Theor Appl Genet (2014) 127:339–347
DOI 10.1007/s00122-013-2222-6

ORIGINAL PAPER

Natural mutations in two homoeologous *TT8* genes control yellow seed coat trait in allotetraploid *Brassica juncea* (AABB)

Lakshmi K. Padmaja · Parul Agarwal · Vibha Gupta · Arundhati Mukhopadhyay · Yaspal S. Sodhi · Deepak Pental · Akshay K. Pradhan

Received: 13 February 2013 / Accepted: 18 October 2013 / Published online: 19 November 2013

Abstract Identification of the candidate gene responsible for the seed coat colour variation in *Brassica juncea* was undertaken following an earlier study where two independent loci (*BjSc1* and *BjSc2*) were mapped to two linkage groups, LG A9 and B3 (Padmaja et al. in Theor Appl Genet 111:8–14, 2005). The genome search from BRAD data for the presence of flavonoid genes in *B. rapa* identified three candidate genes namely, *DFR*, *TT1* and *TT8* in the LG A9. Quantitative real-time PCR revealed absence of transcript for the late biosynthetic genes (LBGs) and showed significant reduction of transcript in the *TT8* from the developing seeds of yellow-seeded line. While mapping of two *DFR* genes, the *BjuA.DFR* and *BjuB.DFR* did not show perfect co-segregation with the seed coat colour loci, that of the two *TT8* genes, *BjuA.TT8* and *BjuB.TT8* showed perfect co-segregation with the seed coat colour phenotype. The *BjuA.TT8* allele from the yellow-seeded line revealed the presence of an insertion of 1,279 bp in the exon 7 and did not produce any transcript as revealed by reverse transcriptase PCR. The *BjuB.TT8* allele from the yellow-seeded line revealed the presence of an SNP (C→T) in the exon 7 resulting in a stop codon predicting a truncated protein lacking the C-terminal 8 amino acid residues and produced significantly low level of transcript than its wild-type counterpart. Hence, it is hypothesized that the mutations in both the *TT8* genes are required for inhibiting the transcription of LBGs in the yellow-seeded mutant of *B. juncea*.

Communicated by R. G. F. Visser.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2222-6) contains supplementary material, which is available to authorized users.

L. K. Padmaja · P. Agarwal · V. Gupta · A. Mukhopadhyay · Y. S. Sodhi · D. Pental · A. K. Pradhan (✉)
Centre for Genetic Manipulation of Crop Plants, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India
e-mail: pradhancgmcp@gmail.com

D. Pental · A. K. Pradhan
Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India

Introduction

Yellow-seeded varieties are preferred over brown-seeded varieties in oilseed *Brassica* species due to their superior meal quality (Slominski 1997; Rehman and McVetty 2011). Among the oilseed *Brassica* species, varieties with natural mutations for the yellow-seeded trait are under cultivation in diploid *B. rapa* (AA) (Yellow Sarson type) and allotetraploid *B. juncea* (AABB). In *B. rapa* (Yellow sarson cv. *trilocularis*), yellow seed coat trait has been shown to be controlled by one recessive gene (Ahmed and Zuberi 1971; Li et al. 2012). In *B. juncea*, the seed coat colour trait has been shown to be controlled by two independently assorting duplicate genes where brown seed coat shows dominance over yellow seed coat (Vera et al. 1979; Vera and Woods 1982; Padmaja et al. 2005). Molecular markers such as AFLP (Negi et al. 2000; Lionneton et al. 2004; Sabarwal et al. 2004), RFLP (Mahmood et al. 2005), RAPD (Yan et al. 2009) and SSR (Padmaja et al. 2005) have been used to map two loci controlling seed coat colour in *B. juncea*.

In Brassicaceae, the brown seed coat colour is due to the deposition of oxidized form of a flavonoid, proanthocyanidin (PA), in the endothelium layer of the inner integument of seed coat (Lepiniec et al. 2006). The pathways have been extensively characterized in the model species,

Arabidopsis thaliana. It is synthesized through a common phenylpropanoid biosynthetic pathway. In *Arabidopsis*, many transparent testa (*tt*) mutants affecting flavonoid accumulation and altering seed pigmentation have been identified. Several of these genes have been characterized at the molecular level encoding two groups of structural proteins; the early biosynthetic genes (EBGs) [chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavanone-3-hydroxylase (*F3H*) and flavanone-3′-hydroxylase (*F3′H*)] and the late biosynthetic genes (LBGs) (dihydroflavonol reductase [*DFR*], leucocyanidin dioxygenase [*LDOX*] and anthocyanidin reductase [*ANR* (*BAN*)]), encoding regulatory proteins (*TT1*, *TT2*, *TT8*, *TT16*, *TTG1* and *TTG2*) and encoding protein probably involved in flavonoid compartmentation (*TT12*, *TT19* and *AHA10*) (Lepiniec et al. 2006). It has been established that the regulation of the expression of structural genes particularly the expression of LBGs (*DFR* and *BAN*) is orchestrated by a ternary complex (called MBW complex) involving transcription factors from *TT2* (*R2R3-MYB*), *TT8* (basic helix–loop–helix, bHLH) and WD40 regulatory protein encoded by *TTG1* leading to the biosynthesis of PAs (Baudry et al. 2004; Hichri et al. 2011).

There is limited information about the candidate gene(s) controlling seed coat colour in oilseed *Brassica* species. In diploid *B. rapa*, it was shown that the yellow-seeded trait in the Chinese cabbage is due to a 94-base deletion in the *TTG1* gene leading to loss of function of this gene (Zhang et al. 2009). In a recent study, it was shown that the yellow-seeded trait in yellow sarson of *B. rapa* is due to the loss of function of *TT8* gene caused by an insertion of 4,320 bp in its second intron (Li et al. 2012).

In an earlier study in allotetraploid *B. juncea*, Padmaja et al. (2005) mapped the two loci controlling the seed coat colour trait in *B. juncea* and observed that microsatellite markers Na10-A08 in the linkage group (LG) A9 and Ni4-F11 in the LG B3 co-segregated without any recombination with the seed coat colour genes *BjSc1* and *BjSc2*, respectively. In this study, the work was further extended in search of the candidate gene responsible for the seed coat colour variation in *B. juncea*. We report here the involvement of *TT8* gene at both the loci in LG A9 and LG B3 containing two different types of mutations affecting the transcription of *DFR* and *BAN* leading to the development of yellow seed coat. Following comparative genomics, fine mapping and cloning, the alleles of *TT8* gene from yellow-seeded variant revealed the presence of an insertion of 1,279 bp in the LG A9 and an SNP creating nonsense mutation in the LG B3 both in exon 7 of the respective gene. Validation of *TT8* gene for its involvement in yellow-seeded trait at both the loci was done by genetic means through phenotype–genotype association and gene expression study by RT-PCR.

Materials and methods

Plant material and mapping populations

Four mapping populations were used in the present study. Mapping population 1 consisting of 123 DH lines was derived from the F1 of cross between a brown-seeded Indian variety, Varuna and a yellow-seeded line, Heera. This population has been extensively used for the development of high density and comparative map (designated as VH map) in *B. juncea* (Ramchiary et al. 2007; Panjabi et al. 2008). The mapping population 2 (130 DH lines) and 3 (103 DH lines) were derived from F1 of two independent crosses, involving two brown-seeded DH lines (derivatives of mapping population 1)—designated as 4-4-2 and 16-4-8, respectively—crossed to a common yellow-seeded DH line, DH 26. Detailed pedigree and genesis of mapping populations 1, 2 and 3 have been shown in Padmaja et al. (2005). While the mapping population 2 has been used to study the co-segregation between the candidate gene and the seed coat colour gene, *BjSc1* located in the LG A9, the mapping population 3 has been used to study the co-segregation between the candidate gene and the seed coat colour gene, *BjSc2* located in the LG B3. Mapping population 4 is a RIL population consisting of 494 lines in F9 generation developed by single seed descent method from F2 of cross between Varuna and Heera. This population was used for fine mapping and validation of the candidate genes. In addition, a set of 20 *B. juncea* and 10 *B. rapa* genotypes consisting of both brown and yellow-seeded were also used for the validation purpose (Table 1).

Mapping and phenotyping

DNA was isolated from well-expanded leaves following the methodology of Rogers and Bendich (1994). Markers and genes were mapped to the existing VH map (Panjabi et al. 2008) using the mapping population 1 following the mapping criteria described by Pradhan et al. (2003). Seed coat colour was assessed by visual observation of seeds harvested from open-pollinated field-grown plants.

PCR amplification, cloning and sequencing

PCR amplifications were carried out in a 20-μl reaction containing 50 ng of genomic DNA, 800 μM dNTPs, 10 pmol of each primer and 1 U TaKaRa Ex *Taq* Hs DNA Polymerase (http://www.takara-bio.com). PCR profile included a two-step PCR with initial denaturation at 94 °C for 5 min followed by 35 cycles with denaturation at 94 °C for 30 s, annealing and extension at 62–68 °C for 1–4 min depending on the size of the amplicon. Final extension was at 62–68 °C for 10 min. The amplified products were cloned

Table 1 *Brassica* germplasm used for validation of two *TT8* genes (*BjuA.TT8* and *BjuB.TT8*) identified in *B. juncea* (present study) and *BrTT8* identified in *B. rapa* (Li et al. 2012)

Sl no.	Name of the line/cultivar	Origin	Type of allele for the *BrTT8* of *B. rapa*[a]	Type of allele for *BjuA.TT8* of *B. juncea*	Type of allele for *BjuB.TT8* of *B. juncea*
B. juncea brown-seeded					
1	RLM 619	India	a	a	a
2	BJATC-94394	Australia	a	b	a
3	RL 1359	India	a	b	a
4	Acc no. 404	Yugoslavia	a	a	b
5	Kranti	India	a	a	a
6	Pusa bold	India	a	a	a
7	RH 30	India	a	a	a
8	Varuna	India	a	a	a
9	Shi Yian Ku youCai	China	a	a	a
10	RC 781	India	a	a	a
Yellow-seeded					
11	Zem 1	Australia	a	b	b
12	Cutlass	Poland	a	b	b
13	Malopoloska	Russia	a	b	b
14	Skorospieka II	Russia	a	b	b
15	Heera	Canada	a	b	b
16	Chang Yang Huanzai 37	China	a	b	b
17	Kranodraskaja	Russia	a	b	b
18	Donskaja	Russia	a	b	b
19	Zem 84-500	Australia	a	b	b
20	DYJ 1	India	a	b	b
B. rapa brown-seeded					
21	Brown sarson	India	a	a	–
22	Pant Toria	India	a	a	–
23	BSH 1		a	a	–
25	Puas Kalyani	India	a	a	–
25	Chiffu	Korea	a	a	–
Yellow-seeded					
26	YID 1	India	b	a	–
27	Tetralocular	India	b	a	–
28	YSPB	India	b	a	–
29	DYS 1	India	b	a	–
30	R 500	India	b	a	–

a Wild-type allele of *TT8* controlling brown-seeded phenotype, *b* mutant allele of *TT8* controlling yellow-seeded phenotype

[a] Genotyping of *BrTT8* gene was done by using TL2–TR1 primers of Li et al. (2012)

into pGEM T Easy vector (http://www.promega.com) and sequenced on 3730 DNA analyzer.

Detailed schemes of cloning of genes undertaken in the study are shown in supplementary Figure S1 and Figure S2 and their genomic sequence in supplementary Figure S3 and Figure S4. Details of the primer sequences for gene amplification and genotyping are shown in supplementary Table S1 and Table S2.

RNA extraction, quantitative real-time PCR (qRT-PCR) and reverse transcriptase PCR (RT-PCR)

Seeds were harvested from three plants, 15 days after pollination (DAP) and were stored in liquid nitrogen. Total RNA was isolated using the Spectrum Plant RNA isolation kit (Sigma) following the manufacturer's instructions. The DNase treatment was performed on-column to remove any DNA contamination (RNase-free DNase, Qiagen). The cDNA synthesis was carried out from 2 μg of total RNA using the High Capacity cDNA reverse transcription kit (Applied Biosystems) in a volume of 20 μl, as per the manufacturer's protocol.

Endogenous *ubiquitin* gene was used as the internal control for qRT-PCR. Reactions were performed in a 96-well optical reaction plates with barcode, and monitored with a 7900HT Fast Real-Time PCR System (Applied Biosystem, USA). qRT-PCR for each sample was performed, in triplicate, in a total volume of 10 μl, consisting of 5 μl of 2× SYBR Green Power master mix (ABI, USA), 0.5 μl of 10 mM

primer mix, 2 µl diluted cDNA (1:20) and 2.5 µl of sterile water. The cycling parameters were as follows: one cycle at 50 °C for 2 min, one cycle at 95 °C for 10 min and 40 cycles of 95 °C for 15 s (denaturation) and 60 °C for 1 min (annealing and extension). Melting curve of each PCR amplicon was obtained by adding the following cycling condition: 95 °C for 15 s followed by a constant increase of the temperature between 60 and 95 °C at an increment of 0.3 °C/cycle. Relative gene expression analysis was carried out using comparative C_T value method (Schmittgen and Livak 2008).

RT-PCR was set up using 2 µl of (1:5) diluted cDNA in 25 µl reaction. The temperature and amplification conditions were: initial denaturation at 95 °C for 5 min, 28 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s and polymerization at 72 °C for 30 s with a final extension at 72 °C for 7 min. Amplified PCR products were analysed by gel electrophoresis on 1.0 % agarose gel. The amplification of actin transcript was carried out as an internal control. List of primers used for the transcript analysis is shown in supplementary Table S3.

Results

Expression level of flavonoid genes in brown and yellow-seeded *B. juncea*

The expression level of seven flavonoid genes consisting of three structural genes *F3H* (belonging to EBG group), *DFR* and *BAN* (belonging to LBG group) and four regulatory genes (*TT1*, *TT2*, *TT8* and *TTG1*) were analysed from developing seeds (15 DAP) in brown-seeded line Varuna and yellow-seeded line Heera. Majority of these genes are present as single copy with an exception of *F3H* and *TTG1* which are present in three and two copies, respectively, in diploid *B. rapa* [BRAD (http://brassicadb.org/brad)]. RNA-seq data of *B. juncea* for each of the seven genes available in the lab (unpublished) were aligned with the CDS sequences of genes downloaded from the BRAD database. A forward and reverse primer was designed for each gene from the consensus sequence (Supplementary Table S3).

qRT-PCR analysis revealed that among the structural genes, the amount of mRNA in the EBG gene F3H was similar between Varuna and Heera. Conversely, significant difference was observed in the transcript amount between Varuna and Heera for LBG genes, *DFR* and *BAN*. No transcript was detected in the yellow-seeded line Heera for these two genes (Fig. 1a). Of the four regulatory genes subjected to qRT-PCR analysis, no significant difference in transcript level between Varuna and Heera was observed for the genes, *TT1*, *TT2* and *TTG1*. On the other hand, the transcript level of the *TT8* gene was significantly lower in yellow-seeded Heera than that of brown-seeded Varuna (Fig. 1b).

Mapping of candidate genes and their association with seed coat colour in *B. juncea*

Padmaja et al. (2005) mapped the two loci of seed coat colour, *BjSc1* and *BjSc2* in the LG1 (now described as LG A9) and LG 2 (now described as LG B3), respectively. The genome sequence data of LG A9 of *B. rapa* in BRAD was scanned for the presence of candidate genes involved in the flavonoid pathway. Three candidate genes, *DFR* (Bra027457, Scaffold000056) (*At5g42800*), *TT1* (Bra028067, Scaffold000059) (*At1g34790*) and *TT8* (Bra037887, Scaffold000135) (*At4g09820*) were identified in the LG A9 of *B. rapa* (Table 2). A comparison with the closest mapped markers in the scaffold harbouring these candidate genes (Table 2) on the VH map (Fig. 2) revealed that two candidate genes *TT1* and *TT8* were very close (<1.0 cM) to the SSR marker, Na10-A08 (Scaffold 000040) that showed tight linkage with seed coat colour gene, *BjSc1* (Padmaja et al. 2005). On the other hand, the mapped marker At5g43430 located in the same scaffold

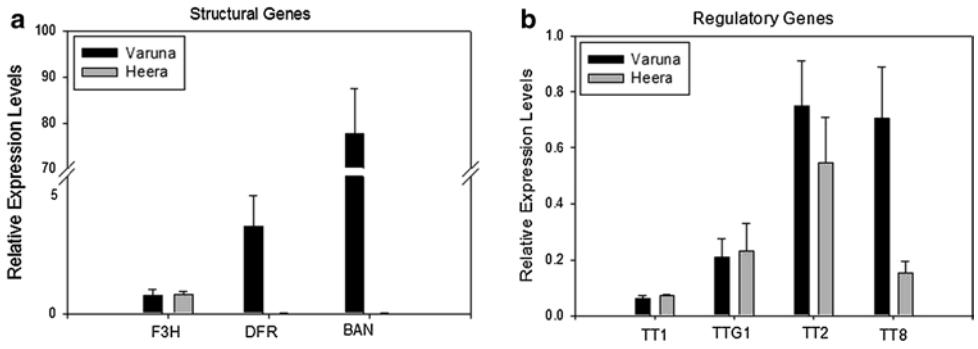

Fig. 1 Relative expression levels of flavonoid biosynthetic genes in developing seeds (15 DAP) of *B. juncea* cv. Varuna (brown seeded) and Heera (yellow seeded). The expression level of three structural genes one EBG (*F3H*), two LBGs (*DFR* and *BAN*) (**a**) and four regulatory genes (*TT1*, *TTG1*, *TT2* and *TT8*) (**b**) of flavonoid biosynthetic genes was detected by quantitative real-time PCR (QRT-PCR). Data were normalized to an internal control, *ubiquitin* gene

where *DFR* is located mapped around 4.0 cM away from the SSR marker Na10-A08 (Table 2; Fig. 2). Based on the above genomics and mapping information, cloning and mapping was undertaken for the two candidate genes, *DFR* and *TT8* showing difference in the transcript level between the brown-seeded Varuna and the yellow-seeded Heera.

Cloning and mapping of DFR gene and determining its association with seed coat colour

Primers for the amplification of *DFR* gene were synthesized based on the sequence information of *DFR* gene in *B. rapa* and *A. thaliana*. The gene(s) were initially isolated from the brown-seeded diploid progenitor species *B. rapa* cv Pusa Kalyani (AA) and *B. nigra* cv IC 257 (BB). Three different sequences were isolated from *B. rapa* cv Pusa Kalyani (designated as *BraA.DFR.r1*, *BraA.DFR.r2* and *BraA.DFR.r3*) and one from *B. nigra* cv IC 257 (designated as *BniB.DFR.n*). Based on the genome specificity and sequence divergence, *DFR* genes were amplified from *B. juncea* cvs Varuna and Heera. Cloning and sequencing of three *B. rapa*-specific genes revealed that *BraA.DFR.r1* was present in Varuna (designated as *BjuA.DFR.r1v*) and *BraA.DFR.r2* and *BraA.DFR.r3* were present

Table 2 BRAD search of genome sequence data of LG A9 of *B. rapa* for flavonoid pathway genes and the position of closest mapped markers on the VH map of *B. juncea* in the same scaffold

Scaffold no.	Flavonoid pathway gene	Mapped marker on the VH map of *B. juncea* in the scaffold and its position (cM)[a]
Scaffold000056	*DFR* (Bra027457, At5g42800)	At5g43430 (78.0)
Scaffold000059	*TT1* (Bra028067, At1g34790)	At1g34380 (82.2)
Scaffold000135	TT8 (Bra037887, At4g09820)	At4g09980 (82.5)
Scaffold000040	Linked SSR marker Na10-A08 to the trait (*BjSc1*)	(82.0)

[a] Mapped markers on the VH map are shown in Fig. 2

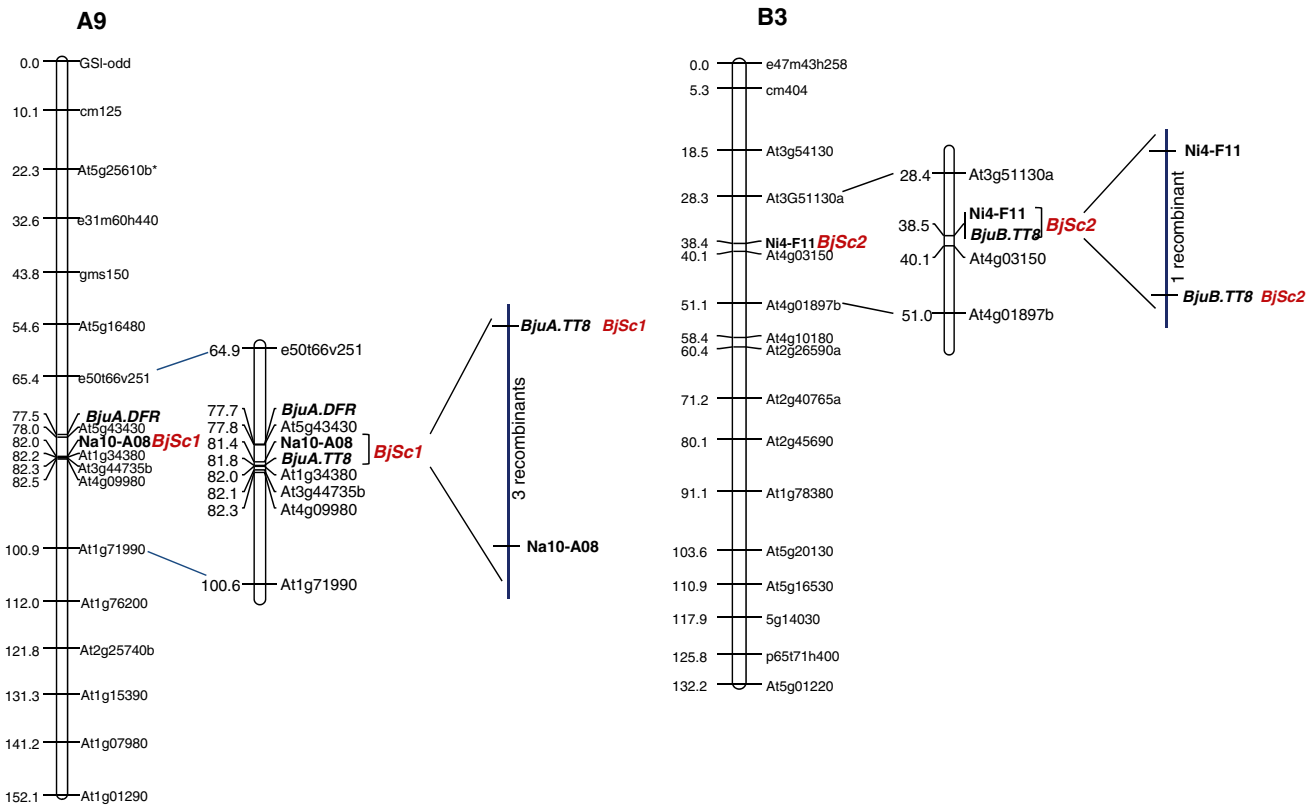

Fig. 2 Fine mapping of the *BjuA.TT8* in the linkage group A9 and the *BjuB.TT8* in the linkage group B3 and their co-segregation with seed coat colour loci, *BjSc1* and *BjSc2*, respectively. Representative gel picture of polymorphism between the two parents used for genotyping is shown in supplementary Figure S6

in Heera (designated as *BjuA.DFR.r2h* and *BjuA.DFR.r3h*, respectively) (supplementary Figure S1 and Figure S3). Genotyping of 123 DH lines of mapping population 1 of *B. juncea* for these three genes (genotyped as dominant markers using the primer pair DFR.r1(V)F and DFR.r1(V)R for *BjuA.DFR.r1.v*, DFR.r2(H)F and DFR.r2(H)R for *BjuA.DFR.r2.h* and DFR.r3(H)F and DFR.r3(H)R for *BjuA.DFR.r3.h*; Supplementary Table S1), interestingly, generated exactly same genotyping data and consequently, mapped to the same position of VH map in the LG A9. The locus was designated as *BjuA.DFR*. However, the map position of *BjuA.DFR* in the LG A9 was observed to be 4.5 cM away (six recombinants) from the microsatellite marker Na10-A08 that was shown to be tightly linked to the *BjSc1* in the earlier study of Padmaja et al. (2005) (Fig. 2). Association of *BjuA.DFR* with the trait variation was also verified by genotyping the mapping population 2 that segregates for *BjSc1* gene and recessive homozygous for the second (*bjsc2*/*bjsc2*) gene. Of 130 DH lines of mapping population 2 genotyped by *BjuA.DFR*, 14 recombinants were identified between *BjuA.DFR* and the *BjSc1* indicating no association of *BjuA.DFR* with the seed coat colour phenotype.

Sequencing of *B. nigra*-specific *BniB.DFR.n* gene from Varuna (designated as *BjuB.DFR.v*) and Heera (designated as *BjuB.DFR.h*) revealed the presence of a polymorphic microsatellite ($[AT]_6$ in Varuna and $[AT]_{24}$ in Heera) between the two alleles (supplementary Figure S1 and Figure S3). Using this polymorphic microsatellite, the *BjuB.DFR* was mapped to the LG B6 of VH map (map not shown). It confirmed that *DFR* gene is not involved in the genetic control of seed coat colour in *B. juncea*.

Cloning and mapping of TT8 gene and determining its association with seed coat colour

Isolation of the *TT8* gene from *B. juncea* cvs Varuna and Heera was done following the similar procedure that was followed for the *DFR* gene. One amplified product each was obtained from *B. rapa* cv Pusa Kalyani (~3,800 bp) and *B. nigra* cv IC257 (~2,800 bp) and upon cloning and sequencing, revealed the presence of a single sequence in both *B. rapa* and *B. nigra*, respectively (supplementary Figure S2 and Figure S4).

Genome-specific amplification, cloning and sequencing of A-genome-specific *TT8* gene from *B. juncea* (designated as *BjuA.TT8*) identified a sequence of 3,550 bp from Varuna (designated as *BjuA.TT8.v)* and 4,829 bp from Heera (designated as *BjuA.TT8.h)*. Alignment of *BjuA.TT8* gene from Varuna and Heera revealed the presence of an insertion of 1,279 bp in the exon 7 of the Heera allele (*BjuA.TT8.h*) (Fig. 3 and supplementary Figure S4 and Figure S5). Genotyping of mapping population 1 using this polymorphism mapped the *BjuA.TT8* (primer pair used was TT8-RsF1 and TT8-RsR4; supplementary Table S2) 0.4 cM away from the microsatellite marker Na10-A08 (one recombinant) on VH map (Fig. 2). However, the *BjuA.TT8* showed perfect co-segregation with *BjSc1* locus (phenotype) without any recombinant in the mapping population 2 as was earlier observed between Na10-A08 and *BjSc1* (Padmaja et al. 2005). Since Na10-A08 and *BjuA.TT8* both showed perfect co-segregation with the phenotype in the mapping population 2, fine mapping of the *BjSc1* region was undertaken by using mapping population 4 consisting of 377 brown and 117 yellow RIL segregants. The genotyping was confined to only yellow-seeded RILs as the brown-seeded RILs would be non-informative for the detection of recombinant where two genes encoding for brown seed coat are segregating. Genotyping of 117 yellow RILs was undertaken both by microsatellite marker Na10-A08 and *BjuA.TT8*. The microsatellite marker Na10-A08 did not show perfect co-segregation as three out of 117 yellows RIL segregants showed wild-type allele for the marker. On the other hand, perfect co-segregation was observed between the *BjuA.TT8* gene and the yellow seed coat phenotype as all the 117 yellow RIL segregants showed the presence of mutant allele for the *BjuA.TT8* gene (Fig. 2).

Cloning and sequencing of B-genome-specific *TT8* gene from *B. juncea* (designated as *BjuB.TT8*) identified an ORF of 2,768 bp from brown-seeded cv Varuna and 2,744 bp from yellow-seeded cv Heera. Sequence alignment between these two homologs revealed the presence of an SNP (C→T) in the last exon (exon 7) resulting in the creation of a stop codon (TAA) in Heera (Fig. 3 and supplementary Figure S4 and Figure S5). This nucleotide substitution also resulted in the creation of a *Mse I* site in the Heera allele of *BjuB.TT8* and was used as a CAPS marker in genotyping. Genotyping of the mapping population 1 (primer pair used was TT8-NsMPF and TT8-NsMPR; supplementary Table S2) mapped the *BjuB.TT8* gene along with the microsatellite marker Ni4-F11 and *BjSc2* in the LG B3 as no recombinant was identified between the *BjuB.TT8* and Ni4-F11 (Fig. 2). Perfect co-segregation was also observed between the *BjuB.TT8* and the *BjSc2* locus in the mapping population 3 as no recombinant was identified as was earlier observed between Ni4-F11 and *BjSc2* (Padmaja et al. 2005). However, fine mapping of the region by genotyping 117 yellow segregants of 494 RILs identified one recombinant between the Ni4-F11 and *BjSc2* whereas perfect co-segregation was observed between *BjuB.TT8* and *BjSc2* (phenotype) (Fig. 2).

These results of co-segregation between the two homoeologous *TT8* genes with the corresponding phenotypic locus (*BjSc1* in the LG A9 and *BjSc2* in the LG B3, respectively) suggested that the candidate gene was *TT8*. This fact

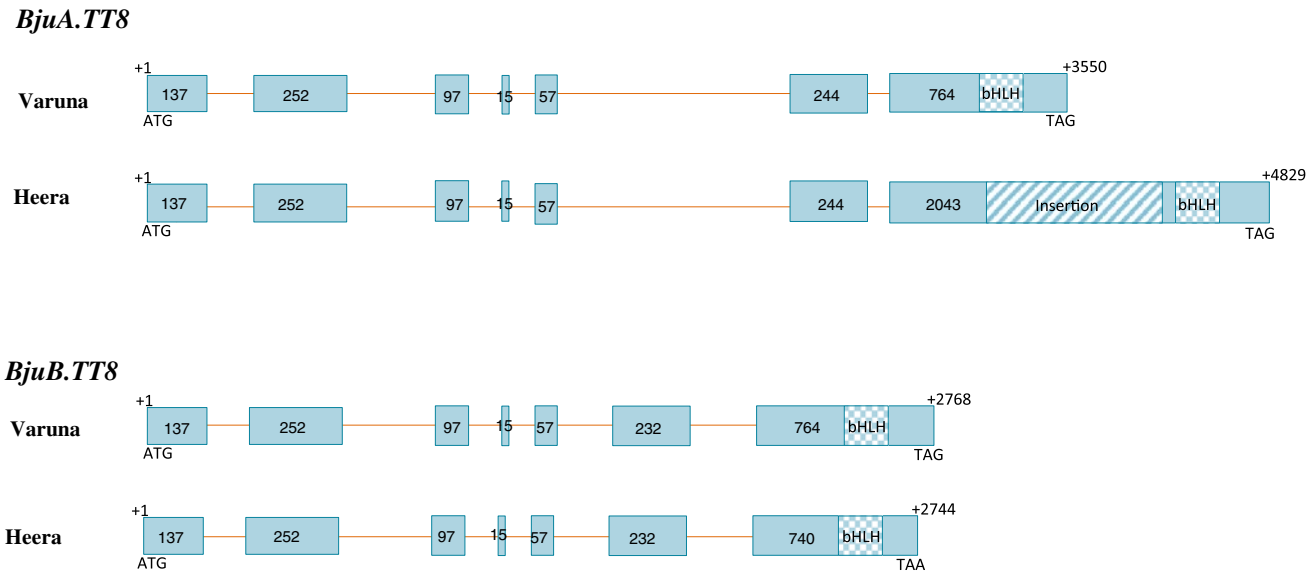

Fig. 3 Structural organization of the *BjuA.TT8* and *BjuB.TT8* from brown-seeded Varuna and yellow-seeded Heera of *B. juncea*. *Rectangles* represent the exons and *numbers in the rectangles* represent the size of the exons in bp. The location of the insertion in the *BjuA.TT8* is shown in a *rectangle* marked as '*insertion*'. The location of the basic helix–loop–helix region is shown as 'bHLH' in the exon 7. Details of the sequence and structural organization of both *BjuA.TT8* and *BjuB.TT8* is shown in supplementary Fig S4 and supplementary Fig S5, respectively

was further validated by genotyping 20 natural *B. juncea* germplasm consisting of both brown and yellow-seeded lines (Table 1). Of the ten brown-seeded germplasm genotyped for *BjuA.TT8* and *BjuB.TT8*, seven lines showed the presence wild-type alleles for both the TT8 genes and three lines showed the presence of at least one wild-type allele for either *BjuA.TT8* or *BjuB*.TT8 gene. On the contrary, all the yellow seed lines showed the presence of mutant alleles for both the *TT8* genes (Table 1).

Expression of two homoeologous *TT8* genes in *B. juncea*

We determined the level of mRNA expression of both *BjuA.TT8* and *BjuB.TT8* genes separately from the developing seeds of brown-seeded Varuna and yellow-seeded Heera. Genome-specific primers for these two genes were designed for the cDNA amplification. RT-PCR was performed to study the level of mRNA expression. While there was no transcript detected for the *BjuA.TT8* gene, little amount of transcript was detected for the *BjuB.TT8* gene in the yellow-seeded Heera. In brown-seeded Varuna, normal amount of transcript was detected for both the genes (Fig. 4). From this study it is apparent that the significantly lower level of transcript observed in qRT-PCR of yellow seed line Heera (Fig. 1) could primarily be due to transcription of *BjuB.TT8* gene. These expression studies (qRT-PCR and RT-PCR) largely indicated that these two mutations in the two homoeologous *TT8* genes have disturbed the normal transcription of *TT8* genes in the yellow-seeded line.

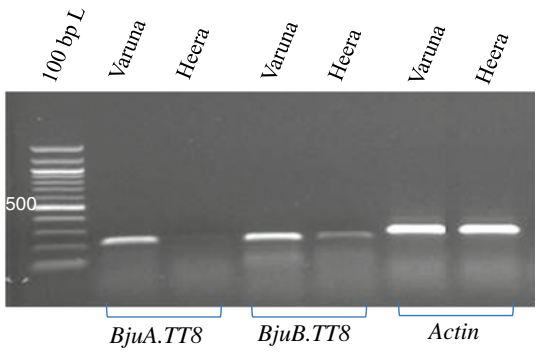

Fig. 4 Comparative transcript levels of *BjuA.TT8* and *BjuB.TT8* in Varuna and Heera as revealed by RT-PCR. *Actin* transcript was used as an internal control

Discussion

Prior mapping information of seed coat colour trait in *B. juncea* (Padmaja et al. 2005) that placed one of the two loci (*BjSc1*) to the LG A9 helped us to search the genome sequence data of LG A9 of *B. rapa* in BRAD (http://brassicadb.org/brad) for the presence of putative candidate genes involved in PA pathway. We identified two regulatory genes, *TT1* and *TT8* and one structural gene, *DFR* in the LG A9 from the BRAD *B. rapa* sequence. This genomics information along with the qRT-PCR data that showed complete absence of transcripts in LBGs (*DFR* and *BAN*) and significant reduction of transcript of *TT8* gene in

the yellow-seeded line suggested the involvement of *TT8* in the genetic variation of seed colour trait. The involvement of *TT8* in regulating the expression of LBGs in PA pathway has been shown in *Arabidopsis*. It affects the transcription of both *DFR* and *BAN* (Nesi et al. 2000), possibly through the formation of a ternary complex involving the products of other two regulatory genes *TT2* and *TTG1* (Baudry et al. 2004; Hichri et al. 2011). Similar observation of involvement of *TT8* gene affecting the transcript level of LBGs has been observed in yellow-seeded line of *B. rapa* (Li et al. 2012).

Of the two regulatory genes identified in the LG A9, the involvement of *TT1* in regulating the expression of LBGs was ruled out. We observed similar level of transcript in both brown and yellow seed lines (Fig. 1). In an earlier study in *B. juncea*, Yan et al. (2010) observed similar expression level of *BjTT1* in both black and yellow-seeded line by RT-PCR and based on the several SNPs between the brown and yellow allele of the *BjTT1* gene, they concluded that *TT1* was not the crucial gene for the seed coat colour. In *Arabidopsis* it has been shown that there is accumulation of *BAN* mRNA in the young siliques of wild type and *tt1-1* mutant plants. The study also shows that *TT1* is necessary for *BAN* promoter activity only in small group of cells at the chalazal boundary of endothelium (Sagasser et al. 2002).

Mapping of two *DFR* genes, the *BjuA.DFR* to the LG A9 4.5 cm away from the *BjSc1* and the *BjuB.DFR* to the LG B6 gave further credibility to the *TT8* as a probable candidate gene for seed coat colour variation in *B. juncea*. In allotetraploid *B. juncea* where two duplicate genes control the seed coat colour, it has to be the same candidate gene of PA pathway at both the loci otherwise no yellow-seeded mutant could arise due to complementation. This fact was established through mapping of the two homoeologous *TT8* genes wherein the *BjuA.TT8* mapped to LG A9 and the *BjuB.TT8* to LG B3 and showed perfect co-segregation with *BjSc1* and *BjSc2* locus, respectively (Fig. 2). This validation was agreed to by using four mapping populations and trait contrasting natural *B. juncea* germplasm (Table 1).

The mutation in the *BjuA.TT8* of yellow-seeded line was due to presence of a 1,279-bp insertion in the exon 7 of the gene between the amino acids 354 and 355 that is 7 amino acids prior to the commencement of the basic helix–loop–helix structure of the *TT8* protein. This insertion disturbed the transcription of *BjuA.TT8* gene as no transcript (very little) was seen for this gene in RT-PCR (Fig. 4). Disruption of transcription due to an insertion in the *TT8* gene that results in yellow seed coat colour was recently reported in *B. rapa* (Li et al. 2012). In *A. thaliana*, no mRNA expression was detected in the mutant *tt8-3* that was caused due to insertion in the second intron of the *TT8* gene (Nesi et al. 2000). However, in the present study in *B. juncea*, the insertion was in the exon 7 of the *BjuA.TT8*.

The mutation in the *BjuB.TT8* of the yellow-seeded *B. juncea* was a single C-to-T base transition in the C-terminal region that introduced a termination codon TAA in place of CAA coding for glutamine. The predicted protein would be truncated, lacking the C-terminal 8 amino acid residues. We also detected some transcript, though significantly lower than the transcript of the wild type, for this mutant gene (Fig. 4). Detection of transcript for the mutants coding for truncated proteins has been observed for *TTG1* gene in *A. thaliana*. The mutant *ttg1-1* responsible for glabra and transparent testa phenotype in *A. thaliana* is due to a transition mutation creating a stop codon resulting in a truncated protein product lacking C-terminal 25 amino acid residues. Although synthesis of transcript has been reported for the *ttg1-1* mutant (Walker et al. 1999), the expression of *DFR* and *BAN* could not be detected in the siliques of *ttg1-1* mutant (Nesi et al. 2000). It has been shown that mutations in any of these three regulatory genes (*TT2*, *TT8* and *TTG1*) affect the transcription of LBGs (Nesi et al. 2000) and produce seeds lacking PAs as they act in concert to regulate the LBGs expression (Baudry et al. 2004). Hence, in the present study in *B. juncea* the mutations in both the *TT8* genes, *BjuA.TT8* in LG A9 and *BjuB.TT8* in LG B3, are required for inhibiting the LBGs expression at the transcription level in the yellow-seeded mutant. Furthermore, the polymorphisms detected in the *BjuA.TT8* and *BjuB.TT8* genes could be used as more reliable markers than the earlier reported linked SSR markers (Padmaja et al. 2005) for the marker-assisted foreground selection of the yellow-seeded trait in *B. juncea*.

Among the *Brassica* species, candidate genes responsible for the seed coat colour variation have been reported in diploid *B. rapa* and mutations in the regulatory genes of PA pathway have been implicated. The yellow-seeded trait in Chinese cabbage is due to a deletion in the *TTG1* (Zhang et al. 2009) and in yellow sarson it is due to an insertion of transposable element in the *TT8* gene (Li et al. 2012). In *B. juncea*, we also observed involvement of *TT8* at two loci containing different types of mutations than the one reported for the yellow sarson. Yellow sarson of *B. rapa* and *B. juncea* are grown in India as oilseed crop. We undertook the screening of yellow-seeded *B. rapa* and *B. juncea* germplasm for the presence of mutant allele of yellow sarson *TT8* gene. Genotyping data revealed that all the yellow-seeded *B. rapa* germplasm and none of the yellow-seeded *B. juncea* possessed *B. rapa* yellow allele (Table 1; supplementary Figure S7). This indicates that yellow seededness in these two *Brassica* species have evolved independently and diversification of the trait has not crossed the species boundary.

Acknowledgments This work was supported by the National Dairy Development Board (NDDB) and the Department of Biotechnology (DBT) under the Centre of Excellence programme on genome mapping and molecular breeding of Brassicas.

References

Ahmed SU, Zuberi MI (1971) Inheritance of seed coat color in *Brassica campestris* L. variety Toria. Crop Sci 11:30

Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant J 39:366–380

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V (2011) Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. J Exp Bot 62(8):2465–2483

Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57:405–430

Li X, Chen L, Hong M, Zhang Y, Zu F, Wen J, Yi B, Ma C, Shen J, Tu J, Fu T (2012) A large insertion in bHLH transcription factor BrTT8 resulting in yellow seed coat in *Brassica rapa*. Plos one 7:e44145

Lionneton E, Aubert G, Ochatt S, Merah O (2004) Genetic analysis of agronomic and quality traits in mustard (*Brassica juncea*). Theor Appl Genet 109:792–799

Mahmood T, Rahman MH, Stringam GR, Raney JP, Good AG (2005) Molecular markers for seed colour in *Brassica juncea*. Genome 48:755–760

Negi MS, Devic M, Delseny M, Lakshmikumaran M (2000) Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. Theor Appl Genet 101:146–152

Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L (2000) The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in Arabidopsis siliques. Plant Cell 12:1863–1878

Padmaja KL, Arumugam N, Gupta V, Mukhopadhyay A, Sodhi YS, Pental D, Pradhan AK (2005) Mapping and tagging of seed coat colour and the identification of microsatellite markers for marker-assisted manipulation of the trait in *Brassica juncea*. Theor Appl Genet 111:8–14

Panjabi P, Jagannath A, Bisht NC, Padmaja KL, Sharma S, Gupta V, Pradhan AK, Pental D (2008) Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using intron polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. BMC Genomics 9:113

Pradhan AK, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D (2003) A high density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. Theor Appl Genet 106:607–614

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 major oilseed crop of dryland areas. Theor Appl Genet 115:807–817

Rehman M, McVetty PBE (2011) A review of *Brassica* seed colour. Can J Plant Sci 91:437–446

Rogers SO, Bendich AJ (1994) Extraction of total cellular DNA from plants, algae and fungi. In: Gelvin SB, Schilperoot RA (eds) Plant molecular biology manual. Kluwer, Dordrecht, pp 1–8

Sabarwal V, Negi MS, Banga SS, Lakshmikumaran M (2004) Mapping of AFLP markers linked to seed coat colour loci in *Brassica juncea* (L.) Czern. Theor Appl Genet 109:160–166

Sagasser M, Lu GH, Hahlbrock K, Weisshaar B (2002) *A. thaliana TRANSPARENT TESTA 1* is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. Genes Dev 16:138–149

Schmittgen TD, Livak K (2008) Analyzing real-time PCR by the comparative C_T method. Nat Protoc 3:1101–1108

Slominski BA (1997) Development in the breeding of low fibre rapeseed/canola. J Anim Feed Sci 6:303–317

Vera CL, Woods DL (1982) Isolation of independent gene pairs at two loci for seed coat colour in *Brassica juncea*. Can J Plant Sci 62:47–50

Vera CL, Woods DL, Downey RK (1979) Inheritance of seed coat colour in *Brassica juncea*. Can J Plant Sci 59:635–637

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC (1999) The *TRANSPARENTTESTAGLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell 11:1337–1350

Yan M, Liu Z, Guan C, Chen S, Yuan M, Liu X (2009) Inheritance and molecular markers for the seed coat color in *Brassica juncea*. Front Agric China 3:1–6

Yan ML, Liu XJ, Guan CY, Liu LL, Lu Y, Liu ZS (2010) Cloning and SNP analysis of *TT1* gene in *Brassica juncea*. Acta Agron Sin 36:1634–1641

Zhang JF, Ying L, Yuan YX, Zhang XW, Geng JF et al (2009) Map-based cloning and characterization of a gene controlling hairiness and seed coat color traits in *Brassica rapa*. Plant Mol Biol 69:53–63