	VH-Tsw-A10-2	Varuna	10.7			
	VH-Tsw-A10-3	Varuna	13.9			
B3	*DE-Tsw-B3-1*	EH-2	5.4	*Tsw-C-B3-1*	46.6 (35.7–57.5)	O, P
	EPJ-Tsw-B3-3	Pusajaikisan	14.5			
	VH-Tsw-B3-1	Varuna	12.5			
	VH-Tsw-B3-2	Varuna	12.5			
	EPJ-Tsw-B3-1	Pusajaikisan	6.5	*Tsw-C-B3-2*	93.6 (79.4–107.8)	R, J
	EPJ-Tsw-B3-2	Pusajaikisan	8.4			
	VH-Tsw-B3-3	Varuna	5.6			
	VH-Tsw-B3-4	Varuna	6.8			

[a] Interval implies the confidence interval of *C-QTL* on the consensus map

Table 4 Estimation of allelic effects (measured in grams) of six consensus QTL for seed weight trait involving four bi-parental populations of *B. juncea* and determination of allelic status of six parents

Consensus *Tsw* QTL	Allelic effect of consensus *Tsw* QTL as estimated from original bi-parental mapping population				Comparative allelic status of parents
	DE	EPJ	TD	VH	
Tsw-C-A3-1	0.3	0.6	–	0.2	PJ > D = T > V > E = H
Tsw-C-A3-2	0.2	0.4	–	–	PJ > D = T > E = H = V
Tsw-C-A7-1	0.2	0.3	0.2	0.5	V > T > PJ > D > E = H
Tsw-C-A10-1	−0.2	0.2	0.2	0.6	V > PJ > E = T = H > D
Tsw-C-B3-1	−0.1	0.5	–	0.6	V > PJ > E = H > D = T
Tsw-C-B3-2	–	0.4	–	0.4	PJ = V > E = H = D = T

V Varuna, *PJ* Pusajaikisan, *T* TM-4, *D* Donskaja-IV, *E* EH-2, *H* Heera

Table 5 Epistatic (additive × additive) effects for seed weight in four bi-parental populations (EPJ, DE, TD and VH)

Mapping population	LG	Flanking markers	Interval (cM)	At block	LG	Flanking markers	Interval (cM)	At Block	Type of interaction[a]	AA[b,c]	H²AA (%)
DE	A8	BrgMS609–BrgMS18(329)	16.5–22.1	U	A9	UGM0322–nia_m026	60.6–62.3	B	III	0.07***	2.0
EPJ	A5[d]	UGM1584–At2g32920(2G415)	24.3–28.4	J	A7	At3g53190a(3g178)–UGM2219	44.6–50.8	N	I	0.1197***	1.6
	A5	UGM1584–At2g32920(2g415)	24.3–28.5	J	B3	At2g45690b(2g169)–UGM1330	48.7–51.5	J	I	0.0957***	1.3
	B3	At2g45690b(2g169)–UGM1330	48.7–51.5	J	B4	At1g75330(1g194)–UGM1854	50.9–54.9	E	I	0.1041***	1.1
	B4	At1g75330(1g194)–UGM1854	50.9–54.9	E	B8	UGM1691A–UGM2260B	112.0–113.6	N	I	−0.1141***	0.6
	A8	At1g23180a(1g33)–UGM1077	62.2–64.0	A	A10	At5g04920c(5g7)–UGM1190A	83.6–89.3	R	III	−0.0978***	1.6
TD	A8	nia_m050a–e48t78v400	61.4–77.2	A	A10	p63t71d210–p79t78d67	12.7–12.9	R	I	−0.0717***	0.7
VH	A3	cnu_m466a–At2g28520(2g401)	48.3–50.6	I	B3	At2g38810(2g354)–At2g36530a(2c22)	69.9–73.9	J	I	0.0774***	1.4
	A4	p53m47h285–At2g25737(2C10)	44.3–48.1	I	A10	At5g14105a(5g416)–BrgMS86	35.8–38.3	R	I	0.1088***	1.9
	A7	At3g55005c(3g173)–BJ_VH_0127	31.7–42.7	N	B3	At2g38810(2g354)–At2g36530a(2c22)	69.9–73.9	J	I	0.0623*	0.9
	A1	BrgMS7–At3g51610(3g185)	56.7–60.7	N	B2	e31m62v199–At5g20010a(5g313)	53.9–57.0	R	III	0.1474***	5.4
	A9	At1g34350(1g42lp)–At4g09630(4g180)	66.7–67.8	B	B3	At5g14670(5g403)–At2g45690(2g191–2)	60.3–64.5	J	II	0.1252***	4.4

*, **, *** $P < 0.05$, $P < 0.01$, $P < 0.005$ respectively

[a] Type I interaction: epistatic interaction between two loci with additive effects, Type II interaction: epistatic interaction between a additive QTL and a modifier locus, Type III interaction: epistatic interaction between two modifier loci

[b] AA additive × additive effect

[c] Positive values of AA indicate trait values are increased by parental association of the alleles for the two interacting loci and vice versa

[d] QTL detected as main effect QTL in QTL Network 2.0, but not in Windows QTL Cartographer 2.5

Table 6 Candidate genes for seed weight showing *in silico* co-localization with *C-Tsw* QTL and the corresponding A subgenome seed weight QTL identified in *B. rapa*

S. no.	Gene	At ID	Crucifer genomic block	Consensus *Tsw* QTL showing *in silico* co-localization	Flanking markers	*Brassica rapa* paralog		Corresponding A genome QTL[b] in *B. rapa* (Li et al. 2013)
						LG	Bra ID	
1	*ATGRF3*	At2g36400	J	*Tsw-C-A3-1*	At2g33120–At2g45690	A3	Bra023066(MF2)	17.0 (J)
2	*CKX1*	At2g41510	J	*Tsw-C-A3-1*	At2g33120–At2g45690	A3	Bra000229(MF2)	
3	*HAIKU1*	At2g35230	J	*Tsw-C-A3-1*	At2g33120–At2g45690	A3	Bra023004(MF2)	
4	*TTG2*	At2g37260	J	*Tsw-C-A3-1*	At2g33120–At2g45690	A3	Bra023112(MF2)	
5	*ARR1*	At3g16857	F	*Tsw-C-A3-2*	At3g06480–At3g21400	A3	Bra001641(MF2)	–
6	*SGT*	At3g07020	F	*Tsw-C-A3-2*	At3g06480–At3g21400	A3	Bra001230(MF2)	
7	*BZR1*	At1g75080	E	*Tsw-C-A7-1*	At1g67230–At1g79970	A7	Bra015868(LF)	–
8	*DA2*	At1g78420	E	*Tsw-C-A7-1*	At1g67230–At1g79970	A7	Bra035036(LF)	
9	*MYB56*	At5g17800	R	*Tsw-C-A10-1*	At5g13440–At5g18650	A10	Bra002108(LF)	7.1–9.0 (R)
10	*TT4*	At5g13930	R	*Tsw-C-A10-1*	At5g13440–At5g18650	A10	Bra008792(LF)	
11	*GASA4*[a]	At5g15230	R	*Tsw-C-B3-2*	At5g14030–At5g17410	–	–	–

LF least fractionated, *MF2* most fractionated 2

[a] *GASA4* shows physical co-localization with *C-Tsw* QTL in LG B3

[b] Phenotypic variance explained in percentage

QTL × environment interactions in any of the bi-parental populations (data not shown). Analysis of digenic interactions (A × A) in these four populations revealed five interactions each in the EPJ and VH populations and one interaction each in the DE and TD populations (Table 5). It was also observed that in the EPJ and VH populations, where the populations were developed with high phenotypic contrasts between the parents, a greater number of Type 1 digenic interactions (four out of five interactions in EPJ and three out of five in VH) were detected compared to that in the populations with low phenotypic contrasts between the parents (DE and TD).

In silico co-localization of candidate genes

An *in silico* search for candidate genes for seed weight trait from the heterologous systems identified a total of 77 genes from Arabidopsis and 11 from other agricultural crops (rice, maize and tomato) (Online Resource 3 and Online Resource 4). The orthologs of 11 crop genes in Arabidopsis were established using BLASTX tool of TAIR database (http://www.arabidopsis.org/Blast/index.jsp). The probable

positions of these candidate genes were predicted in *B. juncea* on the basis of the syntenic blocks of *C-QTL*. *B. rapa* paralogs of these genes were identified by search for the 'syntenic gene' in BRAD (http://brassicadb.org/brad/searchSyntenytPCK.php).

In silico co-localization of candidate genes with *C-QTL* was carried out based on the arrangement of Brassica building blocks on the *B. juncea* consensus genetic map. There were ten candidate genes which showed *in silico* co-localization with four A subgenome *C-QTL* (Table 6). In the absence of B genome sequence of *Brassica* species, no candidate genes could be co-localized to the B subgenome *C-QTL*. However, the candidate gene *GASA4* (At5g15230) mapped in the EPJ and VH populations due to a polymorphism in the gene and was found to co-map to the *Tsw-C-B3-2* region on LG B3. Among the ten candidate genes co-localizing to four A subgenome *C-QTL*, four candidate genes (*ATGRF3*, *CKX1*, *HAIKU1* and *TTG2*) could be located in *Tsw-C-A3-1*, two genes (*ARR1* and *SGT*) in *Tsw-C-A3-2*, two (*BZR1* and *DA2*) in *Tsw-C-A7-1* and two (*MYB56* and *TT4*) in *Tsw-C-A10-1*. While all the six candidate genes in *Tsw-C-A3-1* and *Tsw-C-A3-2* QTL

correspond to the MF2 (most fractionated 2) region in the 'J' block, other candidate genes co-locating to the QTL regions of LG A7 (*Tsw-C-A7-1*) and A10 (*Tsw-C-A10-1*) correspond to the LF (least fractionated) regions in the 'E' and 'R' blocks of *B. rapa*, respectively. Comparisons of identified *C-QTL* in the present study with QTL data for seed weight in *B. rapa* (Li et al. 2013) revealed that *C-QTL* in LG A3 (*Tsw-C-A3-1*) (J block) and in A10 (*Tsw-C-A10-1*) (R block) could be common between the two species (Table 6).

Discussion

Seed weight is an important yield component trait in oilseed *Brassica* species and is quantitative in nature. Among the cultivated Brassicas, *B. juncea* holds the maximum genetic variation for the trait. While extensive breeding efforts in Indian *B. juncea* gene pool lines have resulted in the development of varieties with higher seed weight, the east European gene pool lines are small seeded due to lack of breeding efforts in this direction (Ramchiary et al. 2007). Although there are several studies on genetic analysis of seed weight in *B. rapa* (Li et al. 2013; Basnet et al. 2015), *B. napus* (Yang et al. 2012; Li et al. 2014; Fu et al. 2015; Geng et al. 2016) and in *B. juncea* (Ramchiary et al. 2007; Yadava et al. 2012), no systematic study on the trait harvesting maximum available genetic variation has been undertaken in any of these *Brassica* species. The present study, therefore, has been attempted at understanding the genetics of this important quantitatively inherited trait in *B. juncea* describing the maximum genetic variation in the species. As the *B. juncea* germplasm is highly structured (Srivastava et al. 2001) and might run the risks of association mapping, the genetic dissection of the trait was undertaken using four bi-parental populations involving six lines from both Indian and east European gene pools. The seeds of the six parents involved in the present study showed a wide range in weight varying from 2.0 to 7.6 g per 1000 seeds in different environments (Table 1).

We observed higher frequencies of transgressive segregants in the populations derived from parents with low contrast (DE and TD) than those derived from the parents with high contrasts (EPJ and VH) in seed weight (Fig. 1). It clearly indicates that the positive alleles for seed size are still dispersed among parental lines with low contrast and the transgressive segregants have inherited combinations of alleles acting in the same direction contributed by both the parents. This was substantiated by the QTL analysis of the DE and TD populations wherein both the parents contributed the positive alleles for increased seed weight (Table 2). This observation also indicates that breeding efforts are lacking for improving seed weight in

these lines. On the contrary, the occurrence of very few transgressive segregants in populations derived from the parents with high phenotypic contrasts (EPJ and VH) could be due to accumulation of all positive alleles in the high parents. QTL analysis in these two populations confirmed this hypothesis, as no positive allele was identified from the phenotypically inferior parents in EPJ and VH populations (Table 2). It reaffirms the strong breeding efforts made in the Indian germplasm lines for higher seed size.

The present study identified 65 *Tsw* QTL from four bi-parental populations distributed over 13 of 18 LGs, of which 75 % were minor QTL (explaining <10 % of the phenotypic variance). It indicates that in *B. juncea* also a sizable portion of the genetic variation of seed weight is controlled by a large number of loci with minor effects as has been reported in *B. napus* (Li et al. 2014). The study also identified around 50 % of the population-specific QTL as environmental QTL (*E-QTL*). The occurrence of environment-specific QTL is also a characteristic of quantitative traits. It has been proposed that they may be exhibiting 'conditional neutrality' (they affect the trait in one environment but not in the other) or 'pleiotropy' (they affect *Tsw* in one environment, but affect another trait in a different environment because of pleiotropic nature of the underlying genes). The appearance of QTL which show conditional neutrality and antagonistic pleiotropic effects and their importance in adaptation to different environments have been shown in *Boechera stricta* (Anderson et al. 2013), *Arabidopsis* and other species (EI-Soda et al. 2014).

Quantitative trait variation is also influenced by epistatic and QTL × environment interactions, which contribute to the observable phenotypic variation (Mackay 2014). Our results did not indicate significant QTL × environment interactions. It could primarily be due to the fact that seed weight is a highly heritable trait in *B. juncea* and has a strong additive component. However, the observation of a higher number of epistatic interactions (additive × additive) that were primarily of Type 1 in populations derived from parents with high phenotypic contrasts (EPJ and VH) than the populations with lesser parental contrasts (DE and TD) (Table 5) is suggestive of the fact that the Type 1 interactions have been fixed during the selection and fixation of additive loci for higher seed weight in the Indian gene pool lines. The effects attributed to epistasis are deemed important while conceiving marker-assisted breeding programs for QTL introgression.

Meta-QTL analysis that merges overlapping QTL of a trait detected in multiple populations is a synthetic genetic model based on the hypothesis that different QTL that are common between the populations are controlled by the same gene(s) (Chardon et al. 2014). The analysis results in the identification of 'consensus' QTL which could have broader application in marker-assisted breeding as

compared to a QTL detected in a single bi-parental population. Meta-analysis also precisely quantifies the allelic effects at the QTL participating in the formation of consensus QTL. The estimated allelic effects could then be further used to establish the diversity among the parental alleles contributing to a consensus QTL in a particular case where all the bi-parental populations have been derived using common parents.

The meta-QTL analysis undertaken in the study identified a total of six consensus QTL (Table 3) by merging 33 of the 65 component QTL. The estimated allelic effects on seed weight showed maximum effects from Indian gene pool lines, Pusajaikisan or Varuna, in all the six consensus QTL. However, in the DE population (derived from two east European lines), Donskaja-IV showed maximum allelic effects (Table 4). Further comparative analysis of allelic diversity using EH-2/Heera as the common parent revealed that either Pusajaikisan or Varuna hold the most positive allele among the six parents (Table 4). Hence, the allelic diversity identified in the present study from two gene pool lines reveals the scope for genetic improvement of seed size by allele replacement between and/or within the two gene pools of *B. juncea*. However, the decision on inter-gene pool introgression of QTL for a component trait related to yield should only be made based on the genetic analysis of other yield component traits, as the majority of these traits are significantly correlated due to pleiotropy and/or close linkage (Yadava et al. 2012; Yang et al. 2012; Li et al. 2014).

Due to allelic diversity, QTL identified in a population may not be useful in other populations without marker validation and/or fine mapping (Xu and Crouch 2008; Nicholas 2006). Candidate gene analysis in the QTL regions that are shared across populations would permit the development of polymorphic markers, speeding up marker-assisted selection. In the present study, the synteny of candidate genes with the *Tsw* QTL localized 11 genes mapping to the regions harboring *C-Tsw* QTL (Table 6). These candidate genes could be validated and the polymorphism detected in the candidate genes be used for the precise transfer of QTL.

The comparative QTL map alignment for seed weight between the A subgenome of *B. juncea* (present study) with *B. rapa* (Li et al. 2013) identified two conserved QTL blocks, one in the J block of LG A3 (*Tsw-C-A3-1*) and the other in the R block of LG A10 (*Tsw-C-A10-1*). It indicates that these QTL blocks are evolutionarily conserved in the A subgenome of *B. juncea*. The *in silico* co-mapping study also identified four candidate genes, *ATGRF3, CKX1, HAIKU1* and *TTG2*, locating to the J block of QTL *Tsw-C-A3-1* whose paralogs were detected in the J block in LG A3 of *B. rapa* (Table 6). The J block in LG A3 in *B. rapa* has been shown to be a most fractionated block (MF2)

which is more ancient than the other two blocks LF and MF1 (LF: least fractionated and MF1: medium fractionated) (Wang et al. 2011). We, therefore, believe that one or more of these genes might be functionally conserved across the *Brassica* species and play an important role in determining seed size.

In conclusion, this study addresses a very important aspect of genetic improvement of *B. juncea* by identifying consensus seed weight QTL based on data from multiple segregating populations developed by utilizing the maximum genetic variation existing in the two gene pools of *B. juncea*. Meta-analysis simultaneously revealed six consensus QTL and uncovered the allelic diversity of the trait in *B. juncea*. The candidate genes which show *in silico* synteny to the genomic blocks containing consensus *Tsw* QTL in *B. juncea* would be physically mapped and functionally validated in future. In the context of heterosis breeding in *B. juncea*, the Indian parents having positive alleles for seed weight trait may find application in improving the east European gene pool lines.

Author contribution statement The experiment was conceived and designed by ND, DP and AKP and was performed by ND, KR, SKY and YSS. The data were analyzed by ND, KR, SKY, VG and AKP, and the manuscript was written and edited by ND, SKY, DP and AKP.

Compliance with ethical standards

Ethical standards The experiments included in the manuscript entitled "Genetic dissection of seed weight by QTL analysis and detection of allelic variation in Indian and east European gene pool lines of *Brassica juncea*" comply with the current laws of India for academic and research purposes.

Conflict of interest The authors declare that they have no competing interests.

References

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Koornneef M (1999) Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc Natl Acad Sci USA 96:4710–4717. doi:10.1073/pnas.96.8.4710

Anderson JT, Lee CR, Rushworth C, Colautti R, Mitchell-Olds T (2013) Genetic trade-offs and conditional neutrality contribute to local adaptation. Mol Ecol 22:699–708

Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics 20:2324–2326. doi:10.1093/bioinformatics/bth230

Basnet RK, Duwal A, Tiwari DN, Xiao D, Monakhos S et al (2015) Quantitative trait locus analysis of seed germination and seedling vigor in *Brassica rapa* reveals QTL hotspots and epistatic interactions. Front Plant Sci 6:1032. doi:10.3389/fpls.2015.01032

Bennett EJ, Roberts JA, Wagstaff C (2014) Use of mutants to dissect the role of ethylene signalling in organ senescence and the regulation of yield in *Arabidopsis thaliana*. J Plant Growth Regul 33(1):56–65

Chardon F, Jasinski S, Durandet M, Lécureuil A, Soulay F, Bedu M, Guerche P, Masclaux-Daubresse C (2014) QTL meta-analysis in *Arabidopsis* reveals an interaction between leaf senescence and resource allocation to seeds. J Exp Bot 65:3949–3962. doi:10.1093/jxb/eru125

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971. doi:10.1534/genetics.107.080101

Dhaka N (2015) Genetic dissection of some agronomically important characters with special emphasis on seed size in *Brassica juncea*. Ph.D. Thesis, University of Delhi, New Delhi

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294. doi:10.1111/j.1369-7625.2010.00632.x

Doganlar S, Frary A, Tanksley SD (2000) The genetic basis of seed-weight variation: tomato as a model system. Theor Appl Genet 100(8):1267–1273. doi:10.1007/s001220051433

Doughty J, Aljabri M, Scott RJ (2014) Flavonoids and the regulation of seed size in *Arabidopsis*. Biochem Soc Trans 42:364–369. doi:10.1042/BST20140040

EI-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MGM (2014) Genotype × environment interaction QTL mapping in plants: lessons from *Arabidopsis*. Trends Plant Sci 19:390–398. doi:10.1016/j.tplants.2014.01.001

Fan C, Cai G, Qin J, Li Q, Yang M, Wu J, Fu T, Liu K, Zhou Y (2010) Mapping of quantitative trait loci and development of allele-specific markers for seed weight in *Brassica napus*. Theor Appl Genet 121(7):1289–1301. doi:10.1007/s00122-010-1388-4

Fu Y, Wei D, Dong H, He Y, Cui Y, Mei J, Wan H, Li J, Snowdon R, Friedt W, Li X (2015) Comparative quantitative trait loci for silique length and seed weight in *Brassica napus*. Sci Rep 5:14407. doi:10.1038/srep14407

Garcia D, Fitz Gerald JN, Berger F (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in *Arabidopsis*. Plant Cell 17:52–60. doi:10.1105/tpc.104.027136

Geng X, Jiang C, Yang J, Wang L, Wu X, Wei W (2016) Rapid identification of candidate genes for seed weight using the SLAF-Seq method in *Brassica napus*. PLoS One 11(1):e0147580

Goffinet B, Gerber S (2000) Quantitative trait loci: a meta-analysis. Genetics 155:463–473

Herridge RP, Day RC, Baldwin S, Macknight RC (2011) Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery. Plant Methods 7:1–11. doi:10.1186/1746-4811-7-3

Hossain S, Ford R, McNeil D, Pittock C, Panozzo JF (2010) Inheritance of seed size in chickpea (*Cicer arietinum* L.) and identification of QTL based on 100 seed weight and seed size index. Aust J Crop Sci 4(2):126–135

Jiang WB, Lin WH (2013) Brassinosteroid functions in *Arabidopsis* seed development. Plant Signal Behav 8:4–6. doi:10.4161/psb.25928

Jiang WB, Huang HY, Hu YW, Zhu SW, Wang ZY, Lin WH (2013) Brassinosteroid regulates seed size and shape in *Arabidopsis*. Plant Physiol 162:1965–1977. doi:10.1104/pp.113.217703

Li N, Li Y (2015) Maternal control of seed size in plants. J Exp Bot 66:1087–1097

Li X, Ramchiary N, Dhandapani V, Choi SR, Hur Y, Nou IS, Yoon MK, Lim YP (2013) Quantitative trait loci mapping in *Brassica rapa* revealed the structural and functional conservation of genetic loci governing morphological and yield component traits in the A, B, and C sub-genomes of *Brassica* species. DNA Res 20(1):116

Li N, Shi J, Wang X, Liu G, Wang H (2014) A combined linkage and regional association mapping validation and fine mapping of two major pleiotropic QTLs for seed weight and silique length in rapeseed (*Brassica napus* L.). BMC Plant Biol 14:114. doi:10.1186/1471-2229-14-114

Luo M, Dennis ES, Berger F et al (2005) *MINISEED3 (MINI3)*, a WRKY family gene, and *HAIKU2 (IKU2)*, a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. Proc Natl Acad Sci USA 102:17531–17536. doi:10.1073/pnas.0508418102

Mackay TFC (2014) Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat Rev Genet 15:22–33. doi:10.1038/nrg3627

Nicholas FW (2006) Discovery, validation, and delivery of DNA markers. Aust J Exp Agric 46:155–158

Pradhan AK, Pental D (2011) Genetics of *Brassica juncea*. In: Schmidt R, Bancroft I (eds) Genetics and genomics of the Brassicaceae. Springer, New York, pp 323–345

Pradhan AK, Sodhi YS, Mukhopadhyay A, Pental D (1993) Heterosis breeding in Indian mustard (*Brassica juncea* L. Czern & Coss): analysis of component characters contributing to heterosis for yield. Euphytica 69(3):219–229

Qi L, Mao L, Sun C, Pu Y, Fu T, Ma C, Shen J, Tu J, Yi B, Wen J (2014) Interpreting the genetic basis of silique traits in *Brassica napus* using a joint QTL network. Plant Breed 133:52–60. doi:10.1111/pbr.12131

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. doi:10.1007/s00122-007-0610-5

Rout K (2014) Genetic analysis of oil content and seed glucosinolate trait in *Brassica juncea* Ph.D. thesis, University of Delhi, New Delhi

Rout K, Sharma M, Gupta V, Mukhopadhyay A, Sodhi YS, Pental D, Pradhan AK (2015) Deciphering allelic variations for seed glucosinolate traits in oilseed mustard (*Brassica juncea*) using two bi-parental mapping populations. Theor Appl Genet 128(4):657–666. doi:10.1007/s00122-015-2461-9

Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci 11:535–542

Sodhi YS, Chandra A, Verma JK, Arumugam N, Mukhopadhyay A, Gupta V, Pental D, Pradhan AK (2006) A new cytoplasmic male sterility system for hybrid seed production in Indian oilseed mustard *Brassica juncea*. Theor Appl Genet 114(1):93–99. doi:10.1007/s00122-006-0413-0

Srivastava A, Gupta V, Pental D, Pradhan AK (2001) AFLP-based genetic diversity assessment amongst agronomically important natural and some newly synthesized lines of *Brassica juncea*. Theor Appl Genet 102(2):193–199. doi:10.1007/s001220051635

Sun YN, Pan JB, Shi XL, Du XY, Wu Q, Qi ZM, Jiang HW, Xin DW, Liu CY, Hu GH, Chen QS (2012) Multi-environment mapping and meta-analysis of 100-seed weight in soybean. Mol Biol Rep 39:9435–9443. doi:10.1007/s11033-012-1808-4

Sundaresan V (2005) Control of seed size in plants. Proc Natl Acad Sci USA 102(50):17887–17888

Wang S, Basten C, Zeng Z (2006) Windows QTL cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh **(2010)**

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F, Huang S (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43:1035–1039. doi:10.1038/ng.919

Xu Y, Crouch JH (2008) Marker-assisted selection in plant breeding: from publications to practice. Crop Sci 48:391–407

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. doi:10.1007/s00122-012-1934-3

Yang J, Zhu J, Williams RW (2007) Mapping the genetic architecture of complex traits in experimental populations. Bioinformatics 23:1527–1536. doi:10.1093/bioinformatics/btm143

Yang P, Shu C, Chen L, Xu J, Wu J, Liu K (2012) Identification of a major QTL for silique length and seed weight in oilseed rape (*Brassica napus* L.). Theor Appl Genet 125:285–296. doi:10.1007/s00122-012-1833-7

Zhou Y, Zhang X, Kang X, Zhao X, Zhang X, Ni M (2009) *SHORT HYPOCOTYL UNDER BLUE1* associates with *MINISEED3* and *HAIKU2* promoters in vivo to regulate *Arabidopsis* seed development. Plant Cell 21:106–117. doi:10.1105/tpc.108.064972