Theor Appl Genet (2009) 118:413–421
DOI 10.1007/s00122-008-0907-z

ORIGINAL PAPER

Fine mapping of loci involved with glucosinolate biosynthesis in oilseed mustard (*Brassica juncea*) using genomic information from allied species

N. C. Bisht · V. Gupta · N. Ramchiary · Y. S. Sodhi · A. Mukhopadhyay · N. Arumugam · D. Pental · A. K. Pradhan

Received: 8 July 2008 / Accepted: 27 September 2008 / Published online: 21 October 2008

Abstract Fine mapping of six seed glucosinolate QTL (*J2Gsl1*, *J3Gsl2*, *J9Gsl3*, *J16Gsl4*, *J17Gsl5* and *J3Gsl6*) (Ramchiary et al. in Theor Appl Genet 116:77–85, 2007a) was undertaken by the candidate gene approach. Based on the DNA sequences from *Arabidopsis* and *Brassica oleracea* for the different genes involved in the aliphatic glucosinolate biosynthesis, candidate genes were amplified and sequenced from high to low glucosinolate *Brassica juncea* lines Varuna and Heera, respectively. Of the 20 paralogues identified, 17 paralogues belonging to six gene families were mapped to 12 of the 18 linkage groups of *B. juncea* genome. Co-mapping of candidate genes with glucosinolate QTL revealed that the candidate gene *BjuA.GSL-ELONG.a* mapped to the QTL interval of *J2Gsl1*, *BjuA.GSL-ELONG.c*, *BjuA.GSL-ELONG.d* and *BjuA.Myb28.a* mapped to the QTL interval of *J3Gsl2*, *BjuA.GSL-ALK.a* mapped to the QTL interval of *J3Gsl6* and *BjuB.Myb28.a* mapped to the QTL interval of *J17Gsl5*. The QTL *J9Gsl3* and *J16Gsl4* did not correspond to any of the mapped candidate genes. The functionality and contribution of different candidate genes/QTL was assessed by allelic variation study using phenotypic data of 785 BC_4DH lines. It was observed that *BjuA.Myb28.a* and *J9Gsl3* contributed significantly to the base level glucosinolate production while *J16Gsl4*, probably *GSL-PRO*, *BjuA.GSL-ELONG.a* and *BjuA.GSL-ELONG.c* contributed to the C3, C4 and C5 elongation pathways, respectively. Three A genome QTL: *J2Gsl1*harbouring *BjuA.GSL-ELONG.a, J3Gsl2* harbouring both *BjuA.GSL-ELONG.c* and *BjuA.Myb28.a* and *J9Gsl3*, possibly the 'Bronowski genes', were identified as most important loci for breeding low glucosinolate *B. juncea*. We observed two-step genetic control of seed glucosinolate in *B. juncea* mainly effected by these three A genome QTL. This study, therefore, provides clues to the genetic mechanism of 'Bronowski genes' controlling the glucosinolate trait and also provides efficient markers for marker-assisted introgression of low glucosinolate trait in *B. juncea*.

Communicated by C. F. Quiros.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0907-z) contains supplementary material, which is available to authorized users.

N. C. Bisht · V. Gupta · Y. S. Sodhi · A. Mukhopadhyay · N. Arumugam · D. Pental · A. K. Pradhan
Centre for Genetic Manipulation of Crop Plants, University of Delhi South Campus, Benito Juarez Road, New Delhi 110 021, India

N. Ramchiary · D. Pental · A. K. Pradhan (✉)
Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110 021, India
e-mail: pradhanakshay@hotmail.com

Present Address:
N. C. Bisht
National Institute of Plant Genome Research, JNU Campus, Aruna Asaf Ali Marg, New Delhi 110067, India

N. Ramchiary
Department of Horticulture, Chungnam National Universiy, 220 Gung-dong, Yuseong-gu, Daejeon 305-764, South Korea

N. Arumugam
Department of Biotechnology, Pondicherry University, Pondicherry 605 014, India

Introduction

Improvement of quality of the meal through the development of low glucosinolate lines (<18 μmol/g of seed) has

been a major objective in the breeding of oilseed rape (*Brassica napus*) and mustard (*Brassica juncea*). Low glucosinolate lines of *B. napus* have been successfully cultivated in Europe and Canada for more than 20 years. Although low glucosinolate lines have been reported in the East European type germplasm of *B. juncea* (Potts et al. 1999), no productive and commercially viable low glucosinolate lines have yet been reported in the Indian type germplasm of *B. juncea*. Previous works have shown that the aliphatic glucosinolates (major types of glucosinolates present in *Brassica* species) profile of east European type mustard lines is less complex than that of the Indian types as the former contain mostly 2-propenyl glucosinolate. In comparison, the Indian types contain both 2-propenyl and 3-butenyl glucosinolates (Gland et al. 1981; Love et al. 1990a, b; Sodhi et al. 2002). The difference in the glucosinolate profile is also reflected at the genetic level; aliphatic glucosinolates in the east European lines have been shown to be controlled by two loci (Love et al. 1990b) while the cross involving low glucosinolate east European and high glucosinolate Indian type lines revealed the contribution of six to eight loci to the glucosinolate content (Stringam and Thiagarajah 1995; Sodhi et al. 2002).

QTL mapping of aliphatic glucosinolate loci has been carried out in *B. juncea* using different F_1DH mapping populations (Cheung et al. 1998; Mahmood et al. 2003; Lionneton et al. 2004; Ramchiary et al. 2007a); each study reporting variable number of QTL. However, the study of Ramchiary et al. (2007a) provided an insight into the complexity of the trait and the importance of the epistatic interactions and context by comparing the QTL data from early (F_1DH) and advanced backcross (BC_4DH) generations. The most significant observation was the disappearance of a major QTL detected in the F_1DH and appearance of a new major QTL in the BC_4DH generation. Another significant observation of the study was the presence of negative linkage between QTL alleles of low glucosinolate and yield related QTL in a few linkage groups (Ramchiary et al. 2007a, b).

Extensive analysis of glucosinolate pathways in the model species, *Arabidopsis thaliana* (Kliebenstein et al. 2001) and *Brassica* species like *B. oleracea* (Li and Quiros 2002) and *B. napus* (Magrath et al. 1994) has led to a basic scheme of the metabolic pathway for glucosinolate biosynthesis (Fig. 1). Aliphatic glucosinolates are controlled by two different sets of genes, one set controlling the side chain elongation (*GSL-ELONG* and *GSL-PRO*) and another set controlling the modification of side-chain carbons (*GSL-OXID*, *GSL-ALK* and *GSL-OH)* (reviewed by Wittstock and Halkier 2002; Halkier and Gershenzon 2006). In a recent study in *A. thaliana*, Hirai et al. (2007) have shown that the basal-level of aliphatic glucosinolate is also controlled by R2R3 transcripton factors, *Myb28* and *Myb29*.

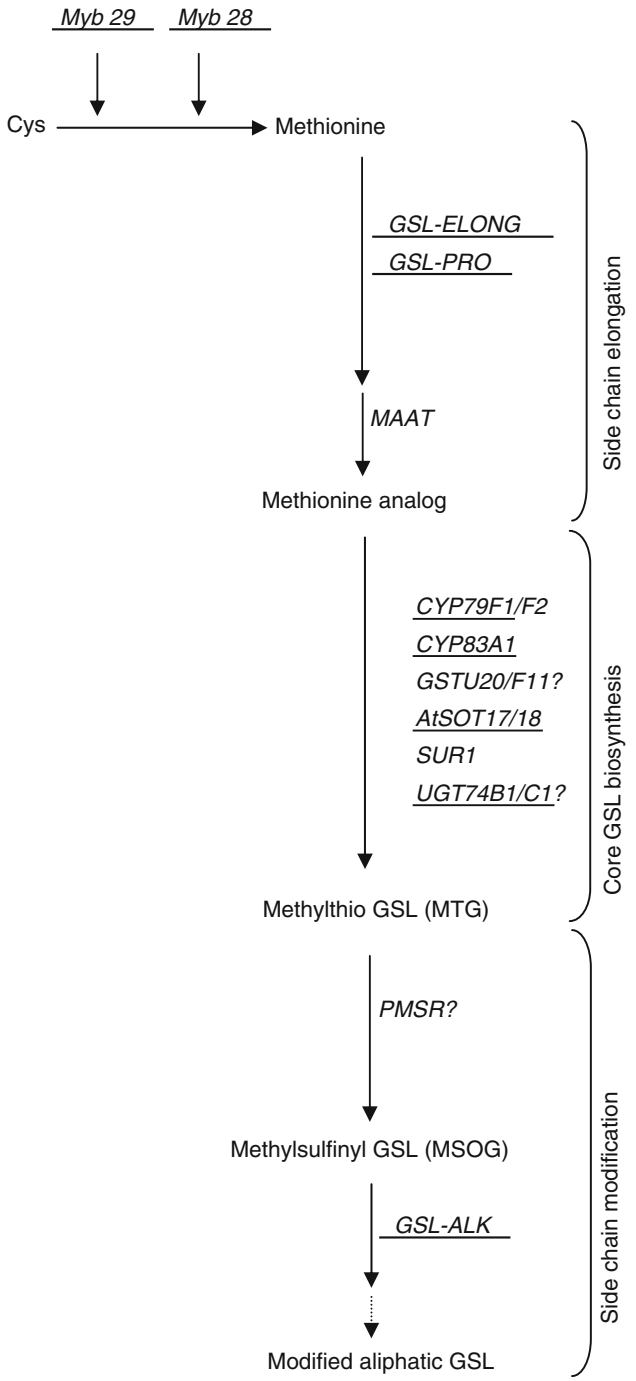

Fig. 1 Aliphatic glucosinolate biosynthetic pathway in *Arabidopsis* and *Brassica* (Hirai et al. 2007). Genes isolated and analyzed in the present study have been *underlined*

In the present work, we report fine mapping of loci involved with the low glucosinolate trait in *B. juncea* through the use of candidate gene-specific markers. These markers were developed using the sequence information available from comparative genomics studies of *Arabidopsis* and *Brassica* species. Fine dissection of the loci through the identification of candidate genes involved with seed

glucosinolate content has allowed the study of genetic mechanisms underlying this complex trait in *B. juncea.* This has also resulted in the development of accurate and easy-to-use markers for the diversification of low glucosinolate trait in *B. juncea* lines belonging to the Indian gene pool. We show that manipulation of three gene specific loci and two anonymous loci can result in *B. juncea* lines with as low as 4 µmol of glucosinolates per gram of seed.

Materials and methods

Plant materials and mapping populations

Gene isolation was carried out from two contrasting lines of *B. juncea* (AABB) viz., Heera (low glucosinolate) and Varuna (high glucosinolate) as well as from the diploid progenitor species of *B. juncea* namely, *B. rapa* (AA) cv. YID1 and *B. nigra* (BB) line IC257.

Two populations were used for mapping the genes. Population 1 consisted of 123 DH lines derived from the F_1 of a cross between a high glucosinolate variety, Varuna (glucosinolate content ~110 µmol/g of seed) and a low glucosinolate line, Heera (glucosinolate content ~12 µmol/g seed). This population has been used earlier for the development of a high-density linkage map in *B. juncea* (Pradhan et al. 2003; Ramchiary et al. 2007b), QTL mapping of glucosinolate content (Ramchiary et al. 2007a) and other agronomically important traits (Ramchiary et al. 2007b) and more recently, in the development of a comparative map of *B. juncea* and *Arabidopsis* (Panjabi et al. 2008). Population 2 consisted of 1160 BC_4DH lines derived from a recurrent selection backcross (RSB) scheme for the transfer of low glucosinolates trait from Heera to Varuna. Details of the RSB breeding scheme and the field experimental design for growing the populations have been described earlier (Ramchiary et al. 2007a).

Cloning, sequencing and mapping of candidate genes

DNA was isolated from the leaves of parental lines and individuals of the mapping population following the protocol of Rogers and Bendich (1994). DNA sequences of candidate genes involved with the aliphatic glucosinolate biosynthesis pathway (Fig. 1) in *Arabidobsis* and *Brassica* species were aligned and used to design primers for the isolation of corresponding genes from *B. juncea*. Primers were designed from regions that were highly conserved and flanked the variable regions of the genes. The list of starting primers used for the amplification of various candidate genes is given in Table 1. Isolation of the candidate genes was carried out in two steps. In the first step various homologs of some of the candidate genes viz., *GSL-ELONG*, *GSL-PRO* and *GSL-ALK* were amplified from the two diploid progenitor parents viz., *B. rapa* (AA) and *B. nigra* (BB) and sequenced. In the second step, primers specific to the A and the B genome were used to amplify candidate genes from *B. juncea* lines Varuna and Heera. Polymorphism between Heera and Varuna for specific genes was established either through length polymorphism or by SNPs in the coding region or the 5′ or 3′ flanks of the gene. Wherever polymorphism was not observed in the structural part of the gene, genome walking was performed to find polymorphism in the 5′ and 3′ regions of the gene using the PCR genomic libraries developed in *B. juncea* cvs. Heera and Varuna, using Genome Walker Kit (Clontech). Allele-specific primers for the two parents were subsequently used for mapping the candidate genes.

All PCR amplifications were carried out using Titanium *Taq* DNA polymerase (Clontech) with amplification cycle of 95°C (30 s), 68–72°C (2–4 min) for 35 cycles. Amplified fragments were cloned in the pGEMTEasy TA cloning vector (Promega). To exclude the possibilities of PCR based mutations giving false polymorphism, cloning was done from three independent PCR amplifications. Sequence matching in at least two of the three independent amplifications was used as the correct sequence of the target gene or its 5′ or 3′ regions. DNA sequencing of the clones was performed using the ABI Prism 310 Genetic Analyzer (Applied Biosystems). Sequence assembly and analysis was done using the DNASTAR (Madison, WI) software package.

Candidate genes showing polymorphism (length and/or SNP) between Heera and Varuna were mapped to a recently published map of *B. juncea* based on intron-polymorphism (IP) markers (Panjabi et al. 2008) using the software JoinMap version 2.0 (Stam 1993; Stam and Van Ooijen 1996). The map described by Panjabi et al. (2008) is an extension of the map of Ramchiary et al. (2007a, b). Classification of linkage groups follows the nomenclature proposed by Panjabi et al. (2008) and gene nomenclature is that followed by Østergaard and King (2008).

Glucosinolate analysis

Seed glucosinolate profiles of parents and different mapping populations were determined by high pressure liquid chromatography (Shimadzu HPLC model LC-10) following the protocol of Kraling et al. (1990). The percentage of propyl, butyl and pentyl were calculated following Magrath et al. (1993) and propenyl, butenyl and pentenyl glucosinolates are expressed as µmol/g of seed.

Genotype–phenotype association study

Genotyping was carried out by the candidate gene markers for the co-localized QTL (where a candidate gene(s)

Table 1 List of starting primers and their sequence used to amplify various candidate genes of the glucosinolate biosynthesis pathway in *B. rapa* (*AA*), *B. nigra* (*BB*) and *B. juncea* (*AABB*)

S. no.	Candidate gene	*Arabidopsis* homolog(s)	Primers used
1	*GSL-ELONG*	At5g23010, At5g23020	F 5′-tctstmcgcctgammcgtccgtacaaca-3′
			R 5′-gccatcttcgcamccaaaytsgatgtc-3′
2	*GSL-ALK*	At4g03050, At4g03060	F 5′-ygstttygaggtgtcatatgatagagtg-3′
			R 5′-ctcagtgtgatagaaactcatcaagtca-3′
3	*GSL-PRO*	At1g18500, At1g74040	F 5′-gactacatccccaaymgnathwsnga-3′
			R 5′-aagcttccccaaaacdarnccngcrtc-3′
4	*Myb28*	At5g61420	F 5′-aaggggcatggaccaccga-3′
			R 5′-tatcctcttcattgmsaatctgctcag-3′
5	*Myb29*	At5g07690	F 5′-ggagggtggcgtgacattc-3′
			R 5′-gcttgttgatgagacatcagaca-3′
6	*CYP83A1*	At4g13770	F 5′-gagcttctcaagacgcaaga-3′
			R 5′-tcgtaaccagcgatcttggt-3′
7	*CYP79F1*	At1g16410	F 5′-tcatgaagatgaaaagagtgatca-3′
			R 5′-acggcctacatgaatgtgg-3′
8	*CYP79C2*	At1g58260	F 5′-gaggagattagrgctcaatgcaag-3′
			R 5′-cccatatcttaggrttccgaccaag-3′
9	*AtSOT17*/19 (*sulphotransferase*)	At1g18590, At1g74090	F 5′-aacgacttcttcgtctgyagctacc-3′
			R 5′-aatacgagccaccatctccg-3′
10	*UDP-glycosyltransferase*	At1g24100	F 5′-cctactccgaatccttcaagctc-3′
			R 5′-ctacctcagcgagttgcttctcaa-3′
11	*ESP* (epithio specifier proteins)	At1g54040	F 5′-tgcaaggcgagtggatcaag-3′
			R 5′-aggtccctctcctcctttctgc-3′
12	*TGG*	At5g25980, At5g26000	F 5′-aaagacttcatcttcggtgttgcatc-3′
			R 5′-tttgcgakgagkscatctatgag-3′

F forward primer, *R* reverse primer

mapped to the genetic interval of a QTL) and the flanking anonymous markers for the other QTL. The average allelic effect of individual locus was estimated from the mean phenotypic value and the test of significance ($P < 0.05$) was done by *t* test.

Results

Isolation and mapping of candidate genes involved with glucosinolate biosynthesis in *B. juncea*

GSL-ELONG gene family

PCR amplification of *GSL-ELONG* identified four different paralogues in *B. rapa* and two in *B. nigra*. Based on the sequence data of the *GSL-ELONG* paralogues from the two progenitor genomes, four A genome paralogues designated as *BjuA.GSL-ELONG.a*, *b*, *c* and *d* and two B genome paralogues designated as *BjuB.GSL-ELONG.a* and *b* were isolated from *B. juncea* cvs. Varuna and Heera. In the A genome specific genes, length polymorphism was detected in *BjuA.GSL-ELONG.a*, *c* and *d and* a SNP in the *BjuA.GSL-ELONG.b* between Heera and Varuna [Table 1S of Electronic supplementary material (ESM)]. The B genome specific paralogues did not show polymorphism in the transcribed regions of the two genes. Genome-walking in the 3′ flank revealed SNP in both *BjuB.GSL-ELONG.a* and *BjuB.GSL-ELONG.b* (Table 1S of ESM). These polymorphisms were used to map all the six paralogues of the *GSL-ELONG* family in *B. juncea*.

While the two A genome specific genes *BjuA.GSL-ELONG.a* and *b* mapped to linkage groups A2 and A4, respectively, the remaining two A genome specific genes *BjuA.GSL-ELONG.c* and *d* mapped to the same position in the linkage group A3 (Table 1S of ESM; Fig. 2). *BjuA.GSL-ELONG.d* from both Heera and Varuna was observed to contain a stop codon in the exon 7 of the gene suggesting that the gene is encoding a truncated protein which in all probability is nonfunctional. The two B genome specific genes *BjuB.GSL-ELONG.a* and *b* mapped to linkage groups B6 and B8, respectively (Table 1S of ESM and Fig. 2).

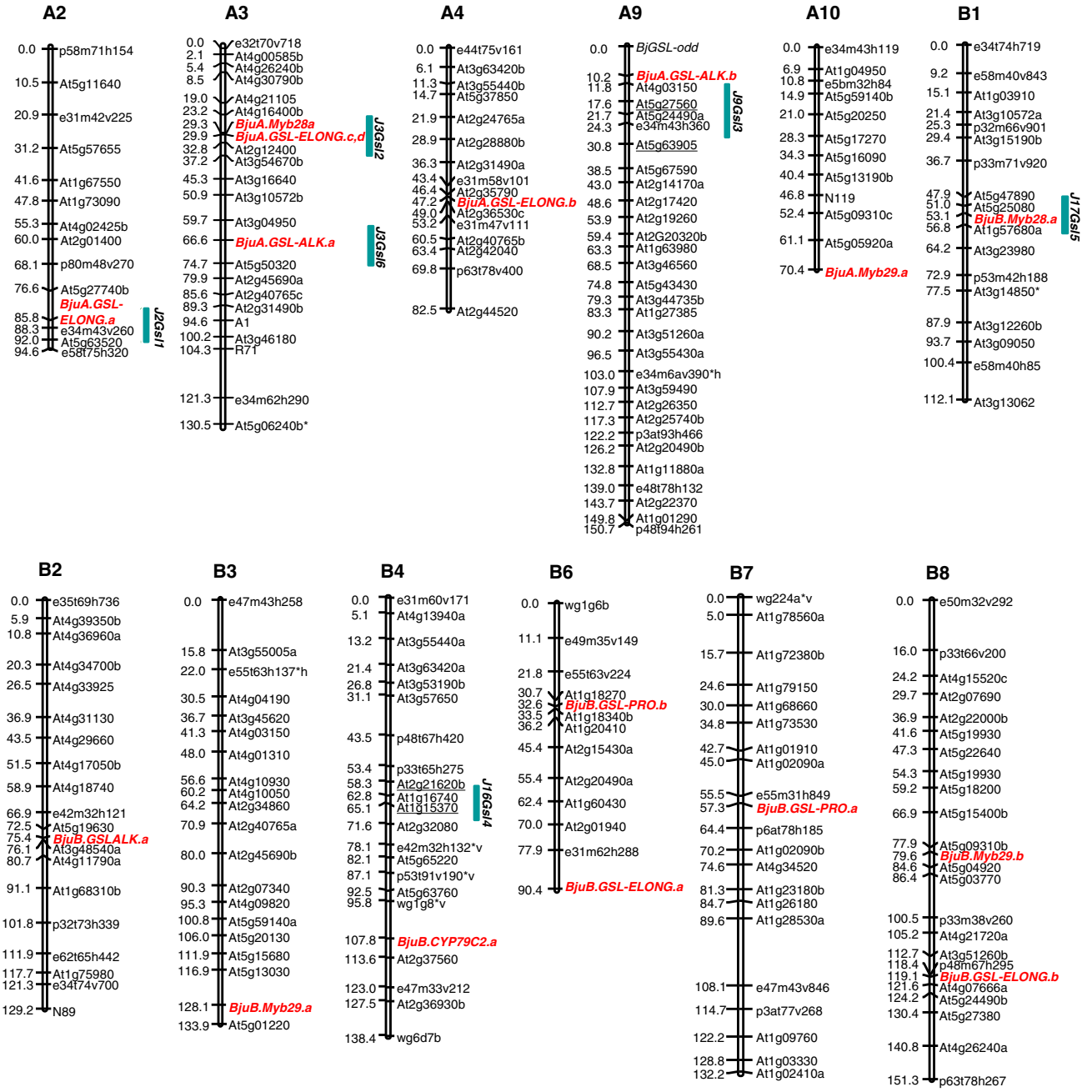

Fig. 2 A framework map of *B. juncea* based on the study of Panjabi et al. (2008) showing positions of 17 paralogues of different glucosinolate biosynthesis pathway genes (in *bold italics*) along with the six QTL for various glucosinolates as reported by Ramchiary et al. (2007a). Gene nomenclature is according to Østergaard and King (2008)

GSL-PRO gene family

The primers for *GSL-PRO* amplified a total of four different paralogues, two each from *B. rapa* and *B. nigra*. Sequencing of these paralogues from *B. juncea* cvs Varuna and Heera revealed SNP between alleles of the two B genome specific paralogues designated as *BjuB.GSL-PRO.a* and *b* and were mapped to linkage groups B7 and B6, respectively, using CAPS markers (Table 1S of ESM; Fig. 2). No polymorphism was observed between the alleles of the two A genome specific paralogues (*BjuA.GSL-PRO.a* and *b*) (Table 1S of ESM).

GSL-ALK gene family

Degenerate primers for *GSL-ALK* amplified two paralogues each from *B. rapa* and *B. nigra*. On the basis of the sequence information from the two progenitor species, all

the four full length genes (*BjuA.GSL-ALK.a* and *b* from the A genome and *BjuB.GSL-ALK.a* and *b* from the B genome) were isolated and sequenced from *B. juncea* lines Varuna and Heera. No polymorphism was detected between Varuna and Heera in the transcribed regions of any of the four genes. Genome walking upstream of the start codon identified SNPs in *BjuA.GSL-ALK.a* and *BjuB.GSL-ALK.a* and an in/del in *BjuA.GSL-ALK.b*. These polymorphisms in the 5′ flank of the genes *BjuA.GSL-ALK.a, BjuA.GSL-ALK.b* and *BjuB.GSL-ALK.a* were mapped to three independent linkage groups A3, A9 and B2, respectively (Table 1S of ESM; Fig. 2). After extensive chromosome walk (up to −3,864 bp upstream of the start codon and up to +1,440 bp downstream to the stop codon), no polymorphism could be detected in the flanks of *BjuB.GSL-ALK.b*.

Myb transcription factor gene family

Primers for *Myb28* amplified two bands in Varuna and Heera. Sequencing of these two bands revealed the presence of two homologues, *BjuA.Myb28.a* and *BjuB.Myb28.a* and showed length polymorphism between Varuna and Heera (Table 1S of ESM). Using these allelic length polymorphisms, *BjuA.Myb28.a* was mapped to linkage group A3 showing tight linkage with *BjuA.GSL-ELONG.c* (at a distance of 0.6 cM) and *BjuB.Myb28.a* was mapped to linkage group B1 (Table 1S of ESM; Fig 2).

Amplification of *Myb29* was carried out on *B. juncea* lines Varuna and Heera by use of intron spanning primers (Table 1). Three bands were observed in Varuna and two in Heera. All the five bands were polymorphic and detected three loci which were mapped in the linkage groups A10 (*BjuA.Myb29.a)*, B3 (*BjuB.Myb29.a*) and B8 (*BjuB.Myb29.b*) (Table 1S of ESM; Fig. 2).

Other glucosinolate related genes

Genes implicated in core glucosinolate biosynthesis such as CYP450 gene family—*CYP79F1* (At1g16410), *CYP79C2* (At1g58260) and *CYP83A1* (At4g13770), *sulphotransferase* (AtSOT17/18) (At1g18590, At1g74090) and *UDP-gylcosyltransferase* (At1g24100) were amplified from *B. juncea* using gene specific primers (Table 1). Only one orthologue of *CYP79C2* showed polymorphism between Varuna and Heera and was mapped (*BjuB.CYP79C2.a*) to linkage group B4 (Table 1S of ESM; Fig 2).

Co-localization of candidate genes with seed glucosinolate QTL

Figure 2 shows the location of six seed glucosinolate QTL (*J2Gsl1*, *J3Gsl2*, *J9Gsl3*, *J16Gsl4*, *J17Gsl5* and *J3Gsl6*) identified by Ramchiary et al. (2007a). Of the 17 candidate genes of glucosinolate biosynthesis pathway mapped to *B. juncea* genome in the present study, map positions of six candidate genes corresponded to the position of four QTL. *BjuA.GSL-ELONG.a* mapped to the QTL interval *J2Gsl1*, *BjuA.GSL-ELONG.c*, *BjuA.GSL-ELONG.d* and *BjuA.Myb28a* mapped to the QTL interval *J3Gsl2*, *BjuA.GSL-ALK.a* mapped to the QTL interval *J3Gsl6* and *BjuB.Myb28a* mapped to the QTL interval *J17Gsl5*. The remaining two QTL, *J9Gsl3* and *J16Gsl4* could not be co-localized with any of the mapped candidate genes (Table 2; Fig. 2).

Contribution of candidate genes/QTL to seed glucosinolate phenotype through genotype–phenotype association

Of the 1,160 BC_4DH lines used in the present study, complete genotype and phenotype data could be obtained only for 785 lines. Out of the expected 64 genotypic combinations for six segregating QTL, only 57 genotypic combinations could be recovered from the 785 lines. Details of genotypic and phenotypic data of 57 genotypic combinations have been provided in Table 2S of ESM. The frequencies of different genotypic classes were highly variable, ranging from 1 to 31. In general, at all the loci, the Varuna alleles (scored as '*a*') were represented at higher frequencies than the Heera alleles (scored as '*b*') with the exception of QTL *J9Gsl3* (Table 3S of ESM). This bias could be due to selective advantage of Varuna alleles over the Heera alleles in microspore embryogenesis used for the production of mapping populations (Mukhopadhyay et al. 2007).

Analysis of data revealed that the alleles contributed by the low glucosinolate parent Heera reduced the aliphatic glucosinolate content with significant reductions contributed by QTL *J3Gsl2*, *J9Gsl3* and *J3Gsl6*. The locus *J3Gsl2* was found to be most important as the Heera allele at this locus contributed more than twofold reduction in the level of aliphatic glucosinolates (Table 3S of ESM). This QTL harbours two candidate genes, *BjuA.GSL-ELONG.c* and

Table 2 Summary information on co-localization of candidate genes with glucosinolate QTL in *B. juncea*

Name of the QTL[a]	Linkage Group[b]	Genetic interval of QTL (cM)	Co-localized candidate gene(s) (map position in cM)
J2Gsl1	A2	82.2–92.6	*BjuA.GSL-ELONG.a* (85.8)
J3Gsl2	A3	23.2–37.2	*BjuA.Myb28.a* (29.3) *BjuA.GSL-ELONG.c,d* (29.9)
J9Gsl3	A9	5.6–26.5	–
J16Gsl4	B4	56.3–73.6	–
J17Gsl5	B1	45.5–58.1	*BjuB.Myb28.a* (53.1)
J3GSl6	A3	58.6–69.5	*BjuA.GSL-ALK.a* (66.6)

[a] Ramchiary et al. (2007a)

[b] Panjabi et al. (2008)

BjuA.Myb28.a at a distance of 0.6 cM (Fig. 2). The functionality and contribution of each gene was deciphered from the analysis of recombinants and is reported in the later part of the result section. Another locus, *J3Gsl6*, showed significant allelic variation for the aliphatic glucosinolate content (Table 3S of ESM).

The relative contribution of all the loci was ascertained by analyzing the QTL-NIL for these loci. We identified all the six QTL-NILs for the Heera allele and five QTL-NILs for the Varuna allele (defined as the line having the Varuna allele for one locus while others have Heera allele) from the genotyping data of the 785 lines (Table 4S of ESM). Varuna allele QTL-NIL for *J9Gsl3* could not be found in the population. The analysis of QTL-NILs for the Heera allele confirmed the contribution of *J3Gsl2* and *J9Gsl3* in reducing the aliphatic glucosinolate content. On the other hand, the significant allelic variation shown by *J3Gsl6* was found to be the effect of linkage with *J3Gsl2* (Fig. 2) as the Heera allele QTL_NIL for *J3Gsl6* did not contribute significantly in reducing aliphatic glucosinolate content. In addition, *J16Gsl4, BjuA.GSL-ELONG.a* in *J2Gsl1* and *J3Gsl2* were found to be involved in C3, C4 and C5 elongation pathways in which Heera alleles are weak alleles in the C3 and C4 pathways and strong alleles in the C5 pathway.

The analysis of Varuna allele QTL-NILs (Table 4S of ESM) revealed that all the low glucosinolate lines contained the Heera allele for the loci *J2Gsl1*, *J3Gsl2* and *J9Gsl3* and the presence of the Varuna allele either in *J2Gsl1* or *J3Gsl2* made the lines high glucosinolate. The analysis also revealed that although the Heera alleles of *J16Gsl4* (Heera allele QTL-NIL for *J16Gsl4* in Table 4S of ESM) and *J2Gsl1* (Heera allele QTL-NIL for *J2Gsl1* in Table 4S of ESM) are weak alleles in the C3 and C4 elongation pathway, respectively, these two alleles when present together with the Varuna allele of *J3Gsl2* led to the production of equal amounts of propenyl and butenyl glucosinolates (Varuna allele QTL-NIL for *J3Gsl2* in Table 4S of ESM and genotypic class from 37 to 40 in Table 2S of ESM). It suggests that the Varuna allele of *J3Gsl2* is epistatic to the Heera alleles of *J16Gsl4* and *J2Gsl1* when present together in the same individual.

High-resolution mapping of the QTL *J3Gsl2* and analysis of recombinants between *BjuA.GSL-ELONG.c* and *BjuA.Myb28a*

High resolution mapping of the QTL *J3Gsl2* was carried out by genotying 1718 DH lines from BC_4 and BC_5 generations (Ramchiary et al. 2007a) using markers for both the candidate genes. A total of 17 recombinants were recovered between *BjuA.GSL-ELONG.c* and *BjuA.Myb28.a* mapping these two genes at a distance of 1 cM. Association of phenotype with genotypic data of 17 recombinants (Table 5S of ESM) revealed that the Heera allele of *BjuA.Myb28.a* was responsible for major reduction in the total aliphatic glucosinolate content by proportionately reducing both the propenyl and butenyl glucosinolates without any influence on the chain elongation pathways. This was observed prominently in the Heera allele NIL for *BjuA.Myb28.a* (Recombinant no. 3, Table 5S of ESM) which reduced the total aliphatic glucosinolate content to 78.9 µmol from the wild type content of 128.6 µmol (Table 4S of ESM). On the other hand, the *BjuA.GSL-ELONG.c* was observed to be functional in the C5 chain elongation pathway in which the Heera allele is responsible for the formation of pentyl glucosinolates (Recombinant no. 4, 5, 9, 10, 13, 14, 15 and 16 of Table 5S of ESM) and the epistatic genetic control of *J3Gsl2* on *J2Gsl1* and *J16Gsl4* was due to *BjuA.GSL-ELONG.c* located in this QTL (Recombinant no. 12 of Table 5S of ESM). Analysis of recombinants between *BjuA.GSL-ELONG.c* and *BjuA.Myb28.a* further revealed that the total glucosinolate content could not be brought down below 18–20 µmol without additional contribution of the Heera allele of *BjuA.GSL-ELONG.c* (Recombinant no. 2 of Table 5S of ESM).

Discussion

The Polish spring rape variety ‘Bronowski’ (Kondra and Stefannson 1970) is regarded as the sole donor source of low glucosinolate trait to the present day ‘canola’ quality *B. napus* and *B. rapa* (Howell et al. 2003). In *B. juncea*, the first low glucosinolate breeding line BJ 1058 was developed by interspecific cross between an Indian type 3-butenyl glucosinolate containing *B. juncea* selection and a ‘Bronowski gene(s)’ containing low glucosinolate *B. rapa* followed by backcrossing to the *B. juncea* parent (Love et al. 1990a). The low glucosinolate parent Heera used in the present study was developed through pedigree breeding using BJ 1058 as a donor source for low glucosinolate (Srivastava et al. 2001). The genetic control of seed glucosinolates has been studied in detail in both *B. napus* and *B. juncea* through QTL analysis. In *B. napus*, the studies have primarily laid emphasis on the genetics of total glucosinolate content (Uzunova et al. 1995; Toroser et al. 1995; Howell et al. 2003; Quijada et al. 2006; Hasan et al. 2008). In *B. juncea*, though some of the studies have also laid emphasis on QTL analysis of different types of aliphatic glucosinolates along with the total glucosinolate content (Cheung et al. 1998; Mahmood et al. 2003; Lionneton et al. 2004), no detailed studies on elucidating the genetic mechanism of aliphatic glucosinolate accumulation are available in any of the *Brassica* species. The present study is an extension of the work of Ramchiary et al. (2007a) and has provided detailed dissection of individual QTL of seed glucosinolates through candidate genes association.

The comparative analysis of QTL controlling seed glucosinolates between *B. napus* and *B. juncea* by aligning the published genetic maps revealed that the QTL detected in linkage groups N2 and N9 (Quijada et al. 2006) correspond to QTL detected in LG16 and LG18 of Uzunova et al. (1995) and LG1 and LG20 of Toroser et al. (1995) of *B. napus* are homoeologous to A2 and A9 of *B. juncea* (Panjabi et al. 2008). On the basis of the comparative map based on common RFLP markers, it was observed that the QTL detected in LG2 of Uzunova et al. (1995) corresponds to LG18 of Toroser et al. (1995) and are homoeologous to A3 of *B. juncea* (Fig. 1 of ESM). In the present study, we observed that three QTL: *J2Gsl1*, *J3Gsl2* and *J9Gsl3* located in the linkage groups A2, A3 and A9 of *B. juncea*, respectively are the most important loci for breeding low glucosinolate *B. juncea*. These three A genome loci in all probability could be the 'Bronowski gene(s)' which have been introgressed from *B. napus* to *B. juncea* through interspecific hybridization. Hence, our work, in addition to developing efficient markers for marker-assisted introgression of low glucosinolate trait, provides clues to the genetic mechanism behind the historically important 'Bronowski gene(s)' controlling low glucosinolate trait in oilseed *Brassica* species.

The results of the present study show that the genetic control of seed glucosinolates in *B. juncea* operates at two levels, mainly controlled by the three A genome QTL discussed above. At the first level, there is allelic variation, which contributes to the overall production of aliphatic glucosinolates. This overall control of glucosinolate level was found to be primarily effected by *BjuA.Myb28.a* contained in QTL *J3Gsl2* and some sequence contained in QTL *J9Gsl3*. Presence of the Heera alleles at these two loci significantly reduced the glucosinolate content without interfering with the chain elongation pathways (Recombinant no. 3 of Table 5S of ESM, Heera allele QTL-NIL for *J9Gsl3* of Table 4S of ESM and Table 3S of ESM). We have not been successful in mapping any candidate gene to the QTL locus *J9Gsl3*. However, mapping of some of the *Arabidopsis* genes (At5g67590 and At5g63950) proximal to *Myb28* (At5g61420) in this QTL region (Fig. 2 and Panjabi et al. 2008) indicates that one more *Myb28* functional paralogue or a regulatory gene having function similar to *Myb28* might be present in this region. The R2R3-Myb transcription factors *Myb28* and *Myb29* have been shown to control the biosynthesis of aliphatic glucosinolates in *A. thaliana*. Analysis of knockout and knockdown mutants demonstrated that *Myb28* is a positive regulator for basal-level production of aliphatic glucosinolate and *Myb29* presumably plays a role in the induction of aliphatic glucosinolate biosynthetic genes in response to plant hormone methyl jasmonate (Hirai et al. 2007). It has been shown that double knockout of *Myb28* and *Myb29* leads to complete absence of aliphatic glucosinolates in *A. thaliana* (Beekwilder et al. 2008). It is fortuitous that a dysfunctional *Myb28* does not create any negative pleiotropic effects on the plant phenotype. Otherwise, productive low glucosinolate lines would not have been possible.

The second level of genetic control was observed in the chain elongation pathways. *BjuA.GSL-ELONG.a* possibly controls the C4 elongation pathway and *BjuA.GSL-ELONG.c* controls the C5 elongation. The wild type parental alleles at these two loci act in opposite ways for their respective pathways. For example, the Heera allele at *BjuA.GSL-ELONG.a* reduced the butyl glucosinolates while the Heera allele at *BjuA.GSL-ELONG.c* increased the pentyl glucosinolates (Table 4S). Although we could not recover any individual having the Heera allele at both these loci, the epistatic control of the Varuna allele of *BjuA.GSL-ELONG.c* together with the Heera alleles of *BjuA.GSL-ELONG.a* and *J16Gsl4* indicated that the Heera allele of *BjuA.GSL-ELONG.c* would be essential for lowering the glucosinolate content below 20 µmol (Recombinant 2 of Table 5S of ESM).

QTL *J16Gsl4* was found to control the C3 elongation pathway. A Heera allele at this locus reduced the formation of propyl glucosinolates. *GSL-PRO* has been shown to control the C3 elongation pathway in *A. thaliana* (At1g18500; Field et al. 2004) and *B. oleracea* (Gao et al. 2006). Although we could map two B genome specific paralogues of *GSL-PRO* to linkage groups B6 and B7 in the present study, no *GSL-PRO* paralogue could be mapped to the region of QTL *J16Gsl4*. The synteny information from the comparative map of *B. juncea* (Fig. 2 and Panjabi et al. 2008), however, indicates that *J16Gsl4* could be one of the potential locations for *GSL-PRO*. We observed that around 30% of the low glucosinolate individuals (18 out of 58 in Table 2S of ESM) did not have the Heera allele at this locus indicating that QTL *J16Gsl4* is not required in breeding low glucosinolate *B. juncea*. In Indian *B. juncea* types, propyl glucosinolates contribute marginally (around 18–20%) to the total glucosinolate content (Sodhi et al. 2002). Hence, the two-step genetic control enumerated above, (1) an overall decrease by *BjuA.Myb28.a* and *J9Gsl3* and (2) control of the elongation process by *BjuA.GSL-ELONG.a* and *BjuA.GSL-ELONG.c*, would be enough to reduce the seed glucosinolate content below 18 µmol without involvement of the low allele from locus *J16Gsl4*. However, manipulation of this locus will be essential for developing low sinigrin (2-propenyl) *B. juncea*.

Ramchiary et al. (2007a,b) highlighted negative linkages between some yield component QTL and glucosinolate QTL *J2Gsl1*and *J3Gsl2*. With the development of candidate gene markers in these loci, marker-based precision breeding would be possible for the development of productive low glucosinolate *B. juncea* cultivars belonging to the Indian gene pool. It may also be interesting to further study the role of 'Bronowski genes' on low glucosinolate content at the molecular level in other oilseed *Brassica* species.

Acknowledgments This work was supported by the Dhara Vegetable Oil and Food Company Ltd (DOFCO), a fully owned company of the National Dairy Development Board (NDDB) and the Department of Biotechnology (DBT). Partial support came from UGC-SAP programme.

References

Beekwilder J, van Leeuwen W, van Dam NM, Bertossi M, Grandi V, Mizzi L, Soloviev M, Szabados L, Moltoff JW, Schipper B, Verbocht H, de Vos RCH, Morandini P, Aarys MGM, Bovy A (2008) The impact of the absence of aliphatic glucosinolates on insect herbivory in *Arabidopsis*. PLoS ONE 3:e2068

Cheung WY, Landry BS, Raney P, Rakow GFW (1998) Molecular mapping of seed quality traits in *Brassica juncea* L. Czern. and Coss. Acta Hort 459:139–147

Field B, Cardon G, Traka M, Botterman J, Vancanneyt G, Mithen R (2004) Glucosinolate and aminoacid biosynthesis in *Arabidopsis*. Plant Physiol 135:828–839

Gao M, Li G, Potter D, McCombie R, Quiros CF (2006) Comparative analysis of methylthioalkylmalate synthase (MAM) gene family and flanking DNA sequences in *Brassica oleracea* and *Arabidopsis thaliana*. Plant Cell Rep 25:592–598

Gland A, Röbbelen G, Thies W (1981) Variation of alkenyl glucosinolates in seeds of *Brassica* species. Z Pflanzenzüchtg 87:96–110

Halkier BA, Gershenzon J (2006) Biology and biochemistry of glucosinolates. Ann Rev Plant Biol 57:303–333

Hasan M, Friedt W, Pons-Kühnemann J, Frietag NM, Link K, Snowdon RJ (2008) Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). Theor Appl Genet 116:1035–1049

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, Goda H, Nishizawa OI, Shibata D, Saito K (2007) Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc Natl Acad Sci USA 104:6478–6483

Howell PM, Sharpe AG, Lydiate DJ (2003) Homoeologoue loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). Genome 46:454–460

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D (2001) Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. Plant Physiol 126:811–825

Kondra ZP, Stefannson BR (1970) Inheritance of major glucosinolates in rapeseed (*Brassica napus*) meal. Can J Plant Sci 50:643–647

Kraling K, Robbelen G, Thies W, Herrmann M, Ahmadi MR (1990) Variation in seed glucosinolates in lines of *Brassica napus*. Plant Breed 105:33–39

Li G, Quiros CF (2002) Genetic analysis, expression and molecular characterization of *BoGSL-ELONG*, a major gene involved in the aliphatic glucosinolate pathway of *Brassica* species. Genetics 162:1937–1943

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

Love HR, Rakow G, Raney JP, Downey RK (1990a) Development of low glucosinolate mustard. Can J Plant Sci 70:419–424

Love HR, Rakow G, Raney JP, Downey RK (1990b) Genetic control of 2-propenyl and 3-butenyl glucosinolate synthesis in mustard. Can J Plant Sci 70:425–429

Magrath R, Herron C, Giamoustaris A, Mithen R (1993) The inheritance of aliphatic glucosinolates in *Brassica napus*. Plant Breed 111:55–72

Magrath R, Bano F, Parkin I, Sharpe A, Lister C (1994) Genetics of aliphatic glucosinolates I. Side chain elongation in *Brassica napus* and *Arabidopsis thaliana*. Heredity 72:290–299

Mahmood T, Ekuere U, Yeh F, Good AG, Stringam GR (2003) Molecular mapping of seed aliphatic glucosinolates in *Brassica juncea*. Genome 46:753–760

Mukhopadhyay A, Arumugam N, Sodhi YS, Gupta V, Pradhan AK, Pental D (2007) High frequency production of microspore derived doubled haploid (DH) and its application for developing low glucosinolate lines in Indian *Brassica juncea*. In: Proceedings of the 12th international rapeseed congress, Wuhan, pp 333–335

Østergaard L, King GJ (2008) Standardized gene nomenclature for the *Brassica* genus. Plant Methods 4:10

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

Potts D, Rakow G, Males DR (1999) Canola quality *Brassica juncea*, a new oilseed crop for the Canadian prairies. In: Proceedings of the international rapeseed congress, Canberra, 26–29 September (CD-ROM)

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

Quijada PA, Udall JA, Lambert B, Osborn TC (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 1. Identification of genomic regions from winter germplasm. Theor Appl Genet 113:549–561

Ramchiary N, Bisht NC, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D, Pradhan AK (2007a) QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed *Brassica juncea*: importance of recurrent selection backcross scheme for the identification of 'true' QTL. Theor Appl Genet 116:77–85

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

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

Sodhi YS, Mukhopadhyay A, Arumugam N, Verma JK, Gupta V, Pental D, Pradhan AK (2002) Genetic analysis of total glucosinolate in crosses involving a high glucosinolate Indian variety and a low glucosinolate line of *Brassica juncea*. Plant Breed 121:508–511

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 102:193–199

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

Stam P, Van Ooijen JW (1996) Joinmap version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Stringam GR, Thiagarajah MR (1995) Inheritance of alkenyl glucosinolates in traditional and microspore-derived doubled haploid populations of *Brassica juncea* L. Czern and Coss. In: Proceedings of the 9th international rapeseed congress, Cambridge, pp 804–806

Toroser D, Thormann CE, Osborn TC, Mithen R (1995) RFLP mapping of quantitative trait loci controlling seed aliphatic-glucosinolate content in oilseed rape (*Brassica napus* L.). Theor Appl Genet 91:802–808

Uzunova M, Ecke W, Weissleder K, Robbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L). 1. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor Appl Genet 90:194–204

Wittstock U, Halkier BA (2002) Glucosinolate research in the *Arabidopsis* era. Trends Plant Sci 7:263–270