RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
55 (3) 681-694, 2023
http://doi.org/10.54910/sabrao2023.55.3.7
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

GENOME-WIDE ANALYSES OF CORE REGULATORY MODULE SHATTERING CASCADE GENES IN CANOLA (*BRASSICA NAPUS* L.)

M. YASIN[1, 2*], G.M. ALI[3], M. RIAZ[4], S. ALI[3], H.U. RAHMAN[5], A. IQBAL[6], S.U. KHAN[7], M. SHAKEEL[1], M. MUNIR[8], M. MOHIBULLAH[9], and M.R. KHAN[2,3]

[1]Gomal Centre of Biochemistry and Biotechnology, Gomal University, Dera Ismail Khan, KP, Pakistan
[2]PARC Institute of Advanced Studies in Agriculture, National Agricultural Research Center, Islamabad, Pakistan
[3]Department of Plant Genomics and Biotechnology, National Institute for Genomics and Advanced Biotechnology, National Agricultural Research Center, Islamabad, Pakistan
[4]School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai, China
[5]Vegetable Crop Research Program, Horticulture Research Institute, National Agricultural Research Center, Islamabad, Pakistan
[6]Institute of Biotechnology and Genetic Engineering, The University of Agriculture, Peshawar, Pakistan
[7]Institute of Molecular Biology and Biotechnology, The University of Lahore, Punjab, Pakistan
[8]Institute of Food Science and Nutrition, Gomal University, Dera Ismail Khan, Pakistan
[9]Department of Plant Breeding and Genetics, Gomal University, Dera Ismail Khan, KP, Pakistan
*Corresponding author's emails: drmywazir-biotech@gu.edu.pk, drmrkhan_nigab@yahoo.com
Email addresses of co-authors: drgmali5@gmail.com, mriaz19@sjtu.edu.cn, shaukat_parc@yahoo.co.in, askhafeez594@yahoo.co.uk, aqib72@aup.edu.pk, samiullah.khan@imbb.uol.edu.pk, shakeelimperial611@gmail.com, mmunir92@gmail.com, drmohib@gu.edu.pk

SUMMARY

Premature seed shattering in canola causes massive losses in yield by up to 50% in adverse climatic conditions. In the model plant *Arabidopsis thaliana,* which belongs to the same family as canola, the Brassicaceae, eight genes participate in a shattering cascade. Phylogenetic reconstruction, syntenic relationships, genomics loci, promoter sequences, and identification of transcription factor-binding sites (TFBSs) faced shattering cascade genes' analysis. Among these, three genes, *SHATTERPROOF1*, *SHATTERPROOF2,* and *FRUITFUL (SHP1, SHP2, FUL),* belonged to a MADS-box family implicated in fruit dehiscence zone and valve margin constitute a core regulatory module. But, in *Brassica*, the exact number of genes involved in shattering remained obscure. Grouping *BnSHP1-N, BnSHP2-N,* and *BnFUL-N* into their respective clades was according to phylogenetic reconstruction of core regulatory modules *(SHP1, SHP2,* and *FUL)* and from other species homologs. The eight shattering cascade genes showed no conservation, indicating their involvement in crushing through separate pathways. The increased number of homologs/paralogs in *Brassica* was due to occurrences of genome duplication or a triplication event during evolution. Exonization and intronization could be responsible for a variable number and size of the exons and introns in gene structures. Comparative genome synteny analysis of *SHP1, SHP2,* and *FUL* revealed correlation and evolutionary insights into gene region relationships in all *Brassicaceae*. Study results provided basic information on cloning, phylogenetic reconstruction, genomics loci, and identifying transcription factor-binding sites (TFBSs) of core regulatory module genes that might be helpful for developing shattering-resistant genome-edited plants to prevent future yield losses in canola.

Citation: Yasin M, Ali GM, Riaz M, Ali S, Rahman HU, Iqbal A, Khan SU, Shakeel M, Munir M, Mohibullah M, Khan MR (2023). Genome-wide analyses of core regulatory module shattering cascade genes in canola (*Brassica napus* L.). *SABRAO J. Breed. Genet.* 55(3): 681-694. http://doi.org/10.54910/sabrao2023.55.3.7.

Keywords: *Brassica napus* L, genome analysis, shattering genes, core regulatory module, homologs, exonization, intronization

Key findings: The phylogenetic reconstruction of *SHP1*, *SHP2,* and *FUL* genes and homologs from other species conglomerated *BnSHP1-N*, *BnSHP2-N,* and *BnFUL-N* into their respective clades. The rise in homologs in *Brassica* resulted from genome duplications or triplications that occurred during evolution. When analyzing gene structure, exons and introns may vary in size and number due to exonization and intronization. Comparative genome synteny analysis of *SHP1/2* and *FUL* genes revealed correlation and evolutionary insights throughout the *Brassicaceae*. The promoter analysis unveiled that the expression divergence may correlate with their divergent promoters, where regulatory motifs, particularly CArG-boxes, might have played varying roles in the siliques of these plants.

Communicating Editor: Dr. Sajjad Hussain Qureshi

Manuscript received: April 23, 2023; Accepted: June 21, 2023.

INTRODUCTION

Agriculture is the mainstay of Pakistan's economy. Though self-sufficient in cereals, Pakistan has been constantly and chronically deficient in edible oil production (Economic Survey of Pakistan, 2019–2020). The total production of edible oil was 0.5 million t compared with the consumption of 2.772 million t, thus, leaving a big gap. Edible oil is the greatest (70%) imported food item, involving a vast expenditure in foreign exchange to make up the deficit. Canola (*Brassica napus* L.), with a high oil yield per unit area in the Pakistani climate, has emerged as a chief oilseed crop, with the potential to narrow the gap between the production and consumption of edible oils (Economic Survey of Pakistan, 2019–2020).

The members of this family are famous for their oil contents and importance as a vegetable worldwide. *Brassica napus (B. napus)* is the third-largest source of vegetable oil production globally (USDA Foreign Agricultural Service, 2015). Several varieties of *B. napus* are the product of evolution and genomic complexity due to physio-pathological changes in their biotic environment in response to competition between them and their environment (Yang *et al*., 2016).

Brassica belongs to the family of plants known as *Brassicaceae,* which includes several economically important crops and vegetables. Rapeseed provided 13% of the global economy's requirement for vegetable oil. Erucic acid and glucosinolates in high-quality rapeseed oil are less than 54%, reducing the rapeseed's nutritional value globally. Various edible condiments and other affordable natural products with a great energy source and nourishment are part of the *Brassica* species (USDA Foreign Agricultural Service, 2015).

Brassica species have variable traits and morphological differentiation, revealing that this family genome is very vibrant, having undergone many genomic rearrangements and evolutionary events (Li *et al*., 2016). Many local varieties and canola hybrids with high-yielding potential are under cultivation in Pakistan today. Pod dehiscence causes significant yield damage (Raman *et al*., 2011). The standard yield damages range is 10%–25% (Price *et al*., 1996). When harmful climatic conditions cause late harvesting, seed losses estimate reached as high as 50% of the yield. Resistance to pod shattering is a desired characteristic for worldwide growth of *B. napus* L. plants to minimize grain damage in the mature standing crop and during windrowing and mechanical harvesting.

Shattering resistance is a critical and vital character for improving oilseed rape (Kadkol, 2009). Interspecific hybridization with similar species, such as, *B. nigra, B. juncea,*

and *B. rapa,* served to eliminate pod-shattering; however, it proved difficult, having added many undesirable traits (Kadkol, 2009). *B. napus* is one of the primary crop plants vulnerable to shattering. Unlike the importance of this trait in cereal species, rapeseed's complete stoppage of pod-shattering and seed loss was not a priority during domestication. As a result, excessive pod-shattering continues to occur.

Biotechnology requires a thorough understanding of the molecular mechanism of shattering cascade genes and their interactions controlling pod dehiscence, with the potential to provide a rapid solution to this problem. Research in Arabidopsis revealed that eight prime genes, including *SHATTERPROOF1-2 (SHP1/2), FRUITFULL (FUL), INDEHISCENT (IND), ALCATRAZ (ALC), REPLUMLESS (RPL), NAC SECONDARY WALL THICKENING PROMOTING FACOTR1 (NST), ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1/2 (ADPG1/2),* and *GIBBERELLIN 3-BETA-DIOXYGENASE 1 (GA3ox1)* participated in pod-shattering. Three of these are MADS-box genes, *SHP1, SHP2,* and *FUL,* constituting a core regulatory module in shattering cascade.

The pod-shattering process starts with cellular partitions alongside a few other cells, known as the dehiscence region (Meakin and Roberts, 1990). Within the bicarpellate silique, *FRUITFULL (FUL)* aids in managing valve extension and distinction opposing *SHATTERPROOF1-2* and *INDEHISCENT (IND)* on the dehiscence sector, whereby both restrain regular lignification. Repressing the *SHP1/2* is also done with the aid of *REPLUMLESS (RPL)* responsible for replum formation (Roeder *et al*., 2003). Similarly, *FUL* regulates other factors, *ALCATRAZ (ALC)* and *SPATULA (SPT),* that play a role in the correct formation of the separation layer in a roundabout way (Rajani and Sundaresan, 2001). The MADS-box circle of relatives includes *FUL* and *SHP1/2*. These families have undergone several duplications and losses in plants, many of which transpired by functional changes. In *Arabidopsis*, a member of the same Brassicaceae family, *SHP1* and *SHP2* are MADS-box transcription factors encoding genes that work redundantly to regulate pod dehiscence, as neither mutant differs significantly from wild type in terms of phenotype (Liljegren *et al*., 2000).

The essence of the fruit dehiscent research in Brassicaceae fleshy fruits, and cereals is that a MADS-box core-regulatory module *SHP-FUL* gets conserved in *Arabidopsis* and close relatives through the interplay and interaction of other genes required for DZ expression. MADS-box genes contribute to autoregulation, as also explained by the ABCD floral quarter model. These genes perform functions by binding with a CArG box, which is a consensus DNA sequence with the motif of $CC(A/T)_6GG$ (Smaczniak *et al*., 2012) residing in the promoter sequence with transcriptional regulation functionality. The complete transcription and developmental processes the MADS-box genes orchestrated (Ihsan *et al*., 2015) gain processing in connection with a CArG box, which is the central signaling network for other procedures with expression and to repress a process (Smaczniak *et al.*, 2012). When MADS-box genes start binding, it allows loop formation with the start of transcription of the target gene. The target gene specificity recruited by MADS protein cofactors allow chromatin remodeling proteins to stabilize the start site and trigger transcriptional activation of the target gene. The function depends on chromatin remodeling and loop-forming co-factors (Ng and Yanofsky, 2001).

In *Brassica*, pod dehiscence is a highly significant feature, although the unreliable process and the particular genes responsible for it are yet fragmented. Therefore, this study proposed genome analysis of pod-shattering genes in canola. The objective is to identify genomic loci and regulatory elements for understanding the evolutionary divergence of the core regulatory module (*SHP1*, *SHP2,* and *FUL*) during pod dehiscence in *B. napus*.

MATERIALS AND METHODS

Phylogenetic reconstruction

Phylogenetic analysis stemmed to from the evolutionary relationship of *SHP1, SHP2,* and

FUL genes in different species. For this reason, coding sequences of *SHP1, SHP2,* and *FUL* from different species surfaced on *Brassica napus*, *Arabidopsis thaliana, Arabidopsis Lyrata, and* other related plants earlier collected. A neighbor-joining tree construction used MEGA6 software with P uncorrected distance method. Conducting 1000 replications of bootstraps calculated and validated the reliability of the tree.

Analysis of syntenic relationships

Comparative genomic synteny was performed for *SHATTERPROOF1, SHATTERPROOF2,* and *FRUITFULL* genes in 11 species of the Brassicaceae family using the Circos program to reveal the correlation among shattering and MADS-box gene family; genome visualization tool. http:/circos.ca/ (Krzywinski *et al*., 2009).

Promoter sequences and genomic loci analysis

Retrieving promoter sequences of *SHP1, SHP2,* and *FUL* genes from the two genomes of canola *(B. napus),* i.e., *SHP1/2-AC* and *FUL-AC* used Ensembl genome browser (Hubbard *et al*., 2002) (http://www.ensembl.org). Promoter sequence was around 2kb of each genome. Dot plot analysis aided in comparing promoter sequences to identify genomic loci and regions of close similarity among them. The conduct of dot plot analysis employed the PipMaker program (Schwartz, 2000) (http://pipmaker.bx.psu.edu/pipmaker/).

Identification of transcription factor-binding sites (TFBSs)

Genomic sequences of the shattering cascade genes *SHP1/2*, *FUL, IND*, *ALC*, *RPL*, *NST*, *ADPG1/2,* and *GA3ox1* of the *Arabidopsis'* retrieval continued from The Arabidopsis Information Resource 10 (http://www.arabidopsis.org/). First, predicting putative TFBSs along the genomic sequences had the help of AGRIS (http:/ /arabidopsis.med.ohio-state.edu; Yilmaz *et al*., 2011) and AthaMap (http:/ / www.athamap.de/; Bulow *et al*., 2009), and faced refinement by referring to literature reports of the sequence features and functional properties of cis-regulatory elements (CREs).

The JASPAR 2016, MULAN, and TRANSFAC databases served to find the functionally important transcription factor binding sites in selected plant genomes for comparing core regulatory module (*SHP1/2-FUL)* genes in *(Brassica napus L.)* (Bailey *et al*., 2015).

RESULTS

Phylogenetic relationship of core regulatory module in different plants

The clade or branch length of a tree shows the evolutionary rate. Interestingly, the core regulatory module of *SHP1, SHP2,* and *FUL* constituted a separate clade. The rest of the genes remained clustered in their respective branch. Observations revealed no close relationship among the genes, indicating their functions in different shattering stages. They have diverse families of genes as well (Figure 1).

For a deeper insight into evolutionary relationships, 13 species underwent phylogenetic reconstruction. Figure 2 demonstrates three MADS-box genes clustering in separate clades. *SHP1* and *SHP2* are sister to each other, indicating their overlapping functions. The *FUL* gene makes a distinct clade due to a different role. Having multiple copies of genes refers to their occurrence on separate chromosomes, which depicts that each gene underwent duplication or a whole genome triplication event. The length of branches and their patterns also reveal the age and dynamic behavior of the genome. Environmental, physiological, and chromosomal shuffling brought about dynamicity or genome evolution during development. Phylogenetic reconstruction at the Brassica family level revealed three clades (groups) and five sub-clades (Figure 3, AB). It shows the conservation of the core regulatory module in *Brassicaceae*.

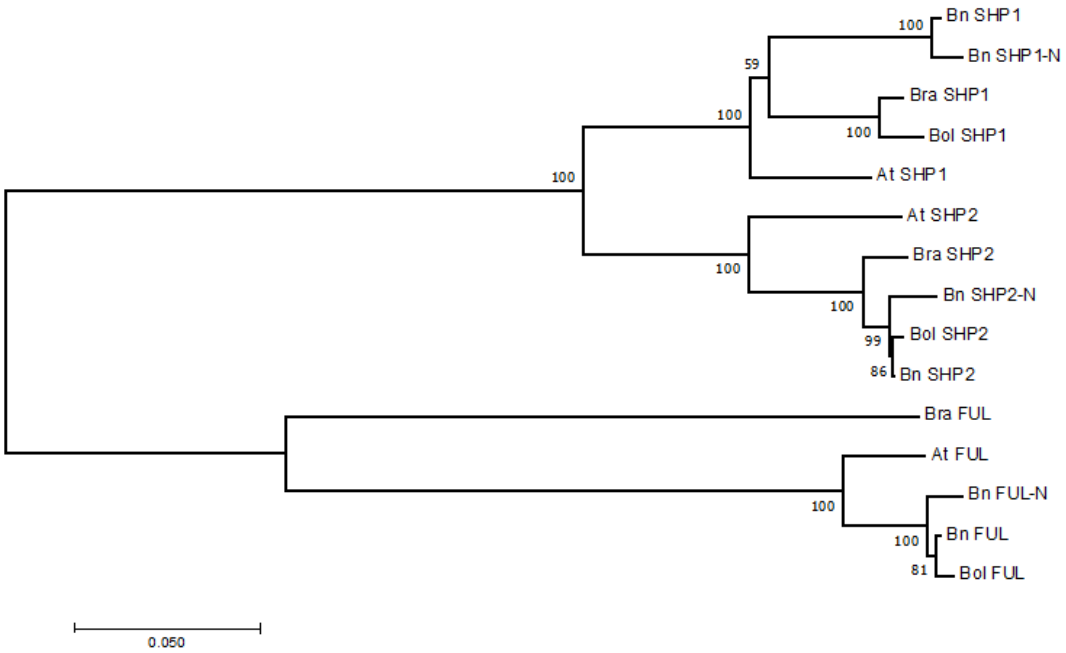

Figure 1. Phylogenetic analysis of core regulatory module genes with old and new ones, i.e., N stands for new. *BnSHP1-BnSHP1-N, BnSHP2-BnSHP2-N,* and *BnFUL-BnFUL-N* of closely related species. NJ algorithm was used. The values below nodes show bootstrapping of 1000 replications.

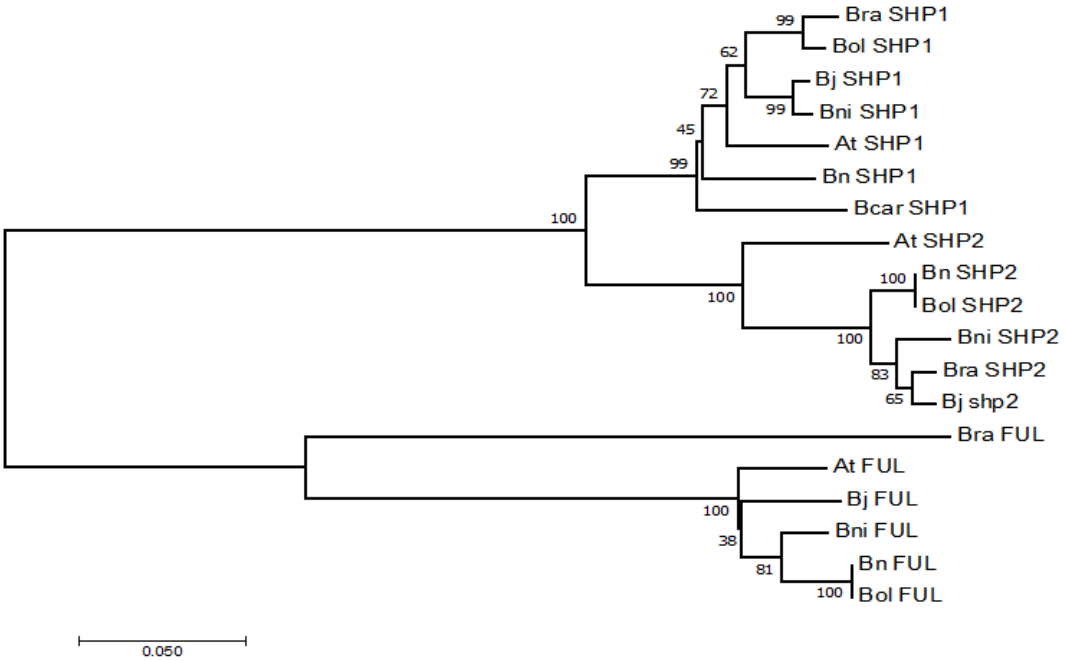

Figure 2. *SHP1, SHP2, and FUL* gene phylogenetic reconstructions in highly associated Brassica plants. MEGA6 software was used to rebuild a neighbor-joining tree with the default P uncorrected width. The values on the nodes represent 1000 bootstrap replications. A scale is also given at the bottom.

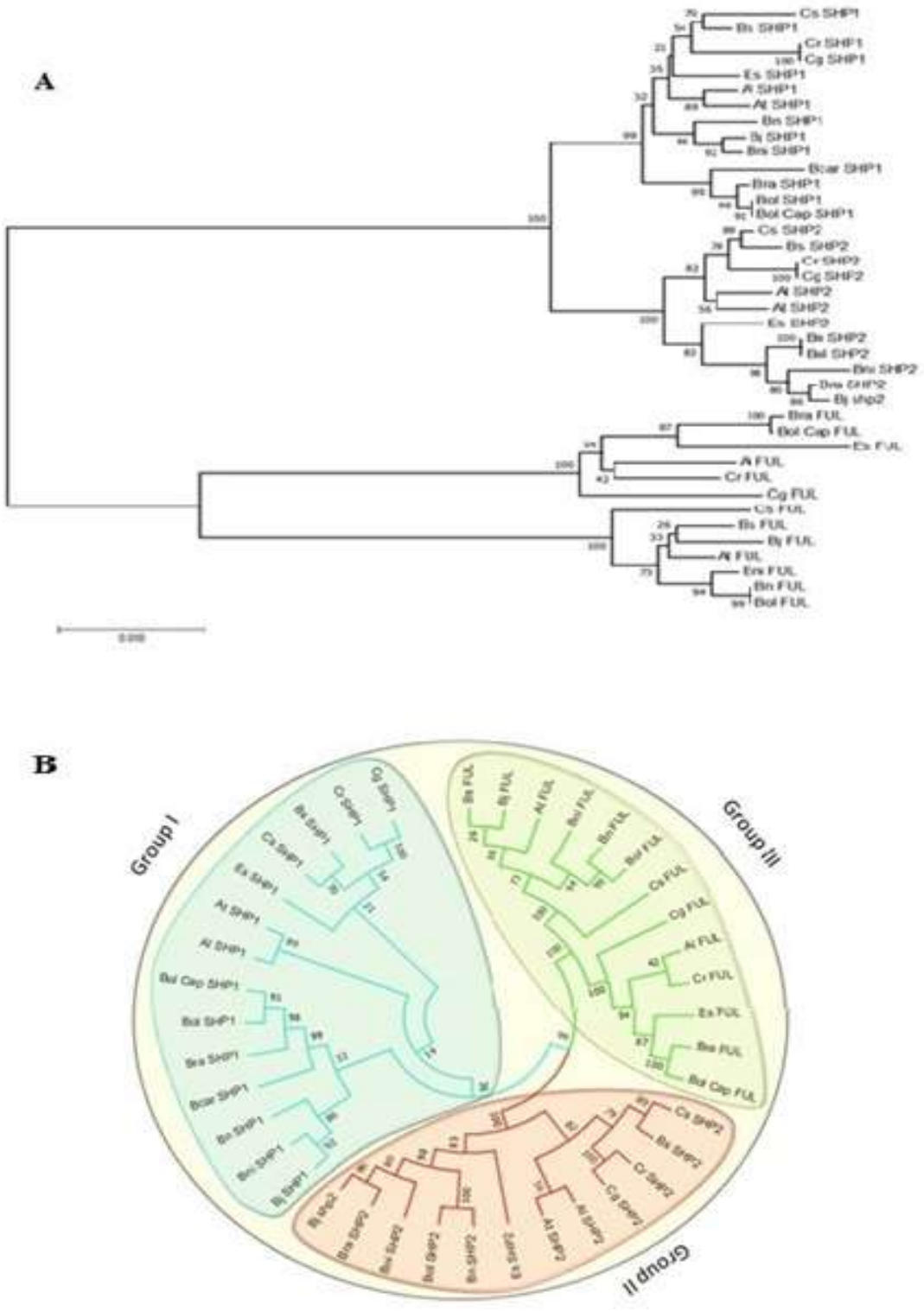

Figure 3. (AB) Phylogenetic reconstruction of *SHP1, SHP2,* and *FUL* genes in Brassicaceae family. MEGA6 software was used to build a neighbor-joining tree. The bootstrap replication of 1000 is designated by the values present on the nodes. Tree (A) and Circular tree (B). A scale is also given at the bottom.

Comparative synteny analysis of *SHPs-FUL* core regulatory module in the Brassicaceae family

The synteny diagram represents a remarkable relationship among these species in the context of evolution, duplication, triplication, function, and expression (Figure 4). The color intensity of inward tangling ribbons represents the conservation rate, and outward tangling of bands corresponds to duplication events in the Brassicaceae family. The syntenic circle plainly defines genomic dynamicity and evolutionary contribution via mobile elements in the genome of the Brassica family. These mobile elements generally play a vital role in the chromosomal rearrangements, duplication, and triplication. These blocks adopt a permanent position in the specific genome region and start expression, which disturbs another biological pathway.

Genomic loci of core regulatory module genes with variable intron-exon structures

The gene structures comprise CDS, upstream/downstream, and introns. Figure 5 details the variable gene structures of all three genes of the core regulatory module. *SHP1* has eight exons in *B. napus* but nine in *Arabidopsis*. *SHP2* harbors six exons in *B. napus* and *Arabidopsis* but six in *B. oleracea.* Unusually, the *FUL* gene exhibits eight conserved exons in all the *Brassica* species included in this study. *SHP2* exhibits the longest first coding exons. Variations in length of the gene structures represent exonization or intronization events during the evolution course.

Accretive data from phylogenetic reconstruction evidence indicate that *SHP1, SHP2*, and *FUL* sequences are all conserved. They have different gene structures and evolutionary associations with their homologs. Moreover, these genes show overlapping in distinctive tissues of canola, but divergent patterns are generally due to divergent promoter and *cis*-regulatory motifs.

As promoter elements underlie the genes' expression differences, the generation of nine combinations in pairwise plots appears in Figure 6 (A-I). The genes with 'N' are newly isolated from *B. napus.* The closeness of sequence similarity determines the direct relationship between a diagonal line and the presented curve and dots. The change or breakdown in the sequences of diagonal lines explains varying features, such as, frameshift, straight, and inverted repeats. The sequence conservation of up to -500 bp occurred, but beyond that, the sequences are divergent for all the genes in all the combinations. The *SHP2* gives significant results with sequence stability with a straight diagonal line, meaning the gene is stable and have more identical region. An observed low level of mutation emerged phylogenetically less dynamic. The sequence similarity diagonal of *SHP1* represents several *cis*-regulatory events (Figure 6, [A-I]).

Mutations, such as insertion, deletion, and inverted repeats, provide the basis for the gain and loss of functional site at a genomic level, which results in differential expression. The promoter sequences of the *FUL* gene reveal a lot of break points in the diagonal line, which means that this gene underwent many sequential changes. Phylogenetic of this gene clearly defines its dynamic nature, as already shown in the phylogenetic tree of *FUL*. This gene displays maximum number of duplication events. Break points and dots in the diagonal results of inverted repeat, and the gain and loss of many crucial sequence patches may be functionally important. Thus, the gain and loss of sequence patches might result in the differential gene expression in the context of evolution of core regulatory module in Brassica plants.

Transcription factor binding sites (TFBS) and motif analysis

The alignment of -2kb upstream promoter highlights mutually conserved motifs, which play a role in developing and bio-processes of shattering genes (Figure 7, [A&B]). There could be changes in their function with the change, mutation, or loss of residues of a significant contribution to the stability of specific motifs. Computational and bioinformatics tools detected regulatory

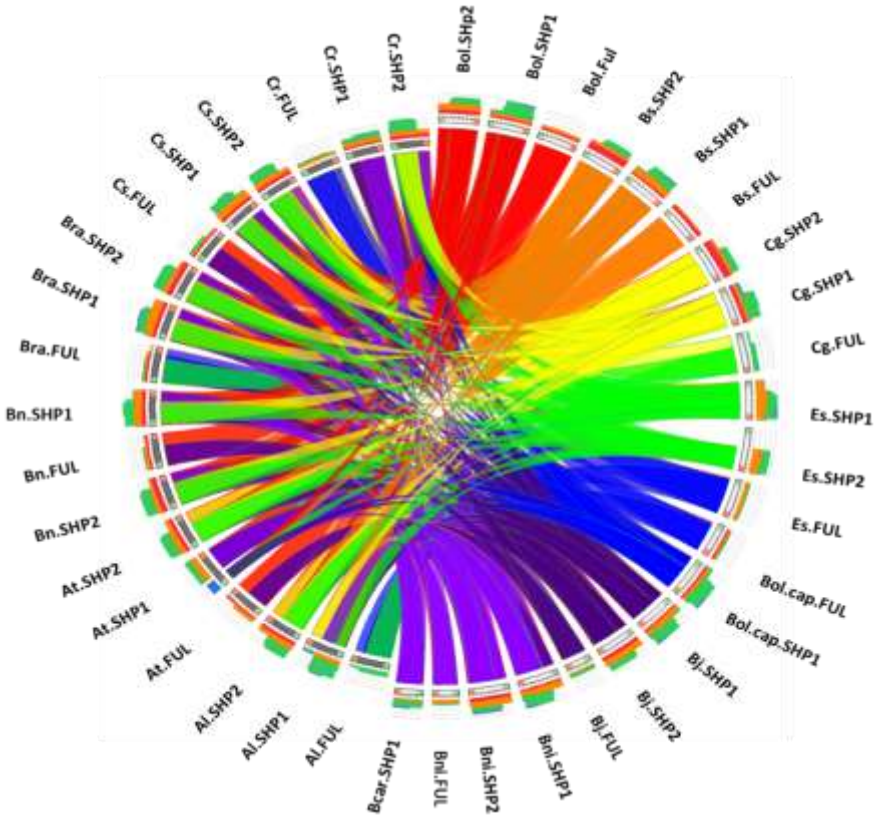

Figure 4. Representation of synteny diagram of Brassicaceae family. The comparative synteny of shattering genes from 14 species of Brassica family showing the level of conservation at sequence level in seven colors. The red, green, blue, yellow colors revealing level and intensity of evolutionary conservation among different shattering genes, e.g., maximum intensity is from red to yellow. The comparative synteny was performed using the Circos program (Krzywinski *et al*., 2009).

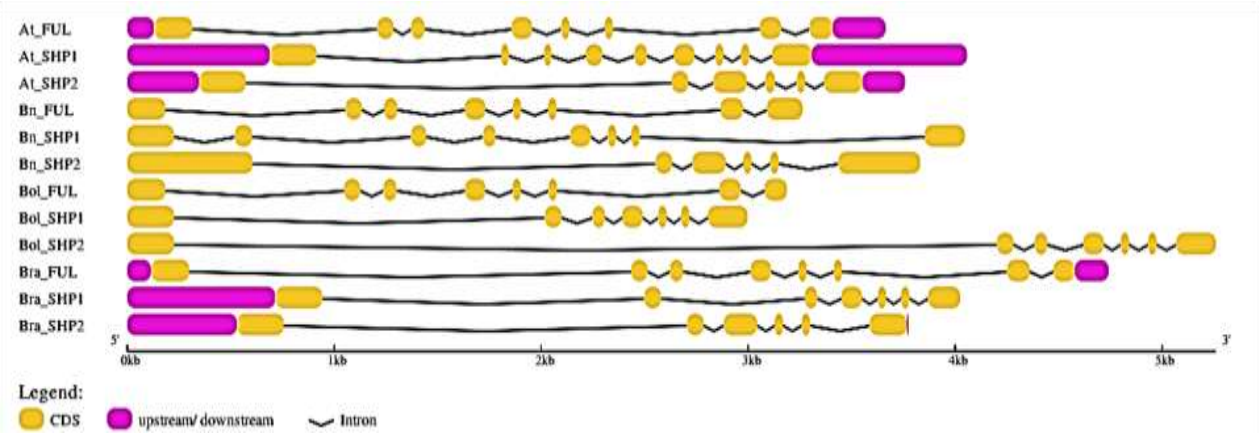

Figure 5. Comparison of gene structures of *SHP1, SHP2,* and *FUL* orthologs. Exons are represented by the loaded yellow boxes in gene structures, while introns are represented by the upward lines between the exonic boxes, which are purple in color. The numbers below in base pairs show the length of exons. At ends, the 5' and 3' UTRs are seen as blank boxes - core regulatory module genes harbor variable promoters elements.

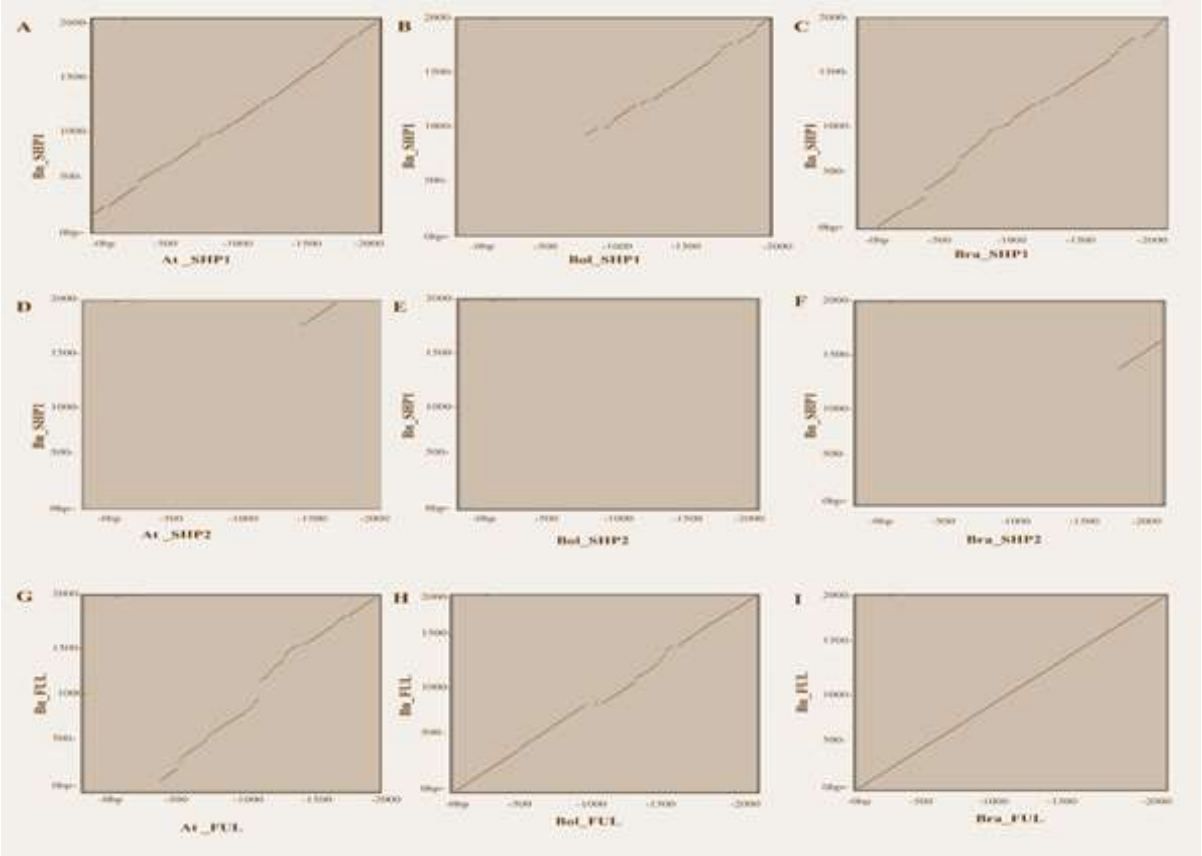

Figure 6. (A-I) Comparison of *SHP1*, *SHP2*, and *FUL* genes promoters of *B. napus*, *A. thaliana*, *B. olereacea*, and *B. rapa*. Pip maker (Schwartz *et al*., 2000) defined convergent and divergent sequences relation among Brassica promoters.

Figure 7. (A) Identification of conserved motifs in the promoter of *SHP1*, *SHP2,* and *FUL* in *B. napus* represented as motif 1, 2, and 3, respectively. The change in conserved motif (Bailey *et al*., 2015) is also represented with newly gained nucleotide that is functionally important. (B) Conservation pattern and substitution with nucleotide lead to change in functionality of coding gene.

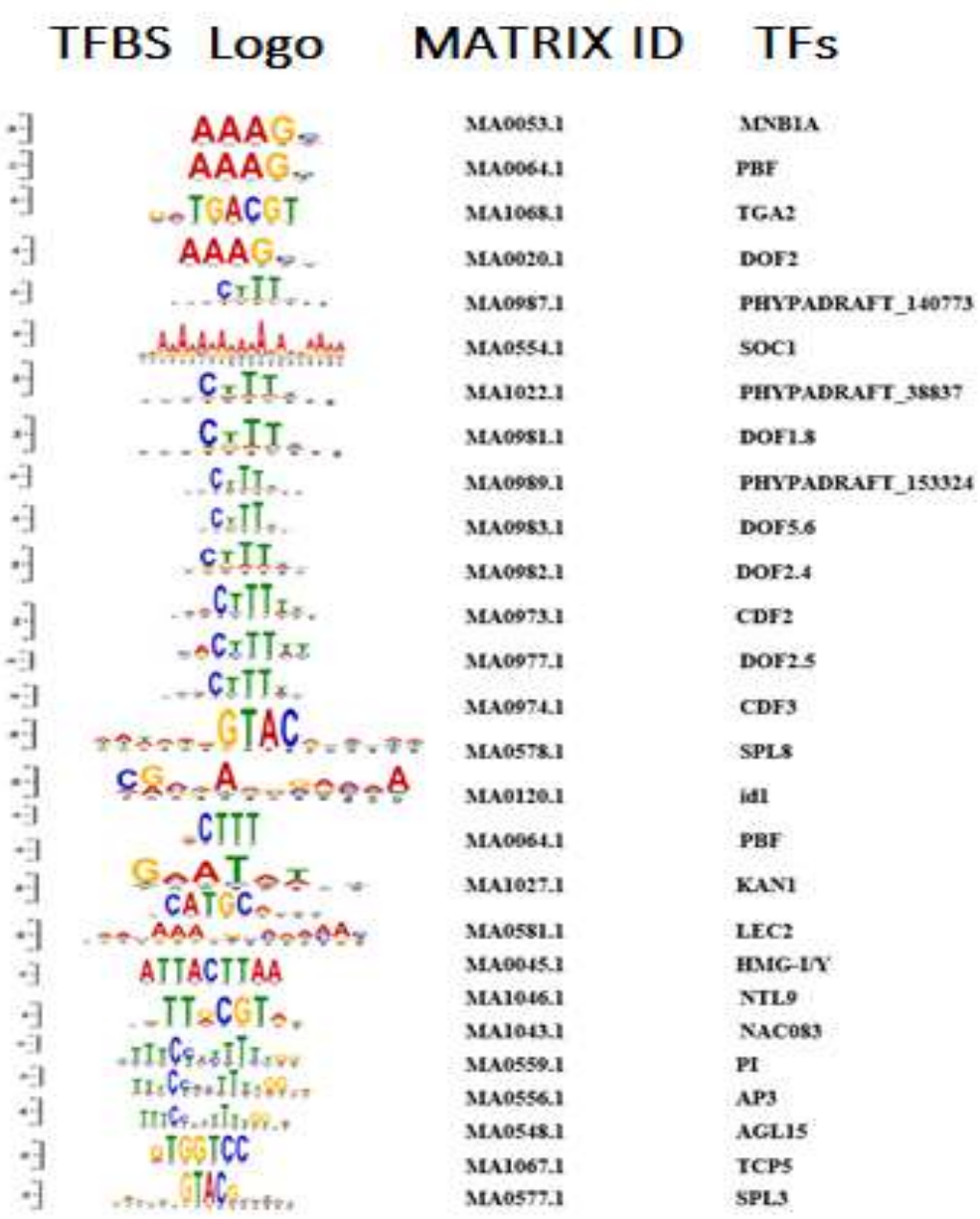

Figure 8. Identification of TFBSs in the promoter sequences of core regulatory module genes involved in shattering. Different transcription factor binding sites are present in the promoter sequences of *SHP1, SHP2,* and *FUL.*

elements involved in cellular differentiation management during ovule development, metabolome, proteomic regulation, and photosynthesis for seed formation in *B. napus.*

Searching for transcription factor binding sites (TFBS) analyzed the promoter regions -2kb upstream sequences using various online tools. A predicted 27 regulatory motifs have a role in the ovule, lipid degradation, valve differentiation, biological nutrients' synthesis, cell elongation, fruit dehiscence zone, and maturation of 72 genes and 11 species of Brassica plants (Figure 8). Each transcription factor (TF) can possess one or more regulatory binding sites.

All these 27 identified regulatory sites belong to different families of TFs implicated in various biological pathways. MADS-box conserved site contributed to the flower development, triglyceride biosynthetic pathway, and fatty acid synthesis. A *DOF* (DNA binding with One Finger) family of transcription factors participates in the developmental routes of seed and germination by regulating a variety of bioprocesses. Other TFBSs, including *SPL3, AGL15, AP3, PI, NAC, TCP, SOC, PHYDRAFT, TGA, PBF,* and *MNB,* belong to a versatile group of transcription factor families that regulate different pathways of germination to seed formation. The AGL15 is a MADS-box TF that CArG-boxes control. Although AGL15 expression is prominent during embryogenesis, it also shows in leaf primordia, shoot apical meristems, and young floral buds, indicating that it can play a role in post-germinative growth. These motifs bind the promoters and other regulatory elements and regulate the expression in flower development and other processes. Each transcription factor manages gene expression during its specific stage in plant development. Some of them influence seed maturation and nourishment, while others are in the cellular process occurring within the seed. Immunity is another critical trait that these TFs plants can maintain during fruit development.

DISCUSSION

Pod shattering is a severe problem of canola that leads to substantial yield loss. Pod shattering is a complex physiological and biological mechanism through which seed disperses effects as waste and yield decline (Dong and Wang, 2015). Shattering of pods starts when the connection between walls becomes weak and internal forces exert them to the moveable situation (Raman *et al.*, 2014). These internal forces and attachments are due to the seed valve that provides necessary protection to seeds (Balanzà *et al*., 2016). The seed of *B. napus* and *B. juncea* have value as they contribute 14% of the global vegetable oil. Different nutrients and the rapeseed's biological molecules contribute to shattering development (Hu *et al.,* 2015). Previous investigation on shattering revealed that it occurred because of molecular components' excess production and enrichment in the valve margin and cellular portion around the pods in siliques. Lignin and cellulose play essential roles in the hardening of pod walls, which lower the water contents during the latter stage of development of *Brassica* species and rapeseed (Schiessl *et al.*, 2017).

The MADS-box transcription factor and genes control the shattering process, including *SHATTERPROOF1-2* (*SHP1/2*), *FRUITFULL* (*FUL*), *INDEHISCENT* (*IND*), *ALCATRAZ* (*ALC*), *REPLUMLESS* (*RPL*), *NAC SECONDARY WALL THICKENING PROMOTING FACOTR1* (*NST*), *ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1* (*ADPG1/2*), *GA3ox1* (Zumajo-Cardona *et al*., 2017), *PISTILATA (PSL)* (Hecht, 2016), and other genes of different classes characterized in *Arabidopsis thaliana.* Various transcription factors binding sites are functionally and structurally crucial for shattering development (Cosio and Dunand, 2010).

Promoter sequences, in general, regulate the expression of genes. *BnSHP1,2* redundant expressions also suggest that their promoter sequences are strongly conserved. The *SHP-FUL* module is thought to have remained conserved in dry fruit populations, and genetic interactions between *SHP* and *FUL* may have developed before the split of rosids and asterids (Dong and Wang, 2015).

Phylogenetic analysis revealed that *SHATTERPROOF1 (SHP1)* has less dynamicity and is stable in the context of genomics but underwent several duplications. Many duplicated genes correlate with a distinct number of chromosomes in rapeseed, suggesting their genomic mobility. Analysis of *SHATTERPROOF2 (SHP2)* provided the same clue to phylogeny analysis but has more level of duplication events undergone by a whole genome triplication (Cheng *et al*., 2015) and duplications (Smaczniak *et al*., 2012). *FRUITFULL (FUL)* is a well-known fruit development MADS-box transcription factor that regulates fruit development in *Brassica* species. The phylogenomic of *FUL* provides remarkably different insights than *SHP1,2.*

Several *FUL* genes show duplication and differential location in the genome (Hu *et al.,* 2015).

The application of comparative and bioinformatics approaches reckoned the genomic conservation for the functional and evolutionary relationship among *SHP1, SHP2,* and *FUL.* Following the earlier investigations to fulfill this purpose, dot plot analysis of 2kb promoter sequences faced scrutiny for existing similar putative functional relevance among shattering genes in *B. napus.* The plot revealed and validated the early results performed in the model plant *Arabidopsis thaliana,* e.g*., SHP1, SHP2,* and *FUL,* which have functional and genetic conservation within them (Hradilová *et al*., 2017). Many observed vibrant evolutionary behaviors resulted in *SHP2* and *FUL* genes in *B. napus*. The conservation patterns among the sequences of 14 *Brassica* species were noticeable. The arrangement of blocks identified among these species came from the comparative synteny approach, according to Krzywinski *et al*. (2009). These blocks indicate the root relationships among them. Numerous conservation and syntenic patterns observed revealed they have functional relevance as they possess conserved sequences. Each species shows duplication events with the intensity of different colors.

Other features of *SHP1, SHP2,* and *FUL* are the gain and loss of putative functional transcription sites. The transcriptional factor binding sites (TFBSs) play a role in initiating a function by transcriptional activation of a start site through the regulation of CArG box, an auto-regulatory domain of conserved sequences in the 2kb promoter region of shattering genes (Pabón-Mora *et al*., 2014). The role of the CArG box is an auto-regulation of MADS-box transcription factors by producing different looping signals that further precede the biological pathways. Different TFBSs identified have an individual role in natural and biochemical pathways, such as, the DOF family of TFBS, often involved in developing paths of ovule and flower (Seymour *et al.*, 2013). AP3 participates in the floral organ identity. SPL has a role in sporogenesis. Huang *et al*. (2019) suggested the role of AGL in consensus binding to DNA, and Hepworth *et al*. (2002) discovered the role of SOC1 in regulating flowering time in Brassica plants.

All these TFBSs reside in the promoter region of *B. napus* (Biewers, 2014). The differential gene expression and shattering can result in the loss and gain of such putative sites during the evolutionary process. The original function of TFBSs retained in the lineage is through duplication events with positive evolution and dynamically positive change. The shattering process starts in the dehiscence zone, with the wall elongation controlled by the *FUL* gene by inhibiting the *SHP1* and *SHP2* expressions (Liljegren *et al*., 2004). *SHP1/2* and *FUL* are the MADS-box TFs family, which underwent numerous duplication events during lineage of evolution (Guo *et al*., 2017).

These genes also have a relationship with the downstream basic helix loop, which acts on abolishing the DZ action to inhibit dehiscence during seed formation, resulting in indehiscence in the existing multiple regulatory genes. The *FRUITFULL (FUL)* gene in *B. napus* and *B. juncea* and in *A. thaliana* differentiates the valve margins and carpel during flower development and become restricted after few stages of carpel development (Ye *et al*., 2016).

CONCLUSIONS

In conclusion, the study alludes that three new sequences, namely, *BnSHP1-N*, *BnSHP2-N,* and *BnFUL-N* genes, occurred in the genomes of local cultivars of *Brassica*. The evolution of the *SHP-FUL* module in *B. napus* paralogs recruited in shattering has attained effects from the gain and loss of TFBSs and other *cis*-regulatory components. Since pod dehiscence is a prime target trait for improving canola yield; therefore, genome analyses of *SHP-FUL* genes may give essential information in canola research to benefit plant breeding for improving crop productivity.

ACKNOWLEDGMENTS

This research received support and acknowledgment from Mr. Mubashir Ahmed Khan (PSO Oilseed Program) at NARC. The authors sincerely appreciate and thank the Institution for providing the plant materials.

REFERENCES

Bailey TL, Jhonson J, Grant CE, Noble WS (2015). The MEME suit. *Nucleic Acids Res*. 43 (W1): W39-W49.

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. *Devel.* 143(18): 3372-3381.

Cheng F, Wu J, Liang J, Wang X (2015). Genome triplication drove the diversification of Brassica plants. *Hortic. Res*. 115-120.

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: 1-16.

Dong Y, Wang YZ (2015). Seed shattering: From models to crops. *Front. Plant Sci.* 6: 476.

Economic Survey of Pakistan. (2020). Government of Pakistan. 2019-2020. Ministry of Finance, Economic Advisor's Wing, Islamabad, Pakistan. pp: 25-26. (http://www.finance. gov.pk/survey_1920.html).

Guo X, Chen G, Naeem M, Yu X, Tang B, Li A, Hu Z (2017). The MADS-box gene SlMBP11 regulates plant architecture and affects reproductive development in tomato plants. *Plant. Sci*. 258: 90-101.

Hecht V (2016). Duplicate MADS genes with split roles. *J. Exp. Bot.* 67(6): 1609-1611.

Hepworth SH, Valverde F, Ravenscroft D, Mouradov A, Coupland G (2002). Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. *EMBO. J.* 21(16): 4327-4337.

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: 542.

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. *J. Plant. Res*. 128(5): 813-819.

Huang J, Wang C, Wang H, Lu P, Zheng B, Ma H, Copenhaver GP, Wang Y (2019). Meiocyte-specific and AtSPO11-1-dependent small RNAs and their association with meiotic gene expression and recombination. *Plant. Cell*. 31: 444-464.

Kadkol GP (2009). Brassica shatter-resistance research update. In: Proceedings of the 16th Australian research assembly on brassicas conference. *Ballarat. Victoria*, pp. 104-109.

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.

Li X, Zhang S, Bai J, He Y (2016). Tuning growth cycles of Brassica crops via natural antisense transcripts of BrFLC. *Plant. Biotechnol. J.* 14(3): 905-914.

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000). SHATTERPROOF MADS box genes control seed dispersal in Arabidopsis. *Nature.* 404: 766-770.

Liljegren SJ, Roeder AH, Kempin SA, Gremski K, Østergaard L, Guimil S, Reyes DK, Yanofsky MF (2004). Control of fruit patterning in Arabidopsis by INDEHISCENT. *Cell*. 116: 843-853.

Meakin PJ, Roberts JA (1990). Dehiscence of fruit in oilseed rape. II. The role of cell wall degrading enzymes. *J. Exp. Bot.* 41: 1003-1011.

Ng M, Yanofsky MF (2001). Function and evolution of the plant MADS-box gene family. *Nat. Rev. Genet*. 2(3): 186-195.

Pabón-Mora N, Wong GKS, Ambrose BA (2014). Evolution of fruit development genes in flowering plants*. Front. Plant Sci*. 5: 300.

Price JS, Hobson RN, Neale MA, Bruce DM (1996). Seed losses in commercial harvesting of oilseed rape. Silsoe Research Institute, Wrest Park, Silsoe, Bredford MK45 4HS, UK. *J. agric. Engng. Res*. 65: 183-191.

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.

Raman R, Raman H, Kadkol GP, Coombes N, Taylor B, Lucket TD (2011). Genome-wide association analyses of loci for shatter resistance in Brassicas. Proceedings of the 11th Australian research assembly on brassicas (ARAB) conference, Wagga Wagga, NSW, pp. 36-41.

Roeder AH, Ferrándiz KC, Yanofsky MF (2003). The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. *Curr. Biol*. 13: 1630-1635.

Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R (2017). Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci. Rep.* 7(1): 41845.

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Miller W (2000). PipMaker. A Web Server for Aligning Two Genomic DNA Sequences. *Genome. Res*. 10: 577-586. (http://pipmaker.bx.psu.edu/pipmaker/).

Seymour GB, Østergaard L, Chapman NH, Knapp S, Martin C (2013). Fruit development and ripening. *Annu. Rev. Plant Biol*. 64: 219-241.

Smaczniak C, Immink RG, Angenent GC, Kaufmann K (2012). Developmental and evolutionary diversity of plantMADS-domain factors: insights from recent studies. *Develo.* 139(17): 3081-3098.

USDA Foreign Agriculture Service (2015). Production, supply and distribution online, Table 03: Major Vegetable Oils: World Supply and Distribution (Commodity View), http://apps.fas.usda.gov/psdonline/psdHome.aspx, accessed January 11 (2016).

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 homoeolog gene expression influencing selection. *Nat. Genet.* 48(10): 1225-1232.

Ye L, Wang B, Zhang W, Shan H, Kong H (2016). Gains and losses of cis-regulatory elements led to divergence of the Arabidopsis APETALA1 and CAULIFLOWER duplicate genes in the time, space and level of expression and regulation of one paralog by the other. *Plant. Physiol.* 171(2): 1055-1069.

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). *Evo. Devo*. 8: 1-18.