Genome Wide Identification, Phylogeny and Expression Profiling Analysis of Shattering genes in Rapeseed and Mustard Plants

Mahideen Afridi (mahidkhanafridi@gmail.com)
Quaid-i-Azam University Islamabad: Quaid-i-Azam University
Khurshid Ahmad
Ocean University of China College of Food Science and Engineering
Shahana Seher Malik
United Arab Emirates University
Nazia Rehman
NARC: National Agricultural Research Centre
Muhammad Yasin
Quaid-i-Azam University Islamabad: Quaid-i-Azam University
Muhammad Talha Yasin
Quaid-i-Azam University Islamabad: Quaid-i-Azam University
Adil Hussain
University of Okara
Muhammad Ramzan Khan
Quaid-i-Azam University Islamabad: Quaid-i-Azam University

Research Article

Keywords: Rapeseed, Indian mustard, Brassica, Fruit, shattering genes, Mads box, Oil seed

Posted Date: June 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-631553/v1

Abstract

Background: Non-synchronized pods shattering in the *Brassicaceae* family bring upon huge yield losses around the world. The shattering process was validated to be controlled by eight different genes in the model plant Arabidopsis thaliana, including *SHATTERPROOF1, SHATTERPROOF2, FRUITFULL, INDEHISCENT, ALCATRAZ, NAC, REPLUMLESS* and *POLYGlACTOURANAZE*. To obtain gene family & examine their expression patterns into fresh & mature silique, then completed genome wide identification, characterization, and expression analysis of shattering genes in *B. napus* and *B. juncea*.

Results: Complete genome analysis of *B. napus* and *B. juncea* revealed 32 shattering genes, which were identified and categorize based on protein motif structure, exon-intron organization and phylogeny. The phylogenetic study revealed that these shattering genes contain little duplications that were determined with a distinct chromosome number. Motifs of 32 shattering proteins were also observed where motifs 6 were found to be more conserved. A single motif was observed for other genes like *BrnS7, BrnS8*, *BrjS23* and *BrjS26*. Comparative genomics for synteny analysis was performed that validated a conserved pattern of blocks among these cultivars. RT-PCR based expressions profile showed higher expression of shattering genes in *B. juncea* as compared to *B. napus*. *FUL* gene was expressed more in the mature silique. *ALC* gene was not expressed in the fresh silique of *B. napus* but highly expressed in the mature silique.

Conclusion: This study authenticates that shattering genes exist in the local cultivars of *Brassica*. *ALC* exhibited strong expression in both the mature and fresh silique of *B. juncea.* Our results showed that shattering genes expression occurred more in *B. juncea* as compared to *B. napus*. It also contributes to the screening of more candidate gene for further investigation and characterization.

Introduction

Brassicaceae is one of the important families, with ~ 360 genera and 3, 700 species around the globe [1]. Species from this family are very significant from and economic, and agricultural point of view. A few examples of species from this family are *Brassica napus*, *Brassica juncea* (oilseed crops); *Brassica oleracea* (cabbage, cauliflower, kale, broccoli); *Brassica rapa* (turnip, leaf vegetable); *Raphanus sativus* (vegetable) and *Arabidopsis thaliana* (model plant). The most cultivated species of the *Brassica* genus includes those with three diploid genomes like *B. nigra* (BB, 2n = 16) *B. oleracea* (CC, 2n = 18) and *B. rapa* (AA, 2n = 20), together with three amphidiploid species like *B. napus* (AACC, 2n = 38) *B. carinata* (BBCC, 2n = 34), and *B. juncea* (AABB, 2n = 36). Hybridization and cytogenetic studies have determined that amphidiploid species are natural hybrids of diploids and these all species are interconnected [2].

Non-synchronous pods shattering remained a major problem of *Brassica* that results in yield loss. It also causes seed loss due to the dispersion of a silique following complicated physiological and biological mechanisms [3]. However, premature and unsynchronized pods shattering like the dehiscence results in a huge loss in crops yield [4]. Pod's shattering occur when the adhesions among walls change into fragile and internal forces apply them to the moveable position [5]. Seed valves are responsible for the attachment and internal force creation that contributes to the necessary protection of seed [6]. Seeds of *B. juncea* and *B. napus* are very important equally 14% of oil around the world is produced by these crops. Moreover, rapeseed is considered as the third most important oilseed crop worldwide [7]. Distinct nutrients and biological molecules are reported to be involved in the evolution of shattering in canola [8].

Previous investigation over shattering revealed that shattering occurred because of molecular components excess production and enrichment in the valve margin and cellar portion around the pods in siliques. Lignin and cellulose play a key role in the hardening of pod walls, which lowers water content during the later development stages of rapeseed and *Brassica* species [9]. The shattering mechanism of *B. napus* and *B. juncea* are controlled by eight different genes, like *SHATTERPROOF1 (SHP1), SHATTERPROOF2 (SHP2) FRUITFULL (FUL), INDEHISCENT (IND), ALCATRAZ (ALC), NAC, (NST1* and *NST2) REPLUMLESS (RPL)* and *POLYGIACTOURANAZE* [10]. For the development of shattering genes, distinct transcription factors binding sites are involved which are important both structurally and functionally [11]. Other genes like *SHP1/2, FUL, IND, ALC, NAC, RPL* and *PG* of canola and Indian mustard were also reported [12]. In one study, a comparative analysis performed to unveil the genomic maintenance for the evolutionary and functional correlation among shattering genes *SHP1/2, FUL, IND, ALC, NAC, RPL* and *PG* [13] having functional and genetic conservation among them. The pattern of conservation in these shattering gene sequences was also found with comparative synteny approach by Krzywinski et al. [14].

The most desirable solution to the shattering problem of *B. napus* and *B. juncea* is to delay pods shattering by knocking out *SHPS* genes and stimulating the expression of *FUL* up to the susceptible crop is ready for harvesting. Therefore, before developing genome modified plants it is essential to study these genes elaborately in local plant *B. napus* and *B. juncea*. Therefore, this study was carried out to identify the orthologous of shattering genes in the local cultivars of *B. napus* and *B. juncea* and to study their expression pattern in fresh and mature siliques. This study further identified the syntenic and evolutionary relationship of shattering genes in the studied cultivars on the basis of phylogenetic analysis with NJ algorithm.

Materials And Methods

Identification of shattering genes

BRAD database (http://brassicadb.org/brad/) was used to retrieve protein, genomic, CDS and cDNA sequences of the 8 shattering genes *SHP1/2*, *FUL, ALC, NAC*, *IND*, *RPL* and *PG* and their orthologuesin *B. napus* and *B. juncea* following [15]. Other databases like NCBI (https://www.ncbi.nlm.nih.gov/), TAIR database (https://www.arabidopsis.org/) and Plants Ensembl (http://plants.ensembl.org/) were also consulted. A web tool from EMBL was used to identify different protein domains (http://smart.embl.de/smart/set_mode.cgi). Basic local alignment search tool (BLAST) (htpp://www.ncbi.nlm.nih.gov/BLAST/) was used to search homology of the shattering genes in *B. napus* and *B. juncea*. ProtParam tool was used to study the primary structure of shattering genes (http://expasy.org/tools/protparam.html). Gene structure display server (GSDS) web tool was used to align the CDS sequences of shattering genes with genomic sequences to identify exons and introns [16].

Phylogenetic analysis of shattering proteins

B. napus and *B. juncea* shattering protein sequences were obtained from the BRAD database using reference sequences of shattering genes obtained from the TAIR database and then aligned using the Clustal X program [17]. Using the Neighbour Joining (NJ) algorithm, a phylogenetic tree was constructed with MEGA 6.06 software [18]. The implication of nodes was calculated using a bootstrap study of 1,000 replicates. For the surety of different domains that shows the topology of NJ tree, pairwise gape deletion mode was used.

Analysis of conserved motifs in shattering proteins

MEME software (Multiple Em for Motif Elicitation, V4.9.0) was used to analyse MADS-box shattering genes protein sequences as described by Bailey et al. (2006) [19]. MEME search was run with the following parameters: (1) maximum number of motif identification = 10; (2) optimum motif width > 6 and < 200.

Analysis of syntenic relationships

The comparative genomic synteny was performed to find relationship among distinct shattering genes like *SHP1/2, FUL, ALC, NAC*, *IND*, *RPL* and *PG* in *B. napus* and *B. juncea* using circoletto program; genome visualization tool circoletto [14].

Primers designing

Primers were designed to assessthe expression of shattering genes*SHP1/2, FUL, ALC, NAC, RPL, PG* and *IND* in fresh and mature siliques by primer3 software [20] then checked by primer stat [21] and Multiple primer analyzers (Thermofisher Scientific). The specificity of primers was checked by UCS PCR at UCSC-Insilco PCR genome browser (https://genome.ucsc.edu/) [22].

Plant collection and sample preparation

Seeds of two *Brassica* varieties canola (Punjab Sarson) and Indian mustard (Super Raya) were collected from the plant Bioresources Conservation Institute (BCI) and Crop Sciences Institute (CSI) of National Agricultural Research Centre (NARC), Islamabad Pakistan. The seeds were cultivated in National Institute for Genomics and Advanced Biotechnology (NIGAB), National Agriculture Research Centre (NARC), Islamabad, Pakistan under pods in a glasshouse. Forty days old samples of pre-mature and mature siliques of *B. napus* (Punjab Sarson) and *B. juncea* (Super Raya) were collected and stocked at -80^{o} C for gene expression analysis. Morphological analysis was performed and the data of plants was recorded in triplicates.

RNA extraction and cDNA synthesis

Total RNA from the fresh and mature silique of *B. napus* and *B. juncea* was extracted using a Pure Link^{TM} RNA Mini kit (Invitrogen). The RNAs were quantified by using BioSpec-nano Micro-volume UV-Vis Spectrophotometer (Shimadzu). The quality and integrity of RNA was checked on 1.5% agarose gel. cDNA was synthesized by using RevertAid^{TM} reverse transcriptase enzyme (Fermentas^{TM} Cat.No. K1621) following the manufacturer's guidelines.

Expression analysis of Shattering genes qRT-PCR

The expression pattern of shattering genes (*SHP1/2, FUL*, *ALC*, *NAC, RPL, PG* and *IND*) was determined in fresh and mature silique of *B. napus* and *B. juncea* using comparative ΔCT method in real-time PCR (Applied Biosystems) with StepOnePlus software. For the execution of a relative expression, the Elongation factor (EF) was used as endogenous control along with no template control (NTC). In total, 10µl reactionvolume, 5µl Maxima SYBER Green (Thermo Fisher) genes specific primers (1 pmol of each) and 1µl cDNA as a template were used. Real-time PCR conditions set were; denaturation at 94ºC for 10 minutes, the second stage followed by 40 cycles at 95ºC for 40 seconds, 58ºC for 32 seconds 72ºC for 32 seconds. Finally, melt curve study was carried out at 52ºC to 95 ºC. The statistical analysis of results was carried out by mean of relative fold expression of transcript ± standard deviation (SD). All theprimers used inqRT-PCR analysis listed in Table 3 were designed used different bioinformatics tools through the conserved region from the genome of *B. napus*. Length of the amplified fragment ranged between 100-150bp.

Results

Identification and sequence analysis of shattering genes

A set of 32 individual orthologues of shattering genes of *B. napus* and *B. juncea* was retrieved from the *Brassica* database (BRAD) (http://brassicadb.org/brad/). These genes were in greater number than those of model plant *Arabidopsis thaliana* as shown in Tables 1 and 2. The domain of these shattering genes was also identified using EMBL (http://smart.embl.de/smart/set_mode.cgi). The first six shattering genes of *B. napus* (*BrnS1-6*) contain the MADS-box domain whereas, 7–10 contain HLH, 11, 12, Pfam, 13, 14 pox/Hox 15–17 contain PbH1 domain. In *B. juncea*, 18–22 contain MADS-box domain while 23–26 HLH, 27, 28 Pfam, 29, 30 pox/Hox and 31, 32 PbH1 domain. Sequence analysis showed all shattering genes of *B. napus* and *B. juncea* have introns. The maximum numbers of introns were identified in MADS-box shattering gene up to 8 and lowest up to 1 in *BrnS9*. Other genes like *BrnS10, BrjS25* and *BrjS26* did not contain any introns. The shattering gene *IND* in both species showed no introns while *BrnS11, BrnS12, BrjS27* and *BrjS28* showed same number of introns. These appearances are persistent with shattering genes previously determined in *Arabidopsis thaliana.*

Table 1
In silico study of 17 shattering genes identified in *B. napus* with their closest *Arabidopsis* homologs and sequence feature

Gene Name	Gene locus	Chr. No	Closest Arabidopsis Homologs	Protein length	Mol. wt. (kda)	PI	Introns
BrnS1	GSBRNA2T00098954001	A07	SHP1/AGL1	348aa	39.77	9.21	8
BrnS2	GSBRNA2T00105875001	C06	SHP1/AGL1	248aa	28.41	9.11	6
BrnS3	GSBRNA2T00132708001	A05	SHP2/AGL5	244aa	28.00	9.21	5
BrnS4	GSBRNA2T00113594001	A03	FUL/AGL8	241aa	27.43	9.37	7
BrnS5	GSBRNA2T00094717001	A09	FUL/AGL8	241aa	27.49	9.31	7
BrnS6	GSBRNA2T00086507001	C02	FUL/AGL8	241aa	27.45	9.36	7
BrnS7	GSBRNA2T00070429001	C07	ALC/AT5G67110/BHLH73	216aa	23.51	9.03	4
BrnS8	GSBRNA2T00063470001	C02	ALC/AT5G67110/BHLH73	98aa	11.12	10.0	5
BrnS9	GSBRNA2T00153545001	C03	IND/EDA33/GT10	178aa	20.36	7.93	1
BrnS10	GSBRNA2T00112126001	A03	IND/EDA33/GT10	182aa	20.60	6.06	0
BrnS11	GSBRNA2T00150558001	A10	NAC/At5g22380/MWD9.18	285aa	32.25	6.91	2
BrnS12	GSBRNA2T00085330001	C05	NAC/At5g22380/MWD9.18	286aa	32.37	7.60	2
BrnS13	GSBRNA2T00069510001	A10	BLH9/RPL/BLR/LSN/PNY	578aa	62.49	7.12	4
BrnS14	GSBRNA2T00088804001	C02	BLH9/RPL/BLR/LSN/PNY	575aa	61.96	6.94	4
BrnS15	GSBRNA2T00064043001	C08	PG/At1g45015	419aa	43.97	8.83	3
BrnS16	GSBRNA2T00052454001	A09	PG/At1g45015	418aa	43.88	8.83	3
BrnS17	GSBRNA2T00089606001	A08	PG/At1g45015	420aa	43.85	8.39	3

(aa, amino acid; kda, kilo Dalton)

Table 2
In silico study of 15 shattering genes identified in *B. juncea* with their closest *Arabidopsis* homologs and sequence feature

Gene Name	Gene locus	Chr. No	Closest Arabidopsis Homolog	Protein length	Mol. wt. (kda)	PI	Introns
BrjS18	BjuB022348	B06	SHP1/AGL1	278aa	31.74	8.49	6
BrjS19	BjuB022350	B06	SHP1/AGL1	247aa	28.20	9.11	6
BrjS20	BjuB001727	B01	SHP2/AGL5	244aa	28.00	9.12	5
BrjS21	BjuB027201	B04	FUL/AGL8	159aa	18.50	9.62	4
BrjS22	BjuB037752	B02	FUL/AGL8	301aa	34.82	9.08	7
BrjS23	BjuB020848	B06	ALC/AT5G67110/BHLH73	222aa	24.59	9.62	4
BrjS24	BjuA011758	A07	ALC/AT5G67110/BHLH73	214aa	23.41	9.37	4
BrjS25	BjuB019604	B08	IND/EDA33/GT10	191aa	21.62	5.97	0
BrjS26	BjuB019326	B08	IND/EDA33/GT10	191aa	21.59	5.97	0
BrjS27	BjuA038017	A10	NAC/At5g22380/MWD9.18	285aa	32.25	6.91	2
BrjS28	BjuB030790	B03	NAC/At5g22380/MWD9.18	293aa	32.96	6.46	2
BrjS29	BjuB001605	B08	BLH9/RPL/BLR/LSN/PNY	577aa	62.00	8.85	3
BrjS30	BjuA040195	A10	BLH9/RPL/BLR/LSN/PNY	586aa	63.15	6.95	3
BrjS31	BjuA029936	A08	PG/At1g45015	420aa	43.74	7.96	3
BrjS32	BjuB032977	B03	PG/At1g45015	421aa	43.74	7.93	2

(aa, amino acid; kda, kilo Dalton)

Phylogenetic Analysis Of Shattering Genes

The identified shattering genes protein sequences were used to analyze the phylogenetic relationship of the shattering gene family in *B. napus, B. juncea* and *Arabidopsis*. The unrooted phylogenetic tree characterizes the length of clades and the level of the evolutionary relationship with well-supported bootstrap values. The sequences of shattering genes *SHP1, SHP2, FUL, IND*, *ALC, NAC, RPL, PG* and their orthologous determined into *B. juncea* and *B. napus* were aligned to generate the NJ phylogenetic tree (Fig. 1). Every individual shattering gene organized in a distinct clade, characterize their functional and sequential conservation. Clade I contains a duplication of *SHP1* genes in *B. napus* and *B. juncea* plants. However, clade II consists of *SHP2* genes where no duplication was observed. This shows that clade I and II are closely related to each other as compared to other clades. In clade III, duplication of *FUL* genes was observed in *B. juncea* and triplication in *B. napus* that indicates divergence in sequences and in clade IV, duplication of *NAC* genes was noticed. It is clear from the resulting tree that clade III and clade IV are closely related to clade I and II. Similarly, clade V and clade IV contains *RPL* and *ALC* genes in a duplicated form in *B. napus* and *B. juncea* plants. However, clade VII and clade VIII comprised *IND* and *PG* genes with duplication. The clade comprising of *FUL* and *PG* genes contain a greater number of genes as compared to others. Genes from these two clades are present on different chromosomes indicate that every individual gene bear

duplication and whole genome triplication events before reaching this level. Environmental, physiological and chromosomal rearrangement at the development level brought changes in the genome. These results authenticate that every individual gene of *B. napus* and *B. juncea* under observation are shattering genes having a close resemblance to each other and with a model plant Arabidopsis thaliana as shown in Fig. 1.

Analysis of conserved motifs in shattering proteins of B. napus and B. juncea

MEME (Multiple Em for Motif Elicitation) motif search tool was used to identify 10 conserved motifs of 32 shattering protein sequences of *B. napus* and *B. juncea* (Fig. 2). Motif 1 and 2 exhibit MADS-box domain which was found in 24 genes whereas other shattering genes did not show motif 1or 2 features. The genes which exhibit the characteristics of motifs 1 or 2 were *BrnS1-BrnS6* and *BrjS18-BrjS22*. These genes did not contain other representative motifs of Mads-box family such as motifs 4, 5, 6, 7, 8, 9, 10. Motif 4 and 5 comprised of *PbH1* domain found in 5 genes which were *BrnS15, BrnS16*, *BrnS17, BrjS31* and *BrjS32. BrnS7, BrnS8, BrjS23* and *BrjS24* genes consists of single motif. Motif 8 and 10 showed pox/Hox domain which was found in *BrnS13, BrnS14, BrjS29* and *BrjS30* gene. *BrnS15, BrnS16, BrnS17, BrjS31* and *BrjS32* comprised PbH1 domain with motif 5 and 6 features. Motif 6 were conserved among all genes which is the characteristic feature of shattering genes. The different motifs are represented by different colours that showed similarities among *B. napus* and *B. juncea* as shown in (Fig. 3). The number of motifs found in both species is similar except for *BrnS7, BrnS8, BrjS23* and *BrjS24* which shows single motif and revealed similarities and differences with other shattering genes among *brassica* species.

Syntenic relationship among shattering genes of B. napus and B. juncea

Comparative genomic synteny analysis was performed by circoletto Tool (tools.bat.inspire.org/circoletto/) for genome conservation visualization. The orthologues relationship and conservation were determined for the shattering gene family in *B. napus* and *B. juncea*. Synteny diagram represents a remarkable relationship among these species in the context of duplication, triplication, evolution, function and expression (Fig. 4) showed a unique relationship among *B. juncea* and *B. napus*. It was observed that *B. napus BrnS13* and *BrnS14* gene sequence showed synteny with *B. juncea* sequence *BrjS29* and *BrjS30*, while *B. napus* gene sequence *BrnS15*, *16* and *17* showed synteny with *B. juncea* gene sequence *BrjS31, 32* and gene sequence *BrnS11* and *12* showed synteny with *BrjS27* and *BrjS28*. In Addition, *BrnS7* and *BrnS8* gene sequence showed synteny with *BrjS23* and *BrjS24* gene sequences while *BrnS9* and *BrnS10* showed synteny with *BrjS25* and *BrjS26* gene sequences. Similarly, *BrnS1* and *BrnS2* showed synteny with *BrjS18* and *BrjS19* gene sequences, while *BrnS3* showed synteny with *BrjS20*. *B. napus* gene *BrnS4, 5, 6* sequences showed synteny with *BrjS21* and *BrjS*22. In comparative synteny analysis inward tangling ribbons colour intensity exhibited the rate of conservation while outward tangling ribbons showed duplication events. Genomic dynamicity and evolutionary improvement along mobile elements in the genome of *B. napus* and *B. juncea* were determined in syntenic circles. In chromosomal shuffling, duplication and triplication mobile elements play an important role. A permanent position was adopted by the blocks at a specific position in genome initiate expression that involve another biological pathway disturbance (Fig. 4).

qRT-PCR expression of shattering genes in fresh and mature siliques

The expression level of shattering genes in fresh and mature siliques of *B. napus* and *B. juncea* was confirmed by qRT-PCR. Our results inferred that the expression level of shattering genes was higher in *B. juncea* as compared to *B. napus* in both fresh and mature siliques. Strong signals of shattering genes were observed in mature siliques in both species, while in fresh silique, the transcripts levels were low (Fig. 5). The correlation is completely noticeable in the evidence that shattering genes play a major role in shattering associated pathways by devoting to developmental

pathways of lignification and valve margin associated transcriptional activity. Moreover, *ALC* gene expression was upregulated in fresh silique of *B. juncea* while down regulation of *ALC* gene was observed in fresh silique of *B. napus*.

The same expression pattern was observed when shattering genes were run on agarose gel. The (Fig. 6) shows that shattering genes were expressed in both plants in mature silique as well as in fresh silique with a little bit difference.

Table 3
The primers used for qRT-PCR analysis

Sr.no	Primer name	Primer sequences 5'–3'
1	SHP1-F	GTAGTCACGACGCAGAGAGTA
	SHP1-R	AACTTCAGCATCACACAAGAC
2	SHP2-F	GTGTAAGAGGAACGATCGAAA
	SHP2-R	TCACCAAGAATGTGTCTGTTC
3	FUL-F	GACTCTTGCATGGAAAGCATA
	FUL-R	TCTTCTCAAGTACCTCAACTC
4	IND-F	GAAACCCTAAGCCACTTCCAG
	IND-R	CTCGCTTATCCTTTCTCTAC
5	NAC-F	GGGCAGCAACTTCTGGTTACT
	NAC-R	TCAGTGAGGCGATATTCATGC
6	ALC-F	GTTTCCTCCGCTGAGATGTTC
	ALC-R	ATGAATTTCGCTGTCTAGCTC
7	RPL-F	GTGTGGGTCATGGTATTTACA
	RPL-R	ATACCTCTTGTAAACCTCGTC
8	PG-F	GTGTGGAAGTCTCTCCAAATC
	PG-R	ACACAGAGGGAGTAGCTTGCC

Discussion

Brassicaceae is a large plant family consists of ~ 338 genera and 3, 700 species, important both economically and agriculturally [19]. In addition to this, plants of this family are grown like a weed in different parts of the world including North America, South America, and Austria [23]. *Arabidopsis thaliana*, a model plant from the family *Brassicaceae* was the first plant to be entirely sequenced [24]. Plants and vegetables from this family offer essential food nutrients to human and other animals. Due to their great importance, all the *Brassica* plants have common and commercial value with a positive influence on earth and manhood. *Brassica* species have inconstant traits and morphological differentiation revealing that the genome of this family is very vibrant and endured a lot of rearrangement and evolutionary measures [25].

In this study, *SHP1, SHP2, FUL, IND, ALC, NAC, RPL* and *PG* when compared at the genomic level showed close similarity. Protein and nucleotide shows an important correlation at the sequence level. It has been showed that these genes are responsible in shattering and seed development of plants [26]. The phylogenetic analysis here showed that *SHP1* as compared to other genes have fever dynamicity which is balanced in the connection of genomics but bear duplication. The duplicated genes determined with a distinct chromosome number in *B. napus* and *B. juncea* which recommended genomic flexibility as previously reported in *Arabidopsis* and *B. rapa* [27, 28] shows similar results with our investigations. *SHP2* shattering gene study uses a novel approach to phylogenetic analysis bears no duplication and triplication as previously reported in other *Brassica* species [29, 30]. *FUL* is known for fruit development in different *Brassica* species. The Phylogenomics of *FUL* affords unusually different results than *SHP1* and *SHP2.* The behavior observed more dynamics among the various species of *Brassica* family. *FUL* genes showing duplication and differential location in the genome of *B. juncea* and *B. napus* also previously described in *B. rapa* further strengthen our results [8].

In current research, we have study 32 MADS box shattering gene of *B. juncea* and *B. napus* which are more in number than the shattering genes reported for *A*. thaliana [31]. The syntenic analysis performed among *B. napus* and *B. juncea* shows the similar sequence feature and whole genome of both species go through triplication events since its divergence from Arabidopsis. The evolutionary and syntenic relationships among *Arabidopsis* and *B. rapa* is also supporting our results [32]. On the other hand, we observed the expression of shattering genes *SHP1, SHP2, FUL, IND, ALC, NAC, RPL* and *PG* in *B. napus* and *B. juncea* like previously reported in *Arabidopsis* [33]. Our result also suggest that these genes are the reputed orthologous of *Arabidopsis* genes AGL1, AGL5, AGL8, AT5G67110, EDA33, At5g22380, BLH9 and At1g45015 might play the similar role, and they are expressed in both plants in fresh and mature siliques.In previous studies, divergence in expression pattern was observed in shattering genes in *B. napus*. Wu et al., [35] determined the expression patterns and evolution of MADS-box TF family in *B. napus*. Becker and Theissen [36] reported that Shatterproof1/2 and genes which are members of MADS box family are engaged in controlling this pod shattering issue. *SHP1* and *SHP2* genes are involved in opening of silique in *B. napus* plants when the expression level is low [37, 26, 38].

The expression of these genes started from developed flower to mature silique with lower expression in the late stage of development of seed [34]. *SHP1, SHP2* and *FUL* showed a relationship with *IND, ALC* that initiate acting to abrogate activity of DZ to forbid dehiscence at the time of seed formation follow indehiscence in the existence of multiple regulatory genes. The present analysis of all shattering genes showed different expression pattern in different tissues such as fresh and mature siliques of both plants as previously reported in *Arabidopsis* and *B. rapa*. These genes were expressed in both plant tissues, although in *B. juncea* they were slightly higher than in *B. napus*. These different expressions of shattering genes shows that they are important for cellular valve and margin evolution [27].

A similar study was conducted by Yasin et al., [39] whose results agree with our results. They demonstrated higher expression of *FUL* gene in mature aerial part silique plant as compared to leaves and flowers of *B. napus* plants. Similarly, *SHP1* and *SHP2* transcripts were expressed in flower silique whereas; no expression was detected in the leaves. Our findings showed basic gene expression information about shattering cascade genes which can be useful for developing genome edited *brassica* plants which are resistant to shattering.

Conclusion

It is concluded that, different orthologous of shattering genes are exists in the local cultivars of *Brassica*. After comparative phylogenetic study, molecular gene characteristics, motifs/domain identifications and comparative expression study, we identified that the sequences were conserved across *B. napus, B. juncea* as well as in *Arabidopsis* plant. The redundant expression was observed in fresh and mature siliques of both cultivars. The different expression patterns of shattering genes are also helpful to study the nature of both plants and their pathways related to transcription and regulation. Further analysis of shattering genes is required to uncover their functions involved in the regulation of different pathways.

Declarations

Acknowledgments

This research work was supported by National Institute of Genomics and Advanced Biotechnology, National Agriculture Research Center, Islamabad Pakistan and National Center for Bioinformatics Quaid-I-Azam University, Islamabad Pakistan.

Conflict of interest All the authors of this study declare that they have no competing interest.

References

1. Warwick SI, Francis A and Al-Shehbaz IA (2006) Brassicaceae: species checklist and database on CD-ROM. Plant Syst Evol 259: 249–258
2. U, N (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Jpn J Bot 7: 389–452
3. Dong Y, Wang Y-Z (2015) Seed shattering: from models to crops. Frontiers in plant science 6.
4. Raman R, Raman H, Kadkol GP, Coombes N, Taylor B, and Luckett D (2011) Genome wide association analyses of loci for shatter resistance in Brassicas. In: Proceedings of the 11th Australian research assembly on brassicas (ARAB) conference, Wagga Wagga, NSW, pp. 36–41
5. Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M (2014) Genome-wide delineation of natural variation for pod shatter resistance in Brassica napus. PLoS One 9 (7): e101673
6. Balanzà V, Roig-Villanova I, Di Marzo M, Masiero S, Colombo L (2016) Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. Development 143 (18):3372-3381
7. Basalma D (2008) The correlation and Path analysis of yield and yield components of different winter rapeseed (Brassica napus ssp. oleifera L.) cultivars. Res. J. Agric. Biol. Sci. 4:120–125.
8. Hu Z, Yang H, Zhang L, Wang X, Liu G, Wang H, Hua W (2015) A large replum-valve joint area is associated with increased resistance to pod shattering in rapeseed. Int J Plant Res 128 (5):813-819
9. Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R (2017) Post-polyploidisation morphotype diversification associates with gene copy number variation. Sci Rep 7:41845
10. Zumajo-Cardona C, Ambrose BA, Pabón-Mora N (2017) Evolution of the SPATULA/ALCATRAZ gene lineage and expression analyses in the basal eudicot, Bocconia frutescens L. (Papaveraceae). EvoDevo 8 (1):5
11. Cosio C, Dunand C (2010) Transcriptome analysis of various flower and silique development stages indicates a set of class III peroxidase genes potentially involved in pod shattering in Arabidopsis thaliana. BMC genomics

11 (1):528

12. Yu K, Wang X, Chen F, Chen S, Peng Q, Li H, Zhang W, Hu M, Chu P, Zhang J (2016) Genome-wide transcriptomic analysis uncovers the molecular basis underlying early flowering and apetalous characteristic in Brassica napus L. Sci Rep 6
13. Hradilová I, Trněný O, Válková M, Cechova M, Janská A, Prokešová L, Aamir K, Krezdorn N, Rotter B, Winter P (2017) A Combined Comparative Transcriptomic, Metabolomic, and Anatomical Analyses of Two Key Domestication Traits: Pod Dehiscence and Seed Dormancy in Pea (Pisum sp.). Front Plant Sci 8
14. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19 (9):1639-1645
15. Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, et al. (2011) BRAD, the genetics and genomics database for Brassica plants. BMC Plant Biol 11(1):136
16. Guo AY, Zhu QH, Chen X, Luo JC, Yi Chuan (2007) GSDS: a gene structure display server. 29(8):1023–6
17. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25(24):4876–82
18. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30:2725–9
19. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34 suppl 2:369–73
20. Rozen S, Skaletsky H (1999) Primer3 on the WWW for general users and for biologist programmers. Bioinformatics methods and protocols:365-386
21. Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28 (6):1102-1104
22. Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ (2009) The UCSC genome browser database: update 2010. Nucleic acids Res 38 (suppl_1): D613-D619
23. Couvreur TLP, Franzke A, Al-Shehbaz IA, Bakker FT, Koch MA, Mummenhoff K (2010) Molecular phylogenetics, temporal diversification, and princliples of evolution in the mustard family (Brassicaceae). Mol Bio Evo 27:55-71
24. Parsons WT, Cuthbertson EG (2001) (2010) Noxious Weeds of Australia., 2nd Edition. CSIRO Publishing (2nd Edition) & Inkata Press (1st Edition), Collingwood, Victoria. Lee J, Lee I. Regulation and function of SOC1, a flowering pathway integrator. J Exp Bot 61(9):2247–54
25. Li X, Zhang S, Bai J, He Y (2016) Tuning growth cycles of Brassica crops via natural antisense transcripts of BrFLC. Plant Bio J 14 (3):905-914
26. Liljegren S.J, Ditta G.S, Eshed Y, Savidge B, Bowman J.L and Yanofsky M.F (2000) Shatterproof MADS box genes control seed dispersal in Arabidopsis. Nat 404: 766-770 https://doi. org/10.1038/35008089
27. Yang J, Liu D, Wang X, Ji C, Cheng F, Liu B, Hu Z, Chen S, Pental D, Ju Y (2016) The genome sequence of allopolyploid Brassica juncea and analysis of differential homeolog gene expression influencing selection. Nat Genet 48 (10):1225-1232
28. Xu Y, Xu H, Wu X, Fang X, Wang J (2012) Genetic changes following hybridization and genome doubling in synthetic Brassica napus. Biochem Genet 50 (7-8):616-624

29. Smaczniak C, Immink RG, Angenent GC, Kaufmann K (2012) Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. Development 139 (17):3081-3098
30. Cheng F, Wu J, Liang J, Wang X (2015) Genome triplication drove the diversification of Brassica plants. In: The Brassica rapa Genome. Springer, pp 115-120
31. Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, et al (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis. Plant Cell Online 15(7):1538–51
32. Song XM, Huang ZN, Duan WK, Ren J, Liu TK, Li Y, et al. (2014) Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (Brassica rapa ssp. pekinensis). Mol Genet Genomics. 289(1):77–91.
33. Lee J, Lee I (2010) Regulation and function of SOC1, a flowering pathway integrator. J Exp Bot 61(9):2247–54
34. Wu Y, Ke Y, Wen J, Guo P, Ran F, Wang M, Liu M, Li P, Li J and Du H (2018) Evolution and expression analyses of the MADS-box gene family in Brassica napus. PLoS One 13(7): e0200762 https://doi.org/10.1371/journal.pone.0200762
35. Becker A and Theissen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol 29: 464–489
36. Ferrandiz C, Liljegren SJ, Yanofsky MF (2000) Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. Sci 289:436–438
37. Wu H, Mori A, Jiang X, Wang Y.M and Yang M (2006) The Indehiscent protein regulates unequal cell divisions in Arabidopsis fruit
38. Khan MR, Ihsan H, Ali GM (2016) WSA206, a paralog of duplicated MPF2-like MADS box family is recruited in fertility function in Withania. Plant Sci 253:215-228
39. Yasin M, Shahzadi R, Riaz M, Afridi M, Ajmal W, Rehman, O, Rehman N, Ali GM, Khan MR (2019) Expression pattern analysis of core regulatory module Shps-Ful transcripts in rapeseed pod shattering. Sarhad J Agri 35(3): 696-707

Figures

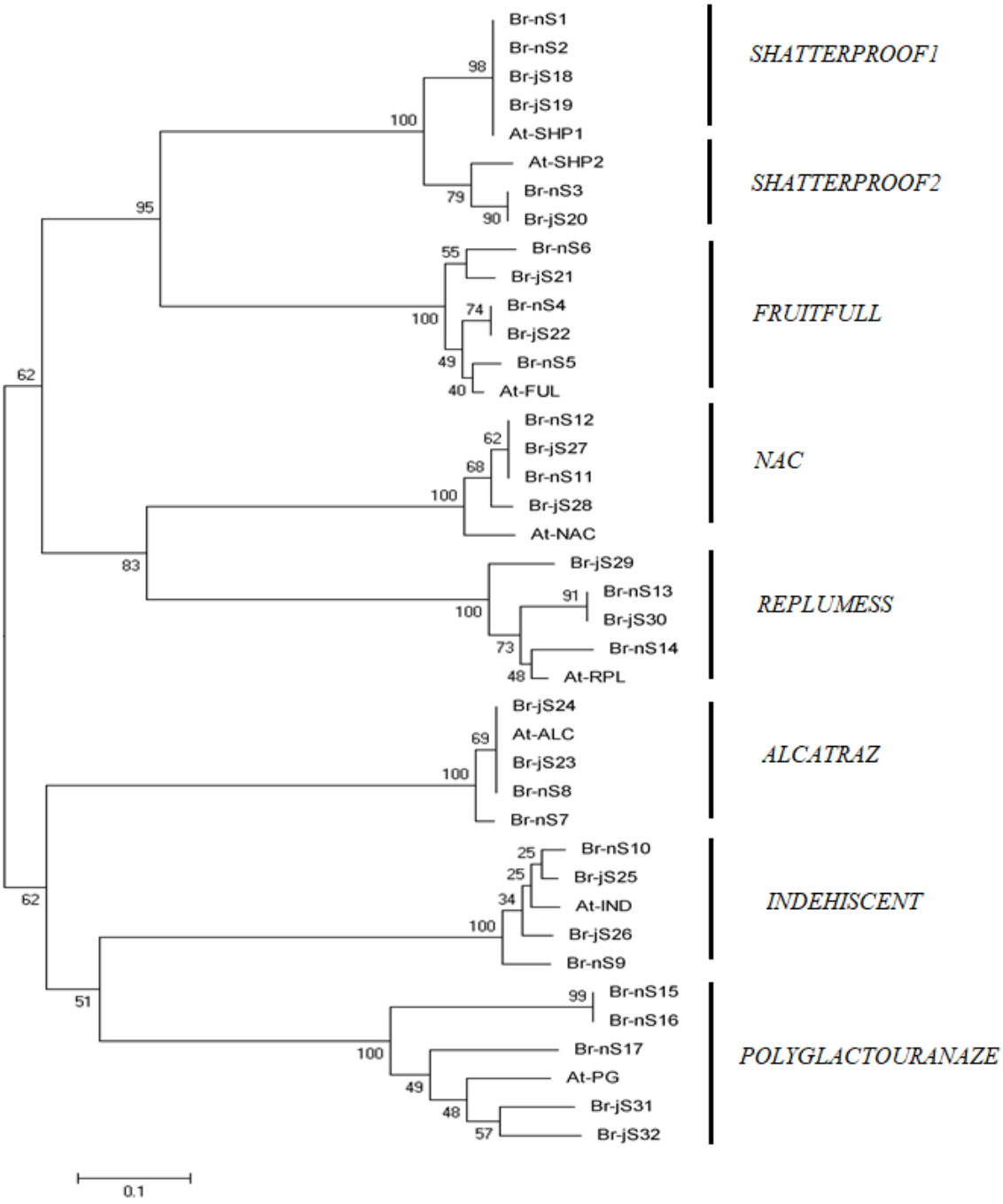

Figure 1

Neighbor Joining consensus phylogenetic tree of shattering genes. The values indicated along branches are the Bootstraps values obtained from NJ analysis with 1000 replicates.

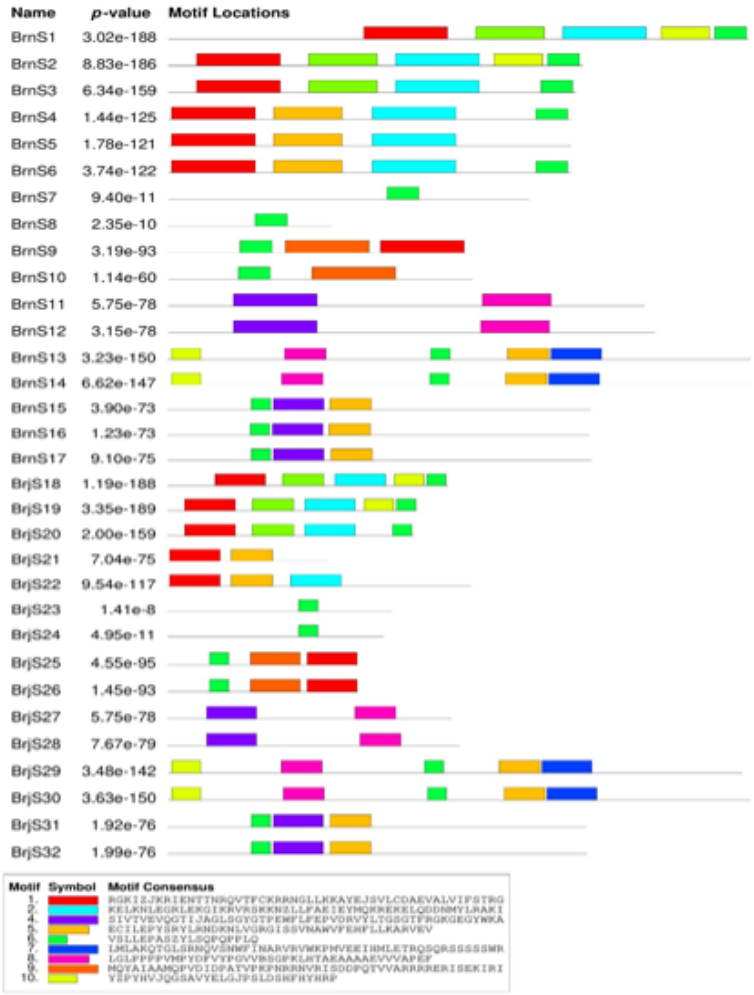

Figure 2

Schematic representations of motifs identified in B. napus and B. juncea using MEME search tool for individual shattering genes. Different motifs were identified by distinct colours and the names of all members are shown at the left side. Motifs order correlates to the position of a motif in specific protein sequences.

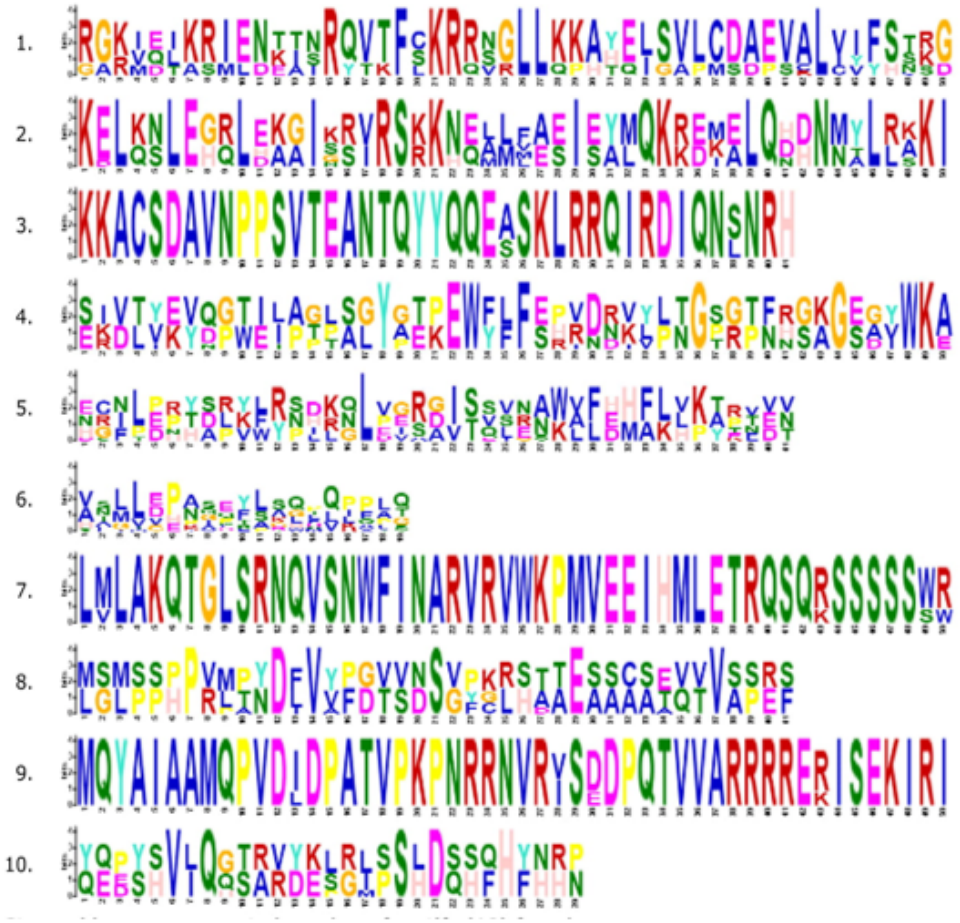

Figure 3

Logos of tens motifs discovered in shattering genes

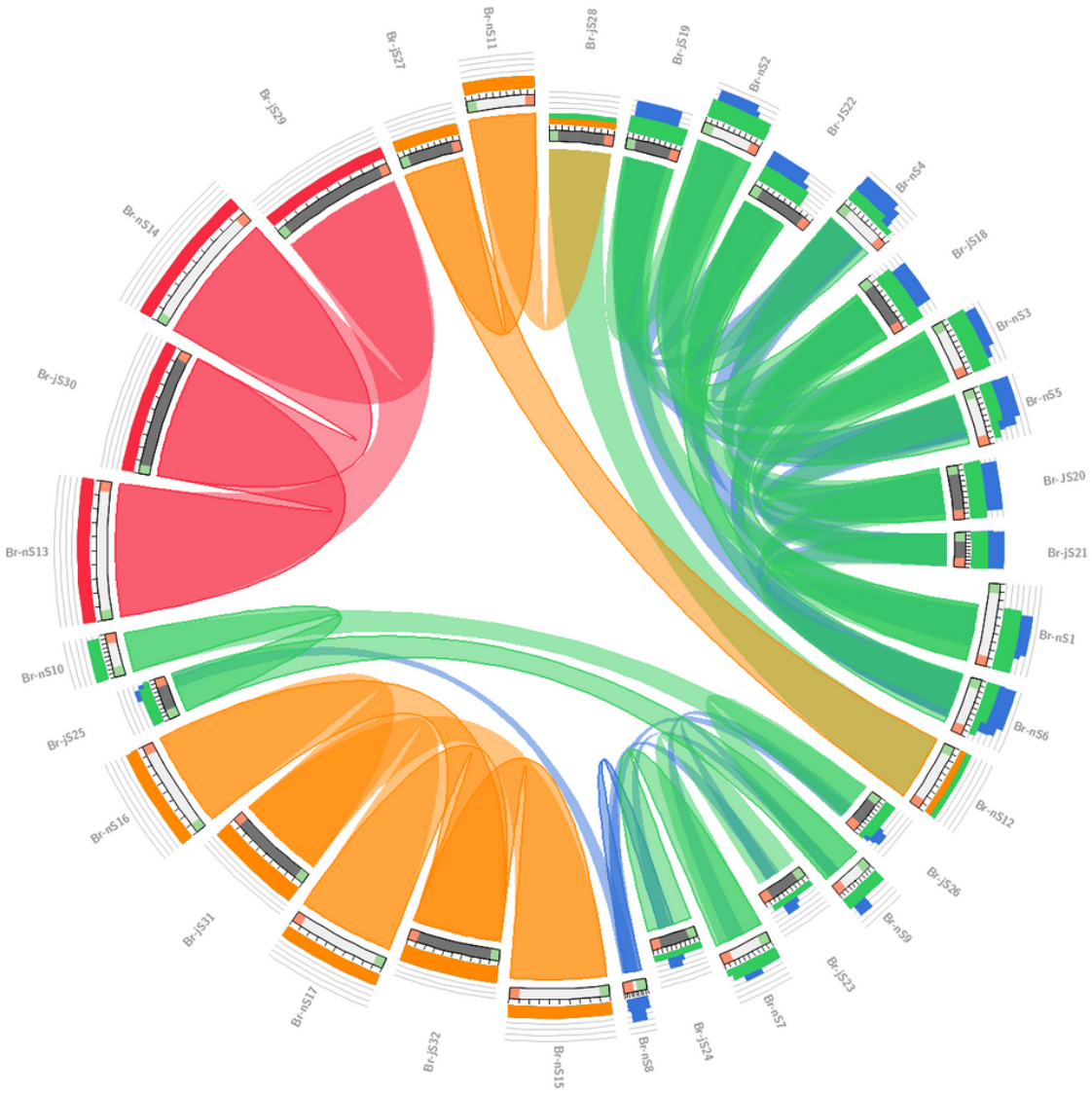

Figure 4

Representation of Synteny of B. napus and B. juncea identifying the level of conservation at the sequence level in 4 colours. The red, green, orange and blue colours signify the level and intensity of evolutionary conservation among distinct shattering genes e.g. maximum intensity is from orange to green.

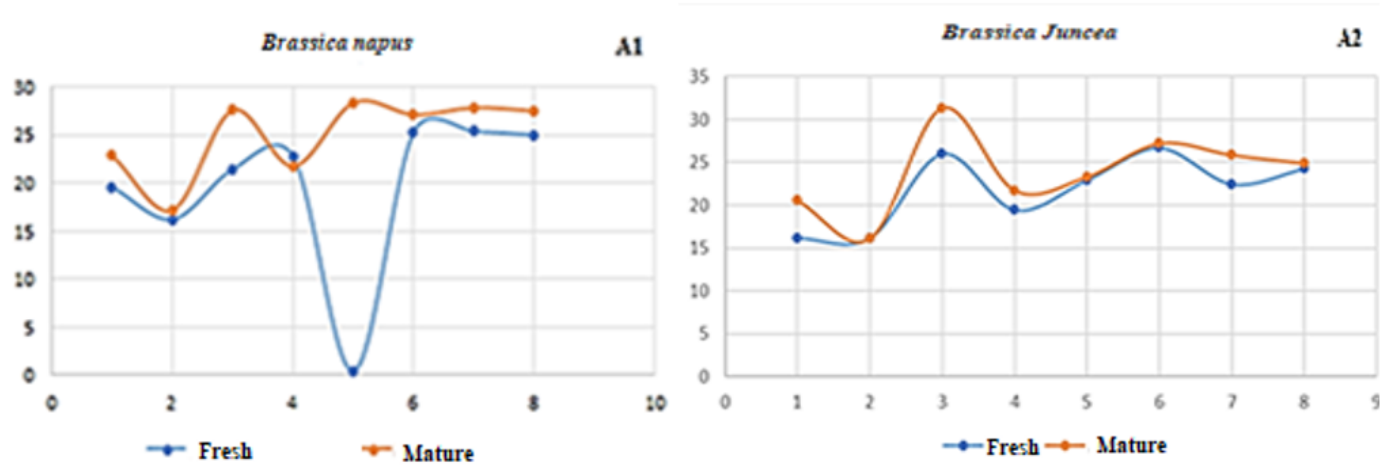

Figure 5

The expression graph of shattering genes in B. napus and B. juncea. Graph dots defining the expression difference and correlation among two tissues of B. napus and B. juncea. (A2) represents a higher level of shattering genes expression in B .juncea than B. napus in the given tissues. ALC gene expressed more in B. juncea. ALC gene was not expressed in fresh silique of B. napus only expressed in mature silique. (A1) represents a lower level of expression of shattering genes in B. napus than B. juncea.

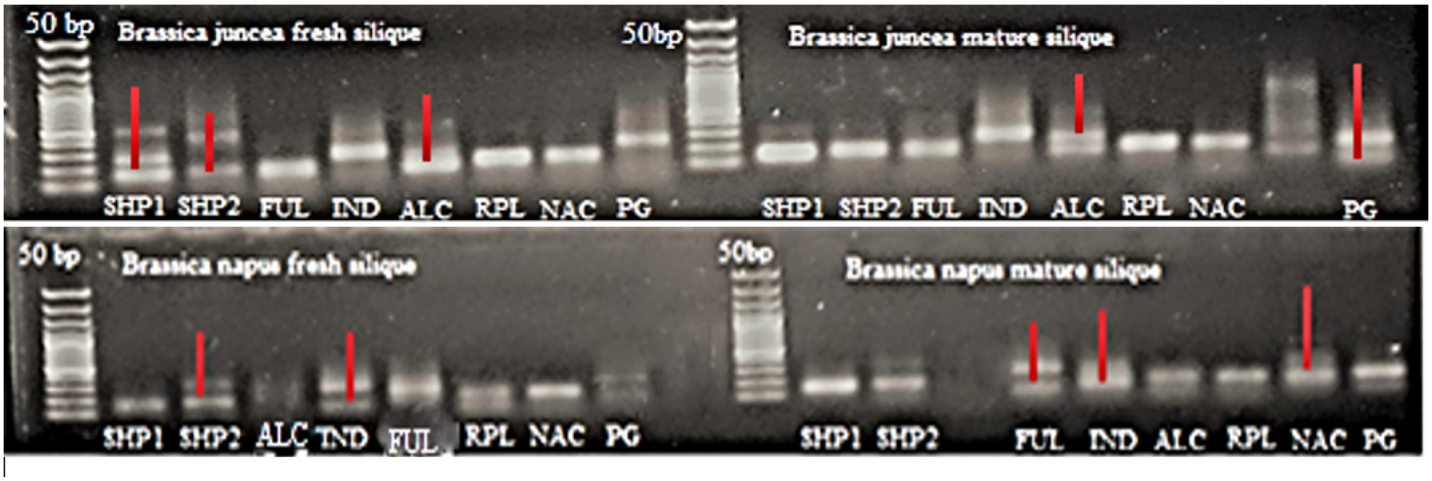

Figure 6

Expression analysis of shattering genes in B. napus and B. juncea fresh and mature silique through RT-PCR. The 50 bp standard ladder was used to confirm the expression of shattering genes.